User Interfaces Design for CVE Software

M.K. Othman
Faculty of Cognitive Sciences and Human Development (FCSHD), Universiti Malaysia Sarawak (UNIMAS), Sarawak, Malaysia

R.J. Lapeer
School of Computing Sciences (CMP)
University of East Anglia (UEA)
Norwich, United Kingdom

Abstract - The project aims to study and design an alternative user interface for Collaborative Virtual Environments (CVE’s) software also known as Networked Virtual Environments (NVE’s). To reduce cost, most current and operative CVE’s use the Internet and standard PC to create a visual Virtual Environment (VE), which can be shared by a large number of users. This project also involves an image processing technique (morphing technique using Thin Plate Spline) for creating a facial expression for the CVE software and involves OpenGL API for implementation. This project discusses communication aspects in the CVE system and suggests the different types of communication that are suitable for the project. It also suggests a suitable user interfaces for the software.

I. INTRODUCTION

The objectives of this research are to study and evaluate the Collaborative Virtual Environment (CVE) also known as the Networked Virtual Environments (NVE) software, and suggest a good user interface for the software. It includes a software analysis and evaluation, updating the software if and when necessary or possible. This research also involves an image processing technique (morphing technique using Thin Plate Spline) for creating a facial expression for the CVE software. In this project, user interface not only refers to the graphical interface (Graphical User Interface (GUI)) between the user and the system, but also includes other interfaces for communication such as aura communication, voice or video interface communication, facial expression, and also standard tools used for communication (e.g. keyboard, mouse, etc.). The CVE software in this report refers to the software designs by Oliver Wennel at School of Computing Sciences, University of East Anglia United Kingdom.

II. PROJECT DESIGN

Project design and implementation was divided into a few phases, including current software evaluation, the development and implementation (updating the current software) phase, and the final phase of testing and evaluation. The software evaluation phases were conducted to determine the weakness of the current system and how to improve it. After some recommendations have been made on how to upgrade the current CVE’s system, the storyboards for the system will be built based on the project scenario. Project scenario was designed based on some requirements for the update version of CVE’s software. The development and implementation phase was divided into three main components. The first component is a facial expression development. The second component of the development and implementation phase is the communication protocol aspects between multiple computers within the same network. The communication protocol is called the “handshake protocol”.

III. PROJECT IMPLEMENTATION

A. How to Improve the Current CVE System

The success of CVEs is not only determined by a number of users using the system or a good technical aspect of the system, but also by a good interface for communication between the multiple users in the VE. In order to achieve this goal, new architectures of the CVEs must be developed to provide a new module to be integrated into the CVEs, and a new interface for communication must be introduced to provide functionality and to exploit the main characteristics of CVEs as much as possible.

1) Multimodal Communication

Designing a good interface for multimodal interaction in CVEs is not an easy task. Many options must be considered, such as how reliable the system could be, how to design such system that will suit all users (not only a majority user, but the minority as well), was it worth designing such a system, or how bad the system could be if designed in such way. A system is considered as multimodal if it meets or supports human modalities such as gestures, written or spoken language [2].

Natural ways of human communication are not only limited to verbal communication (speech) [3]. In real life, when communication takes place between two or more individuals, it does not only involve verbal communication. It also involves non-verbal communication such as eye-gaze, body postures and language, and facial expressions [3]. Haptic and force feedback [4], lip movement and facial expressions [5]. When talking to each other, people tend to convey body gestures and postures to make it more realistic.
This improves the communication realism, as it will show the user’s emotional state of mind. This type of natural human communication (verbal and non-verbal) should be mapped into the CVE’s to enhance the interaction and communication between users-users or users-objects within the system. Non-verbal communication not only involves the communication between users but, also between users and objects, or between user and the environment [6]. To date, there is substantial research that focuses on the important aspect of non-verbal communication; see [3], [7-11]. Most of the research focuses on developing a model of rich interaction for CVE systems.

Research shows that communication between multiple users in a (CVE) must be a multimodal interaction [2], which includes gestures communication, facial expressions and lip movement [5], audio communication [12], body language, postures, facial expressions, eye-gaze and text messaging [3].

Audio communication (voice-to-voice) could be used within the CVE using an Audio link [12]. This feature needs a user to have a microphone attached to the system. The microphone is used to capture audio data and it is compressed using a standard compression code before it can be sent to the server and then distributed to other users within the network.

When designing a user interface for communication, the designer should not ignore minority users such as people affected by deafness. The user interface must be user friendly and allow sign language as a mean of communication between users in the CVE’s. The sign language could be used with the aid of avatars or using real time video integrated with the environment. Gestures communication could be used as a means of communication for people affected by deafness using rule-based sign language to define some actions [5], [13].

