Camera Set-up, Image Placement and Interactive Devices in Virtual Learning Environments

Increasing the Feeling of Co-Presence and Interactivity in Access Grid mediated sessions

Christine Rivers*, Christina Silver*, Nigel Woodger+

*Department of Sociology, University of Surrey +Department of Psychology, University of Surrey

Guildford, GU2 7XH, United Kingdom
c.rivers@surrey.ac.uk, c.silver@surrey.ac.uk, n.woodger@surrey.ac.uk

Abstract—Virtual learning environments such as Access Grid Node technology offer new ways of teaching to multiple remote sites. This paper investigates the potential of using Access Grid technology for delivering training sessions to remote and co-located learners. Qualitative thematic-based analysis of textual and audiovisual data reveals that co-presence and interactivity are related to camera positioning and image placement. This paper proposes alternative set-ups for large and small virtual teaching sessions and the use of interactive devices to facilitate the feeling of co-presence, support interactivity and encourage active participation. After contextualizing the research by introducing general issues with respect to Access Grid technology this paper describes the methodology and research design including teaching session formats, the AG technology used, participants and data collection methods. It suggests ways to increase the feeling of co-presence and interactivity in remote teaching sessions.

Keywords—Applications for Information and Communication Technologies in Education and Training, Computer-based Training, Distance Learning, Virtual Classroom

I. INTRODUCTION

Technological developments such as Access Grid Nodes (AGNs) add new perspectives to the field of distance learning, potentially changing the way students and tutors experience distance learning. AG-based technology offers alternative ways of delivering teaching sessions remotely by enabling high-fidelity visual images and audio streams to be exchanged in real time amongst multiple AG sites [1,2,3]. Images are projected on node room walls with the intention of replicating the feeling of being in the same room, or co-presence, as much as possible. This aspect is one of the advantages of using AG technology as opposed to earlier virtual learning and reality environments such as Active Worlds [4] or ConferenceXP [5]. Figure 1 illustrates an AG room node wall displaying different sized images.

The ability to display life-size video and to stream files (e.g. PowerPoint presentations, web pages etc.) has several advantages, including increasing virtual awareness [2,7,12], human-to-human interaction, verbal and non-verbal interaction (e.g. eye-contact, turn-taking [2,3,11]), the effectiveness of the learning experience [9,11], the performance of the tutor [9] and the feeling of co-presence [2,3]). However, co-presence can only be emulated by using camera angles and image positioning that adequately captures and displays participants at host and remote sites. Using multiple cameras to capture the room from different angles [2,7,11] can address this issue as participants can orient themselves similarly to a face-to-face situation. This in turn affects the level and nature of interaction, in particular active participation such as initiating discussion and providing both verbal and non-verbal feedback. Although AG technology is largely seen as an effective virtual collaboration tool [2,3,6,7,8], previous research also emphasizes that the level of interaction is still not comparable to traditional face-to-face meetings [2,7,11,12]. In particular it may not adequately facilitate active participation or interactivity because images display participants from a distance. This can result in more difficulty in detecting subtle physical cues than is the case in face-to-face interactions [2]. In teaching settings this may include visual indicators that students are confused or struggling to keep up which, in a face-to-face session, would result in the tutor slowing down or asking if a concept...
requires further elaboration. This sort of subtle and informal feedback is important for tutors in assessing audience engagement [3] and understanding. In order to afford the feeling of co-presence collaborative systems therefore need to be able to communicate as many such visual cues as possible [3,8,11], whilst simultaneously giving participants the option to actively participate and interact with the tutor at any time. The concept of engagement is discussed in [3], including the tutor-participants relationship.

This raises two questions which the research reported here addresses: 1. How can the feeling of co-presence be increased in order to support mixed presence groups (multiple participants at different locations) [13] in virtual learning environments via AG technology? 2. How can interactivity be increased in order to diminish the gap between virtual and face-to-face classroom environments by means of AG technology? This paper identifies aspects of co-presence that hinder participants in establishing a feeling of co-presence during AG-mediated teaching sessions and comments on ways interactivity between all participants can be increased. This includes the use and considered placement of multiple cameras, feeds and angles [2,3,7] as well as additional interactive devices [14] e.g. handheld devices or tablet computers. Findings contribute to the development and evaluation of AG technology as a medium for delivering teaching sessions to multiple remote sites as well as to the fields of virtual collaboration and learning, and human–computer interaction.

