Ontario Telepresence Project

FINAL REPORT

INFORMATION TECHNOLOGY RESEARCH CENTRE
TELECOMMUNICATIOSN RESEARCH INSTITUTE OF ONTARIO

MARCH 1, 1995

Contact: John Chattoe, (416) 978 7203
Peter Leach, (613) 592 9211
Ron Riesenbach, (416) 482 6698

Executive Summary

The Ontario Telepresence Project (OTP) was a three year, $4.8 million pre-competitive research project whose mandate was to design and field trial advanced media space systems in a variety of workplaces in order to gain insights into key sociological and engineering issues. The OTP, which ended December, 1994, was part of the *International Telepresence Project* which linked Ontario researchers to counterparts in four European nations. The Project's major sponsor was the Province of Ontario through two of its Centres of Excellence -- the Information Technology Research Centre (ITRC) and the Telecommunications Research Institute of Ontario (TRIO).

The Ontario Telepresence Project, was a tight partnership of academic and industry researchers including faculty, students and professional staff from Engineering, Computer Science, Psychology and Sociology from the University of Toronto and Carleton University as well as staff located at Industry Partner sites. To gain first hand experience with the media spaces being prototyped, all aspects of work were conducted without regard for geographic location of the project participants. Experimental versions of a media space system, employing a variety of communications protocols (including ATM) were used to link collaborator's desk-tops and conference rooms. Versions of the systems prototyped were also deployed at arms-length sites to study implementation issues from a sociological and psychological aspect.

In all, the project levered $4.8 million in cash and in-kind resources over 3-years of which almost $2 million (41%) was contributed by Ontario firms. Over the course of the three-years, twelve industrial partner personnel received extensive training/experience through day-to-day interaction in the advanced research laboratories of the project. Of the seven Ontario firms actively partnered with the project, three were actively commercializing products based, in part, on the insights gained through their partnership; Applied Silicon International ("Video Vise"), Arnott Design Group ("Active Desk"), and Corel Corporation ("Corel Video"). Details of these products, the scientific accomplishments and the implications for future computer supported collaborative work products and services are presented in the report.

SECTION 1: MANAGING DIRECTOR'S REPORT

1. MANAGING DIRECTOR'S REPORT

Ron Riesenbach

1.1. MISSION

The mission of the Ontario Telepresence Project was to conduct user-centered, pre-competitive research into the design of computer/video supported collaborative work systems - providing Industrial Partners with valuable outputs which they may use to market new products and services.

1.2. PROJECT OUTPUTS

The outputs of the project manifested themselves in four ways (Figure 1-1):

1. Scientific Research Leadership
2. Technology Transfer from Academia to Industry
3. Professional Development of those affiliated with the project
4. Knowledge and lessons gained through field trials, prototypes and test-beds

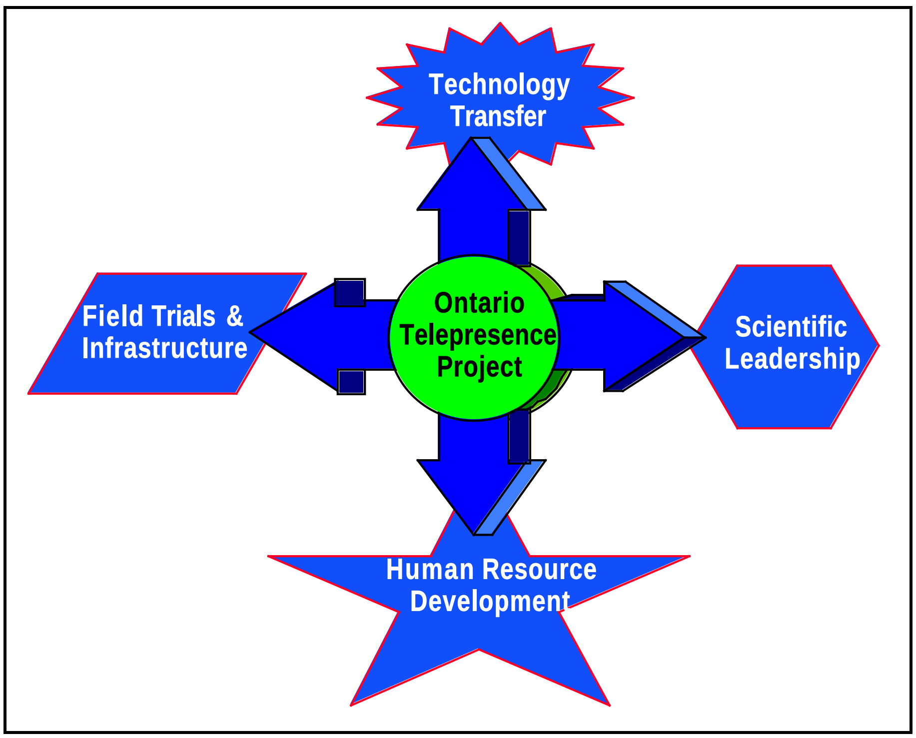


Figure 1-1 Outputs of the Project

Performance measures for each of these four outputs are presented below and in the sections of this document containing the reports of the Scientific director, Engineering Director and Chief Social Scientist.

1.3. Background

The Ontario Telepresence Project spawned from a 1990 proposal from ITRC and TRIO to Technology Ontario. The proposal called for the creation of a multi-disciplinary project to explore the application of user-centred design techniques to Computer Supported Collaborative Work.

Shortly after submitting the proposal, an opportunity arose to expand the scope of the project by linking it with initiatives in four European countries (the "4-Motors of Europe"). A conference was held in Toronto[i] with government and scientific delegates from France, Germany, Spain and Italy were details of cooperation were outlined and agreement in principle was reached to collaborate. Project funding ($2.6M) was approved in the fall of 1991 and the project commenced on January 1, 1992.

1.4. Management

The Scientific Co-Directors of the project, at it's inception were Bill Buxton, Computer Science, University of Toronto and Dr. Morris Goldberg, Electrical Engineering, University of Ottawa. Morris Goldberg resigned as Principle Investigator in April of 1992 to take a full-time position at Eurecom, in the south of France. Bill Buxton remained the sole Scientific Director from that point on.

To provide managerial leadership, Ron Riesenbach was seconded from ITRC to assume the role of Managing Director in February of 1991. Drs. Gale Moore and Gerald Karam were brought in to oversee the Sociological and Engineering aspects of the project, respectively.

Financial, human-resource, industry/government interactions and other operational decisions associated with the project were made by Ron Riesenbach. Scientific direction and strategy were set by Bill Buxton. Policy was set by a committee chaired by the president of ITRC President (John Chattoe), the president of TRIO (Peter Leach), the Managing Director (Ron Riesenbach) and the Scientific Director (Bill Buxton). Gerald Karam (and prior to him, Morris Goldberg) had supervisory responsibilities for Ottawa staff and facilities. Bill Buxton had supervisory responsibilities for Toronto technical staff.

1.5. Structure

The project was administratively organized around a set of linked contracts (Figure 1-2). ITRC was "prime" contractor with the province and assumed overall responsibility for the performance of the project. TRIO was contracted to be the main sub-contractor. Each of ITRC and TRIO then negotiated further sub-contracts with the University of Toronto, the University of Ottawa and Carleton University. Industry partners negotiated contracts with either ITRC or TRIO as appropriate for the kind of interaction they sought to pursue.

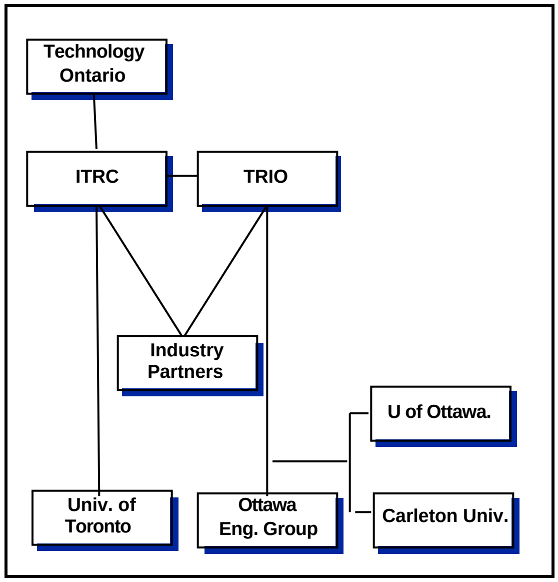


Figure 1-2 Contractual Structure

1.6. *Staffing and Facilities*

Research was conducted at a number of sites in Toronto and Ottawa:

- Computer Systems Research Institute (CSRI) and Faculty of Sociology, University of Toronto - 4.5 full time equivalent staff positions as well as the offices of the Managing Director, the Scientific Director, 5 affiliated faculty and their graduate students and laboratories[ii].
- Ottawa Engineering Group (OEG) located at Hewlett-Packard - Queensview Drive, Ottawa - 4 5 full time equivalent staff occupying 3600 sq. feet of laboratory space.
- Electrical Engineering, University of Ottawa - Dr. Morris Goldberg and Dr. Ahmed Karamoush maintained offices and research activities which supported a number graduate students and facilities[iii]
- Department of Psychology and Department Of Systems and Computer Engineering, Carleton University - Dr. Richard Dillon and Dr. Gerald Karam maintained offices, graduate students and facilities in these two Departments[iv].
- Bell Ontario Network Services, Trinity Square – Mr. Ken Schuyler, Bell's representative to the project in 1993, maintained an installation of prototype Telepresence equipment in his office for local experimentation and research.

The sites with the largest complement of research staff were CSRI and OEG. The OEG staff consisted primarily of systems engineers while the CSRI staff consisted primarily of applications developers. Initially, all Ottawa research staff reported to Morris Goldberg (and later, Gerald Karam) while all Toronto research staff reported to Bill Buxton[v].

To accommodate the distribution of researchers over a number of sites, project staff rapidly prototyped and deployed an advanced communications infrastructure – the Telepresence Media Space (TMS)[vi]. Once the TMS was in place (late 1992), reporting structures were adjusted so that staff were organized along research areas (Field Studies, Engineering and Applications) regardless of their geographic location (Figure 1-3).

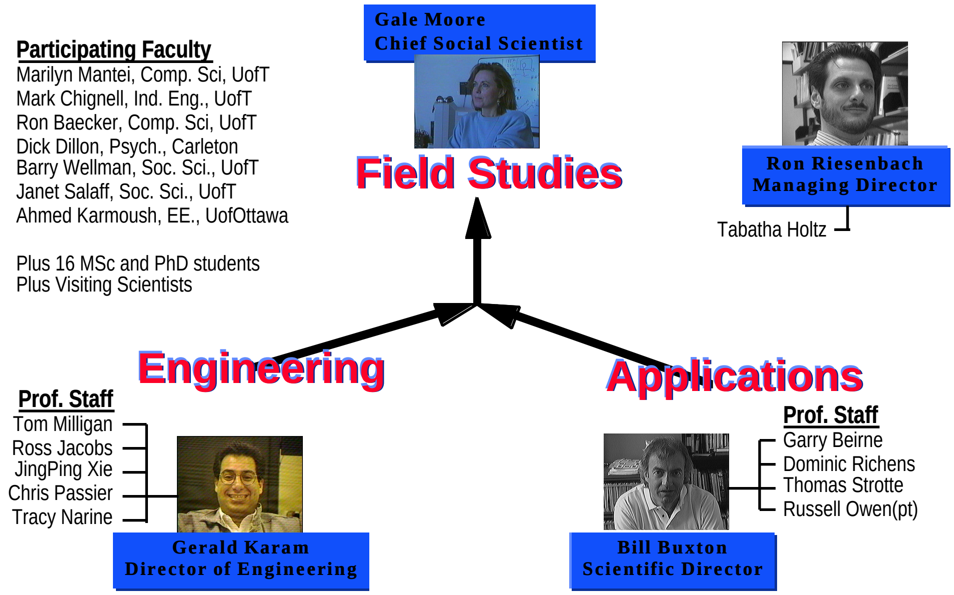


Figure 1-3 Three Research Areas

1.7. SCIENTIFIC LEADERSHIP

The project distinguished itself as a world focal point of research and practical experience in advanced systems and methodologies for the support of time/space separated collaborating groups. New models for the design and implementation of collaborative work environments were developed including the notion of *Ubiquitous Media* - new framework for human-centred distance collaboration system design.

New software architectures were designed, built, deployed and studied in arms-length work environments. The Telepresence Media Space and the Network Software Infra-structure (NSI) were the foundations of many of the prototypes built and studied. The NSI was the first novel use of ITU Generic Conference Control model - an emerging standard of the world telecommunications body. The project's software/systems are cur-rently in use in 12 advanced research labs Internationally including Xerox Rochester and PARC laboratories, USA; Univerisity of Flinder, Austrailia; Eurecom University of Paris – SUD, France; CEFRIEL, Italy; UBC andU. Calgary, Canada and IRPEACS/INRS, France.

While many groups around the world were building laboratory prototypes. the Ontario Telepresence Project was the only group in the world conducting systematic arms-length studies of the effect of media-space systems on organizations. The studies included an ex-amination of the sociological and organizational impacts of group collaboration systems.

Distance education experiments were also conducted. The Ontario Telepresence Project conducted the world's first Transatlantic graduate-level course taught through advanced video/data conference systems (1993) in a link between Toronto, Ottawa and Eurecom in the south of France.

