Implementing Three-Party Desktop Videoconferencing

Jian Sun
University of Otago
Information Science
Dunedin, New Zealand
++64 3 4738687
sunji568@student.otago.ac.nz

Holger Regenbrecht
University of Otago
Information Science
Dunedin, New Zealand
++64 3 4798322
holger@infoscience.otago.ac.nz

ABSTRACT

We describe the implementation and test of a novel three-party desktop videoconferencing system. To allow for gaze and workspace awareness between the participating partners a special quasi-spatial arrangement of cameras and graphical user interface elements is chosen. We informally tested the system setup with a usability evaluation presented at the end of this paper.

Our prototypical solution is a customizable off-the-shelf, affordable way of supporting mutual awareness in three-party videoconferencing.

Categories and Subject Descriptors

H.4.3 Communications Applications - Computer conferencing, teleconferencing, and videoconferencing

General Terms

Design, Experimentation, Human Factors

Keywords

Two-camera systems, Three-way communication, gaze awareness

1. INTRODUCTION

Desktop videoconferencing systems are increasingly used around the world thanks to the affordability of hard- and software as well as the availability of high bandwidth Internet connections. The majority of the systems are used for video chatting and for the communication of two parties. The use in “more serious” application fields is not very well supported with hindering factors being missing information integration (e.g. text documents, presentation slides, calculation sheets, photos) and also limitations in establishing communication between more than just two parties. Even if systems like Apple’s iChat or daviko support multi-party videoconferencing sessions, gaze and workspace awareness as well as aiding non-verbal communications cues are poorly supported. As always with multi-party systems, a general question is: Who is talking to whom about what? There are different approaches addressing this issue (e.g. [37])

We present a possible solution, which makes use of affordable, customizable off-the-shelf hard- and software components, but integrate them in a particular way to support these awareness requirements. We propose a solution, which allows three parties to communicate with gaze awareness using a setup illustrated in figures 1 and 2. Each party’s computer is equipped with two web cameras, attached to the sides of the monitor screen. The video streams of the other two participating persons are displayed in (1) close proximity to the camera attached and (2) in a spatially correct way, simulating a triangle (meeting table) setup. Therefore the camera video streams are linked to the output windows in a certain way shown in figure 2.

Figure 1: The two-camera system concept

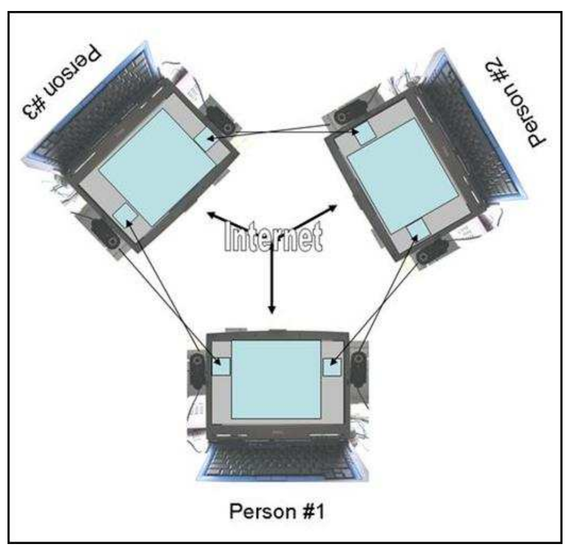

Figure 2: Virtual spatial arrangement of systems

As shown in figure 1, the center of each screen is used as a shared collaboration area to which the video stream output windows are attached.

OzCHI 2007 Proceedings, ISBN 978-1-59593-872-5

With this hard- and software setup it is possible for each user to estimate where the other people are looking: either at each other, the shared workspace in the center of the screen or at the user (me). Figure 3 shows two example views. As the objective is Integrating Communication and Collaboration among three parties so we call this tool IC3.

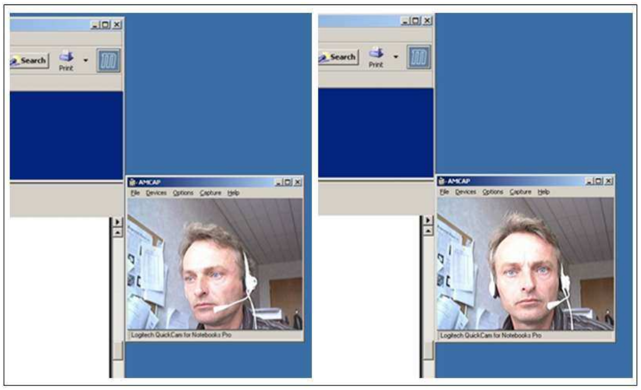

Figure 3: Different appearances of other party according to look-at perspective

This paper is based on a thesis work of the first author under supervision of the second author. It first highlights some related work in the field, describes the design and implementation of the system in some detail, including the component alternatives considered. Finally a heuristic evaluation of our prototype is presented.