II. METHODOLOGY & RESEARCH DESIGN

This paper draws on a research project investigating the potential of using AG technology for delivering Computer Assisted Qualitative Data Analysis (CAQDAS) training sessions to remote learners as well as on previous findings of AG data from student and criminal justice practitioner discussion groups [3]. A range of software applications have developed since the mid-1980s specifically designed to facilitate the systematic analysis of qualitative (initially textual, more recently also multimedia) data. The CAQDAS Networking Project, based in the Department of Sociology at the University of Surrey, UK, provides information, advice, training and ongoing support in the use of a range of such software packages. Training sessions have historically been offered in traditional face-to-face settings, via seminars or hands-on computer workshops. Due to the high demand for CAQDAS training it is relevant to test the potential of AG technology to deliver content to a wider population. A comparative methodology is being employed to investigate two types of face-to-face session with the remote equivalent.

1) Session Formats

Teaching sessions were held in 2007 and 2010 at the University of Surrey, in remote collaboration with three other UK AG sites: Manchester, Southampton and Bristol. Technical and pedagogical aspects raised during the 2007 sessions were accounted for in designing the 2010 session, influencing session set-up in order to further investigate identified themes.

Two types of CAQDAS training sessions were offered: awareness raising sessions (AWR) and intermediate support sessions (ISS). AWR sessions are essentially lecture-style sessions comprising demonstration and discussion of several CAQDAS packages in order to enable participants to make informed choices between software packages. ISS sessions are interactive hands-on teaching sessions in which participants have a laptop loaded with their work in progress. The tutor provides project-specific support to individual participants and general advice on more advanced software features to the group as a whole. In 2007 AWR and ISS teaching sessions were run via the AG, and in 2010 an AWR was offered. A face-to-face AWR session was also recorded for comparative purposes.

2) Access Grid Technology

The AG technology used consisted of hardware and software resources including audio and video equipment [2,3,14,4], four Canon VC-C4 communication cameras (three in the front of the room facing the tutor and one capturing the room from the back). The teaching sessions were recorded using two methods. In addition to two miniDV format camcorders, in the Surrey node room, a further node was adapted as an additional window for the teaching session. Video and audio data from the SGW node was carried via a conventional SVHS display, whilst simultaneously being fed to a VCR or hard disk video recorder/DVD writer, and from there to a video monitor. In the remote sites, Manchester, UK (ISS) and Southampton, UK (AWR). Drawing on the insights gained in the 2007 sessions the group size was increased for the 2010 session and mixed presence groups [13] were formed; six participants in the host location (Guildford) including the tutor and five participants in the remote location (Bristol). All participants were recruited through the project website or local contacts at the remote sites. Participants received an incentive payment of £10 and completed an informed consent form which allowed them to specify whether they agreed to their image being transmitted for presentational purposes.

III. DATA COLLECTION

Several different types of data were collected, including AG node and video recordings of each training session, structured debriefing sessions with participants following each session, observational field notes and tutor reflections.
This approach was taken in order to provide an holistic representation of the sessions, reflecting their inherent multimodal nature.

1) Video Recordings and Debriefing

Video recordings show the training session as well as the subsequent debriefing sessions. This paper reports a thematic analysis conducted on the full verbatim transcripts of the debriefing sessions, and the direct visual analysis of the video of the session content. The significance of using debriefing data, which reflects users’ experiences, to analyse and understand concepts of engagement, has been pointed out by [3].

Each training session took between 30 and 60 minutes and different camera angles show the interaction from different perspectives. The debriefing sessions were approximately 30 minutes long and covered questions regarding the teaching session in general, the technical and interactional aspects of the AG as well as questions regarding participants’ willingness to attend such training events in the future.