The project's scientific output, as measured by scholarly publications and information dissemination, is given in Figure 1-4:

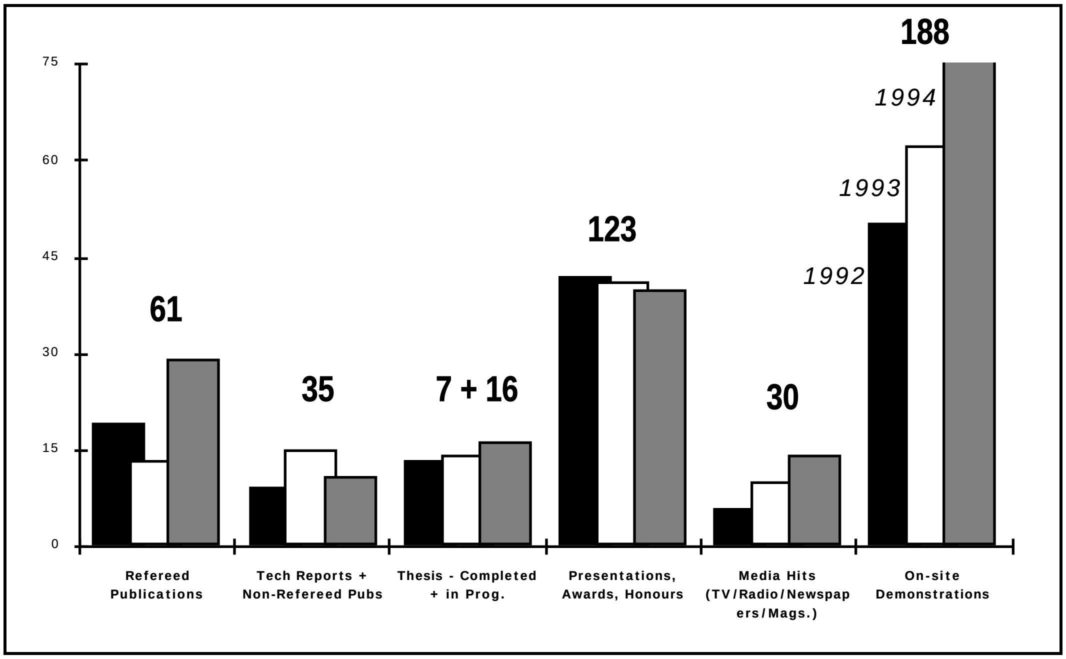


Figure 1-4 Scientific Output Measures

Details on the Scientific outputs of the project and on the Field Trial results are presented in the following three sections of this report.

1.8. INDUSTRY PARTNERSHIP AND TECHNOLOGY TRANSFER

The Ontario Telepresence Project was a partnership between academic researchers and a number of Ontario-based companies to share knowledge and technologies gained through the course of the research. There was a commitment by the Project and the companies to support each other's interests in the research. Firms joining the Project were called “Industrial Partners” according to the definition in the “International Strategic Framework” document struck in negotiations with the Four Motors.

OTP was unique in the way that it orchestrated collaboration between academic and industry participants. A precondition to industrial participation in OTP was that all Industry Partners commit people to the Project in addition to cash and other in-kind support. Experience has shown that technology transfer between academic researchers and industry professionals happens best through an ongoing process of bi-directional interchange and learning. Project management firmly believed that it was through the active participation of individuals from partner firms that company-specific benefits can be identified and extracted from the on-going work.

The benefits of partnership are shown in Figure 1-5.

BENEFITS AVAILABLE TO INDUSTRIAL PARTNERS

1. Education, Information and Expertise
    - Bi-directional exchange of know how through active participation
    - Personnel attend, participate and report on all key conferences world-wide
    - Project attracts visits by world's leading experts
    - Regular seminar series and working group meetings
    - Early access to scientific publications, journals and technical report series
    - Annual report, annual general meeting
    - Relationship with European Partners provides intelligence about world-wide activities
    - Video, computer & paper library of key internal and external research work and results
2. Demonstrations of and Experimentation with Advanced Systems
    - Dedicated demonstration facilities at UofT and OEG for prototyping of new systems
    - Field trials provide hands-on experience with advanced products & prototypes
3. Technology Acquisition, Intellectual Property
    - Intellectual property policy enabling and encouraging technology transfer to Partners
    - Build prototypes and deploy in Industry Partner sites
4. Access to Experts
    - Recruitment of graduating Masters and PhD students in Industry partner sites
    - Opportunity to involve leading academics in industry lead projects
5. Alliance with European and Industrial Partners
    - Build business alliances with other Partners having complementary skills

Figure 1-5 Industrial Partnership Benefits

Some Industry Partners chose to have their researchers reside on campus with the university researchers. Others chose to equip company resident researchers with Telepresence prototype systems to enable collaboration to take place with their employees remaining at their home offices. At it's conclusion in 1994, the project had seven industrial partners

Industrial Partners as of Dec. 31, 1994	Other Contributing Firms	
1. Applied Silicon, International, Ottawa	Adcom Electronics Ltd	Newbridge Networks Corp.
2. Arnott Design Group, Toronto	3-D Interactive Tech.	Nippon Tele. & Teleg
3. Bell Canada, Toronto	Alias Research, Inc.	Sun Computers
4. Corel Corporation, Ottawa	Apple Computers	Hitachi Corporation
5. Hewlett Packard, Toronto	Bell Northern Res.	Worldlinx/MediaLinx
6. Object Technology International, Ottawa	Digital Network Services	IBM Canada Labs[vii]
7. Xerox Corp., Toronto and Palo Alto, Ca.	Dave Dunfield & Assoc	

and 12 other firms who contributed to and benefited from the research (Figure 1-6):

Figure 1-6 Industrial Partners and Other Contributing Firms

The strategy for technology transfer followed the rational that firms will have the most success adopting outputs which were based on sound theoretical foundations, prototyped, deployed and whose utility was demonstrated. The project did not restrict the definition of outputs to software and hardware. Significant outputs of the project included trained graduate students, methodologies for technology introduction and adoption in organizations, education, field trail results and other non-tangible but highly valuable outputs. Figure 1-7 pictorially shows the strategy and process of technology transfer used in the Ontario Telepresence Project.

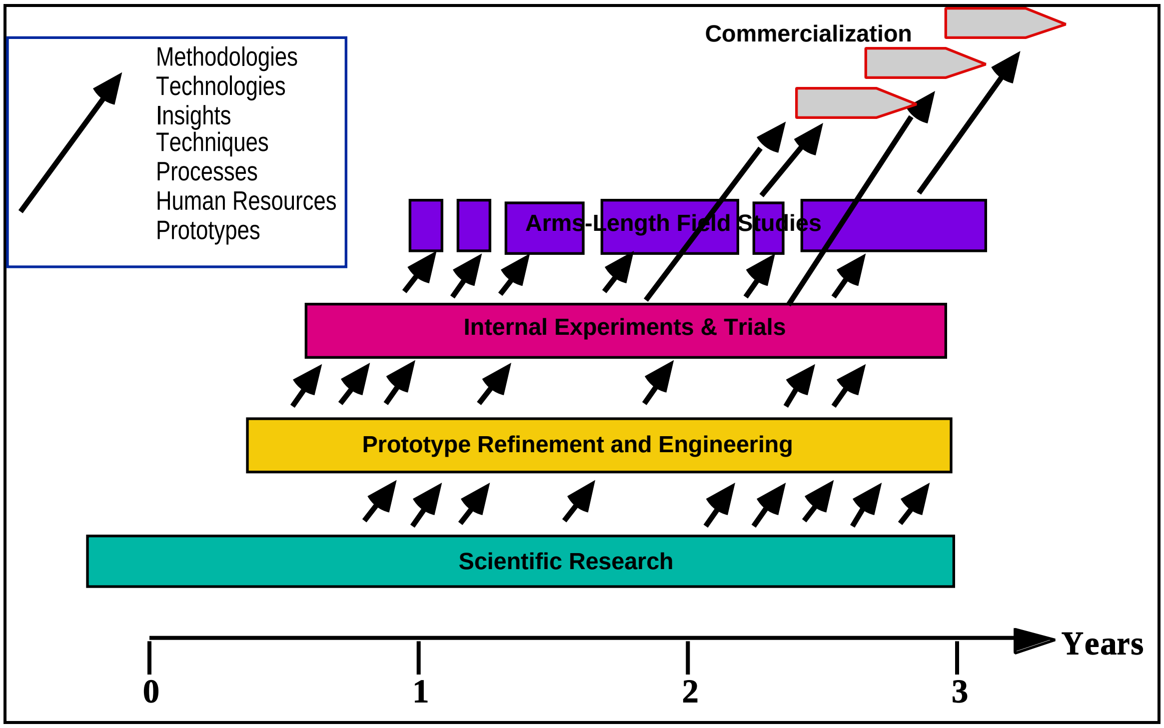


Figure 1-7 Technology Transfer Strategy

Industrial Partners contributed cash, equipment, facilities and people to the research effort. Over the 3-years of the project, these totaled $1,996,200 of which about $300,000 was cash and the rest in-kind. The firms representative(s) took an active role in the research activities by negotiating the terms of participation, deliverables, time-frames, and the division of responsibilities amongst the participants. The Intellectual Property framework used was one of open sharing of research results tempered by special non-disclosure agreements were necessary.

By the end of 1994, 3 Industrial Partners were in advanced development of products stimulated buy their interaction with the project. These were:

1.8.1. Arnott Design Group – "Active Desk"

The Active Desk is a hardware platform designed jointly by the Ontario Telepresence Project and th Arnott Design Group of Toronto. The Active Desk is essentially a large electronic drafting table. It employs a 100x66 c.m. rear projection display. One interacts with the system with a high-resolution stylus (4000x4000 points resolution), or a keyboard. This novel user-interface device hides technology behind the familiar furnature artifact of a drafting table. Industrial design has been used to strong effect to render the technology transparent to the task. In essence, there is no desktop computer and no desktop metaphor. The desktop is the computer.

At the end of 1994, John Arnott, President of th arnott Design Group, was under a contract with Alias Research corporation to build a more refined prototype of the desk for internal trials and expimentation. As well, in late 1994, Mr. Arnott was putting together the financing and business plans to take this prototype, and others designed for the Ontario Telepresence Project commerical. More detailed descriptn of hte prototype is presented in the section titled "Scientific Director's Report".

1.8.2. Corel Corporation – "CorelVideo"

On March 2, 1995, Michael Copland, President of Corel Corporation, formally announced the new product "Corel Video". This new product, which will be widely availible in the summer of 1995, is based in large part on the architecture of TMS -- the prototype system pioneered by Ontario Telepresence Project. Corel Video will put full-motion, full-color video on the desktop for under $2000 US. The product consists of a mixture of software, communications hardware and video hardware to do this. The feature list includes:

- Desk-to-desk direct calling, both within the building, and through the public switched network
- Multi-party connections
- Flexible phone books
- Full motion, full color video for inside calls

Corel's system works under Windows in using a point-and-click approach to person-to-person connections. The system is built on the video PBX model using analog phone lines; there is also a broadcast channel and a "window" channel. All of these features were adopted from TMS.

Insight into successful designs of desk-top videoconferencing systems was transfered to Corel through intense consultations between Telepresence staff, professors and students and Corel marketing, engineering and executive personel. This interchange included multiple day-long design sessions, experience reviews and design critiques as well as on-site demonstrations and the transfer of technical documentation.

Information on Corel Video is presented in Appendix 1.

1.8.3. Applied Silicon International – "Video Vise"

ASI and OTP worked together intensely in the spring/summer of 1994 to exchange hardware/software and system design expertise in a program of prototyping and evaluation of video systems to support group collaboration. Members of OTP's Ottawa-based team worked side-by-side with ASI staff to transfer designs specifications and prototype software assisting them in reaching an important product delivery milestone.

The result of this collaboration was that launching of ASI's ***Video Vise*** **–** a PC based systems capable of continual logging of video data from one or may camera sources. It provides the ability to capture, digitize and compress real-time video images for processing, storage and transmission purposes. Some typical applications include remote monitoring and playback capability for:

- Security and surveillance
- Medial Imaging
- Road Traffic Management
- Quality Control

As a result of this collaboration, ASI was able to deliver engineering prototypes of this advanced communications system to Newbridge Networks Corp. in time for integration into that companies product lines. The Newbridge product, which incorporates the Video Vise, is called ***2611 Main Street.*** Information on these two products is available in the Appendix 1.

1.9. FOUR-MOTORS INTERACTION

Collaboration with the 4-Motors scientists consisted of an exchange of information, software components, and researchers. The four Motors scientific partners were primarily concerned with the application of Telepresence technology to distance education and tele-medicine. This collaboration established an international research network with the Ontario Telepresence Project as Ontario's representative (Figure 1-10).

Interactions among the five regions were governed by an International Memorandum of Understanding which was negotiated by a steering committee consisting of a government representative from each region together with a senior scientist involved in the work. The Memorandum of Understanding included the details of the agreed Intellectual Property framework and dispute resolution mechanisms. A committee of researchers was also formed to coordinate the scientific direction and information exchanges among partners.

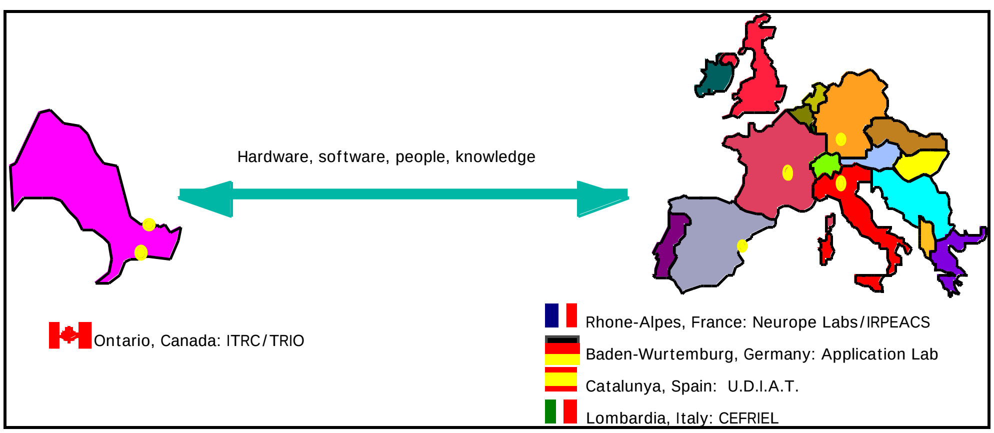


Figure 1-10 International Telepresence Project

Since each region was individually funded and mandated, it was difficult to coordinated and consolidate work being done into a cohesive program. Instead, a loose-collaboration model emerged with bi-lateral experiments and exchanges as it's central model of operation.

The bi-lateral exchanges the Ontario Telepresence Project had with it's Four Motors collaborators are given below:

1.9.1. Interactions with Catelunya

The Barcelona project coincided with the start of OTP. Their area of interest is in medical applications particularly tele-radiology. Dr. Morris Goldberg's IRIS project (in cooperation with BNR/ Northern Telecom), was a source of much inspiration to their work. A large-scale visits by Catalonians to Ottawa was arranged in November of 1991. This included a series of interviews with technical and medical staff at Ottawa Civic and as a result, the Catalonians adopted many of the approaches pioneered in this work. Dr. Goldberg's specifications were in fact used in the multi-media radiological communications system they have developed and the user interface design was also adopted and extended to included video communications. In early 1993, a transfer took place of the entire documentation set from the IRIS project to Catalonian researchers.

