Presented at WOFIEE'13, Athens, Greece, 16-17th July 2013

Improving Communication and Presence in Online Telepresence Systems

Alfonso TORREJON, Prof. Vic CALLAGHAN and Prof. Hani HAGRAS

Intelligent Environments Group, Department of Computer Science and Electronic Engineering, University of Essex, UK

Abstract. This describes research into a mechanism that enables geographically dispersed participants in telepresence system to be placed in artificial views that better mimics a real life lecture or meeting room situation. In support of this we present a novel solution for collaborative activity that converts participant's into Mixed Reality's visual objects allowing them to interact more naturally from any location, enhancing their communication.
The system enables people to logon, choose a seat or location (which also assigns them a virtual presence in that location) thereby enabling them to talk and interact, with their virtual neighbors. The sound volume can be controlled so as to produce a kind of virtual bubble that surrounds the participants allowing the audio to be directional and selective. In addition, attendees to the real environment can interact with the remote attendees.

Keywords. Directional audio, telepresence, immersive presence, MR directional audio, panoramic audio, immersive reality.

Introduction

Augmented Reality (AR) is a technology that enables users to interact and combine 3D virtual objects with the physical world in real-time applications. Until now AR technology has mostly been used to overlay virtual objects on physical scenes to in a way that augments the information value of the scenes. Most AR system used in education are used to visualize objects and processes that are invisible to the naked eye, s such as molecules, animals, body parts and similar items. By way of an example, Cofin et-al [1] proposed a system that addressed two main goals: the use of Augmented Reality for virtual annotations and illustrations which allowed students to interact with objects fellow students and the lecturer over great distances.

Augmented Reality and other Virtual Reality work largely focuses on the visual domain, with the audio side being given little or even no consideration at all. It is not always clear why that is but one reason might be that it was regarded as less important. The communication between objects and viewers is normally either via controlling an animation or auxiliary information such as text boxes or other overlays.

Some VR environments have tried to integrate audio communication, perhaps the most well regarding being Open Wonderland ®, where it is one of their main competitive advantages. In their system they use directional audio with privacy spheres. However, the system is avatar based rather than video and so the sense of realism

suffers as a consequence. Our work goes beyond this by, in addition to using real video, providing the user with more control over the audio environment.

In this paper we present a similar strategy to that used in AR, but instead of using graphical objects, we use video of people in order to provide a more realistic mixed reality experience for participants. This type of mixed reality is often referred to as ***Immersive Reality***.

The motivation for our thread of work can be traced back to Billinghurst and Kato's 1991 paper [2] which pointed out the merit of "beyond being there" which mostly related to earlier Computer Supported Collaborative Work.

All setups used for telepresence have used similar methods to provide remote presence, some with more or less success. The main intention always has been to achieve feeling that is indistinguishable between attending locally or remotely.

When people communicate in a face-to-face conversation in the real world, there is a dynamic and easy interchange of information between the listener and the speaker's interpersonal space and a sense of sharing valuable information related to their personal goals. Equally the value of the information shared between individuals is related to the distance between them. People who "feel" closer together will share more intimate information.

It has been demonstrated that search interests of people who already connected via a social network, are more similar than random individuals. Moreover it was found that the longer they were sharing messages the closer they were [5]. Our system supports these ends by providing virtual bonds between people that are geographically dispersed. In our system those bonds are derived from their common goals, shared knowledge and desire for success. In addition, technical affordances such as synchronous communication which gives the opportunity for more intimacy, leading to a better understanding of each other, providing the opportunity to balance the groups weaknesses in order to achieve their goals.

1. Social Interaction and Social Presence

What makes students engage in an online virtual learning? Beyond the chosen subject itself, there is a social view that emerges in support of online virtual learning that argues the individual joins a community group and which gives rise to the potential for building and maintaining that sense in virtual classrooms [6][7].

There are 3 interrelated elements to be considered: ***Social presence***, ***teaching presence*** and ***cognitive presence***.