2. RELATED WORK

2.1 Groupware

Groupware is defined by Ellis, Gibbs and Rein [14] as a computer-based system that supports a group of people engaged in a common task that provides an interface to a shared environment. With recent advances in network infrastructure and computing power, desktop video conferencing and groupware systems are rapidly evolving into technologically viable solutions for remote communication, collaboration and coordination. Groupware not only reflects a change in emphasis from using the computer to solve problems to using the computer to facilitate human interaction, it is also regarded as a highly demanding tool since a significant portion of a person's activities occurs in a group, rather than an individual context [13]. Conventional communication systems such as email, newsgroups, document management systems and defect tracking systems are asynchronous groupware. They involve neither immediate user interaction nor object sharing. This contrasts with a synchronous groupware system, which allows physically separated users to interact with one another with shared computational objects in real time [30]. Therefore, it is argued that successful synchronous groupware should be able to provide effective communication. The effectiveness of these communications can be further improved by proper seating arrangements and better gaze contact support. Similar to communication, collaboration is a cornerstone of group activity. Effective collaboration demands that people share information. For that reason, a shared workspace that offers a group context and explicit notification of each other's actions (gestures) becomes very necessary. The effectiveness of communication and collaboration can be enhanced if a group's activities are coordinated. Without coordination, a team will easily engage in conflicting actions [14].

In a real working situation, people are kept aware of what others are doing by just seeing it or by notification from other people. This helps to organize people's own work and co-operation with others. This important aspect of groupware is referred to as workspace awareness [19]. Information available in and throughout the workspace allows people to maintain awareness of others locations, activities and intentions. According to Gutwin et al [19], when shared activity moves from face-to-face to distributed groupware, many things change that impair people's ability to maintain workspace awareness. For instance, the perceivable environment shrinks, the communication means are weakened, the physical interactions need to be replaced by a computerized way and so on. They further point out that the groupware designer must try to recreate the conditions and cues that allow users to keep up a sense of workspace awareness.

2.2 VoIP, VCoIP

In recent years, VoIP (Voice over IP) and VCoIP (Videoconferencing over IP) software have become more and more popular as a communication tool among geographically distributed people. VoIP (also called IP Telephony, Internet telephony, or Digital Phone) is the routing of voice as data packages over the Internet or any other IP based network. VCoIP is used for routing both audio and video data over the Internet or IP based network. Because many people still see video images as an additional feature on top of the VoIP system, frequently the term VoIP along is used to refer both of them.

There are many challenges faced by VoIP. Most critically, VoIP relies heavily on the internet. Consequently, this leads to issues of latency (network conjunction and loss of packet), network reliability, security and emergency calls (as not all VoIP networks are connected to emergency services and are unable to identify the location of the calls due to the inherent mobility of IP end points). The quality of service also relies on the hardware environment such as the computer, webcam and microphone. Finally, as computers require local electricity to work, not like PSTN (Public Switched Telephone Network) devices can be powered by the telephone line, which is a more reliable separate power source, the VoIP service is more vulnerable to any sort of power outage. However, some of these issues can be resolved. For example, the security problem can be eliminated by better system design: the data can be encrypted at the endpoints and throughout the network, and an additional level of identity authorization features can be introduced through the use of passwords [24]. Because VoIP is part of the proposed system, the limitations VoIP is currently facing also apply to our prototype system.

2.3 Seating arrangement

Parsons [31] summarizes several key aspects affecting effective communication. The sending and receiving skills are two important ones. The ability to express yourself clearly and efficiently, and the ability to understand other people have always been associated with success, whether in business or life in general. He also identifies that the setting and environment in which the communication is taking place must be given serious consideration. For instance, the room you are in may be too big, too small, too cold, too hot, too noisy, or for various reasons, uncomfortable. If the purpose of the communication is giving an information session, then theatre style seating is appropriate. If the communication will involve participation and collaboration, then a round table seating arrangement is more suitable; as everyone can be seen and heard more readily. Further, group size and atmosphere are also two important considerations identified by Parsons [31]. Many of these factors cannot be controlled or improved by a communication support system. What can be controlled for however are group size and seating arrangement. As the proposed system is aiming to provide a communication and collaboration environment for three people, it is also suggested that a round table seating arrangement could be simulated through special camera and video stream placement.

Barbour and Barbour [3] suggest that teams can use seating arrangement to establish a sense of equity in a meeting. Dabkowski [9] supports this view and further points out that using a round conference table eliminates a sense of hierarchy among participants that may otherwise exist when particular team members are seated at the head of a rectangular table. Wilson [38] investigates the practices of discussions with students via videoconferencing, and concludes that the specific seating arrangement is important for improving eye contact and interaction.