A document repository was created on a computer in Barcelona to which all regions have access for drop-off and pick-up of documentation relating to each others projects.

1.9.2. Interactions with Rhone Alpes

Throughout the first two years of OTP, there was an intense consultative relationship (through Morris Goldberg) in project definition and review of proposals. Over half a dozen meetings and working visits on a number of topics including joint research opportunities in tele-education have taken place. This consultation resulted in the successful funding of the Rhone Alpes project in the spring of 1993.

In July, 1993, Drs. Mantei and Dillon participated in the Eurecom/Carleton/UofT tele-teaching experiment which saw the Canadian researchers instruct a series of classes at the French university by remote visual and computer presence.

Claire Belisle, a senior researcher at IRPEACS research centre in Rhone Alpes has partnered with Morris Goldberg to evaluate Telepresence software in a tele-teaching application. A three-day experiment (trial) resulted which involved the first transborder ATM service in Europe was the result. The Ontario Telepresence Project "IIIF" server was a central server in this environment. This was the first use of IIIF in a digital environment and provided us with valuable insights. "Timelines" (software developed in-part by Telepresence) was used to evaluate the video-tape collected at the trial.

1.9.3. Interactions with Baden Wurtemburg

A number of visit by various German ministers took place throughout 1992/93 to OTP sites. Because of internal funding problems, no significant scientific contacts were established although several opportunities were explored over the years.

1.9.4. Interactions with Lombardia

Lombardia, also interested in tele-medicine applications, sent a 7-person delegation (including a cardiac surgeon) to Ontario in 1992 to learn from our experiences in the IRIS project and in Telepresence.

Software and documentation from the Cardiology Workstation project (URIF, Bell, Ottawa Heart institute) was transferred to them. Cefriel re-implemented the software on SUN computers as a prototype to show local doctors (at the Brecia Hospital) the potential functionality of the system they hoped to build. Dr. Morris Goldberg and Dominic

Richens (a technical person attached to the Ottawa Engineering Group of OTP) acted as consultants during the Lombardy re-development.

Morris Goldberg was asked to consult to CEFRIEL on their Telepresence Project proposal to the Lombardy government. He visited with them several times including a site visit to the Hospital of Brecia where the trials were slated to take place. At this site visit, he met with Chief of Staff, Chief of Radiology and Chief of Cardiology. Partly as a result of these consultations, Cefriel was successful in it's funding proposal to the Region of Lombardia.

From July to Sept., 1993, Luca Giachino, a researcher with Cefriel, visited with the UofT Telepresence group. His stay was arranged under the International Strategic Framework. He resided at the UofT and undertook various research activities of mutual interest. He created several software prototypes, some of which are still undergoing refinement under the direction of Bill Buxton.

In July of 1994, a software package was transferred from the UofT to Cefriel under the International Strategic Framework. The package, VANNA - Video Annotation and Analysis System, is a system for video annotation or "coding" video tapes. It is software that runs on Apple MACINTOSH computers which controls a VCR providing the user with the ability to associate text or tags with particular segments of the tape.

Stemming from these interactions, a proposal was put forward by Cefriel to it's funding agencies that would see "Telepresence" research become a part of Cefriel's mandated research and not simply an externally funded research project. At the writing of this report, it is unknown whether this proposal was accepted or not.

In November of 1993 a formal presentation of the first prototype tele-medicine system was made to hospital staff, Italian Minister of Health and senior bureaucrats to get approval of funding for the second phase. Dr. Goldberg was requested to attend to support the presentation.

1.10. HUMAN RESOURCE DEVELOPMENT

Over the course of the three-years, twelve industrial partner personnel received extensive training/experience through day-to-day interaction in the advanced research laboratories of the project. Five of these twelve were resident on-site full-time for periods varying between 4-months and 14-months. The other seven were either resident for shorter periods (3-months or less) or maintained their base-of operations at their company work-site.

Eight full-time and one part-time professional research staff were dedicated to the project. In addition, 3 other computer professionals were employed as contractors at various times. These staffers participated in all major international Scientific/Industry conferences in the field. The staffers supported nine faculty members and approximately 16 graduate students/year building and sustaining a state-of-the-art research environment in Toronto and Ottawa. The staff worked closely with industrial partners during the course of the project, and, at it conclusion, most of them moved over to positions with affiliated firms. Seven staff alumni are now working in Ontario firms –3 BNR, 1 Alias Research, 1 Newbridge Networks and two have formed independent start-up companies.

16 Graduate students were directly involved in the project. Of the seven graduates that completed by December, 1994, six are now working in related industries including:

- Gordon Kurtenbach, Director of User Interface Design, Alias Research

- Michael Sheasby, Director of User Interface, SoftImage

1.11. Information Dissemination

The project was hugely successful at disseminating the results of it's work. Through an energetic communications program, the project featured prominently in scholarly publications, trade journals, public press, television, radio and at major trade shows and events.

Figure 1-11 provides a three-year synopsis of media-hits, programs and information dissemination measures resulting from Ontario Telepresence Project activities:

Media Hits, Programs and Information Dissemination Measures
55+ publications in refereed Journals or conferences
30 Technical reports and non-refereed publications
25 Popular Newspaper & Magazine articles (including Globe&Mail Toronto Star, Financial Times, Canadian Bus., Byte)
6 TV spots (including CTV Towards 2000, CBC Prime Time, CNN)
105 industry demonstrations/tours (25-1992, 45-1993, 35-Q3/94)
500 Information packages & 2,400 colour brochures distributed
69 feature presentations at major industrial gatherings
18 Video tapes produced featuring Telepresence research
WWW Site providing on-line access to documents, photos, etc.
Annual General Meeting

Figure 1-11 Information Dissemination

1.12. Financial Resources

Figure 1-12 provides a detailed annual break-down of cash revenue and other resources employed during the course of the project. In all, the project levered $4.8 million in cash and in-kind resources over 3-years of which almost $2 million (41%) was contributed by Ontario firms.

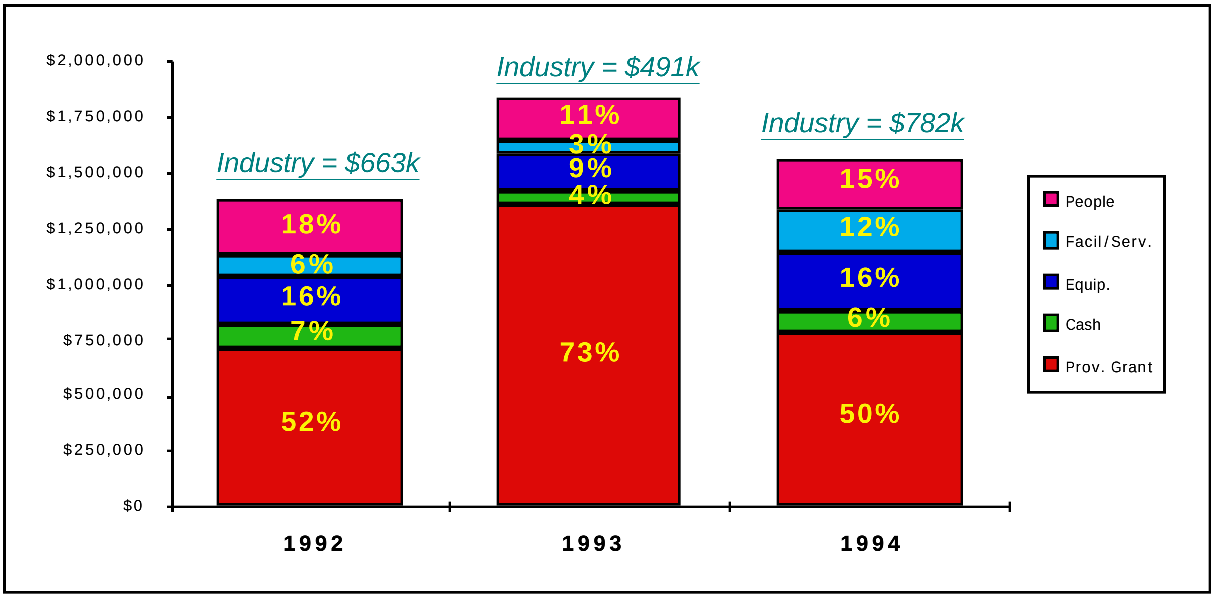


Figure 1-12 Resource Contributions

1.13. Conclusions

The Ontario Telepresence Project was a $4.8M 3-year R&D project that demonstrated world leadership in both user-centered design and deployment methodologies of CSCW systems. It's inter-disciplinary and university-industry research program made huge in-roads scientifically while linking to similar projects in the four-Motors regions of Europe. Seven Ontario firms actively partnered with the project leading to direct commercialization in three cases.

Through it's support of the Ontario Telepresence Project, the Province of Ontario achieved the following:

- Demonstrated scientific leadership in embracing technological/organizational innovation to improve coordination between geographically distributed workers
- Stimulated partnerships between Ontario-based firms and establish new markets by creating new technologies and testing them in arms-length work environments.
- Supported and encourage spin-offs from world-class research taking place at it's universities
- Levered initiatives in broad-band networks (OCRINET, etc.) by spawning market leading applications which can run on these networks.
- Maintained and extended Ontario profile in the Four Motors by demonstrating the technological leadership of it's universities and firms.

---

[i] "Interregion '90", hosted by Premier David Peterson.

[ii] As well, a full-time administrative assistant was hired and housed at the University of Toronto to support the Managing Director

[iii] Dr. Karamoush was only peripherally involved with the project in it's first year. Dr. Goldberg left the University of Ottawa early in 1992

[iv] In the fall of 1994, the Ottawa Engineering Group offices at Hewlett Packard were moved to the Department Of Systems and Computer Engineering at Carleton to enable the ATM trial to be launched and supported.

[v] The Managing Director and the Scientific Director report to the President of ITRC, while the Director of Engineering reported to the President of TRIO.

[vi] The architectural approach used in the Telepresence Media Space was adopted by Corel – one of the Industrial Partners – and commercially exploited in 1995.

[vii] IBM Canada Labs was an industrial partner from January, 1992 until February 1993. Mr. Rich Helms resided on-campus for 14-months contributing greatly to the definition of the research architecture. IBM Canada Labs was forced to withdraw from participation in the project due to re-focusing of their world-product mandates away from telecommunications and group collaboration systems.

SECTION 2:
SCIENTIFIC DIRECTOR'S REPORT

2. Scientific Director's Report: Living in Augmented Reality

Bill Buxton
Scientific Director
Ontario Telepresence Project

University of Toronto
6 King's College Rd.
Toronto, Ontario

Alias Research Inc.
110 Richmond St. East

Introduction

In 1991, Mark Weiser, of Xerox PARC, published an article that outlined a vision of the next generation of computation (Weiser, 1991). He referred to this model as *Ubiquitous Computing,* or *UbiComp.* UbiComp was based on the notion that it is inappropriate to channel all of one's computational activities through a single computer or workstation. Rather, Weiser argued that access to computational services should be delivered through a number of different devices, each of whose design and location was tailored to support a particular task or set of tasks. It is on this notion of delivering computational services throughout our work, play and living spaces, that the ubiquity in the name is based.

In addition to ubiquity, UbiComp assumes that the delivery of computation should be *transparent.* There is a seeming paradox that arises between the principle of ubiquity and that of transparency. The resolution of this paradox, through the use of examples, will constitute a significant part of what follows.

Figure 1: Xerox PARCtab:
A palm-sized device to support computation, communication and location One of the expanded repertoire of computational resources offered by UbiComp. (Photo Credit: Xerox PARC)

Around the same time that Weiser and his colleagues were developing the ideas that were to emerge as UbiComp, others down the hall at Xerox PARC were developing video-based extensions to physical architecture, so-called *Mediaspaces* (Bly, Harrison & Irwin, 1993). These were systems through which people in remote offices, buildings, and even cities, could work together as if they were in the same architectural space. While prototypes, these systems enabled one to work side by side at one's desk with someone in a remote location. You could call out of your door and ask "Has anyone seen Sara?" without thinking about whether the answer would come from Portland, Oregon or Palo Alto, California. Nor did it matter at which of these two centres either you or Sara were at. The technology supported a sense of shared presence and communal social space which was independent of geographical location. The result can perhaps best be described as a *social prosthesis* that afforded support of the links that hold together a social network — links which are typically only maintainable in same-place activities.

While reading Weiser's paper gave no hint of the activities of the Mediaspace group, and *vice versa*, I increasingly began to see the projects as two sides of the same coin. Consequently, I directed the *Ontario Telepresence Project* research program to apply the tenets of UbiComp to the mediaspace technology. Just as in UbiComp it is deemed inappropriate to channel all of your computational activity through a single workstation, so in *Ubiquitous Video* (*UbiVid)*) is it inappropriate to channel all of your communications through a single "video station" (viz., camera, video monitor, microphone, loudspeaker). As in UbiComp, these technologies are deployed in the locations, scales and forms appropriate to their intended function. And while ubiquitous, access to the services of

In what follows, I discuss our experience living in such an environment over the past seven years (3 with OTP, and preceeding this, with Xerox PARC). From this experience emerge insights that we believe have important implications to the future deployment of media — insights that we feel are doubly important in this period of technology convergence, especially since they are derived from actual experience, rather than theoretical speculation.

Figure 2: Shared Open Office via Mediaspace
Here, Randall Smith and Steve Harrison collaborate in a shared office environment despite their geographical separation. (Photo Credit: Xerox PARC)

these communications technologies must also be transparent.

UbiComp and UbiVid — let us call them collectively *Ubiquitous Media* — represent an approach to design which is radically different than that seen in today's multimedia workstations. In current designs, all activities are centred around a single multipurpose device. Ubiquitous Media, on the other hand, is an architectural concept in that it is concerned with conventions of traditional location-function-distance relationships. Ubiquitous Media says that the box into which we try to fit our solutions is not some box or "super appliance" on the desk; rather, it is the desk itself, or the room, or building, in which the desk is found.