2) Field Notes and Reflections

During the 2010 session two observers were present, one visible to the tutor and participants at the host site but not to the remote participants and another observer not visible to any participants, who also controlled the AG technology from a separate room. Both observers took notes during the sessions. After all training sessions the tutor wrote reflections regarding her perspective on the session. These typically included comments on the impact of the technology on presentational style and the nature of interactions with participants as compared with conducting similar sessions in face-to-face settings.

IV. DATA ANALYSIS

This section explains the analytical and procedural processes taken to identify aspects that hinder or encourage the feeling of co-presence and interactivity amongst participants. Analysis was undertaken in two stages; firstly thematic textual analysis of the debriefing transcripts, field notes and reflective comments; secondly audio-visual analysis of the video data focusing on non-verbal visual cues. Audio-visual analysis served the purpose of validation; investigating whether the aspects identified as important by participants, observers and the tutor are evidenced in the visual record of the training sessions.

A. Analysis of Transcripts, Field Notes and Reflections

A combined inductive-deductive approach was adopted to initially code the textual data, facilitated by the CAQDAS package NVivo. Certain issues such as understanding the effects of the physical and technical setup on aspects of interactivity and seeing patterns of interactive behavior had been identified from literature in the field [2,3,14] and our previous work [2] on AG technology. Fifty-six codes were developed during the first coding phase, which were subsequently grouped into 11 categories, e.g. social interaction, experience, communication, visual effects, motivation, physical effects, etc. The most promising category is social interaction, which includes sub-codes such as aspects of co-presence, aspects of interactivity and lack of social interaction. Lack of social interaction and aspects of co-presence have been coded 17 times and predominately appear in relation to one another. Aspects of interactivity has been coded 12 times across the whole data set, and has also been found to be highly related to aspects of co-presence, which substantiates findings outlined by [3]. This gave reason for more in-depth analysis in order to understand how aspects of co-presence and interactivity unfold and to further identify aspects that hinder or increase co-presence and interactivity within the AG environment. Most of the utterances that were coded as an aspect of co-presence referred to space awareness such as the importance of easily seeing the other remote group which referred mostly to the poor quality of the remote video. Participants identified the importance of displaying both remote and host images as key to emulating a feeling of being in the same room, thereby increasing the level of co-presence, which is also emphasized by [2,3] who analysed group discussions in an AG environment and pointed out the staged behavior participants anticipated during the meeting as people are highly adaptive to new communication media [3]. The following example illustrates the necessity of seeing the remote group:

“But, I agree you do need to know that the other people are present because otherwise, although we know that you are hearing us, you don’t get that same sense of interaction that you do when you can actually see everybody who is participating.” (Debrief Surrey 2010, Guildford participant)

However, as illustrated by the following comment, being able to see one’s own image also contributes to the feeling of co-presence:

“I quite like the way you can see yourself in the room it creates a feeling of being all in the same space ... people in my organisation would quite like that a lot because they tend to complain about technology not giving that feeling of participation that they are used to and I thought that is quite impressive.” (Debrief Surrey 2007, Manchester).

This illustrates that a lack of social interaction may be perceived if participants are not adequately aware of who is taking part in the virtual space [2]. Thus co-presence can be increased by providing participants with images of themselves and all other participants. As well as displaying images of other locations and participants, image quality and placement were highlighted as important in order to further emulate a classroom environment. These two aspects were also emphasized in the field notes taken by the
observers as well as the tutor’s reflection:

“... I usually get quite a lot back from participants and therefore can gauge how the session is going – whether people are ‘getting it’, whether they are engaged or looking bored etc. I use this kind of non-verbal feedback to tweek my presentation accordingly. Lost a lot of this during the session. This may be in part to do with the positioning of our main feed on their wall and the quality of the video streams being able to pick up the subtleties of facial expression etc.” (Tutor Reflections, Surrey 2007)

The two aspects addressed above (quality of images and arrangement of windows) will be addressed in greater detail in the next section but is worth mentioning here as it shows that further analysis of different data is necessary.