Research on the social use of space suggests that spatial arrangements exert an important influence on the course of interactions. Batchelor and Goethals [4] examine the relationship between task instruction and seating arrangement. Their results show that the type of work has a significant effect on the spatial arrangement. When the instruction given to participants is for a collaborative work, a circular seating arrangement is more preferred by participants. Hendrick, Giesen and Coy [21] and Giesen and McClaren [17] also show a clear preference for circular arrangements to facilitate interaction.

2.4 Eye Contact / Gaze Awareness

Many studies address the importance of eye contact or gaze awareness in communication and group interaction. It is estimated that 60 percent of conversation involves gaze, and 30 percent involves mutual gaze [33]. Garau, Slater, Bee, and Sasse [15] summarize that eye contact serves at least five distinct communicative functions: regulating conversation flow, providing feedback, communicating emotional information, communicating the nature of interpersonal relationships and avoiding distraction by restricting visual input. Takeuchi and Naito [34] further demonstrate the importance of eye contact in social interaction by arguing that eye contact can indicate whom a message is directed to.

Lack of eye contact seems to be one of the most difficult problems to overcome in a video conferencing system. In a traditional VCoIP system, the camera is usually placed on top of the monitor. Thus, when a conferee looks at the displayed video image of the interlocutor, an angular deviation between the camera shooting axis and the line of sight of the interlocutor exists. This is recognized as eye contact parallax. This deviation impedes interlocutors in identifying out true viewing directions, and eye contact between conferees will never be experienced [26].

Many studies show that an absence of individual eye contact deteriorates the communicative presence [6; 22]. To eliminate the angular deviation, various systems have been developed and tested. One example is the “Multi-Attendant Joint Interface for Collaboration” (MAJIC) system developed by Okada, Maeda, Ichikawaa, and Matsushita [29]. MAJIC achieves gaze awareness by specially arranging video cameras and projectors to display life-size video images of participants onto a large curved screen without boundaries between them. Another video conferencing example aiming to provide eye contact is documented by Yang and Zhang [39]. They describe an eye-gaze correction system using a pair of calibrated stereo cameras and a personalized face model to track the head pose in 3D. Both of these approaches are expensive to set up. A cheaper, pure software approach was introduced by Gemmell, Toyama, Zitnick, Kang, and Seitz [16]. Their approach is similar to Yang and Zhang’s [39]; however the drawback of this approach is that it uses a synthetic view to simulate a natural gaze awareness video conversation which still seems unnatural.

2.5 Collaboration

2.5.1 Shared Workspace

Many groupware systems provide an environment for collaboration called a shared workspace. This is a bounded space where people can see and manipulate artifacts related to their activities. In the real world, a shared workspace is simply a physical space where people can undertake joint activities.

In a real-time synchronized situation, according to Lauwers and Lantz [25], generally there are two categories of approaches to develop synchronous collaborative applications. In the first category, collaboration awareness, applications are specifically designed for simultaneous use by multiple users. The shared workspace and multiple user collaboration mechanisms are designed at the very beginning and built into the applications. In the second category, collaboration transparency, single-user applications are shared by collaboration-aware mechanisms that are unknown, or transparent to the shared applications and their developers. This is a more generic way of application sharing such as screen sharing and application view sharing. The screen sharing approach replicates one user’s desktop onto multiple workstations so that all users see and interact with the same view. This allows singular user’s applications to be shared to multiple users simultaneously without modification. On the other hand, collaboration awareness tools provide a richer user interface that support collaboration more effectively. For example, these tools might support different views based on the roles of the individual user, or allow users to work on separate portions of the shared artifact while keeping its state consistent [13].

Commercial products that support screen sharing include Microsoft’s NetMeeting, Symantec’s PCAnywhere and most recently the VNC system.

2.5.2 Telepointer

Studies of small face-to-face groups working together over a shared work surface reveal that gesturing comprises about 35% of the group’s activities [35]. A gesture is a rich communication mechanism. Through gesturing, participants are able to indicate relationships between the artifacts, draw attention to particular artifacts, show intentions about what they are about to do, and suggest emotional reactions [18]. In a real-time desktop based video conference, the ability to point at shared objects on the display can also greatly enhance communication between conference participants [25]. Many groupware systems use telepointers (also known as multiple cursors) to provide a simple but reasonably effective mechanism for communicating gestures (Hayne, Pendergast & Greenberg [20]).

According to Lauwers and Lantz [25], telepointing and annotation facilities are most easily implemented if the underlying window system supports transparent windows. In this case, the annotation and telepointing tools can create transparent windows on top of existing windows, and use these transparent windows as their drawing canvases.