2) Avatar Communication

Human representation in the CVE system is represented by an embodiment called an avatar. This embodiment could also be represented by a simple graphical representation (e.g., a textured cube). An avatar is an advanced feature of a human embodiment in a virtual environment and could support facial and gestures communication [5], [14-16].

An avatar is an incarnation of the user within a virtual world. As a human representation within a VE, the avatars can meet and interact with each other in an intuitive way and could sense a presence of each other in the VE. Sense of presence is an important aspect of a CVE system [17-19]. Sense of presence is more important in a multi-user environment to ensure effective interaction among users [20-21].

Communication within a CVE using an avatar requires an efficient and realistic embodment [18], [22]. In order to further improve realism using avatar communication, human behavior could be integrated into the avatar. This could be achieved by:

- Mapping the user’s face picture to the avatar could give an identity to the avatar. Using the user’s own picture will solve the problem of few people using the same avatar. This option does not only improve the current CVE system, which only use the avatar’s name appeared on the top of the avatar head, but also gives a more presentable avatar’s appearance.

- Enabling live facial expressions (live video) to be texture mapped onto the avatar [5], [23]. This could be done using a real time video and texturized map onto the avatar face. This technique requires each user to have a video camera attached to the CVE system. Conventional live videoconferencing is not suitable to be used within the CVE’s because it requires a separate window, which will again lead to divided attention for users. The importance of live facial expression used in a CVE system is that it will improve the realism of communication and interaction through the avatar because it will show the current emotion of the user.

- Apart from using live texture maps of facial expressions, predefined facial expressions could be used to improve the realism of communication and interaction for avatars. The easiest way to implement facial expressions is by using predefined facial expressions. Users simply type the keyboard smiley (e.g., :) , :3, :3, L, etc) to show their expression to others. This predefined facial expression was stored in the CVE’s module engine and could be retrieved when needed. This option is not as good as live expression because users need to remember keys on the keyboards that represent each expression. In some occasions, predefined facial expressions cannot give the exact user facial expression because when the user is talking, they tend to make other kinds of action at the same time such as rubbing the chin.

Little research has been carried out to date to show how important facial expression is as a means of communication within CVE’s because it is useful for representing intentions, thoughts, and feeling [24]. The research shows that facial expression improves communication in a CVE system [3], [24-25]. Methods that could be used to integrate facial expressions in a CVE include video texturing method by continuously texture mapping the user face to an avatar, which needs a user to sit in front of the camera at a particular angle to ensure that the camera could
The aura concept is based on the spatial model [29], communication that could be used within the CVE's. Aura communication is one kind of interaction and spatial awareness. Visual, textual and social communication as well as support various communication types such as aura, the collaborations are effective, the systems should time regardless of their locations. In order to ensure communication between multiple users at the same time, several considerations for the CVE system need to be made. In this project, we will take a look at two: user interaction and spatial awareness.

Facial expressions technique could also be applied using a real-time facial animation system [26]. This technique allows real-time user facial expressions to be used in the CVE system. This technique uses MPEG-4 Facial Animation Parameters (FAP) to extract real time facial expression from the video and send it to the virtual environment via the Internet.

A good avatar does not necessarily require details of facial expression or be photo realistic of human anatomy because it is hard to model the avatar to be like a human. Research has been conducted to see how this affects the communication in CVE. It shows that an attempt to reproduce an avatar in detail physically similar to humans is uneconomical and wasteful [27]. Another study by Hindmarsh et al. [28] shows that direct translations of real human to virtual environment was unsuccessful to some extent. Studies show that a minimalist approach of creating human embodiment in CVE helps to represent humans in a natural way such as using facial expression. It is more understandable that facial expression could provide a good way of non-verbal communication and could be supported by body gestures, eye gaze, etc.

3) User Interaction/Collision Detection

User interaction in a CVE system includes the interaction between user with user, user with the environment or user with the object. The key of this interaction is collision detection between user with user or users with objects. Collision detection is a program script that determines how close a user is to an object and stops their movement when they collide with the object [33], or a more complex response (collision response).

4) Aura Communication

Collaboration in a CVE system involves communication between multiple users at the same time regardless of their locations. In order to ensure the collaborations are effective, the systems should support various communication types such as aura, visual, textual and social communication as well as spatial awareness.

