Synchronous virtual spaces - transparent technology for producing a sense of presence at a distance.

Claus J. S. Knudsen
Media Technology and Graphic Arts, Royal Institute of Technology (KTH),
Drottning Kristinas v.47 D, SE-100 44 Stockholm, Sweden
(Tel: +46-8-790 6042; Fax: +46-8-791 8793; E-mail: clausk@gt.kth.se)

Abstract

As video conferencing is usually being used, the participants are displaying a part of their physical space for each other but do seldom have the feeling of sharing a common "space" as if being in the same room and doing a "handshake". Can the sense of presence at a distance be produced with transparent distance technology? If people treat television and new media like real people and places, it should be possible to produce a non-physical meeting place for people to interact in. Several experiments aiming at this have been carried out and two of them are described in this paper. High quality digital video, audio and data were transformed into light and sound in two installations using broadband communication technology on fiber optic cables. People could walk into the distributed spaces and share a sense of common non-physical space on a large projected display. Content from a computer connected to the Internet could easily add information to the common "space" using a mixer. The main research goal was to explore if the distance technology installation could support a high level of transparency. In addition to this, we wanted to analyze the way people adopted the non-physical space as being real. Recorded video documentation shows that most people intuitively adopted the non-physical space as real and started to interact with people both physically and non-physically present. People experienced a high degree of presence while playing music, dancing, discussing and socializing together at a distance.

1. Introduction

Synchronous Virtual Spaces is a new research effort at the Royal Institute of Technology (KTH), where we wish to explore the possibilities and restrictions of synchronous interaction and cooperation in published virtual spaces. One of the objectives of this research is to produce a sense of presence using distance technology. Presence is defined as the subjective experience of being in one place or environment, even when one is physically situated in another. In earlier times Tele-operators could describe a sensational feeling of being at the remote site rather than being at the operators control station. As applied to the synchronous virtual spaces, presence refers to the sensation or experience of being present in a non-physical space shared with others at a distance [12].

This paper focuses on experiences from two experimental installations for presence production carried out at the Advanced Media Technology Laboratory [2] at the Royal Institute of Technology (KTH) [3], Stockholm.

The first experimental installation was carried out in October 2000 between two physical spaces, the "Provisorium" at one of the KTH Learning Laboratories [4] in Stockholm and the Stockholm Academic Forum [5] situated in a public area in downtown Stockholm [6]. The second, quite similar, experimental installation was carried out in March 2001 between, again, the KTH "Provisorium" and the Stockholm City Hall during a EU minister meeting. At the Stockholm Academic Forum ad hoc visitors and students from the Media Technology program at KTH tested various ways of communicating in the synchronous virtual spaces. In the City Hall [7], physical and non-physical actors were doing a performance together on a "experimental stage". The basic technology being used was based on broadband telecommunication carrying digital audio, video and data and quite simple camera and projection technique. Physical and non-physical participants were interacting synchronously with, for example, music playing, dance, games, discussions, teaching and learning. The synchronous virtual spaces were transmitted on the Internet in real time and visitors with video and audio from the Internet could join the virtual spaces. To achieve a qualitatively good sense of presence, people could establish face-to-face contact at a distance different from virtual reality (VR) applications using helmets, glasses [8], avatars [9], animated heads or stereoscopic cameras on top of remote controlled robots [10].

The name used for this installation is synchronous virtual spaces but in the documentation on the Internet [1] you may also find names like communicative spaces or ePresence meaning the same type of installation.

2. Methodology

To measure the experienced degree of presence is not easy. Many factors like the installation characteristics, individual preconditions, sensory environment and content characteristics influences the attention to the mental world. The human factor is essential. According to Sheridan (1992) [11], presence is a subjective sensation or mental manifestation that is not easily amenable to objective physiological definition and measurement. He indicates that "subjective report is the essential basic measurement" (Sheridan, 1992, p. 121). Witmer and Singer have been doing research on the measuring of presence in virtual environments. They presented a Presence Questionnaire (PQ) in the MIT Presence Journal June 1998 [12]. In addition to this they developed an immersive tendencies questionnaire (ITQ) to measure differences in the tendencies of individuals to experience presence. The term "immersive" often refers to certain types of sensory reproduction systems used in VR and telepresence were the user actually becomes part of the experience to the exclusion of their immediate reality, as opposed to being a mere observer. For example, a user using a head mounted display or CAVE [8] system would be immersed in the experience, whereas someone viewing a remote location on a simple computer monitor would not be.