The CoI framework (Community of Inquiry) is a process model of online learning [8]. It assumes that effective online learning, especially higher order learning, requires the development of community, and that such development is not a trivial challenge in the online environment.

Figure 1. Community of Inquiry

1.1. Social Presence

Social presence is defined as the degree to which the student feel socially and emotionally connected to others in the group, projecting themselves as "real" people, independent of the communication medium used [15]. Socio-cognitive theorists describe learning as an interactive group process in which learners actively construct knowledge and then build upon that knowledge through the exchange of ideas with others [9]. If the students are distant in traditional distance learning scenario will be condemned to isolation and loneliness of their own space without any possibility of synchronously exchanging ideas, and missing information beyond words and text that sometime are milestone in the learning process.

Studies on social presence with a large set of participants [11] determined that social presence was composed of three subjective elements: co-presence, intimacy, and immediacy, none of them naturally available at distance learning, so the need to build a model that takes those into consideration. In this 'value social interaction' and 'learning interaction 'are the key factors needed to achieve learner satisfaction and goal achievement, which is basically the final goal of any teaching institution.

1.2. Cognitive Presence

Cognitive presence describes the extent to which the student is able to construct meaning through sustained communication, reflection and discourse.

Poor critical thinking has been found in some online discussions between students [10]. Results indicate that participants of some courses did not value the forum which was reflected in the quality of posts submitted. One of the reasons found is that, rather than critically analyzing each other's posted comments, they tend to discuss the student-content interaction, failing to engage one another's ideas (i.e. they are not recognizing the value of student-student interaction above the content) [12]. This can be illustrated using the Practical Inquiry Model (Fig.1) where there is a triggering event

on a shared world, that began in the form of an issue or dilemma that needs resolution. As a result of this event, there is a natural shift to exploration in search for relevant information that can provide insight into the problem. As ideas grow, connections are made between all elements found and possible explanations, where finally, there is a selection and application of such solutions and resolution of the issue. This is the complex path of any critical analysis. The more participants and elements that are available to the deliberation, the higher the number of possible explanations that could be found, confirming the hypothesis that discussion enriches all path of exploration.

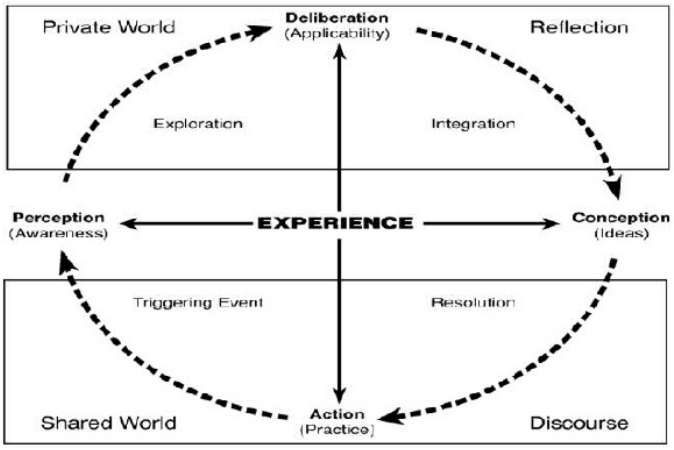

Figure 2. The Practical Inquiry Model (PI Model)

1.3. Teaching Presence

The glue to bring all together is teaching presence. By designing, facilitating and providing direction to the cognitive and social presence so as to allow students to achieve their full potentials. In traditional class teaching, the educator is responsible for providing emotional presence that will lure and engage the student into being inquisitive, in order to progress. The question arises as to the amount of support that teachers can provide to unknown remote students, with no physical links or contacts, rather barriers that limit their teacher-student contact. How can teachers replace the information provided in a face-to-face teaching sessions for students that are not present?

Emotions are present throughout our lives, and obviously are present in any online community [14][15]. Research on emotional presence, within online communities, demonstrates the existence of emotion in online learning. Given this reality, emotion must be considered, if not a central factor, at least as a ubiquitous, influential part of learning online [15].