A number of prior studies in this area focus on the design, implementation and usefulness discussion of multiple mouse cursors working on a single display device. This kind of application is termed as Single Displayed Groupware (SDG). Early efforts to implement multiple input devices include Bier's MMM [5], Hourcade's MID [23], Tse's SDGTookit [36], and the open source project CPNMouse [8]. The former three implementations face the same problem, that the Windows operating system (OS) elements only provide "listeners" for a single mouse pointer, thus these SDG implementations involve the creation of own applications and elements that facilitate listening to multiple mice. CPNMouse uses a common mouse driver to cover original mice drivers and register multiple mice to a small application. Although ordinary Windows OS elements such as scroll bars and buttons are able to respond to each of the registered mice, the performance is very poor and unstable. Most importantly, these implementations focus on supporting multiple devices on a single display or a single system rather than displaying and managing multiple mouse cursors from clients over the network. Nevertheless, researchers in this area struggle with the potential value of their work as they face the difficulty of implementing an application that supports multiple input devices. And for those who do have this kind of application, they often come up with positive results on the tasks performed among the collaborative parties [5; 23].

2.5.3 Floor Control

Floor control is often recognized as one of the most important attributes or requirements for groupware. What it concerns is who takes control of the shared object in a collaborative shared workspace environment. Dommel and Garcia-Luna-Aceves [10] define floor control as a technology that deals with conflicts within shared workspaces. It helps coordinate and mediate access to a shared workspace, by regulating turn taking in conversations or access of controls over shared objects. They further summarize that:

"Floor control encompasses the following tasks: …, granting or denying floor requests according to the enacted service policy, tracking the status of shared resources at all connected sites, relaying floor requests between sites, managing temporary access rights to data, authorizing or denying usage according to session control information, and broadcasting or multicasting changes of floor control states to collaborating members." [11, p. 27]

The benefits of floor control from a user and system perspective are identified by Dommel and Garcia-Luna-Aceves [12]. They state that, from user level, misbehaviors are avoided "by shifting the risk of collisions from data to signaling" (p. 20). From system level, floor control can "orchestrate intelligent allocation of scarce shared resources based on user input" (p. 20). They also point out that social and cultural issues of communication may be involved in floor control deployment.

Dommel and Garcia-Luna-Aceves [11] particularly address the significance of floor control for "tightly coupled" sessions, which often involve ample interactions and collaborations. They further point out that floor control can either be deployed by human moderators or by the system through using prediction, filtering and reservation.

2.6 Combination of video and shared work space

Overall we are trying to achieve a low cost system that supports VCoIP and a shared workspace that would enable more natural and effective communication and also enhance the feeling of sitting around a (virtual) table. Not surprisingly, some researchers questioned the value of such systems. Masoodian and Apperley [27] carried out an experiment which compared 4 environments: face-to-face, video with shared workspace (full motion), video with shared workspace (slow motion) and audio only. They found that adding a visual channel to an existing voice channel does not increase the sense of co-presence and the meetings supported by an audio link create the same sense of co-presence as face-to-face meetings, when shared workspace software is used. Based on their result they concluded that "little if anything is to be gained by adding video communication when using shared work-space CSCW systems" [27]. However there are three aspects of their evaluation we would not agree with. Firstly, they chose a problem solving task which we believe, to a certain degree, impairs the needs for eye contact. Secondly, the design of their system does not suggest a fair distance between the shared workspace and each user which further impairs the needs for communication. Thirdly, their conclusion was mainly drawn from verbal communication analysis. Results from visual communication analysis and user satisfaction analysis are equally important which might show a different result on the sense of co-presence. On the other hand, Buxton [7] supports the combination of video conference and shared workspace. According to him by using a video "…, one can at least maintain an awareness of who is present and get a general reading of their body language, for example. The absence of checks like, 'Are you still there Marilyn?' that are characteristic of telephone conferences is an example of what video contributes to maintaining a sense of personal presence" [7]. He also discovered an interesting pattern which is as one person can not look at the shared object and the collaborator's face at the same time, the eye contact was established especially when testers were initially negotiating how to proceed and at the end when checking results. At other times the video is supplemented by the voice channel. However he further addressed that the reason such a system is so effective is because the methods and overhead in switching contexts (such as from computer screen to eye contact) had the same overhead and action as is used in analogous work-a-day tasks. Finally he concludes that the integration of these two types of spaces (video and shared workspace) is important. In order to achieve a natural flow of interaction the smoothness of transitions between them is critical.

3. DESIGN AND IMPLEMENTATION

3.1 Software

As indicated earlier, our system consists of two major parts: the VCoIP part for communication and the shared (web browser) part for information sharing and collaboration. This section describes the choices of software solutions for both parts and the integration into the final system. We are providing a rather detailed description of the implementation because this is needed to successfully replicate our prototype. Also, it highlights the actual real-world difficulties in setting up a system, which does not follow standard hard- and software requirements.

3.1.1 VoIP Solution Evaluation and Implementation

The VCoIP part of this study focuses on finding a solution that satisfies the following requirements. Firstly, the solution has to be developed on an open standard such as ITU's H.323 or IETF's SIP standards. Secondly, it has to be a Microsoft Windows operating system based solution. As Microsoft still has the largest market share in PC operating systems, a communication and collaboration tool should target as many users as possible. Thirdly, there should not be any critical bugs existing in the VCoIP system which prevents the whole system functioning, or at least the system must be able to be fixed up to an acceptable level in order to be integrated into the system. Lastly, the solution has to be successfully built in the current software development environment.