In this paper, the subjective report method (Sheridan) will be used to analyse the empirical recorded video material. I will also describe the installation developed for the study. The analyses will be based on the model described by Enlund. [13].

3. Earlier research

Previous to these case studies, a set of artistic installations had been carried out to explore distributed virtual spaces. These experiments include a test installation in a course in "Physical Computing" at CID, Center for User Oriented IT Design [15], February 1998, performance installations in the project "Cave Experience" [16] at the "R1, KTH Experimental Scene" [17] May 1998, the mobile performance installation and concerts in the Spindelvev@ project [18] for the Norwegian Concert Institute [19], 1999 – 2000 and a performance installation at the TET1999 conference [19] in Gjovik, May 1999.

The installation and test case study "Synchronous virtual spaces" described in this paper was based on results from earlier research works. Still the main concept was based on the "invitation installation" in the Spindelvev@ project [18] applied with broadband technology to support the creation of a sense of presence and reality at a distance using data, video and audio streaming on fiber optics.

4. The "synchronous virtual spaces" installation design

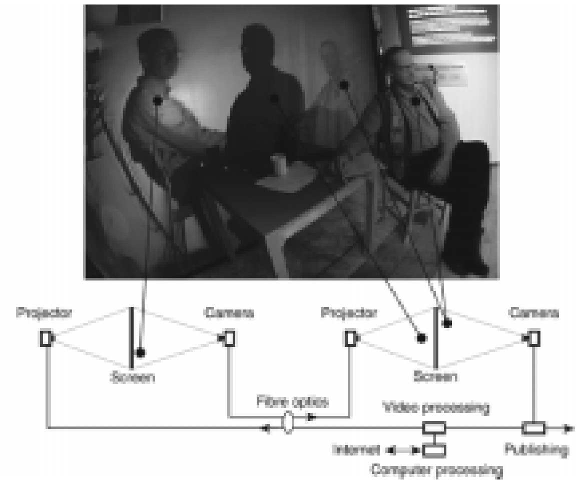

Figure 1. *Design diagram for the synchronous virtual spaces.*

The diagram in figure 1 visualizes the basic design principles in the "synchronous virtual space" installation. The main goal of the design was to develop a installation supporting synchronous intuitive interaction and cooperation in published virtual spaces using transparent technology. As Wan and Mon-Williams (1996, p845) [20] state, in such cases " the goal is to build virtual environments that minimize the learning required to operate within them but maximize the information yield".

The chosen solution was based on simple technique in a fixed calibrated video loop. Each physical space had a similar combination of equipment installed, consisting of projector, screen and camera. The two physical spaces were connected with broadband technology on fiber optic cables. A video processing unit was installed as a active storytelling tool for manipulating video and adding information from a computer connected to the Internet. The publishing unit transmitted the activity in the synchronous virtual spaces to different external audiences on the KTH campus and on the Internet. These audiences had the possibilities to communicate back to the actors in the installation by using audio, chat and video on the Internet and on intranet.

4.1 Continuously streaming video – the "video-loop"

The main design principle for the installation functionality was to develop "a continuous stream of video running through two physical spaces" using video input sensors such as cameras and video output devices such as projectors combined with projection screens. The large back projection screen (2,30 x 3 meters) was placed in the center of the rooms with 6 meters free space on each side. On the one side the video projector was mounted 6 m from the screen, projecting the incoming video on to the screen. The space between the projector and the screen will from now on be called the “projector space” (PS). On the other side of the screen, the camera was mounted 6 meters from the screen for framing the screen. This space between the camera and the screen will from now on be called the ”framing space” (FS).

The cameras were framing the pictures projected on the screens in the two physical spaces and the video information was sent to the projectors at the distant location using fiber optic transmitters and receivers. This high quality video link generated a continuously flow of video through the two physical spaces. We may say that the installation consisted of a “video loop” (VL). An element, someone or something, placed in the projection spaces (PS), would immediately appear as a shadow on the screens both locally and at a distance. Likewise would an element in the framing space (FS) immediately appear on the local and distant screens. The installation was passive, that means that no-one had to serve the basic technique after it was started.