By way of another contrast, Ubiquitous Media can also be understood in relation to Artificial Reality. Rather than turning inward into an artificial world, Ubiquitous Media, encourage us to look outward. It expands our perception and interaction in the physical world. (For example, in the attempt to find Sara, consider the augmentation of the social space to include the physical space of both Palo Alto and Portland. The augmentation was socially transparent. There was no "user interface" other than that used in conventional architecture: one just called blindly out the door.) In contrast to "virtual" or "artificial" reality, we consider our use of Ubiquitous Media as *Augmented Reality* (Wellner, Mackay, & Gold, 1993).

UbiComp: A Brief Overview

Introduction

As described by Weiser, UbiComp can be characterized by two main attributes:

- *Ubiquity:* Interactions are not channeled through a single workstation. Access to computation is "everywhere." For example, in one's office there would be 10's of computers, displays, etc. These would range from watch sized Tabs, through notebook sized Pads, to whiteboard sized Boards. All would be networked. Wireless networks would be widely available to support mobile and remote access.
- *Transparency:* This technology is non intrusive and is as invisible and as integrated into the general ecology of the home or work place as, for example, a desk, chair, or book.

These two attributes present an apparent paradox: how can something be everywhere yet be invisible? Resolving this paradox leads us to the essence of the underlying idea. It is not that one cannot see (hear or touch) the technology; rather, that its presence does not intrude into the environment of the workplace (either in terms of physical space or the activities being performed). Like the conventional technology of the workplace (architecture and furniture, for example), its use is clear, and its physical instantiation is tailored specifically for the space and the function for which it is intended. Central to UbiComp is a break from the "Henry Ford" model of computation which can be paraphrased as:

You can have it in any form you want as long as it has a mouse, keyboard and display.

Fitting the square peg of the breadth of real needs and applications into the round hole of conventional

Figure 3: Xerox Liveboard and PARCpads

The photo shows two other components of UbiComp: a wall-sized interactive surface, called a "Liveboard" and lap-sized interactive pen-based displays called PARCpads. (Photo Credit: Xerox PARC)

designs, such as the GUI, has no place in the UbiComp model.

Technology Warms Up

We can most easily place Weiser's model of computation in historical perspective by the use of an analogy with heating systems. In earliest times, architecture (at least in cold climates) was dominated by the need to contain heat. Special structures were built to contain an open fire without burning down. Likewise, in the early days, special structures were built to house computation. These were known as "computer centres."

As architecture progressed, buildings were constructed where fires were contained in fireplaces, thereby permitting heat in more than one room. Nevertheless, only special rooms had fire since having a fireplace required adjacency to a chimney. Similarly, the next generation of computation was available in rooms outside of computer centres; however, these had to have special electrical cabling and air conditioning. Therefore, computation was still restricted to special "computer rooms."

In the next generation of heating system, we moved to Franklin stoves and even to radiators. Now we could have heat in every room. This required the "plumbing" to distribute the system, however. The intrusion of this "plumbing" into the living space was viewed as a small price to pay for distributed access to heat. Again, this is not unlike the next generation of computation, (the generation in which we are now living), where we have access to distributed computation everywhere, as long as we are connected to the "plumbing" infrastructure. And like the heating system, this implies both an intrusion into the space and an "anchor" that limits mobility.

This leads us to the next (today's) generation of heating system: climate control. Here, all aspects of the interior climate (heat, air conditioning, humidity, etc.) is controllable on a room-by-room basis. What actually provides this is invisible and is likely unknown (heat-pump, gas, oil, electricity?). All that we have in the space is a control that lets us tailor the climate to our individual preference. This is the heating equivalent of UbiComp: the service is ubiquitous, yet the delivery is invisible. In this mature phase, the technology is seamlessly integrated into the architecture of the workplace.

By analogy, within the UbiComp model, there is no computer on my desk because my desktop *is* my computer. As today, there is a large white board on my wall, but with UbiComp, it is active, and can be linked to yours, which may be 3000 km away. What I see is way less technology. What I get is way less intrusion (noise, heat, etc.) and way more functionality and convenience. And with my Pads and Tabs, and the wireless networks that they employ, I also get far more mobility without becoming a computational "orphan."

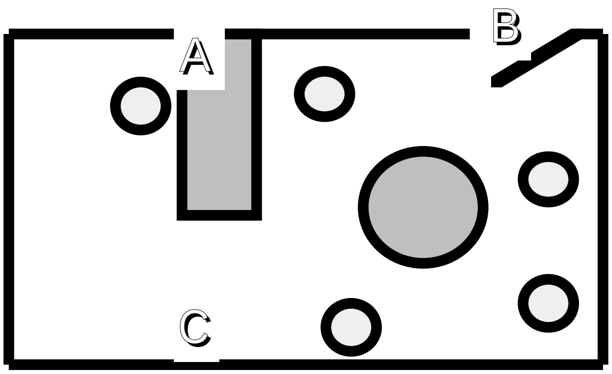


Figure 4: My Office Layout:
People visiting electronically can assume one of three locations (A, B or C). Location "A" on my desk is for close one-on-one work. Location "B" is above the door. This affords people glancing into my office without disturbing me and, on entering, making a gradual approach from a distance. Location "C" is a virtual seat around the coffee table. This affords the remote person to participate in informal round-table conversations.

Mediaspaces and Ubiquitous Video

Introduction

UbiVid is the video compliment to UbiComp in that it shares the twin properties of *ubiquity* and *transparency*. In "desktop videoconferencing," as it is generally practiced, what we typically see is a user at a desk talking to someone on a monitor that has a video camera placed on top. This is illustrated in Figure 2. Generally, the video interactions are confined to this single camera-monitor pair.

In UbiVid, we break out of this, just as UbiComp breaks out of focusing all computer-mediated activity on a single desk-top computer. Instead, the assumption is that there are a range of video cameras and monitors in the workspace, and that all are available. By having video input and output available in different sizes and locations, we enable the most important concept underlying UbiVid: *exploiting the relationship between (social) function and architectural space.*

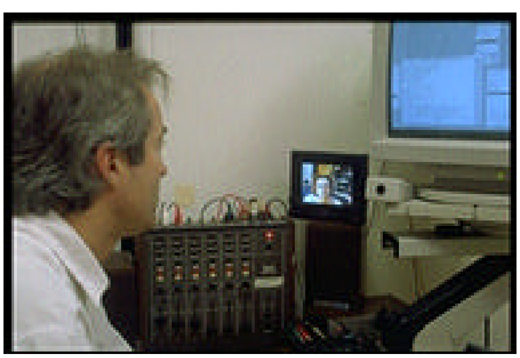
In what follows, we explore the significance of this relationship. We start by articulating some of the underlying design principles, and then proceed to work through a number of examples.

> *Design Principle 1: Preserve function/location relations for both tele and local activities.*
> ***Design Principle 2:*** *Treat electronic and physical "presences" or visitors the same.*
> ***Design Principle 3:*** *Use same social protocols for electronic and physical social interactions.*

Example: My Office

Let us work through an example to illustrate how these principles apply in a specific context, namely my office. A floor-plan of my office is illustrated in Figure 4.

There is a desk where I work and a coffee table, around which I have small informal meetings. There are five chairs, one of which is normally behind my desk. The others are around the coffee table. There are three distinct locations where remote visitors can appear. If they we working closely one-on-one, they appear on my desk. (This is shown as location "A" in Figure 4.) Here, they appear right beside my computer screen (which might contain information that we are both



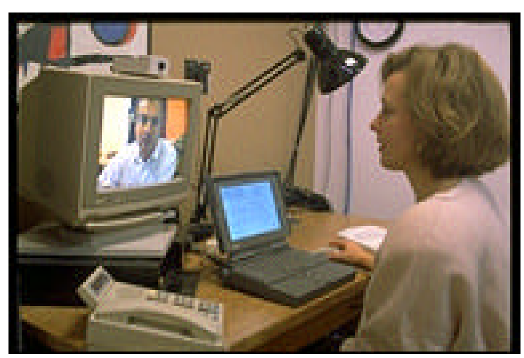
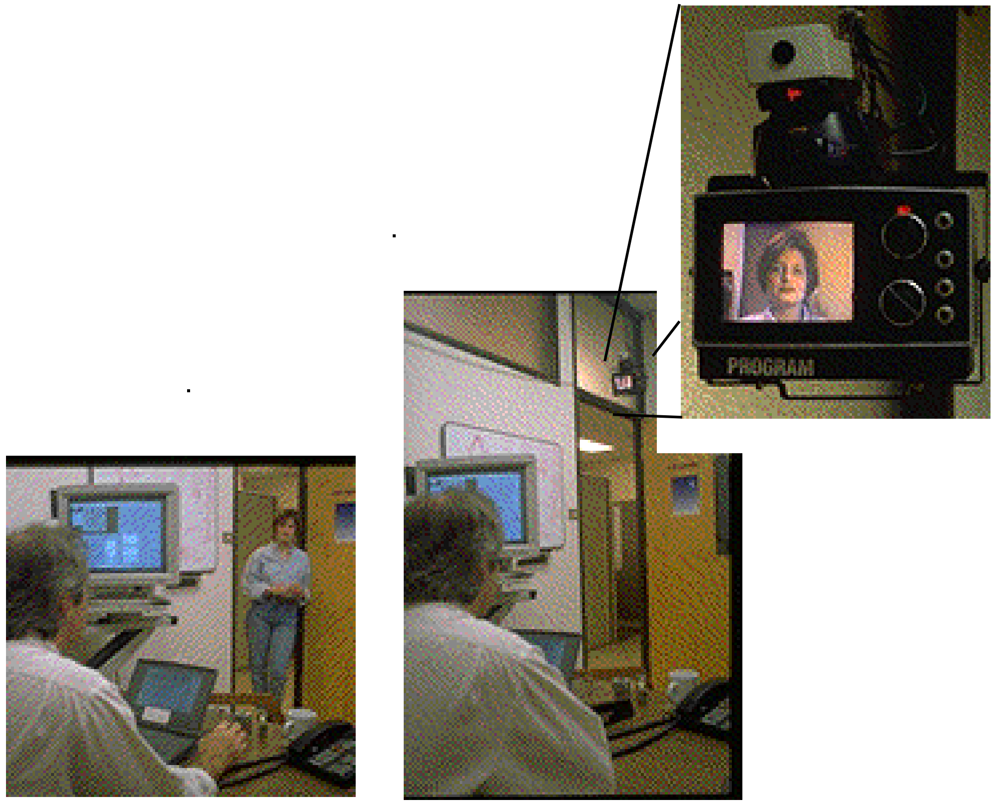


Figure 5: Close Collaboration at the Desktop
Here we see the author engaged in a face-to-face collaboration with a colleague Gale Moore. Through the close proximity of each participant's video display to their computer display, a shared presence of both person and shared document is provided. The interaction here is akin to same place intimate side-by-side work at the desk.



(a) **(b)**

Figure 6: Interactions at the Door
When someone comes to my office, they come via the door, whether they come physically or electronically. This is shown in the figures. Figure 6.a illustrates my assistant Tabatha Holtz having a temporary conversation from my doorway. Figure 6.b illustrates the electronic equivalent. Regardless of how she comes, conventions of social distance and approach are respected.

viewing simultaneously). An example of this type of meeting is illustrated in Figure 5.

If someone wants to glance into my office to see if I am available, they can do so from the door (location "B" in Figure 4), whether they come physically or electronically. A camera mounted above the door gives them approximately the same view that they would have if they were glancing through my physical door. This is illustrated in Figure 6. I can see who is "at the door" on a small monitor mounted by the camera, and — as in the physical world — I can hear their approach by means of an auditory icon, or *earcon.*

Likewise, when I'm engaged in a meeting in my office, if someone comes by the door to see if I'm available, this same arrangement provides me with the same options regardless of whether the person comes electronically or physically. I can ignore them if I don't want to be interrupted, and due to their position and distance, they don't intrude on my meeting. If I want, I can glance up and discretely determine who is there. If it is someone that I don't want to speak to at the moment, I can then glance down and continue my meeting. The person at the door is aware that I know of their presence, and by my action, they know that I can't see them at the moment. On the other hand, if it is someone who could contribute to the meeting, I invite them in. Finally, if it is someone that I know needs urgent attention, I will suspend the meeting and deal with the issue (hopefully briefly).

While some may claim that this additional technology is superfluous or an added "luxury", we believe that it may well make the difference between success and failure of a system. We can illustrate this with an example. In 1993/4, Hiroshi Ishii visited us from NTT for a year. When he first came, this "door cam" was not deployed. After he had been with the project for a while, he explained to me that when he first came he was reluctant to use the system to contact me because he felt that it was rude to just "arrive" on my desktop. His reasons were partially due to not knowing me that well at the time, and partially out of "respect" for my position as director of the project. To him, the distance and means of approach afforded by the "door-cam" was an important affordance to his making effective use of the system. Our claim is that the need for such social sensitivities is not rare.

In addition to working at my desk and interactions at the door, there is a third location-sensitive function that takes place in my office: informal meetings. These normally take place around the round coffee table, and may involve up to five or six people. Frequently these include a participant from a remote site. In order to enable them to do so from an appropriate location, a special "seat" is reserved for them around the table. This is located in position "C" in Figure 4, and is shown in Figure 7.

Figure 7: Informal Meetings
Here, a group of colleagues are having an informal meeting in my office. One is attending from a remote site. All take their place around the table, regardless. The video "surrogate" affords the remote party to have a sense of gaze awareness much like the other participants.

By appearing in a distinct and appropriate location, participants physically in my office are able to direct their gaze at the remote participant just as if they were physically present. Likewise, the remote participant has a sense of gaze awareness, that is, who is looking at whom, and when. The reason is that the remote participant *has* a physical presence in the room — a presence afforded by the location of the video *surrogate* through which they communicate.

In our discussion, we have mainly dealt with social function and distance in relation to fixed locations. These are issues, however, which normally have a strong dynamic component. People move. In so doing, functions change. In this regard, our system is still lacking. One can move from location to location within the room, but the transitions are awkward. This is an area that needs improvement. But before one can work on movement, one has to have places to move to. This has been our main focus to date.