As VCoIP is a branch of the VoIP industry, the investigation of the communication part of the IC3 starts from exploring the potential VoIP solutions. It is found that many of the current open source VoIP solutions fail to meet the above requirements in order to be integrated into the system. Problems encountered are summarized as follows:

1. Platform incompatible. Much current open source VoIP software only runs on an operating system other than Microsoft Windows. For example, Ekiga, known as the second version of Gnomemeeting, is a very well developed and supported open source VoIP solution, however it only runs on X windows OS. Other examples include Kphone and Linphone.

2. Stopped developing a long time ago and critical bugs exist. Some open source VoIP solutions are developed as sample applications to demonstrate the early version of the H.323 standard. Thus these projects are either unfinished, finished with limited functionalities or finished with usage problems. Examples of this are openPhone and Cphone

3. Build failure. Some solutions failed to be built by the development tool – MSVS 2005. Examples include openPhone, Freeswitch and Cphone. However this could due to limited knowledge about how to build the solutions. Also it addresses one important open source software problem-lack of documentation.

4. Backwards incompatible. Old version programs such as myPhone are not compatible with the new version of the run time libraries (openH323.dll & PWLib.dll). Furthermore, the related old versions of libraries are not available any more.

5. No video support. Video is seen as an additional feature upon existing VoIP infrastructure. Many VoIP solutions are designed as voice only, therefore do not support video features such as ZAP, sipXphone and YATE.

6. Open source client with proprietary server. This is a particular problem for SIP protocol based VCoIP solutions, because SIP uses a client-server architecture. One example is openWengo.

7. Still under heavy development. There are a couple of open source solutions still under development, therefore desired features are not available yet. Examples are openWngo and YATE.

8. Poorly designed solutions. This is another common open source software problem. For example, VideoNet, an open source VoIP solution, is released with no audio, video or internet related configuration at all.

Because of these problems, this study uses an open source solution called VIS [1]. VIS has been released as a sample application for the demonstration of the VIS H.323 library. It is a MS Windows OS based solution that supports both audio and videoconference through the H.323 standard. It has a basic user interface with sufficient functionalities. More importantly, VIS seems to be more stable than many other VoIP solutions, and can be successfully compiled in Visual Studio 2005. Moreover, VIS has some important features such as audio and video configuration, network configuration, peer-to-peer connection and remote/local video display. The audio codecs VIS supports include: ITU's G.723.1 (default), G.729, G.729A, GSM-06.10, G.711 A-law and G.711 µ-law. The video codec used by VIS is H. 263/261. Despite of these advantages, one shortcoming of VIS is that it has no documentation except the comments in source code. Another limitation is that the VIS application is published as open source but the VIS H.323 library is not. Without purchasing the authentication code, the VIS H.323 library will automatically shut down the video and audio connection after 3 minutes.

One vital problem identified in the implementation of VCoIP is that normal VoIP software is only designed to manage one communication channel, while the proposed system aims to send two audio and video streams (captured by the left and right cameras) to separate parties, as well as receive two audio and video streams from these two parties. This requires two separate communication channels for simultaneous data distribution. In order to achieve this, the VIS application end point has been replicated and further assigned to listen to different network ports. Each client communicates to the other two clients via two separate channels.

3.1.2 Shared Browser Design and Implementation

For the shared web browser part, VNC (Virtual Network Computing) is one of the solutions. VNC is a desktop view sharing and remote control system that uses the RFB (Remote Frame Buffer) protocol. VNC has two parts, a client and a server. The server is the program on the machine that shares its screen, and the client (or viewer) is the program that watches and interacts with the server. In addition, multiple client computers can share the view and control of one server computer simultaneously, thus supporting application sharing in the Computer Supported Cooperative Work (CSCW) style [32]. Since VNC uses a lot of network bandwidth, various methods (encodings) are used to reduce the bandwidth consumption.

There are several reasons to use VNC. Firstly, many commercial products announce that they provide application sharing. They are actually using very similar techniques to VNC – remote computer desktop view sharing and remote control. Examples of these products include: NetMeeting 3.0 (stand-alone application), BeemYourScreen (desktop view sharing application very similar to VNC) and Glance (desktop view sharing application built in web browser).

Secondly, because VNC is a desktop view sharing system, thus the integrated system can share the view of applications other than a web browser. The whole system can be easily extended to provide a collaborative environment for purposes other than web information retrieving.

TightVNC is chosen to be integrated into our system, as it is a very well developed open source VNC system. It provides many useful features such as bandwidth-friendly "Tight" encoding, password authentication, window sharing. A particular GUI component of an application you want to share can even be specified, such as a panel inside a window.