4.2 Telecommunication and installation calibration

A serial data interface (RS 232) from a personal computer with an open software protocol made it possible to remotely control camera treatments such as tilt, pan, zoom, and technical adjustments, such as iris, white balance, black/white level and gain on the camera at remote site. It was also possible to add more remote controlled facilities such as light control, power switching etc. This was very useful, because the installation had to be calibrated by adjusting the cameras and projectors in the correct position, choosing the correct lenses and adjust the video signal variables for an optimal picture in the loop.

The two-way audio-, video- and serial data communications were modulated directly onto a multi-modus optical fiber using 1300 nm and 1500 nm wavelengths. The optical fiber also had other external traffic on the same cable, but of course at other wavelengths. The speed of the data stream was about 200 Mbps and had “broadcast quality” standard. Because of the small delay in the video-loop and the audio communication, (less than 10 milliseconds), it was possible to play music, dance and act together in the installation.

4.3 Multidirectional stereophonic audio

A local and a distant bi-directional sound installation were developed consisting of unidirectional microphones and amplified speakers. The stereophonic sound environment should help the participant to orientate to the non-local participants and vice versa. A mix of the two directional audio feed was done for the publishing of the activity and the documentation video recording on tape and on the Internet. In addition to this a public address loudspeaker system was installed on the experimental stage in the Stockholm City Hall to reach the physical audience with sound from the installation in the big hall, see figure 2.

Figure 2. *Physical and non-physical participants communicating on the experimental stage in Stockholm City Hall.*

4.4 Transmitting and recording for the Internet

The video loop passed through a net-based remote controlled audio-video matrix for distribution on KTH campus, recording on videotapes and transmitting on the Internet. The video loop and a mixed audio were transmitted on the Internet through a Polycom [21] stream station at Gjovik College in Norway using the MPEG 4 standard. A Tandberg 6000 system [22] was feeding the broadcast stream station at a distance through the high quality H.263 protocol, also on the Internet. The stream station was automatically recording the event for immediate playback on the Internet also using the MPEG4 standard. Good reports reception was received from the US and from several European countries.

4.5 Lighting for human interaction on projected screens

When lighting for human interaction on projected screens using video was carried out, three variables were critical, all dependent on each other. The three variables were the luminance from the screen, the sensitivity of the camera sensor, and the artificial and natural ambient light in the space. These three variables had to be adjusted independently for optimal results.

In the experiments, spots lighted up the space between the cameras and the screen. The artificial light with fixed color temperature was strictly intercepted from “hitting” the screen. It was also necessary to calibrate the positions of the cameras and projectors in the loop for minimum virtual “echo” effect. In addition to this, one of the cameras had an alternative preset position to actually generate echo effect for artistic use in the distributed installations, see figure 4.

4.6 Tools for (inter) active storytelling and art effects

The installation could run without active participation of a technician or a producer. Still, possibilities were there to actively use a set of tools to plot a “better story” for the physical and non-physical participants and audiences.

The storytelling “tools” consisted of a video mixer, a computer and a video effect generator connected to the video loop. These tools could add video sources such as software, images, text or communication from/to the Internet as a part of the video loop or synchronous virtual spaces in the installation. Video effects, such as wipes appeared like virtual curtains on the screen and limited the interactive space for virtual presence production.

Figure 3. *A person at a distance watches a video clip of two recorded dancers played back from a local personal computer.*

People could easily enter the local and remote projection spaces (PS) and the framing spaces (FS) by just walking into them. As the participants or visitors entered these four spaces they became part of the same video loop as their pictures appears on the screens, again both locally and at a distance as mentioned before. In *figure 2*, you can see four physical and two non-physical visitors having a discussion on the experimental stage. The background for the spontaneous discussion was a non-physical participant on the stage asking people to enter the stage for a handshake and a chat. People intuitively did so.

In *figure 3*, you can see a picture taken from a local point of view. The entire computer screen from a local computer is mixed into the video loop and we can see a person at a distance looking at a video playback window on the desktop. In figure 4 an external photographer is documenting an improvised dance between a physical and a non-physical dancer in the installation. A video effect generator tool and the video mixer are being used to create the artifacts combined with small preset adjustments on the camera to achieve visual “echo” effects, as mentioned in chapter 4.5. This is an example of using the storytelling tools actively for artistic performance.