Aura communication is one kind of interaction and communication that could be used within the CVE’s. The aura concept is based on the spatial model [29], e.g. in Distributed Interactive Virtual Environment (DIVE) system [30] or Model, Architecture and System for Spatial Interaction in Virtual Environments (MASSIVE) [31]. Aura is an area or a region surrounding the user’s avatar [29], [32]. If the user’s aura intersects or collides with another user’s aura, they could sense the presence of other people within the environment. Users not only could sense the presence of other users or be aware of other users around, but they could start a communication between each other.

5) Multiple Views

When collaboration work takes place in the environment, the user not can only see from his view/angle, but also can see what other users are seeing. This is helpful when collaboration work involves a changing of a particular part of the scene and needs other users to comment on the changes. Research shows that multiple perspectives viewing display are useful for navigational purposes [22], [34].

6) Human Factors Issues

The human factor issues which VE researchers are interested in could be divided into three main categories: health and safety, human performance and efficiency, and social implication [35]. The design of the CVE interface system should maximize the efficiency of human task performance within the environment. The problem with the current system is that there is no standard guideline to design an interactive interface and this defeats the purpose of the CVE system. The goal of designing such interfaces (user friendly) is to minimize learning time for the user and maximize collaboration between users.

B. Facial Expression Implementation

In this project, avatar facial expression is a main component and was implemented using the warping technique (Thin Plate Spline). The facial expression was created offline using this technique and mapped into the avatar’s face when needed. Furthermore, the facial expression implementation created is in 2D mode. In order to make a desired predefined facial expression, two images of the avatar were needed. The first image shows the image before a facial expression had been applied and the second image shows the image after it had been altered with a facial expression.

The warping technique code used in this project enables the creation of facial expressions needed, using landmarks on the avatar’s face. The picture in Fig. 1 shows how a landmark (the red dot on the face) was applied to an original image (left image) and it changed the original image to a desired image (right image).
Corresponding landmarks on both images are required. If too many landmarks are used, the avatar’s face will be squashed and will not look natural, which will destroy the purpose of creating a natural facial expression. Moreover, the locations to place a landmark also needs to be determined correctly; wrong placements could more or less destroy the original image affinity.

In this project, only six basic facial expressions were created and used in the chat system; this includes facial expressions for smiling, being bored, laughing, being surprised, being frustrated and sad. Fig. 2 shows all the facial expressions that have been used in this project. The reason for using only six basic facial expressions in the project is because it’s not easy to create a complex facial expression. A complex facial expression requires more complex changes of facial structure and this is not an easy task using this warping technique, as it would completely change the face features.

The creation of a facial expression needs humanoid facial features, which means that only selected avatars will have facial expressions because not all avatars have complete features to create good facial expressions.

The system only enables six basic facial expressions and this could be improved by creating more facial expressions. The use of facial expressions will help the user to communicate better in a VE. Other techniques for creating facial expressions could also be used to create more facial expressions (e.g. cry, shame, blush, etc.) and could probably produce a more realistic facial expression.

A more realistic facial expression could be created by using live facial expressions captured with a camera where a live human face can be mapped to the face of a human avatar. Another option that could be considered when using live facial expressions is using the video where the user’s facial expression can be directly displayed on the facial expression window. This option enables users to see real facial expressions of other users not the expressions on the avatar.

Communication using live facial expressions will improve the communication between users in the CVE system. As we know, this option is not always a good option as it also has the weakness when it is implemented to the CVE system. The main problem when using this option is that there could be problems with the network, as it requires a sufficient bandwidth between users.
Text-based communication in a 3D environment, especially in a shared or CVEs, is not realistic because there are other kinds of interaction or communication (e.g. voice, face expressions, gestures communication, avatar, aura, etc) that could be used in 3D as a means of interactive and functional communication. Voice communication could solve the divided attention problem with text-based communication. This is because the user could talk and at the same time could focus on what other users are doing in the environment.

V. CONCLUSIONS AND FUTURE WORK

This research has been carried out to design a user interface for Collaborative Virtual Environment (CVE) software and some components have been implemented to the existing software. If the desired user interface that has been suggested to improve the existing software were successfully implemented, this will produce a very good piece of work in the CVE field. Future research work could also be extended to improve the facial expression techniques by increasing the number of facial expressions used within the system. The future work must also not limit the communication between avatars in the CVE by the facial expressions, but must also consider other options such as audio, gestures and body postures, etc.

The impact of the social interaction within CVE should be studied in detail in the future to improve the interaction between users in the system. This includes the impact of such interfaces to the users and it is important to know what are the specific tools needed to make a CVE which will be valuable in different research disciplines and applications. Social interaction using aura must be taken into consideration when improving current system or when designing a new CVE system.

REFERENCES