One of the goals of our system is to achieve the display of multiple mouse pointers and to distinguish them, for example in different colors or links to the video window of the user. At present, Windows based VNC services only support multiple clients sharing the same mouse pointer on the server because the standard Windows system provides only one cursor. Users either have to share the concurrent control of the server mouse pointer or use some floor control mechanisms to take turns to control the server mouse pointer.

3.1.3 The Final System

The target environment is accomplished by integrating VIS and TightVNC, and the integration is achieved by modifying the source code. The VIS for VCoIP communication remains a peer-to-peer architecture, and the TightVNC for application view sharing remains in client-server architecture. In order to manage two separate VCoIP communication channels, the VIS part has been replicated and then integrated with either TightVNC server or TightVNC client. Due to the limitation of VNC itself, the telepointers and floor control are not supported (figure 4).

Figure 4: The Final IC3 System (one client view)

Since this study focuses on a low-tech setup, our system requires only the following hardware: a server computer with Microsoft OS, a tag supported web browser application (such as Mozilla Firefox) installed (to be the server), two client computers with Windows OS, two web cameras (supporting QCIF 176 x 144 video format) and one headset for each participant and network connection.

With the developed solution the general goal of providing a three-party conferencing system could be achieved. Due to the low bandwidth components in our setup a reasonably good overall quality could be achieved (e.g. regarding delay). But there are still some major limitations: The goal of integrating telepointers into the IC3 system has not yet been achieved due to software programming detail issues (need significant modifications on TightVNC). The setup of the current system is complex. At present, each VCoIP call as well as client-to-server VNC connections need to be made manually. This process involves finding remote peers' IP addresses, making 2 separate VCoIP calls to each remote peer, setting up the VNC server and initiating VNC connection. A single and simpler system workflow needs to be cautiously designed and implemented for such a system.

4. USABILITY EVALUATION

A preliminary trial is conducted to get a sense for the overall performance and usability. Our system evaluation helps answer questions such as: does this system improve users' gaze contact and increase the feeling of sitting around a table? Is the system an effective communication and collaboration tool?

4.1 System Evaluation Setup

As mentioned earlier, the integrated VIS has a "3 minutes" limitation. Therefore, in order to successfully carry out the trials, a system hi-fi prototype is used. This hi-fi prototype is a combination of the unlimited time version of VIS and TightVNC. In terms of the hardware, the testing environment is set up exactly as described in the former section, except that the tests are carried out in a local network environment.

4.2 Method

For the sake of process and time economy and because of the prototypical stage of our system development we decided for a heuristic evaluation. The original use of heuristic evaluation such as Nielsen's [28] methodology was to evaluate the usability for single user applications. However more recent studies have applied the technique successfully to groupware applications. Baker, Greenberg & Gutwin [2] adapt Nielsen's [28] heuristic evaluation methodology to help inspectors identify usability problems within groupware systems, and their framework has proven to be effective at identifying collaboration and communication usability problems. This study uses a modified Baker et al [2] heuristic evaluation to assess the usability issues for our system. There are, in total, 12 evaluation criteria using 5 unit scales and 1 free style question. The first 8 criteria are directly selected from Baker et al's [2] groupware heuristic evaluation. Because these criteria are general evaluation criteria for groupware, the importance of the criteria is firstly evaluated. From this how the targeting system satisfies these criteria is assessed. The following 4 criteria specifically assess the potential usability issues of the proposed system.

In order to initiate meaningful communication among testers (participants) we have developed a real life scenario for them to carry out which is described as following: Participant A wants to buy a car from an online trading website - trademe (www.trademe.co.nz), and would like his friends (participants B and C) to provide some suggestions. A has a budget of $5000, and could borrow another $1000 from his parents if he really likes the car. Participant B and C will ask relevant questions to A to get more information on his taste of cars. The possible topics that could be discussed are car make, model, color, year, mileage, and other features.

4.3 Procedure

In total nine evaluators conducted the system evaluation. They were split into three groups (three people in one groupasked to follow the given scenario. As the purpose of this evaluation is focused on the effectiveness of the communication therefore each evaluators were only assigned to play one of the three roles in the scenario (either A, B or C). Each of the test session lasts at least 20 minutes but no more than 30 minutes.

4.4 Results and Discussion

First, we made the observation that there are more head turnings and eye contacts at the start and the end of each test session. Beginning midway through the test, the sense of co-presence established by the audio channel performs more effective than the high fidelity video display. This observation confirms Buxton's findings [27].

The results of the evaluation can be categorized into two parts and are analyzed separately. The first part contains the results obtained from Baker et al.'s groupware heuristic evaluation criteria. The second part illustrates the results for 4 more questions, which are specifically designed for this study.