Recent research in social presence [14][16] has suggested that an increase in social presence is important, not only for its impact on the quality of collaboration, but also for the development of shared group identity. Asynchronous methods lack the shared context, body information and timely feedback during collaboration, slowing down the whole process of creativity and leading to misinterpretation, reducing the perception of the common interest.

A recent development in collaborative working and learning has been the use of synchronous tools such as Web videoconferences, whereby learners meet online at a fixed time (synchronous) in an online classroom. While Web videoconferencing is not a new phenomenon, tools like Skype®, MSN Web Messenger ® and Acrobat Connect® allow learners to communicate efficiently, they still create virtual barriers, as the distant students are limited by the amount of communication that they may have with peers and teacher. Equally the lack of a sense of being socially embedded or immersed into the real space.

Until recently, basic technology would allow only for asynchronous communication, as in discussion groups, message boards and forums. Some courses have made use of those boards but with very poor response and results. For example, some learners reported a lack of spontaneity and improvisation that resulted from asynchronous communication, where frustration built up on the learners side while they were forced to wait for answers to their questions before being able to move on to the next point in the discussion [7].

Web and video conferencing have opened new dimensions for collaborative work. While the majority of students' time is spent in asynchronous activities (e.g. reading at library, preparing coursework and essays, searching for more information), the synchronous part of students' life is crucially important, such as attending debates, face-to-face tuitions and lectures, joining sporting groups and socializing at the University events. Web and Video based courses are often disappointing experiences for all ages groups but particularly for younger generations for the lack of those synchronously activities.

2. Approaches and Scenarios for Local and Distant

We now analyze a simple scenario based on traditional education. Student A, student B and student C arrive to their seminar room and seat contiguously.

As their lecture begins there are some questions that student B wants to confirm from student A, related to the taught subject. Naturally he/she will come closer to his/her ear and whisper the question. If the question stays unresolved then he/she might approach other people within close area (such as student C) or even will request attention from the tutor expecting a more public reply.

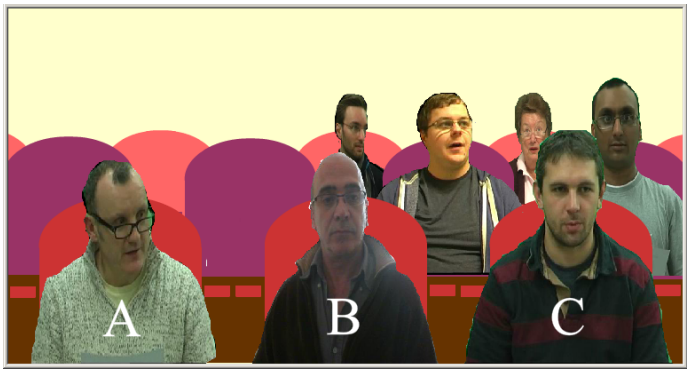

Figure 3. Traditional classroom seating

Now let's translate this scenario into the proposed distant learning scene. Let's consider student A and C arrive to their local seminar room, while the student B is attending remotely the lecture. Student A and C take a seat and student B remotely choose to sit on an empty chair between A and C.

Figure 4. Distant students immersed in the real classroom

The first question to arise might be; how will student A, B, lecturer and other attendees (including distant learners), be aware that student B is "present"? While this question can be easily solved by using similar approach to AR using a real video stream (that is visible to all), a new question arises: How can they communication between them providing with the discussion being limited to the same spatial limits as in real life and, where the operation is controlled by the remote student?

Clearly, to be realistic, any comment from student B should only be heard by the person seating next to his/her virtual space. Similarly, if student B's question stays unanswered then the need arises to approach other people near-by and, as last resort, the tutor or lecturer.

How the remote student will communicate their intention to speak? Evidently there is a need to model a system that takes all those issues into consideration.

3. Taking Spatial Ownership

The first problem is to define a system where the remote student can take ownership of a seating space to position him/herself among the other students, either other remote students or local attendees.