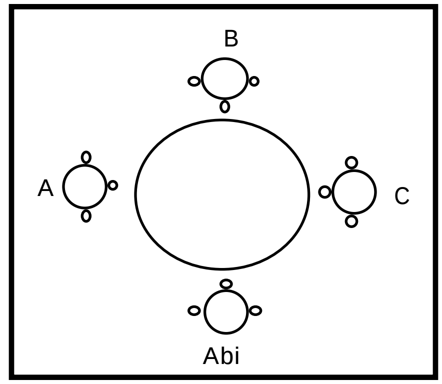
Having lived in this environment in this form for almost two years, perhaps the most striking thing is a seeming paradox. By adding this extra equipment into the room, there actually appears to be less technology and far less intrusion of the technology in the social interactions that it mediates. Our argument is that this is due to the technology being in the appropriate locations for the tasks undertaken in the room. In a single desk-top solution, for example, one would be twisting the camera and monitor from the desk to the coffee table when switching between desk-top and group meetings. As well, due to the multiple cameras and monitors, we avoid the contention for resources that would otherwise result. For example, I can be in a desk-top conference on one monitor, monitor a video which I am copying on another, and still not prevent someone from appearing at my electronic door.

Figure 8: Back-to-Front Videoconference
The remote attendee takes his place at the table by means of a video monitor mounted by the back wall. He sees through the camera above, hears via a microphone, and speaks through his monitor's loudspeaker. The presenter interacts in the same way with those attending physically and those attending electronically. No new skills are required.

As we have pointed out in the examples above, through increased ubiquity, we have achieved increased transparency. This last point is achieved, however, only through the *appropriate* distribution of the technology — distribution whose foundations are the social conventions and mores of architectural location/distance/function relationships.

Example: Back-to-Front Videoconferencing

Another example of using spatially distributed video is the implementation of "back-to-front" videoconferencing at the University of Toronto. In contrast to traditional videoconferencing rooms, the camera and monitors are placed at the back of the room, as illustrated in



(a) (b)

Figure 9: Hydra: Using video "surrogates" to support a 4-way video conference
Figure (a) shows Abi Sellen in a 4-way video conference where each of the three remote participants attends via a video "surrogate." By preserving the "round-table" relationships illustrated in (b), conversational acts found in face-to-face meetings, such as gaze awareness, head turning, etc.. are preserved.

Figure 8.[1] The intent here is to enable remote participants to "take their place at the table."

The scenario shown in the figure illustrates the notion of transparency. Due to the maintenance of audio and video reciprocity, coupled with maintaining "personal space," the presenter uses the same social mechanisms in interacting with both local and remote attendees. Stated another way, even if the presenter has no experience with videoconferencing or technology, there is no new "user interface" to learn. If someone raises their hand, it is clear they want to ask a question. If someone looks confused, a point can be clarified. Rather than requiring the learning new skills, the design makes use of existing skills acquired from a life time of living in the everyday world.

Example: Hydra: Supporting a 4-way Round-Table meeting

In this example, we introduce a technique to support a four-way meeting, where each of the participants is in a different location. It was designed to capture many of the spatial cues of gaze, head turning, gaze awareness and turn taking that are found in face-to-face meetings. Consistent with the design principles outlined above, we do this by preserving the spatial relationships "around the table."[2] This is illustrated in Figure 9.

As seen in the left-hand figure, each of the three remote participants is represented by a small video *surrogate.* These are the small *Hydra* units seen on the desk (Sellen, Buxton & Arnott, 1992). Sitting in front of the desk is a colleague, Abi Sellen. Each unit provides a unique view of her for one of the remote participants, and provides her a unique view of them. The spatial relationship of the participants is illustrated by the "round-table" in the right-hand figure. Hence, person A, B and C appear to Abi on the Hydra units to her left, centre and right, respectively. Likewise, person A sees her to their right, and B to their left, etc.

Collectively, the units shown in the figure mean that Abi has three monitors, cameras and speakers on her desk. Yet, the combined footprint is less than that of

[1] In fact, the room also supports traditional "front-to-back conferencing, which just pushes the issue of ubiquity even further.

[2] This idea of using video surrogates in this way for multiparty meetings turns out not to be new. After implementing it ourselves, we found that it had been proposed by Fields (1983). Later, we found that it had been demonstrated and documented at the University of Toronto in Feb., 1970 (Vowles, 1992).

her telephone. These Hydra units represent a good example of transparency through ubiquity. This is because each provides a distinct point source for the voice of each remote participant. As a result, the basis for supporting parallel conversations is provided. This showed up in a formal study which compared various technologies for supporting multiparty meetings (Sellen, 1992). The Hydra units were the only technology tested that exhibited the parallel conversations seen in face-to-face meetings.

The units lend themselves to incorporating proximity sensors that would enable aside comments to be made in the same way as face-to-face meetings: by leaning towards the person to whom the aside is being directed. Because of the gaze awareness that the units provide, the regular checks and balances of face-to-face meetings would be preserved, since all participants would be aware that the aside was being made, between whom, and for how long.

None of these every-day speech acts are supported by conventional designs, yet in this instantiation, they come without requiring any substantially new skills. There is no "user interface." One interacts with the video surrogates using essentially the same social skills or conventions that one would use in the face-to-face situation.

> ***Concept:*** *Video Surrogate: Don't think of the camera as a camera. Think of it as a surrogate eye. Likewise, don't think of the speaker as a speaker. Think of it as a surrogate mouth. Integrated into a single unit, a vehicle for supporting design Principles 1 & 2 is provided.*

Much of the above is based on the notion that the physical location of participants has an important influence on social interactions in face-to-face meetings. What we are driving at from a design perspective is that these same cues can be used or exploited in telepresence. When we talk about distance between participants, therefore, it is important to distinguish between their physical distance from me, and the distance between their video surrogate and me. The latter, rather than the former, is what determines social distance.

> ***Premise:*** *Physical distance and location of your video surrogate with respect to me carries the same social weight, function, and baggage as if you were physically in your surrogate's location. Furthermore, the assumption is that this is true regardless of your actual physical distance from me.*

> ***Qualification:*** *This equivalence is dependent on appropriate design. It sets standards and criteria for design and evaluation.*

Figure 10: The Telepresence Client
This is the interface to the main Telepresence application. It mediates connections and interactions among users and resources. Connections are made by selecting the person's name from the scrolling list in the main part of the screen, then selecting the "contact" button below. Note that beside each person's name is an icon of a door in one of 4 states. This indicates that user's accessibility. See Figure 11 for more detail.

Ubiquitous Media: UbiComp + UbiVid

Introduction

To this point, we have discussed computation separately from the mediaspaces. Clearly, however, these two classes of technology coexist. They compliment each other in a number of ways. First, there is a dependence relationship: it is only through the computational resources that the control and operation of mediaspaces can be deployed.

Second, there is a cummulative relationship. In collaborative work, the mediaspace technology provides the shared space of the people, and the computers the shared space of electronic documents. Both types of shared space are required to establish a proper sense of shared presence, or *telepresence*.

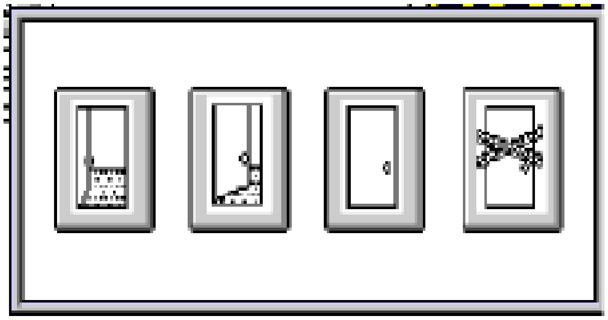


Figure 11: Using "door state" to specify accessibility

The figure illustrates a technique for users in a media space to control their own accessibility following the same approach used in physical space: by the state of their door. Each one of the selectable door states allows a different level of accessibility. The specified door state is visible to potential callers who, likewise, know the implied permissions by analogy with the physical world.

When used together, a sense of awareness of the social periphery is afforded — as sense which would otherwise only be possible in a shared corridor or open concept office.

In the remainder of this section, we will give examples which illustrate each of these three cases.

Making Contact

It is misleading to think about computation as separate from collaboration via mediaspaces. The notions of UbiComp and UbiVid go hand-in-hand. This is most easily seen in the example of how the computer is used to mediate interactions within a mediaspace.

Figure 10 shows the user's view of the main application used to mediate connections among people and resources of our system. The main part is made up of a scrolling list of names. These are the names of the people at the specified site. I can contact someone by selecting their name, then selecting the "Contact" button below the list.

The system enables me to contact people at various sites, not just those who are local. In the example, residents at the Toronto Telepresence site are displayed in the list. This is indicated by the labeled box above the list. Selecting this box enables me to select a different site, with the resultant effect that the list now shows the residents for that new site. The method of contact is the same regardless of the geographical location of the site.

In essence, this technology paves a potentially wide pathway to your door. To balance this, a means is provided to enable the user to control their own degree of accessibility. This is done using a metaphor from the physical world: the door.

Notice that beside each name in the list is an icon of a door. The door can be in one of 4 states. Each indicates a different degree of accessibility for that user. If it is *open*, you are welcome to "pop in." If it is *ajar*, you can glance in and determine if I am busy, but you must "knock" if you want to enter. If it is *closed*, you must knock and wait for a response before entering, and glancing is not possible. Finally, if the door is *boarded shut*, you can only leave a message.

A user can set their door state by selecting the door icon in the top left corner of the panel. This causes the panel shown in Figure 11 to pop up. The state is set by selecting one of the four icons. If I select the "open door" icon, for example, this new state is set and communicated to all other users at all other sites. They need only look at the door icon beside my name to determine my accessibility. Hence, a means is provided to control accessibility which is based upon everyday social protocols.

Shared Presence of Person & Task

In an earlier publication (Buxton, 1992), we argued for the need to support a sense of shared presence of both task and person. The main argument was that being able to move seamlessly from one to the other was important in undertaking a number of common interactions. This point has been argued by others, especially Ishii, Kobayashi, and Grudin (1992).

The next set of examples are examples of how we have combined industrial design and the integration of UbiComp and UbiVid technologies to support such shared person / shared task interactions.

***Example*: Hydra with Shared Computer "Whiteboard"**

The first such example is illustrated in Figure 12. Here visiting scientist, Hiroshi Ishii of NTT, is in the midst of a multiparty meeting. The remote participants appear on the Hydra units, seen previously in Figure 9. Each participant can both see and markup the technical drawing which is the topic of the meeting. In the figure, the drawing is distributed and displayed by the computers and associated network, and displayed using a rear-projection system (thereby emulating a large flat panel display).

In this example, note that the gaze awareness discussed earlier in the Hydra example extends to the document. Because of the use of the Hydra units and the geometry of the set-up, people can follow each other's gaze from person to person, as well as person to document.[3]

***Example:* the Active Desk**

The example in Figure 13 shows another configuration involving collaborative work on a shared document with a remote participant. Again, it is achieved through an integration of UbiComp and UbiVid technologies.

In this case, the document appears on what we call the *Active Desk*.[4] This is a large electronic drafting table. It employs a 100x66 c.m. rear projection display. One interacts with the system with a high-resolution stylus (4000x4000 points resolution), or a keyboard. In the illustration, a Hydra unit is mounted on the desk, through which the operator is communicating with the remote colleague. As with the previous case, the document is visible to all parties, and all parties can interact with it, such as pointing, making annotations, or adding to it.

What is central here is how little technology is actually visible. In the example, industrial design has been used to strong effect to render the technology transparent to the task. In essence, there is no desktop computer and no desktop metaphor. The desktop is the computer.

> ***Design Principle 4:*** *The box into which we are designing our solutions is the room in which you work/play/learn, not a box that sits on your desk.*

***Example:* Sitting Across the Desk**

This final example, illustrated in Figure 14, is the UbiVid equivalent to sitting across the desk from someone. Here, through rear projection, the remote participant appears life-size. What we are trying to capture in this example are two people working together on a joint project, such as a drawing or budget. While not the case in the example, the shared documents would appear on the desktop, using a technology similar to the Active Desk seen in the previous figure.

First, notice that having one's counterpart displayed this way is not like seeing them on a regular video monitor. Because of the scale of the image, the borders of the screen are out of our main cone of vision. Hence, the space occupied by the remote person is defined by the periphery of their silhouette, not by the bezel of a monitor.

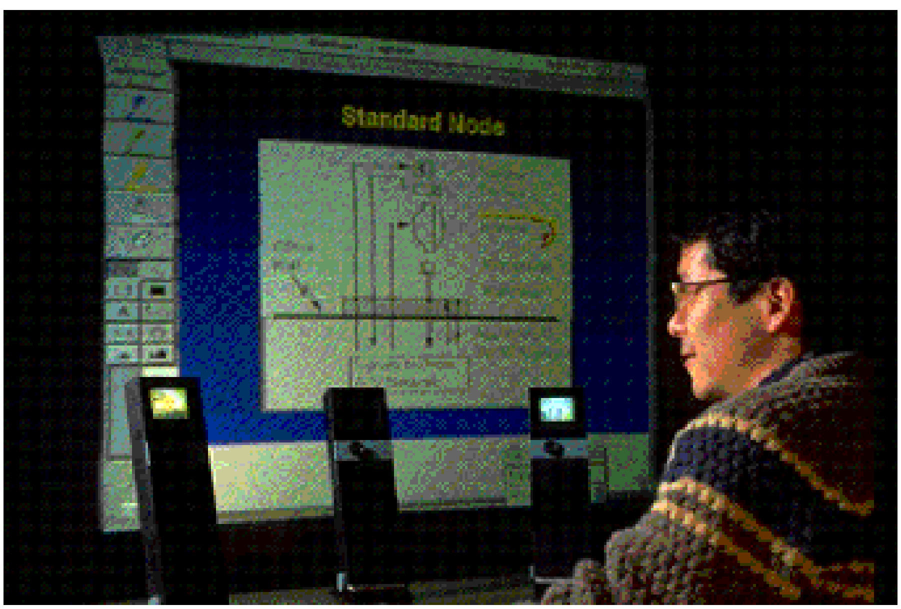


Figure 12: Shared Task & Person Space

A multiparty meeting concerning a technical drawing s illustrated. The technical drawing is displayed on the large screen behind the Hydra units (which are used for the spatial presence of the participants). Each participant can see and markup the technical drawing.

afford the full utilization of gaze stability and gaze cues.

[4] The Active Desk was designed jointly by the Ontario Telepresence Project and the Arium Design Group of Toronto.