The two most important requirements found by evaluators which are also well satisfied by the system are that it "provides intentional and appropriate verbal communication", and "allows people to coordinate their actions". The importance of these two criteria for the system is agreed upon by almost everyone. Criteria such as that it "has to facilitate gestural communication", "provide consequential communication of an individual's embodiment", and also "provide consequential communication of shared artifacts", are assessed by evaluators as being relatively important. Our prototype system has performed well with regard to these criteria. Some criteria are identified to be unsatisfactorily applied. Evaluators have rated "protection", and "finding collaborators and establishing contact" as important criteria, however they believe that these are poorly satisfied by the current system. The importance of a "tightly and loosely-coupled collaboration" criterion, was unclear. Five evaluators have noted it as relatively important, whereas four has rated it as not very important. However, they agree that the targeting system does not satisfy this criterion well.

Two major questions on the evaluation sheet ask evaluators how well they think the system facilitates eye contact and simulates the environment of sitting around a table. The results are consistent, as most evaluators believe these two criteria are satisfied very well. This indicates that system has performed well as an effective communication tool.

Five evaluators answered the open-ended question. They confirmed that the system does provide good interaction between members, however several problems were raised. A common issue identified by three evaluators is the ignorance of video when people are focusing on the collaborated content. Two evaluators point out that the control over a shared mouse pointer among participants is a vital issue. One evaluator identifies that the sound quality needs to be improved (this could be a hardware problem). Another two mention overlap artifacts of images on the VNC client side when one person rolls the scroll wheel on the mouse (this is a software issue).

5. Conclusion

The overall evaluation shows that the proposed system IC3 is probably a suitable tool for communication and collaboration. It achieves most of the goals from the initial design, such as video/audio communication, a shared web browser, gaze awareness support and the simulation of sitting around a table. However, there is still a gap between the existing system and the expected system. Further studies should focus on achieving a simplified system setup, better telepointing and floor control support as well as adding verbal and visual communication analysis to the usability evaluation.

6. ACKNOWLEDGMENTS

We'd like to thank all our participants in the usability study for their time and effort and Martin Purvis for his academic advice. This work was partly funded by an University of Otago research grant (CALT: JDLJ17096)

7. REFERENCES

[1] 115studio. (2006) Retrieved May 15, 2006 from www.115studio.com

[2] Baker, K., Greenberg, S., & Gutwin, C. (2001). *Heuristic evaluation of groupware based on the mechanics of collaboration.* In: Proceedings of the 8th IFIP Working Conference on Engineering for Human-Computer Interaction. May 11-13, Toronto, Canada.

[3] Barbour, C., & Barbour, N. (2001). *Families, schools, and communities: Building partnerships for educating children.* Upper Saddle River, NJ: Merrill.

[4] Batchelor, J., & Goethals, G. (1972). *Spatial arrangement in freely formed groups.* Sociometry, 35, 270-279.

[5] Bier, E., & Freeman, S. (1991). *MMM: A User Interface Architecture for Shared Editors on a Single Screen.* In: Proceedings of the 4th ACM SIGGRAPH Symposium on User Interface Software and Technology, November, 79-87.

[6] Bocker, M., & Mühlbach, L. (1993). *Communication Presence in Videocommunications,* In: Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, Santa Monica, USA.

[7] Buxton, W. (1992). *Telepresence: integrating shared task and person spaces.* In: Proceedings of Graphics Interface '92, 123-129.

[8] CPNMouse. (2006). Retrieved August 5, 2006, from http://cpnmouse.sourceforge.net/

[9] Dabkowski, D. (2004). Encouraging active parent participation in IEP team meetings. *Teaching Exceptional Children*, 36, 34-39.

[10] Dommel, H-P., & Garcia-Luna-Aceves, J. (1997). Floor control for multimedia conferencing and collaboration. *Multimedia Systems*, 5, 23-38.

[11] Dommel, H-P., & Garcia-Luna-Aceves, J. (1999a). Efficacy of floor control protocols in distributed multimedia collaboration. *Cluster Computing*, 2, 17-33.

[12] Dommel, H-P., & Garcia-Luna-Aceves, J. (1999b). Group coordination support for synchronous internet collaboration. *IEEE Internet Computing Magazine,* Special Issue on Collaboration - Internet-style, 3, 74–80.

[13] Ellis, C., & Gibbs, S. (1989). *Concurrency control in groupware systems.* In: Proceedings of the 1989 ACM SIGMOD international conference on Management of data, 18, 399–407.

[14] Eills, C., Gibbs, S., & Rein, G. (1991). Groupware: some issues and experiences. *Communications of the ACM*, 34, 38-57.

[15] Garau, M., Slater, M., Bee, S., & Sasse, M. (2001). *The impact of eye gaze on communication using humanoid avatars.* In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, March 31- April 5, 309-316.

[16] Gemmell, J., Toyama, K., Zitnick, C., Kang, T., & Seitz, S. (2000). Gaze awareness for video-conferencing: A software approach. *IEEE Multimedia*, 7, 26-35.