A spatial area is needed to be defined and coordinates passed onto the system. In our proposed scenario, a camera is kept pointing at the seating area, providing a similar view to that of the teacher. Students can click on the real time video image of the classroom at the available hotspots (real seating area) and, by doing so, will take ownership for that specific point in space. This spatial point will be released as soon as the student virtually leaves the room.

Once a student has taken spatial ownership then they virtually enter the physical room. Accessing the room is though a camera on the remote location (the distant student's location and web camera) that will remove any background information and in real time will stream only the student's body (removing background) onto the real-time video at the specific point chosen.

To do this convincingly it is necessary to consider the depth of view and ensure that closer students will occlude farther students (either remote or local ones).

By taking this spatial ownership we set the matrix where the distant student will remotely execute distant audio, defining who is within the different ranges and boundaries.

From this point onwards, groups are created dynamically to everyone's desire. Virtual and real spaces are synchronously mixed to the convenience of each element in the shared space. This space comes alive, not only for the distant student but also by the local ones, who can make use of this tool for their own internal communication without affecting the surrounding environment.

4. Aural Communication Channel

Real-time communication is a must, in order to support the emotions needed for a distant student to "feel" being part of the group, and for the local student to accept and identify the remote student as part of the group.

In Fig. 5 we illustrate our scheme for differentiating real students (blue dots) from immersive telepresence students (red dots). As any conversation between them should not reach public level, local students are encouraged to make use of their headphones connected to their network devices, such as their phones, tablets or laptops, which display a real time video showing the classroom view. Not only are they able to see their peers seated among them, but they will be able to hear and talk to them too, depending of their seating distance, as the boundaries will be set automatically depending the volume of the remote student (these audio levels are represented as ellipses in Fig. 5).

Natural speech is directional, therefore is necessary to take into account the location in space of the participants. This is easily managed with the built-in magnetic sensors that almost current pads and phones have. Equally this built-in facility is used to eliminate unwanted conversations of students away from the boundaries. When other students, either real or immersive, are talking between them, the audio will be automatically mute, if they are outside boundary thereby avoiding eavesdropping into each other's communication.

Rozier, Karahalios, and Donath [18], presented an augmented reality system using audio as the primary interface. Their intention was to create "audio imprints" at an outdoor location where, at a later stage, other individuals could collect them as they travel through the author path. Our proposed model is somewhat similar, with the variant that an audio imprint is not recorded but, rather, lives audio, which is both directional and selectable by the author at any given point (i.e. the same that natural speech).

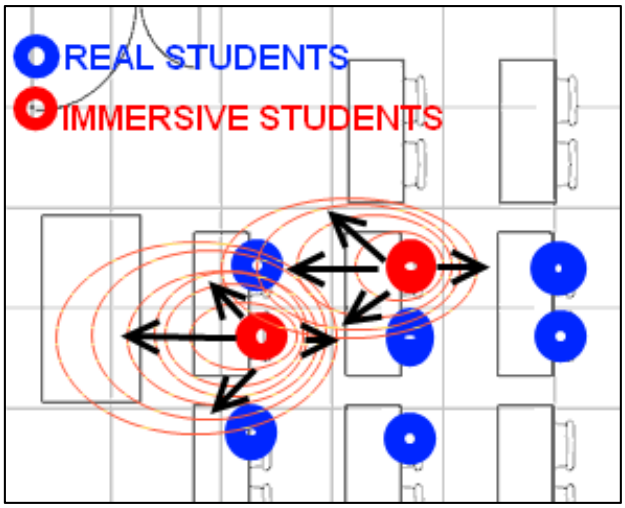

Figure 5. Communication between real and immersed elements

Our working system has been developed with ActionScript 3® (AS3) and Adobe Flash® for reasons of portability. There are some limitations but AS3 is a continually evolving language with new packages and classes becoming available every day.

Our audio system originally was developed for an immersive panoramic environment (360° real time video), positioning audio hotspots on a panoramic video which was presented in another paper at this conference [19]. It has been proven to be effective for most common types of room layout. For this reason the student's location is registered in relation to other elements as "audio hotspots". Local students wanting to use the immersive system will need to register their location by interacting with the video image.