Reflecting upon the former aspect of providing specific visual information to participants might encourage them to interact more actively in the created virtual space. Other participants have expressed this desire for active participation more explicitly:

“I don’t think that there will be any space for social interaction between, like between the lecturer and the other students, ah you know, across the different sites, which is sometimes the way and really, really nice and interesting things also happen. It is supposed to be ah I guess the main value is interactivity.” (Debrief 2007, Manchester participant)

This example suggests that the participant feels an AG teaching session does not allow sufficient space for social interaction, but concurrently, stresses that interactivity is expected from participants. This can be seen as contradictory, prompting questions about what hinders interactivity. The criticism that AG technology in a classroom environment may not offer sufficient space for interactivity highlights the need to identify ways in which the effectiveness of AG-mediated teaching sessions can be increased and supported. Looking at the code ‘social interaction’, in particular co-presence and interactivity and other co-occurring themes and sub codes, revealed that interactivity and co-presence are related to set-up issues, which has also been classified as a barrier for engagement by [2]. Thus a code ‘Set-up issues’ was created to referred to utterances that stress the significance of having a different set-up that provides a ‘space’ for interactivity. This was reinforced by another participant:

“... I did not know how to interrupt to ask questions, I did not know how to engage.” (Debrief, Surrey, 2010)

At this point it became clear that the textual data in themselves were not sufficient to answer the two research questions. Thus, the video recordings showing the main training sessions were analysed in great detail in order to identify aspects of co-presence and interactivity in relation to set-up issues.

B. Analysis of Video Recordings

Recording the teaching session from different camera angles is of great value in investigating set-up issues in relation to co-presence and interactivity. In order to facilitate the analysis, the multiple video recordings were synchronized so that the interaction in both locations could be viewed and interpreted simultaneously. This had the implication of further reducing the quality of the video images but this limitation was outweighed by the possibility synchronicity affords for identifying aspects that hinder co-presence and interactivity at various points during the teaching session. The first aspect of crucial importance is the placement of the AG video images on the node wall. In the Surrey-Southampton session in 2007, multiple windows were displayed in the host location but the participants looked in different directions and away from the tutor when she was speaking. Fig 2 shows the arrangement of the windows and the camera angles which caused this view. This shows that the directions participations are looking in and talking towards can be distracting for other participants, and, as in this case, the tutor. If image placement does not allow participants to look at one another when speaking it is difficult to establish eye contact and therefore the feeling of co-presence [2,7] will be diminished. This may also result in reduced interactivity between participants as they cannot rely on normal visual cues to initiate communication. This example implies that camera angles and AG image placement are of great significance for all sites in remote collaboration and therefore time and effort should be given to these factors in planning for and setting up remote teaching sessions. These factors are the main technical contributors to facilitating high level of interactivity and the feeling of co-presence and can easily be influenced by correct placement.
Related to these issues are the ways in which participants indicate the desire to participate in remote teaching sessions, thereby increasing the level of interactivity. The video recordings show that participants in all teaching sessions either raised their hands, waited for a gap to ask a question or to be approached directly by the tutor. Getting attention seemed to be difficult particularly in the 2010 session due to the mixed presence nature of the groups [13] which increased the level of complexity in terms of interaction. The tutor played a significant role in this session to coordinate participants and thus mediated the level of interactivity. This implies that participants were able to indicate their wish of asking a question but felt they had to wait for confirmation before proceeding. In addition, participants located at the remote sites could not easily see who was asking a question as the images displayed on the wall were very small and thus indistinct.

This validates the findings from the textual data analysis, illustrating how different data types representing the same event combine to provide an holistic and corroborated interpretation. It is clear therefore, that in remote teaching sessions, when a question is asked, it is important that all participants – tutor and students – can see the person asking the question in detail. Therefore a minimum level of video quality, optimal camera positioning and image placement are imperative in increasing the feeling of co-presence and fostering interactivity in AG-mediated teaching sessions.