Figure 4. *Physical and non-physical actors performing together in the installation.*

4.7 Scalability

In the projection and the framing spaces, the participants could scale the size of their virtual presentation by adjusting the distance from the screen. This scalability was also used to enlarge objects like illustrations or handheld dolls for the audience at a distance. A lot of creative use was explored. Monitors behind the cameras and on each side of the camera spaces helped the users to orientate and establish face-to-face contact at a distance.

5. The recorded documentation

The recorded material from the experiments shows a multi-modal interaction between the participants in the installation. Four selected scenes will be described.

In one scene we can see male students on the installation screen calling out to some girls passing by. They stop turn a round and enter the framing space without hesitating. They start to mingle, four non-physical students at the Campus and four physical girls at Stockholm Academic Forum. One of the girls looks behind the screen to see if somebody is there. The male students and the girls shake hands and touch each other, treating the virtual person as being real. They laugh and sit close together, experimenting with the new experience.

In a second scene, two students, one physical and one non-physical, agree upon dancing together in the installation without music. They can see each other but the remote student at Campus acts as “master” and the other as “slave”. Something goes wrong during the dance and they immediately stop and turn around to look into each other’s faces. After arguing for a short while and showing each other the right steps they continue with the dance and this time synchronously without problems. A physical audience watches the two dancers right outside the framing space on the local and the remote side.

In a third scene, two students play jazz together at a distance. The saxophone and guitar player are watching each other carefully on the screen to share the same tempo and emotional musical expression when playing together. Some of the music is played “ad lib”, and the extremely sensitive communication between the musicians at a distance is supported by the perception of their body language on the screen.

In a fourth scene we can see six persons having a meeting. Two persons are in the local and remote projection space (the shadow space), three persons are sitting in the remote framing space and one person is joining the meeting from the local framing space.

6. Results and analysis

The main goal of the experiments was to see if the user using the installation could be immersed in the experience of sharing a synchronous virtual space with someone at a distance, to the exclusion of their immediate reality [12]. A major method was to produce the highest degree of presence as possible for the users acting at a distance by using an immersive transparent distance technology. Enlund [13] is pointing out three major factors for producing a sense of presence and reality experienced by individuals. These three factors are the sensory environment, the individual preconditions and the content characteristics.

For both of the experiments, one of the spaces was public. The sensory environment at these places became crucial for the users because of the loud background noise level and the unpredictable light sources. Another problem of the immersive transparent distance technology being used, was the quality of the picture on the screen and the audio level from the bi-directional sound system. Some of the problems were caused by audio “feedback” problems, visual echo effects and limited level of luminance from the projectors being used. Despite of all these negative factors, the users immersed in the experience of being together with non-physical persons and objects by acting in the installations with natural excitement. Two reasons for the limited consequences for the sense of presence could be the human psychological acoustic phenomena filtering out the desired sound source and the excitement of acting in a new media environment. The audience observed the users, acting both physical and virtual, in the installation. These observers on the Internet or at the KTH Campus included their own immediate reality comparable to watching a movie on the television screen. For those physical present, surrounding the installation, the performance could look a bit surrealistic as people were gesticulating and communicating naturally with a screen.

The student individual preconditions for acting together at a distance were positively influenced by the fact that they all ready knew each other well from “building trust” through physical presence experiences. Those students with experiences from acting on a stage also had the greatest benefits when using the immersive transparent distance technology. They seemed to have no problems with suspending their disbelief when acting together with non-physical persons at a distance. In the City Hall, the stage became a double barrier for sensing presence at a distance because the visitors were afraid of exposing them selves negatively in front of the physical audience and at the same time exposing them self to the unknown persons in the published virtual reality. This important individual negative factor for sensing presence was reduced at Stockholm Academic Forum and at the KTH laboratory, because there were no stage and the installation became a natural part of the physical environment. Here the documentation video shows individuals talking with each other and shaking hands in the virtual space without knowing each other. The youngsters immersed more naturally in the experience of sharing a synchronous virtual space elder ones. These interesting differences should be further investigated.