While the experience has been taken from a panoramic scenario, a new algorithm has been developed to add elliptical behavior, due to the different goal (Fig. 6) of trying to achieve a fixed directional audio. Binaural audio elements can be added to enhance the "feeling of been there".

Those locations are recorded in a database and presented to all users through a web query, making every device aware of the other's location. In normal operation, devices are querying frequently in order to draw their boundaries and find elements within their range (as they will be the only devices receiving the audio stream). Audio stream are dependent of the distance and position from the source, and the manual control of the audio originator.

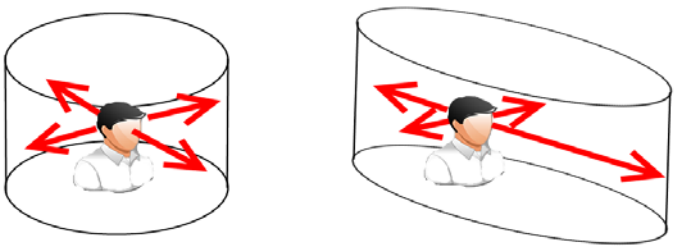

Figure 6. Circular and elliptical audio coverage

5. Visual Communication Channel

The immersive reality embedded into the real space (i.e. available to real people in the real teaching space) can only be provided by the use of networked devices, such as tables, smart phones and laptops (Fig. 7). Immersed students can be seen displayed either on students' tablet or another auxiliary projection system in the classroom (LCD screen or projector). Mixed elements can be tagged to provide extra information, such as the students' names, email contacts, public profiles and groups that belong to, etc… (as approved by the student whose information it is).

By the use of this method, teachers can experiment first-hand using face-to-face communication with remote students (and/or counterpart in collaborative teaching), providing the necessary student-teacher relationship without the content interference. West [20] mentioned that one the biggest challenges for teachers or instructors is the need to adapt to the lack of visual cues that by nature we use when interacting with people face-to-face. Until now preparing content that accommodates that lack of student-teacher interaction has been the main approach adopted but this is clearly not ideal.

Using this model teachers have the opportunity to enhance teaching by the adding aspects of augmented reality as the students already have tablets to facilitate this, without the need of any additional equipment (e.g. augmented glasses, 3D goggles with head-trackers, etc…) The usefulness of AR applied to learning has been demonstrated in many studies were participants where, for example, participants said they can memorize content better [21][23].

Constructivist theory underlines that knowledge is not transmitted from one "knower" to another but is actively built up by the learner, based on the way learners construct their own knowledge by testing ideas, taking different approaches, experimenting with the situation and comparing and contrasting what they already know. The closer that learners are to the reality, the faster they will build their knowledge [23].

Figure 7. Mixed real and Immersive Reality

6. Conclusions

In this paper, we have described a novel use of Immersive Reality based on embedding live video feeds into computer supported environments in a way that facilitates more realistic communication between them and their environment. We introduced the term, Immersive Reality and described our unique variant of it that used embedded real-time video and where the two elements interacting are video overlays on real space. In this world (i.e. the meeting or lecture setting) can be either left un-modeled or partially modeled to accommodate new structures and elements (such as AR and VR).

We provided a new insight to communication between participants in a telepresence session with the capability of applying segregation, occlusion and directionality, controlling the spatial domain for 1-to-1 and 1-to-many communication within the same event. Thus, from this model, individuals have the capability or dynamically extending or shrinking the communication group to accommodate their needs and even creating subgroups, modeling their relationship as they please.

7. Future Work

The proposed model has been tested technically at an experimental classroom level by making use of our iClassroom at the University of Essex. Transfer of this knowledge to a real classroom is needed in order to collect feedback from real students and teachers to get more definitive data.

References

[1] C. Cofin, S. Bostandjiev, J. Ford and T. Hollerer. "Enhancing Classroom and Distance Learning Through Augmented Reality". Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2010 (pp. 1140-1147).

[2] M. Billinghurst, H. Kato. "Collaborative Mixed Reality". Proceedings of the International Symposium on Mixed Reality (MR '99), March 19-21, 1999, Yokohama, Japan.