V. RESULTS AND DISCUSSION

Data analysis revealed that co-presence and interactivity are related to the camera set-up in a node room [2,3] and the choice of image placement on the node wall in all locations. This insight is of particular importance when designing AG-mediated teaching sessions where multiple participants and mixed presence groups take part. Virtual teaching sessions require effective communication and should offer students the opportunity to interactively communicate with the tutor in a similar manner to a face-to-face session [2,3]. The AG system was originally envisaged to facilitate remote communication between groups of less than ten people [2,6]. However, if sessions are delivered to more students, which is a likely scenario if part of the purpose is to reduce the costs associated with travelling and to reach a larger population, addressing camera set-up issues is imperative to satisfy the complexity of the setting. This includes the physical positioning of cameras in all locations, as they relate to the positioning of participants in terms of height; the angles at which the cameras are pointing in order to capture images; and the physical positioning of video images on the node wall. Video images must be general enough to show all participants, yet specific enough to allow the facial expressions and gestures of particular participants to be clearly seen when required [2,3,5]. However, the number of video streams must be kept manageable, both from the point of view of participant attention, and of network traffic. Audio must allow effective communication between the lecturer and any individual participant, and allow any participant to be heard at all sites, which would facilitate the tutor participant relationship [2]. The system should be largely automatic (i.e. an additional operator should not be required to control the cameras and sound during the lecture). This might enable AG-mediated training sessions to be comparable to traditional face-to-face classroom situations [2,3]. Doing so, however, requires achieving a careful balance between supporting co-presence and allowing interactivity.

Drawing on the insights gained from the analysis this paper suggests a system using a combination of interactive devices such as student handsets and pre-programmed camera controllers to address these problems.

This paper suggests equipping lecture theatres with access grid technology, cameras with slightly modified control software. This software can store a pre-programmed pan-tilt-zoom combination for each seat in the lecture theatre. This is a well established feature common in security cameras [16]. The camera is sighted as close as possible to the portion of the display that will be used for the lecturer.

At the start of a class, students will be issued with an interactive device (e.g. handheld device or tablet computer) comprising a microphone, and an alphanumeric keypad with a ‘call’ button. This device may be wireless or plugged into the arm of the student’s seat. The student then uses the device to log onto the system, entering the name and seat number, information which is transferred and shown on the lecturer’s display. During the class, if the student wishes to ask a question, s/he can use the ‘call’ button, which highlights their name on the lecturers display. The lecturer can accept the question, for example by clicking on a ‘button’ next to the students name on the lecturers display. After accepting a question the microphone built into the students interactive device becomes live, and the automatic camera moves to the pre-programmed pan-tilt-zoom settings for that student’s seat. A close up view of the student is displayed on the node room wall such that all other students...
and the tutor can clearly see the student and hear the ensuing discussion. After the question has been answered the camera returns to a general view and the student’s microphone is muted.

Students can withdraw their ‘call’ if, for example, their question has been answered before their call is accepted and the tutor can choose to accept or decline calls as appropriate. This would allow smooth interactive communication with bigger group sizes, enabling students to interact actively at anytime during a teaching session and ensuring other students can see who is asking the question clearly which should increase the feeling of co-presence and stimulate engagement. At the same time, the tutor remains in control of the interaction process which adds a level of authority and leadership to the session which has been found to be a crucial factor for virtual collaboration [17].

VI. CONCLUSION

This paper investigated aspects of co-presence and interactivity in a virtual teaching environment using AG technology drawing on preliminary analysis of ongoing research. Findings reveal that the feeling of co-presence and interactivity can be increased by adjusting the camera set-up to focus on individuals and give the impression that the tutor is looking at a particular remote site during a virtual teaching session by arranging the video windows in a specific way. Additionally, students should be provided with interactive devices that allow them to communicate actively with the tutor. In order to test the effectiveness of the suggestions given, further research is required and the implementation of such interactive devices should be taken into consideration.

ACKNOWLEDGMENT

The work presented here stems from an on-going research project that forms part of the capacity building elements of the Qualitative Innovations in CAQDAS (QUIC) project, a Research Node funded by the ESRC National Centre of Research Methods (NCRM).

REFERENCES