[redacted]

Figure 13: The Active Desk
The Active Desk is a drawing table with a 100x66 c.m. rear projection computer display. The user interacts with the desk with a high-resolution stylus, or keyboard.

Second, by being life size, there is a balance in the weight or power exercised by each participant.

Third, the gaze of the remote participant can traverse into our own physical space. When the remote party looks down on their desk, our sense of *gaze awareness* gives us the sense that they are looking onto our own desktop. Their gaze traverses the distance onto our shared workspace, thereby strengthening the sense of *Telepresence.*

Portholes and Awareness of the Social Periphery

The original *Portholes* system was a joint development between Xerox PARC and Rank Xerox EuroPARC (Dourish & Bly, 1992). A typical Portholes display is shown in Figure 15. What it does is present a tiled view of video snapshots of all members of a distributed workgroup. In the case of the Ontario Telepresence Project, these snapshots are updated every five minutes. Unique to the Telepresence application is the superimposition of the door icon on the snapshot.

The snapshot and door state icon provides information as to both the presence or activities of group members, as well as their degree of accessibility. Furthermore, the snapshots provide a user interface to certain functions concerning individuals. For example, after selecting the snapshot of me on your screen, you can then click on the *Info* button on the top of the frame to get my phone number, name, address and email address. Or, double clicking on my image, or selecting the *contact* button asserts a high bandwidth connection to me (thereby providing an alternative means t make a connection to that illustrated in Figure 10).

Portholes takes advantage of the fact that each office has a video camera associated with it. It goes beyond the stereotyped notion of desktop video as simply a videophone. Rather, it supports a very important sense of awareness of the social periphery — an awareness that normally is only available in shared office or shared corridor situations. It introduces a very different notion of video on demand and delivers its potential with a transparent user interface.

Finally, discussions about Portholes always touch upon the issue of privacy. "How can you live in an environment where people can look in on you like that?" we are frequently asked. There are a couple of responses to this. First, Portholes is not an "open" application. It embodies a sense of reciprocity within a distinct social group. People cannot just randomly join a Portholes group. Members know who has access to the images. Secondly, even within the group, one can obtain a degree of privacy, since the distribution of your image can be controlled by your door state. Finally, remember that the images have no motion and no audio. What is provided is less than what would be available to someone looking through the window of your office door. This is especially true if the snapshot is taken from the "door camera", such as illustrated in Figure 6.

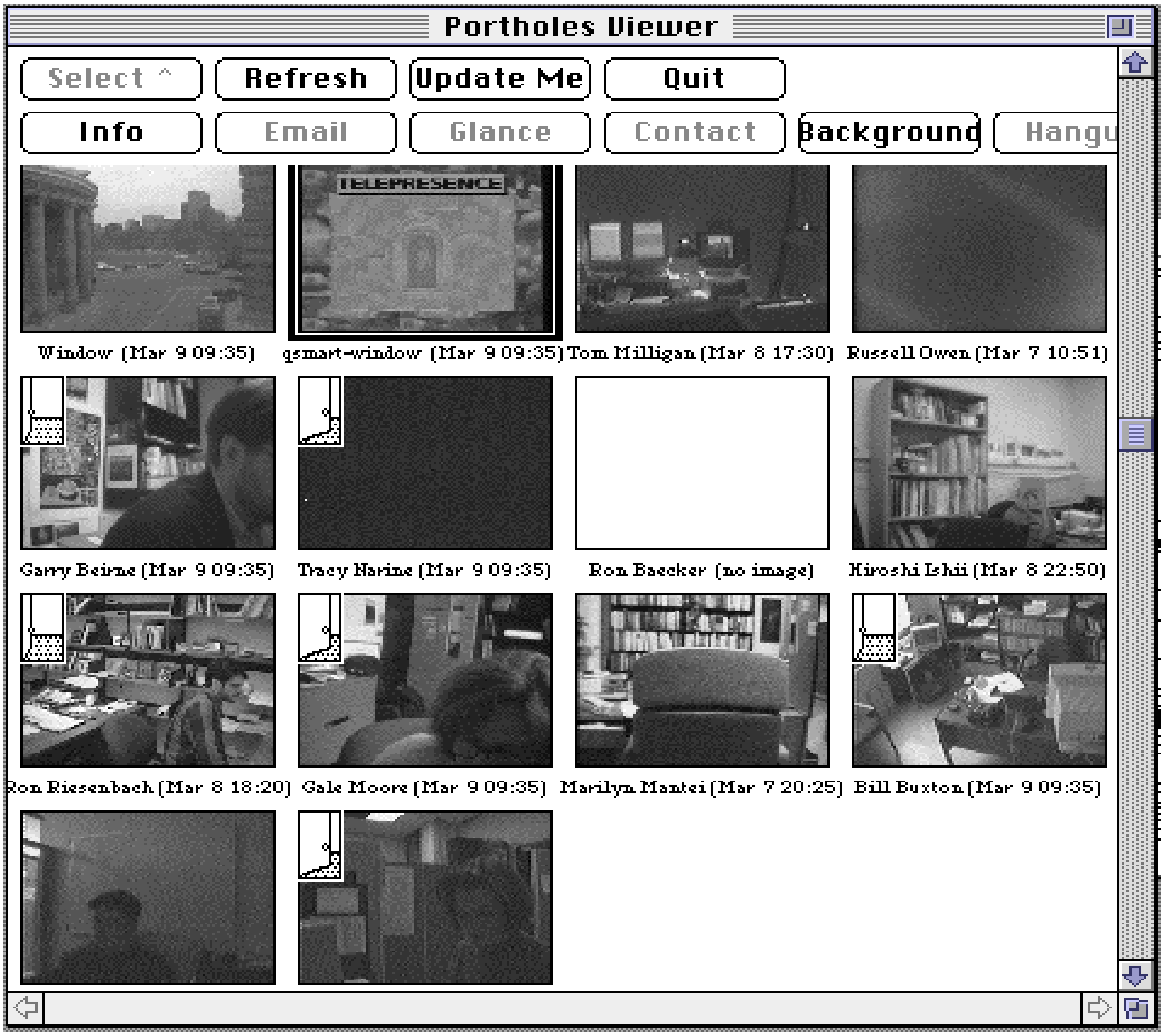


Figure 15: Portholes: Peripheral Awareness of One's Social Network

The original Portholes system was a joint development of Xerox PARC and Rank Xerox EuroPARC. Every 5 minutes, in the case of the Telepresence Project, a snapshot of each member of the workgroup is distributed to all other members of the group. In the Telepresence implementation, this is accompanied by an icon of that member's door icon (see Figure 11). The resulting tiled image of one's workgroup affords a strong sense of who is available when. It also can serve as a mechanism for making contact, finding phone numbers, and avoiding intruding on meetings.

Active Sensing and the Reactive Environment

Introduction

In the examples thus far, the use of computational and video technologies has been complimentary. However, the net effect has been cumulative. Our argument is that the benefits go well beyond this. In this section we will show that there is a synergy that occurs when these two classes of technologies are used together. The result can be far more than the sum of the parts. We will illustrate this in the examples in this section, of what we call *proximal sensing, reactive environments,* and *context-sensitive interaction.*

Video, Portholes and "Call Parking"

We can leverage the video and computational technologies of Ubiquitous Media by recognizing that the same cameras that I use for video conferencing can give my computer "eyes." Furthermore, the same microphone through which I speak to my colleagues can also provide my computer with an "ear."

Design Principle 5: *Every device used for human-human interaction (cameras, microphones, etc.) are legitimate candidates for human-computer interaction (and often simultaneously).*

Krueger (1983, 1991) has demonstrated how video signals of the user can be effectively used in human-computer interaction. By mounting a video camera above the Active Desk, and feeding the video signal into an image processing system, one can use the techniques pioneered by Krueger to track the position of the hands over the desk. This is illustrated in Figure 16, which illustrates a prototype system developed by Yuyan Liu, in our lab. In the example, the system tracks the position and orientation of the left hand as well as the angle between the thumb and forefinger. The resulting signal enables the user to "grasp" computer-generated objects displayed on the desk's surface.

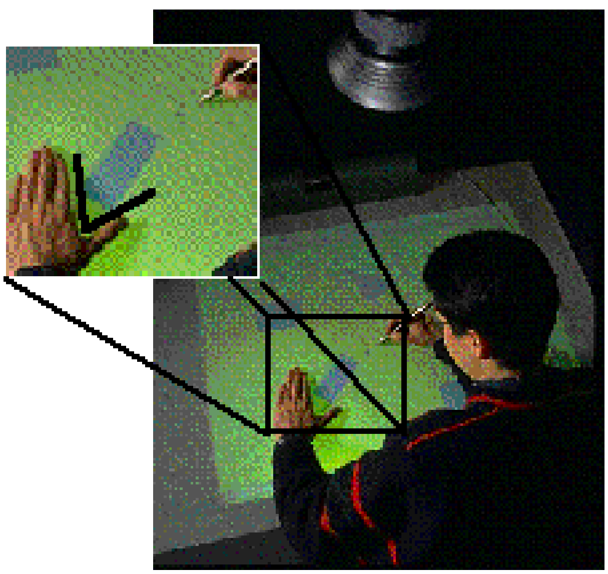


Figure 16: Using the Hand as Input
In this prototype system developed by Yuyan Liu, a video camera mounted above the Active Desk captures the position and orientation of the operator's left hand. It also determines the angle between the thumb and forefinger. The user is able to grasp and interact with computer generated objects displayed on the desk's surface.

Such use of video is relatively non intrusive. One need not wear any special gloves or sensors. The system sees and understands hand gesture much in the same way that people do: by watching the hands or body.

Another simple, yet effective, use of video to support interaction can be demonstrated by an extension of the *Portholes* application. A prototype written by Luca Giachino, a visiting scientist from CEFRIEL in Milan, Italy, demonstrated this. The underlying observation is that two Portholes images in a row constitute a motion detector.

By comparing two frames, if more than 40% of the pixels change, there has been motion. Hence, one can have a rather reliable indication whether there is someone there. By keeping 1 bit of state for each frame, one can determine — within 5 minutes of resolution — if someone is still there, still away, come in or gone out.

With this observation and the resultant code, the mechanism for a new type of "call parking" is provided. If I want to call you, I could look up at Portholes to see if you are there. If so, I could double click on your image to assert a connection. Otherwise, I could instruct the system that I want to talk to you. In the background, while I get on with other work, it could monitor the state of your office and alert me when you appear to be in *and* (by virtue of your door state) when you are available. The benefit of such a utility increases dramatically when it is a conference call that one wants to set up.

The Digital Desk

The information store of most of us consists of two solitudes: the information that is electronic form and that which is not (such as all the paper in our filing cabinet. As was discussed earlier, it is important to be able to maintain a shared space of task as well as person. And, yes, video can transmit images of documents and other nonelectronic artifacts, while the computer can distribute the electronic ones. But the combined forces of the technologies can do even better than that, as has been shown by the *Digital Desk* of Wellner (1991).

With this system, the video camera enables the computer to "see" what is on the desktop. Like the systems of Krueger, it enables the computer to see the actions of the hands on the desk, and to use this as input. It also enables the computer to "see" documents and objects on the desktop. Here again the potential exists for recognition. In Wellner's working prototype, for example, the camera was used to scan alphanumeric data to which optical character recognition techniques are applied, thereby enabling the computer to "read" what is on the desk. Wellner's system is an excellent example of the concepts of Ubiquitous Media in action.

Doors Revisited: the "Door Mouse"

The cameras and microphones found in the office are not the only sensory devices that can be taken advantage of in the domain of Ubiquitous Media. Other alternatives include the full repertoire of motion and proximity sensors used in home automation and security. Let us revisit an earlier example, the specification of door state, as a case in point.

Specifying door state using the mechanism illustrated in Figure 11 preserves the protocols of the physical world *by metaphor*; however, it fails to comply fully with the design principal of using the same mechanism in both the electronic and the physical domain. The reason is that while the protocols are *parallel*, they are not *one.* One still has to maintain two systems: the physical door and the logical one, as represented in the computer application.

Using the *physical* door to control both mans that accessibility for both electronic and physical visitors are handled by the same mechanism. Hence (naturally subject to the ability to override defaults), closing my physical door is be sensed by the computer and prevents people from entering physically or electronically (by phone or by video). One action and one protocol controls all.[5]

Such a system was implemented in a number of rooms in our lab by a student, Andrea Leganchuk. Her simple but elegant solution is illustrated in Figure 17.

> ***Observation:*** *A door is just as legitimate input device to a computer as are a mouse or a keyboard.*

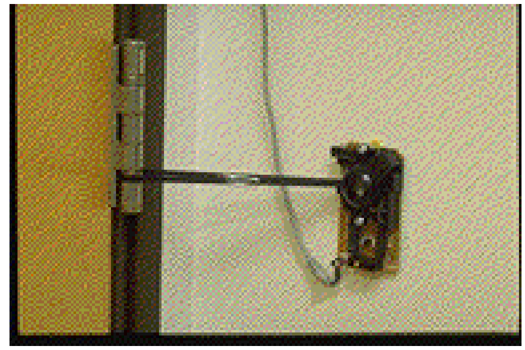


Figure 17: The "Door Mouse"
This is a mechanism for instrumenting the door such that its current state can be monitored by my computer. It was designed and implemented by Andrea Leganchuk. It consists of a Macintosh computer's mouse, with the cover removed, screwed onto the wall. A belt couples the door hinge to one of the mouse's shaft encoders. The mouse is then connected to the computer.

Proximal Sensing and Context

What characterizes the previous examples is the increased ability of the computer to sense more than just the commands that are typed into it. Our experience suggests that computation is moving towards a future where our systems will respond to more and richer input.

One hint of this today is *remote sensing,* the gathering of data about the earth and environment by sensors in satellites. What we are describing is similar, except the sensors are much closer, hence the term *proximal sensing.* In this case, it is the ecology and context of the workspace which is being sensed.

> *When you walk up to your computer, does the screen saver stop and the working windows reveal themselves? Does it even know if your are there? How hard would it be to change this? Is it not ironic that, in this regard, a motion-sensing light switch is "smarter" than any of the switches in the computer, AI notwithstanding?*

We see this transition as essential to being able to deliver the expanded range of functionality being promised as a result of technological convergence. Our perspective is that if considerable complexity is not off-loaded to the system, much (if not most) of the promised functionality will lie beyond the complexity barrier, or the users *threshold of frustration.* Our final example briefly introduces some of our ongoing work which is based on this premise.