The programs, performances and events being produced during the four days of experimental activity had a wide range of content characteristics. Those students acting in the installation for a longer period of time learned the installation possibilities and used them actively when improvising in the immersive transparent distance technology. The distance theater using handheld dolls and the games, combining non-physical and physical performers, were a powerful agent for sensing presence and reality. Also the distance music play, some times with dancers, created a high degree of presence. The more frequently they were acting, the more specialized they became in manipulating the virtual domain as being real. Some of the students even became specialized and succeeded a number of times to attract attention and visitors into the distributed spaces and even produced games at a distance by using simple tools. This knowledge was tacit and based on "learning by doing".

7. Conclusions

Enlund [13] has argued that we can use quite simple media technology to produce a sense of presence that can be delivered over distance in time and space. This is also supported by the studies done by Nass and Reeves [23]. The subjective analyse of the synchronous virtual space installations confirms the theory, that quite simple technology can achieve a high degree of presence and reality at a distance.

8. Future work

After analyzing the experimental installation the following development possibilities came up:

- Better and larger screens for double sided display
- Wireless remote controlled storytelling tools
- Open flexible networking monitoring systems, camera treatments and light systems
- Easy to use telecommunication systems for video loop below 10 mS delay.
- New analyses of the users using PQ and ITQ questionnaires [12].

9. Acknowledgments

I want to thank the creative media technology students at KTH and my colleague Leif Handberg for management and support, my colleague Hilding Sponberg for his support in Norway, and the innovator Mats Erixon and his staff, at the Advanced Media Technology Laboratory, for creative communication and distance technology solutions. I also want to thank Hans Larsson for practical solutions and Tibor Barany [24] for the excellent documentation on the Internet. The research work could not be possible without the continuous support on many levels from professor Nils Enlund, thank you.

10. References

[1] The synchronous virtual space installation, URL:http://www.r1.kth.se/epresence/

[2] The Advanced Media Technology Lab., URL:http://www.amt.kth.se

[3] Royal Institute of Technology, URL: http://www.kth.se

[4] KTH Learning Lab., URL:http://www.learninglab.kth.se/

[5] Stockholm Academic Forum, URL:http://www.stockholmacademicforum.org/ram.html

[6] Stockholm City, URL:http://www.stockholm.se/english/index.htm

[7] Stockholm CityHall, URL: http://www.stockholm.se/cityhall/

[8] KTH, the PDC CUBE, URL:http://www.pdc.kth.se/projects/vr-cube/

[9] CID, Connected Communities, URL: http://cid.nada.kth.se/cc

[10] Transparent Telepresence Research Group (TTRG), URL:http://telepresence.dmem.strath.ac.uk/index.htm

[11] Sheridan, T. B. (1992). Musings on Telepresence and Virtual presence. Presence: Teleoperators and Virtual Environments, 1(1), p120-125.

[12] MIT Presence Journal Volume 7, Number 3 · June 1998, Measuring presence in virtual environments: A Presence Questionnaire by Bob G. Witmer and Michael J. Singer.

[13] Enlund, N.: "The Production of Presence - Distance techniques in Education, Publishing and Art", ACS'2000 Proceedings, Szczecin, 2000, pp. 44-49.

[14] Handberg, L., Knudsen C., Sponberg H.: "New Learning modes in the production of presence – distance techniques for education", Proceedings of the 20th World Conference on Open Learning and Distance Education, ICDE01, Düsseldorf, 2001, on Compact Disc.

[15] CID, Center for user oriented IT design URL:http://www.nada.kth.se/cid/

[16] Knudsen, C.: "The cave experience: People and technology in an experimental performance space", in Roger, J.-Y. et al (eds.): Business and Work in the Information Society: New Technologies and Applications, IOS Press, 1999, pp. 931-938.

[17] R1, KTH experimental scene: URL: http://www.r1.kth.se

[18] Knudsen, C.: "Interaction between musicians and audience in a learning process on the Internet", ISTEP 2000 Proceedings, Kosice, 2000, pp. 159-164.

[19] Norwegian Concert Institute, URL:http://www.rikskonsertene.no

[20] Wann, J., & Mon-Williams, M.: "What does virtual reality NEED? Human factors issues in the design of three-dimensional computer environments", International Journal of Human-Computer Studies, 44, 1996, pp. 829–847.

[21] Polycom, URL: http://www.polycom.com

[22] Tandberg, URL: http://www.tandberg.no

[23] Reeves, B., Nass, C.: The media equation, Cambridge University Press, New York, 1996.

[24] Orbit as, URL: http://www.orbi.se/