[17] Giesen, M., & McClaren, H. (1976). Discussion, distance and sex: changes in impressions and attraction during small group interaction. *Sociometry*, 39, 60-70.

[18] Greenberg, S., & Roseman, M. (1999). Groupware toolkits for synchronous work. In Beaudouin-Lafon, M. (ed.), *Computer Supported Cooperative Work*, Chapter 6, 135-168. John Wiley & Sons.

[19] Gutwin, C., Greenberg. S., & Roseman, M. (1996). *Workspace Awareness in Real-Time Distributed Groupware: Framework, Widgets, and Evaluation.* In: Proceedings of HCI'96: Conference on Computer Human Interaction: People and Computers, 281-298. Springer-Verlag. Conference held at Imperial College, London, August 20-23.

[20] Hayne, S., Pendergast, M. & Greenberg, S. (1994). Implementing gesturing with cursors in Group Support Systems. *Journal of Management Information Systems*, 10, 43–61.

[21] Hendrick, C., Giesen, M., & Coy, S. (1974). The social ecology of free seating arrangements in a small group interaction context. *Sociometry*, 37, 262-274.

[22] Hopf, K., Runde, D., & Borker, M. (1994). *Advanced videocommunications with stereoscopy and individual perspectives.* In: Towards a Pan-European Telecommunication Service Infrastructure - IS&N '94, Kugler et al. (eds.), Berlin, Heidelberg, New York.

[23] Hourcade, J., Iyer, V., & Bederson, B. (1999). *Architecture and Implementation of a Java Package for Multiple Input Devices (MID).* In: Proceedings of User Interface and Software Technology (UIST 99).

[24] Kolbehdari, M., Lizotte, D., Shires, G., & Trevor, S. (2006). Session Initiation Protocol (SIP) Evolution in Converge Cmmunications. *Intel Technology Journal*, 10, 10-19.

[25] Lauwers, C., & Lantz, K. (1990). *Collaboration awareness in support of collaboration transparency: Requirements for the next generation of shared window systems.* In: Proceedings of ACM CHI'90 Conference on Human Factors in Computing Systems, April, 303–311.

[26] Liu, J., Beldie, I., & Wopking, M. (1995). *A Computational Approach to Establish Eye-contact in Videocommunication.* In: the International Workshop on Stereoscopic and Three Dimen-sional Imaging (IWS3DI), 229–234. Santorini, Greece.

[27] Masoodian, M., Apperley, M., and Frederikson, L., Video Support for Shared Work-Space Interaction: An Empirical Study. *Interacting with Computers*, Elsevier Science Ltd, 7, (3), 237-253.

[28] Nielsen, J. (1994). *Heuristic evaluation.* In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods. New York: John Wiley & Sons.

[29] Okada, K., Maeda, F., Ichikawa, Y., & Matsushita, Y. (1994). *Multiparty videoconferencing at virtual social distance: MAJIC design.* In: Proceedings of CSCW'94, Chapel Hill, NC, October, 279-291.

[30] Phillips, W. (1999). *Architectures for Synchronous Groupware*, Technical Report 1999-425, Department for Computing and Information Science, Queen's University, Kingston, Ontario, Canada.

[31] Parsons, G. (1982). *Basics of Communication.* Retrieved October 2, 2006, from www.cedresources.ca/docs/modules/comm.doc

[32] Richardson, T., Stafford-Fraser, Q., Wood, K., & Hopper, A. (1998). Virtual network computing. *IEEE Internet Computing*, 2, 33-38.

[33] Sellen, A. (1992). *Speech patterns in video mediated conversations.* In: Proceedings of ACM CHI'92, 5, 49-59.

[34] Takeuchi, A., & Naito, T. (1995). *Situated facial displays: Towards social interaction.* In: Proceedings of ACM CHI '95- Conference on Human Factors in Computing Systems, 1, 450–455.

[35] Tang, J. (1991). Findings from observational studies of collaborative work. *International Journal of Man Machine Studies*, 34, 143–160.

[36] Tse, E., & Greenberg, S. (2004). *Rapidly prototyping single display groupware through the SDGToolkit.* In: Proceedings of the Fifth Australasian User Interface Conference, volume 28 of CRPIT Conferences in Research and Practice in Information Technology Series, 101–110.

[37] Verteegal, R., Velichkovsky, B., & van der Veer, G. (1997). *Catching the Eye.* SIGCHI Bulletin 29(4), October 1997. Online: http://acm.org/sigchi/bulletin/1997.4/vertegaal.html

[38] Wilson, H. (2004). *Videoconferencing at UNISA: Synchronous real-time discussions for student support.* Paper presented at e-merge 2004 conference, South Africa.

[39] Yang, R., & Zhang, Z. (2002). *Eye Gaze Correction with Stereovision for Video-Teleconferencing.* In: Proceedings of the Seventh European Conference on Computer Vision, 2, 479-494. Copenhagen, Denmark