Reactive Environment

The way in which proximal sensing and context-sensitive interaction can help reduce complexity while supporting new services is illustrated in our final example, an augmented meeting room. Much is promised in the way of meeting support by new technologies. Videoconferencing, electronic whiteboards, audio and video based meeting capture and annotation and electronic presentations that support video and computer graphics are just some

[5] In reality, it is probably wrong to hard-wire such protocols into a system. The meaning of door state is culture specific, for example. As the ability of a system to sense the context within which it is to react increases, so must the quality and flexibility of the tools for user tailoring of those actions. The examples that we give are to establish another way of thinking about systems. They are not intended to provide some dogma as to specific designs.

examples. The components for nearly all of these services are now commercially available. And yet, our ability to deliver them in a way that augments a meeting, rather than intruding upon it, is limited, to say the least. Their being delivered to a techno-novice in a walk-up-and-use conference room is virtually unthinkable.

The reason is the amount of overhead associated with changing the state of the room to accommodate the changing demands and dynamics of a typical meeting. Take a simple example. Suppose that you are in a video conference and someone asks, "record the meeting." This turns out to be nontrivial, even if all of the requisite gear is available. For the meeting to be recorded, the audio from both sites must be mixed and fed to the VCR. Furthermore, the video from each site must be combined into a single frame using a special piece of equipment, and the resulting signal also fed to the VCR. Somehow, all of this has to happen. And recognize that the configuration described is very different than if just a local meeting was to be recorded, a video played back locally, or a video played back so that both a remote and local site can see it.

In each of these cases, let us assume that the user knows how to perform the primary task: to load the tape and hit *record* or *play*. That is not the problem. The complexity comes from the secondary task of reconfiguring the environment. However, if one takes advantage of proximal sensing, the system knows that you put a tape in, which key you hit (play or record), and knows if you are in a video conference or not, and if so, with how many people. Hence, all of the contextual knowledge is available for the system to respond in the appropriate way, simply as a response to your undertaking the simpler primary task: loading the tape and hitting the desired button.

Over the past year, we have been instrumenting our conference room (the one seen previously in Figure 8), in such a way as to react in such a way. Furthermore, we have been doing so for a broad range of conference room applications, in order to gain a better understanding of the underlying issues (Cooperstock, Tanikoshi, Beirne, Narine,Buxton, in press).

Summary and Conclusions

We have hit the complexity barrier. Using conventional design techniques, we cannot significantly expand the functionality of systems without passing users' threshold of frustration. Rather than adding complexity, technology should be reducing it, and enhancing our ability to function in the emerging world of the future.

The approach to design embodied in Ubiquitous Media represents a break from previous practice. It represents a shift to design that builds upon users' existing skills, rather than demanding the learning of new ones. It is a mature approach to design that breaks out of the "solution-in-a-box" *super appliance mentality* that dominates current practice. Like good architecture and interior design, it is comfortable, non intrusive and functional.

The Ontario Telepresence Project has clearly shown that to reap the benefits of ubiquitous media, we require a rethinking of how we define, teach and practice our science. Following the path outlined above, the focus of future research in this area must be to apply our skills in technology and social science to both refine our understanding of design, and establish its validity in those terms that are the most important: human ones.

References

Bly, S., Harrison, S. & Irwin, S. (1993). Media Spaces: bringing people together in a video, audio and computing environment. *Communications of the ACM,* 36(1), 28-47.

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

Cooperstock, J.R., Tanikoshi, K., Beirne, G., Narine, T., Buxton, W. (in press). Evolution of a Reactive Environment. To appear in the *Proceedings of CHI'95,* May, 1995.

Dourish, P. & Bly, S. (1992). Portholes: Supporting Awareness in a Distributed Work Group. *Proceedings of CHI '92*, 541- 547.

Elrod, S., Hall, G., Costanza, R., Dixon, M. & Des Rivieres, J. (1993) Responsive office environments. *Communications of the ACM,* 36(7), 84-85.

Fields, C.I. (1983). Virtual space teleconference system. *United States Patent 4,400,724,* August 23, 1983.

Gaver, W., Moran, T., MacLean, A., Lövstrand, L., Dourish, P., Carter, K. & Buxton, W. (1992). Realizing a video environment: EuroPARC's RAVE System. *Proceedings of CHI '92*, 27-35.

Ishii, H., Kobayashi, M. & Grudin, J. (1992). Integration of inter-personal space and shared workspace: Clearboard design and experiments. *Proceedings of CSCW '92,* 33 - 42.

Krueger, Myron, W. (1983). *Artificial Reality*. Reading: Addison-Wesley.

Krueger, Myron, W. (1991). *Artificial Reality II.* Reading: Addison-Wesley.

Mantei, M., Baecker, R., Sellen, A., Buxton, W., Milligan, T. & Welleman, B. (1991). Experiences in the use of a media space. *Proceedings of CHI '91, ACM Conference on Human Factors in Software*, 203-208.

Russ, Charles (1925). An instrument which is set in motion by vision. Discovery, Series 1, Volume 6, 123-126.

Sellen, A. (1992). Speech patterns in video mediated conferences. *Proceedings of CHI '92, ACM Conference on Human Factors in Software,,* 49-59.

Sellen, A., Buxton, W. & Arnott, J. (1992). Using spatial cues to improve videoconferencing. Proceedings *of CHI '92*, 651-652. Also videotape in *CHI '92 Video Proceedings.*

Stults, R. (1986). Media Space. Systems Concepts Lab Technical Report. Palo Alto, CA: Xerox PARC.

Vowles, Harding (1992). Personal Communication regarding the *TELEMEET* Project, Feb. 1970, United Church Berkely Studio, Toronto.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3), 94-104.

Wellner, P. (1991). The DigitalDesk Calculator: Tactile manipulation on a desktop display. *Proceedings of the Fourth Annual Symposium on User Interface Software and Technology (UIST '91)*, 27-33.

Wellner, P., Mackay, W. & Gold, R. (Eds.)(1993). Computer-Augmented Environments: Back to the real world. Special Issue of the *Communications of the ACM*, 36(7).

SECTION 3:
REPORT OF THE DIRECTOR OF ENGINEERING

Telepresence --- Current and Future Technologies for Collaboration
(Did the Interstate System kill Route 66?)

Gerald M. Karam
Dept. of Systems and Computer Eng.
Carleton University
1125 Colonel By Drive
Ottawa, Ontario, Canada K1S 5B6
613-788-5749 or karam@sce.carleton.ca

Abstract

Telepresence is the art of enabling social proximity despite geographical or temporal distances through the integration of computer, audio-visual, and telecommunications technologies. The *Ontario Telepresence Project (OTP)* has constructed a variety of software and hardware systems to support telepresence applications and it was the mandate of the project to deploy these applications in the field to evaluate their acceptance and use by people in an office environment. The applications that were developed included: (1) the *Telepresence Media Space (TMS),* a system for computer-mediated intra-site and inter-site synchronous audio/video (A/V) communications employing strategically placed cameras and monitors, (2) *Postcards*, a background awareness tool using low frame rate snapshots of users (based on Xerox's Portholes), (3) video and voice mail, (4) receptionist and video automated attendant functions for video call management, and (4) room level A/V management. Beginning with conventional H.261 dial-up A/V codecs for long haul connections, OTP moved its applications to a trial ATM network for long haul A/V and data connections and explored a number of uses of its technologies in the context of this emerging telecommunications capability. This report describes the OTP approach and experiences in the transition from current telecommunications technology to future technologies, and the how the two technologies can co-exist. Numerous observations are made for end user applications such as: point-to-point and multipoint meetings; virtual open office; shared public spaces; and remote lecturing. Suggested deployment ideas and open questions are provided a starting points for further research.

1. Introduction

Telepresence is the art of enabling social proximity despite geographical or temporal distances. It is a set of computer, audio-visual, and telecommunications technologies, which are carefully integrated to permit people to work together using technology as an intermediary [28]. In the activities of the Ontario Telepresence Project (OTP) [28] we have constructed a variety of software and hardware systems to support telepresence applications and it was the mandate of the project to deploy these applications in the field to evaluate their acceptance and use by people in an office environment. The applications that were developed included: (1) computer-mediated intra-site and inter-site synchronous audio/video (A/V) communications employing strategically placed cameras and monitors, (2) background awareness using low frame rate snapshots of users (based on Xerox's Portholes [2]), (3) video and voice mail, (4) receptionist and video automated attendant functions for video call management, and (4) room level A/V management.

In the last phase of OTP research, several of the applications were re-engineered to function over a 45 Mb/s ATM link between the OTP lab at Carleton University in Ottawa, Canada, and the OTP lab at University of Toronto, in Toronto, Canada. As a result, we had an opportunity to "live in the future" and study the ways in which current and future technologies for collaboration might be combined and interact with one another. As our original long distance A/V link was handled by conventional 112 Kb/s dial-up video (H.261 compliant), this change to broadband A/V transmission was akin to moving from an old 4-lane blacktop highway, such as the famed "*Route 66*", to a brand new, high-speed, state-of-the-art Interstate highway, such as I-40 that snakes its way across the American southwest, bypassing much of the roadway originally serviced by *Route 66*.

ROUTE 66

Connoisseurs of American highway history will know that the Interstate system did indeed kill *Route 66*, but this didn't happen quickly, nor did it happen completely. The OTP experience with both current and future A/V transport mechanisms replays much of the history of transition from *Route 66* to interstate highways, and the analogies serve has historical reminders for the telecommunication service providers, application developers, and new businesses that will spring up around the so-called "*Information Superhighway*".

The OTP trials over conventional and ATM networks has led to some understanding of the issues that will make or break synchronous collaborative applications, and in many cases, raised more questions than were answered.

2. Building the Old Road

The goals of the Ontario Telepresence Project were two-fold: (1) to study methods of deployment for telepresence applications in workplace environments, the way in which telepresence technology is incorporated into workpractice, employ iterative, user-centred design, and to recommend design goals and strategies for telepresence technologies based on this work; and (2) develop telepresence technologies; evaluate them in experimental situations, and hopefully in field studies in a workplace environment, and then recommend design goals and engineering trade-offs based on the development of these technologies. The philosophy of application development was to use off-the-shelf components to create a possible future workplace. This was tied to our belief that "in order to design the future we must have lived it in the past". The project was funded by the Province of Ontario and several industrial partners with a total cash budget of about CAN$3 Million over a three year period ending in December of 1994; in-kind contributions of seconded staff, equipment and services added another CAN$1 Million to the value of the project. During its lifetime, OTP employed up to 12 full-time staff (the majority of which were engineers and computer scientists), and involved the efforts of many graduate students and faculty researchers.

ROUTE 66

"66 is the path of a people in flight, refugees from dust and shrinking land, from the thunder of tractors and shrinking ownership ... they come into 66 from the tributary side roads, from the wagon tracks and rutted country roads. 66 is the mother road, the road of flight"
--- John Steinbeck from *The Grapes of Wrath* [6]

The engineering phases of the project produced a number of applications that were tested, deployed for use within the project, and in some cases in field study sites. The internal use of applications often provided a good first order evaluation because the physical organization of OTP had been established to require the use of collaborative tools. Specifically, half of the project was located in Toronto, and the other half was located in Ottawa --- a separation of some 270 miles; furthermore, in Toronto some members of the project were located on the second floor of the building that housed the project, and some on the fourth floor --- this distance was sufficient to make a casual "walk-by" rather tiresome. But distance alone does not justify the necessity for collaborative tools; there must be a need to communicate. Communications among staff members was fostered by splitting the development and management teams horizontally, thus the applications engineers and systems engineers were split across sites, and the various managers who coordinated the research effort were similarly divided. In short, OTP was a testbed for its own tools, by design.

The major applications developed by OTP included: (1) the *Telepresence Media Space (TMS)*; (2) *Postcards* --- a background awareness tool; (3) *Video Mail*; (4) *Ubiquitous Voice Mail*; (5) the *Desk Area Network (DAN)*; (6) the *Video Receptionist*; and (7) a mock-up of a *Video Automated Attendant*. Of these, only TMS, Postcards, and DAN are discussed in detail in this article as they were used most heavily in the deployment over the ATM network.

2.1 The Telepresence Media Space

The phrase "media space" captures the idea that audio and video devices (sources, sinks) and transmission are managed collectively for a site (a "local area", such as a floor or a building). This is different from typical notions of desktop video conferencing through a workstation or conventional conference room video conferencing. OTP uses the concept of a *media ecology*, in which A/V is not locked in a box or room to a specific purpose, but in fact exists where and when it is needed, to best serve the objective of maintaining social proximity. For example, a meeting between three people in a private office, including one remote visitor, is not well served by a 4 sq.in. image of the visitor on someone's workstation that is facing away from the meeting area. Not only is the *video surrogate* (camera, visual display, microphone, and speaker) of the visitor improperly placed in the room, but the image is too small to maintain a "presence" in the room, thus, the visitor quickly becomes a disenfranchised member of the meeting. The same situation is equally poorly handled by most boardroom conferencing systems, which are rarely designed for the type of informal meeting previously described --- often the video surrogate is too far removed from meeting table, or projected with too large an image, or placed in such a way, that all discussion is directed at the remote visitor, rather than the general group in awkward attempts "not to leave the visitor out of the discussion". The OTP approach is to use a separate video surrogate for each distinct social function (such as face-to-face meeting, small group meeting, visitor "dropping in", etc.). This multiplicity of functions and A/V devices thus demands a more sophisticated approach to managing devices, connections and services.