[3] R.L. Draft, R.H. Lengel. "Organizational Information requirements, media richness, and structural design". Management Science, Vol. 32, 1991, pp. 554-571.

[4] J. Short, E. Williams, B. Christie. "The Social Psychology of Telecommunications". London, Wiley 1976.

[5] P. Singla, M. Richardson. "Yes, There is a Correlation - From Social Networks to Personal Behavior on the Web". Proceeding of the 17th International World Wide Web Conference, Beijing, China, 2008.

[6] A. Rovai. "Building Sense of Community at a Distance". International Review of Research in Open and Distance Learning, Vol. 3 No. 1, April, 2002.

[7] E.J. Stodel, T.L. Thompson, C.J. MacDonald. "Learners' Perspectives on What is Missing from Online Learning: Interpretations through the Community of Inquiry Framework". Review of Research in Open and Distance Learning, Vol. 7 No. 3, December, 2006.

[8] D.R. Garrison, T. Anderson, W. Archer. "Critical inquiry in a text-based environment: computer conferencing in higher education". The Internet and Higher Education, 2, 2–3, 87–105, 2000.

[9] A. Caspi , I. Blau. "Social presence in online discussion groups: testing three conceptions and their relations to perceived learning". SocPsycholEduc , April 2008.

[10] L.M Harasim, L.M. "Online education: Perspectives on a new environment". New York: Praeger 1990.

[11] C. Wei, N. Chen, Kingshuk. "A Model for Social Presence in an Online Classroom". Proceedings of the 18th International Conference on Computers in Education. Putrajaya, Malaysia:Asia-Pacific Society for Computers in Education, 2010.

[12] C. Angeli, N. Valanides, C.Bonk. "Communication in a web-based conferencing system: the quality of computer-mediated interactions." British Journal of Educational Technology, 34(1), 31-43, 2003.

[13] P. Gustafson, D. Gibbs. "Guiding or hiding? The role of the facilitator in online teaching and learning. Teaching Education". 11(2), 195-210, 2000.
[14] M. Cleveland-Innes, P. Campbell. "Emotional Presence, Learning, and the Online Learning Environment". Research Articles in The International Review of Research in Open and Distance Learning, Vol. 13, No. 4, October 2012.
[15] G. C. Marchand, A. P. Gutierrez. "The role of emotion in the learning process: Comparisons between online and face-to-face learning settings". The Internet and Higher Education, 15(3), 150–160, October 2011.
[16] R. Pekrun, T. Goetz, W. Titz, R. Perry. "Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research". Educational Psychologist 37, 2, pp-91-106, 2002.
[17] P. Rogers, M. Lea. "Social presence in distributed group environments: the role of social identity". Behaviour & Information Technology, Vol. 24, No. 2, March – April 2005, 151 – 158.
[18] J. Rozier, K. Karahalios, J. Donath. "Hear&There: An Augmented Reality System of Linked Audio". International Conference for Auditory Display 2000, Atlanta, USA, April 2000.
[19] A. Torrejón, V. Callaghan, H. Hagras. "Selectable Directional Audio for Multiple Telepresence in Immersive Intelligent Environments". IE'13, Athens, Greece. July 2013.
[20] G. West. "Teaching and learning adaptations in the use of interactive compressed video". T H E Journal, 21 (9), 71-74, 1994.
[21] D.D. Sumadio, D.R.A. Rambli. "Preliminary Evaluation on User Acceptance of the Augmented Reality use for Education". 2010 Second International Conference on Computer Engineering and Applications. Bali, March 2010.
[22] W.F. Maqbleh, M.S. Sindhu. "From Boards to Augmented Reality Learning". Conference on Information Retrieval & Knowledge Management, (CAMP), 2010 International, March 2010.
[23] H. Kaufmann, M. Csisinko, A. Totter. "Long Distance Distribution of Educational Augmented Reality Applications". Proceedings of EURO-GRAPHICS 2006, pp. 23-33. Blackwell, Oxford. 2006.