The essence of the TMS is a site of A/V devices controlled by an audio/video private branch exchange (A/V PBX). The A/V PBX is composed of software servers running on a host computer, that control a number of devices, most notably the A/V switching fabric. The architecture of the TMS is shown in Figure 1. The key components are a set of conventional off-the-shelf devices, as follows: (1) a *node* --- that includes one or more cameras, monitors, microphones, speakers,

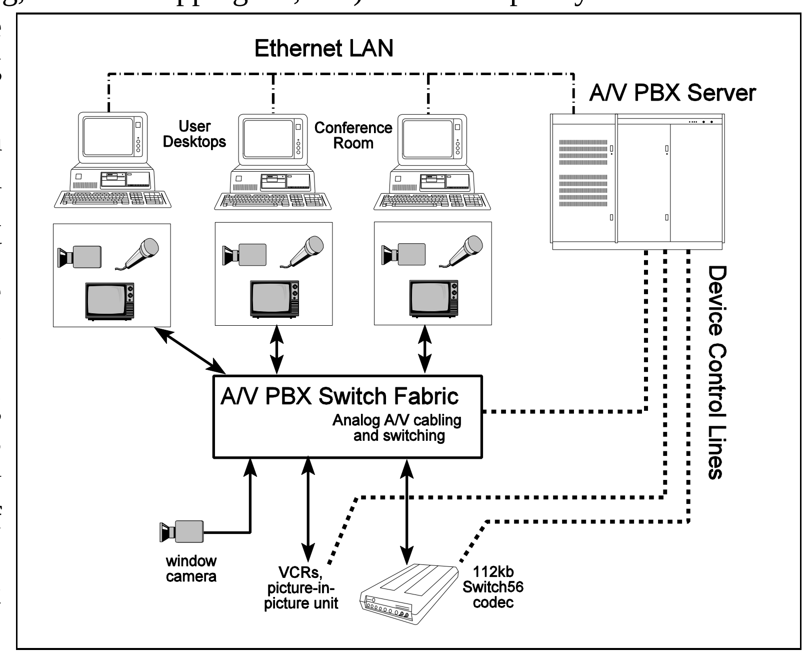


Figure 1: Telepresence Media Space

associated with a *client* computer, that executes the TMS client applications; (2) an Ethernet LAN that interconnects the client computers and the A/V PBX host computer; (3) an analog A/V network connecting all A/V devices (most with bi-directional circuits) to the A/V switching fabric (one or more analog A/V switches); (4) various VCRs, audio mixers, preamplifiers, and a *picture-in-picture* unit (used for multipoint meetings); and (5) an H.261 compatible dial-up-video codec operating over two Switch56 (2 x 56 Kb/s) data lines (ISDN can be used instead) provided by the public switched telephone network (PSTN). The extensive use of analog devices and transmission for audio and video reflects the realities of state-of-the-art A/V --- the only inexpensive way to achieve high quality A/V in a local site is by analog technology. Current digital technology requires specialized hardware in each user's PC, a significant bandwidth from the LAN, and still will not achieve the image size and quality of analog video. Furthermore, it is a requirement that video be projected outside of the client computer, and quite possibly, involve multiple video sources and sinks. Digital A/V networks may become less expensive, but not in the short term, especially as organizations will be reluctant to replace existing computer networks.

The TMS provides a number of different functions to the users of the system. Some of the main services are summarized below. Note that these functions are described first in the context of single site, and then later extended to a multi-site model.

Point-to-point and multipoint meetings. Users can contact one another from a node (which includes both conference rooms and desktops), by using the TMS client application shown in Figure 2. User names appear in a list, that is used to select a person with whom to communicate, and then the *contact* button is used to establish the connection. A *door state* is used to denote a user's accessibility to others; there are four states: (1) *door open* --- others may enter the user's virtual space without explicit confirmation; (2) *door ajar* --- others must "knock" before entering and receive a positive confirmation before a connection is established, however short segments of video-only (*glances* or *Postcards*) could be used without explicit confirmation; (3) *door closed* --- like door ajar, except that no video segments are available; and (4) *door locked* --- no connections are permitted. A user's current accessibility is shown by an icon in the upper left corner of the application and the accessibility of other users are shown by icons beside their names in the name list. Once a connection is established, a user can add other users into a multipoint meeting by highlighting their name and using the contact button again. Multipoint meetings are accomplished by using a picture-in-picture unit to divide the video screen into four quadrants, where each participant (up to a maximum of four) occupies one of the quadrants. The audio signals are mixed so that an audio source receives a combination of all audio sources except it's own. We have found this form of

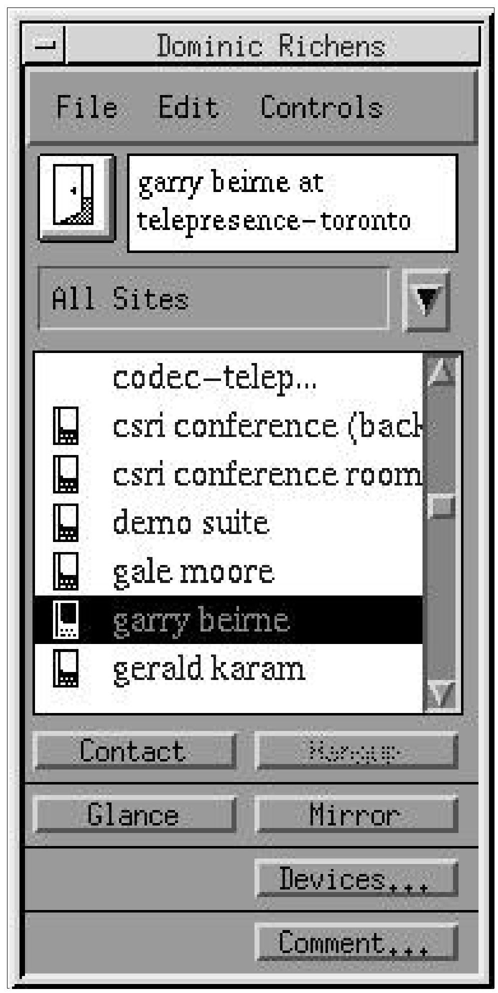


Figure 2: TMS Client Application

spatial multiplexing to be very superior to voice-activated switching, that uses the voice of the current speaker to select the image that is shown.

Shared video devices. Audio/Video devices can be shared by users for a variety of purposes. For example, VCRs can be controlled and shared by users (although this was not built into the simple interface of Figure 2). A video tape could be placed in a common VCR and users can share the VCR to watch the tape as a background activity, and yet still be able to make and accept foreground connections. The simplest and most unassuming shared device is the *window camera*. In this case, a camera is pointed through a window that shows a live outdoor scene to the community. Users, particularly those in windowless offices, can select the window camera as a background shared device, and thus use their video monitors to gain a sense of the weather or time of day, and share a common view with members of the group; again, this would not preclude the use of the video connection for initiating or accepting foreground connections. The *devices* button on the user interface gives access to shared devices.

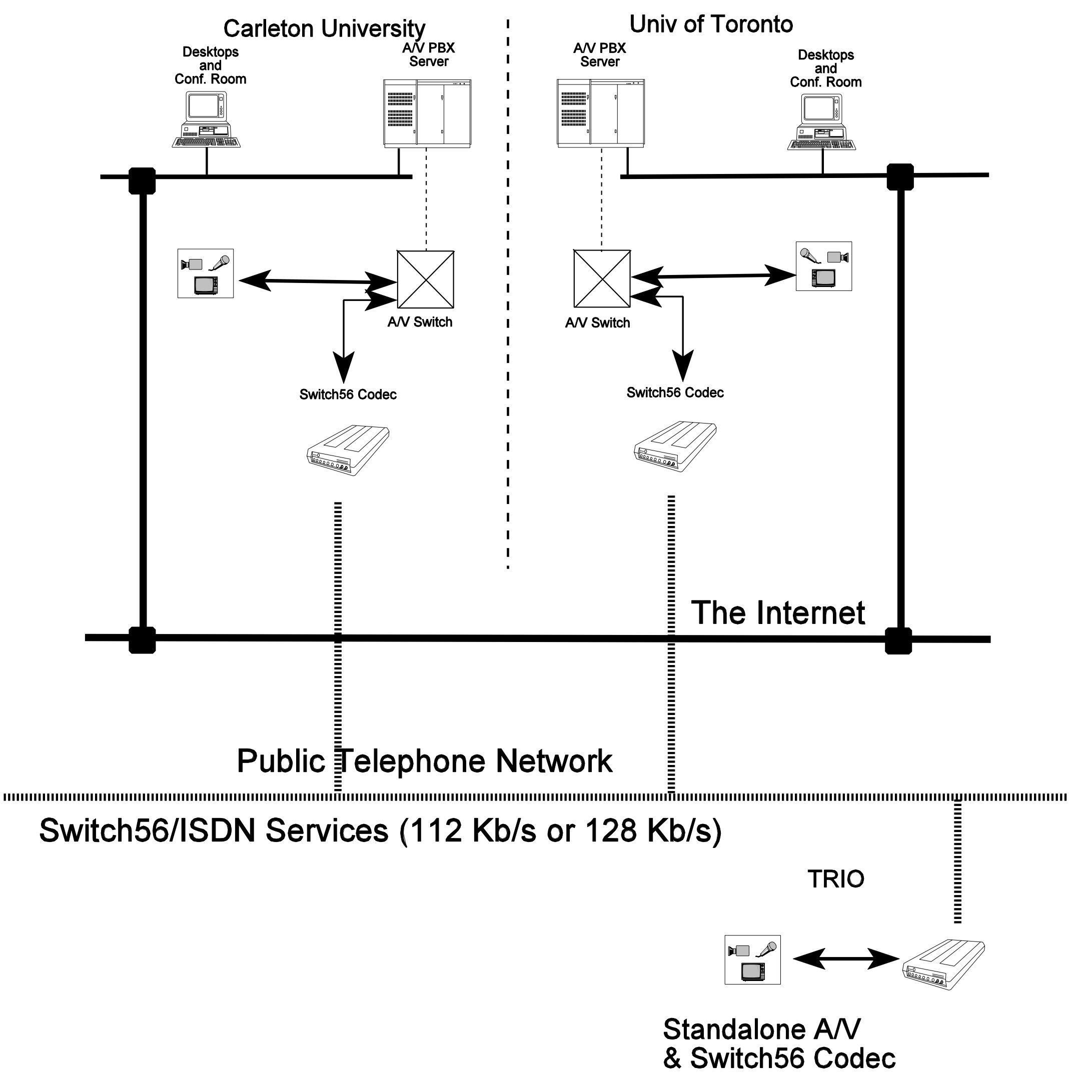


Figure 3: Multisite TMS Installation with Dial-up Video Codecs

Integrated desk-level and room-level interactions. The TMS does not distinguish between nodes in private offices and those in conference rooms (except for names). This allows users to participate in conference room meetings from their desks, either as active members (asking questions, presenting information), or as passive members (listening only, in which case the conference room nodes can be treated as shared devices). Furthermore, users may enter a conference room from different logical positions. For example, the client application in Figure 2 shows two positions: "**csri conference (back)**" and "**csri conference room**"; in the first case, the user enters and views from the back of the room, and in the second case, from the front of the room. In this example, there are at least two separate A/V lines into the room. Once inside a room, a DAN can be used to move a remote visitor around the room (the DAN is discussed later).

The A/V PBX supports all of these functions through its client-server architecture: (a) clients running on user or room workstations converse with the A/V PBX server to declare ownership of devices, share devices, control devices, and establish connections between devices; (b) the server maintains a database of the devices and possible interconnections, user characteristics, and state information for devices and users; and (c) the server constantly updates each user's client regarding state changes in all users (e.g., door state, login state, etc,).

The TMS supports connections to "the outside world" through two approaches: (1) connections to other TMS sites and (2) connections to foreign sites. In both cases, a variety of short and long haul A/V communications technologies can be used. For a foreign site, the most popular type of connection is through the *outside line* feature (shown as "codec-telep" in the user list of Figure 2), in which the user selects an outside line, and is then able to dial an explicit phone number (or use a speed dial) using a conventional H.261 dial-up video codec. Another choice, that was never explored, but is possible, is an outside line that connects to a dedicated analog connection to a foreign site. For a connection to another TMS site, the A/V PBX servers communicate signalling information over the Internet --- this includes the exchange of state information, connection establishment, and accessibility. Intersite A/V links are provided transparently through dedicated analog A/V for nearby sites or dial-up video for distant sites. Once one or more TMS sites have been interconnected (this is done by configuration files), the user names and rooms become seamlessly integrated. For example, Figure 2 shows users from Ottawa (e.g., "**gerald karam**") and users/rooms from Toronto (e.g., "**garry beirne**", "**csri conference room**"). A user at one site can then contact a user at another site, as easily as a user at the same site. An example of a multisite TMS setup is shown in Figure 3. In this illustration, two sites are connected via the Internet (for signalling) and dial-up video for intersite A/V links. The diagram also shows how a simple standalone dial-up video codec at a foreign site ("TRIO") can be accessed through the outside line feature.

Multisite A/V PBX servers communicate as a fully connected network; while this is not the most effective way for them to function, it was relatively simple to construct. Each A/V PBX server is implemented as two Unix processes: the *Telepresence Application Server (TAS)* and the *Integrated Interactive Intermedia Facility (IIIF)*. The IIIF process manages the A/V devices, connections, and a database of users for a local site; it has no knowledge of other TMS sites. The TAS process maintains knowledge of its local users, as well as users at all other TMS sites to which it is connected. TAS processes at different sites query the IIIF servers at their own site and peer sites to build a database that reflects all of the users in its network of sites. Furthermore, whenever the state of a user at a site changes (e.g., the door state is changed), all TAS processes that are connected to the site (including the local TAS process) are automatically informed of the change. In this way all TAS processes at the connected TMS sites are constantly updated so that their databases reflect the same state information; i.e., they are similar to a distributed database in which local sites each maintain a full copy. An intersite connection is established through the TAS process at the originating site, by issuing transactions to the IIIF processes at the originating and target sites to make A/V connections.

2.2 Postcards

The original Xerox *Portholes* system [2] supports background distributed group awareness by providing a palette of continuously updated (about once every 5 minutes), small "postage stamp" images of persons that are part of the user's *community*. A community refers to those users that are capable of providing images to a group of one or more sites, that in turn exchange and distribute these images. Technically, this system had a narrow capability for a limited model of use: (1) it was composed of clients connected to a server that acquired image stills through a frame grabber and distributed them to clients on a periodic basis; (2) servers at different sites could exchange images to build a distributed community, but the software was not sufficiently reliable nor efficient to support more than two sites; (3) users could not select from the set of images to include in their palette --- they had to take all images, thus all images were sent to all users; (4) users could not control who would have access to their image; and (5) anyone with a copy of the Portholes client and Internet access to the server, could connect to the server and view images --- even if they were not part of the community.

The new *Postcards* system implemented at OTP (and shown in Figure 4) revamped both the social model of use and the engineering implementation, in order to provide an efficient, reliable solution that was more sensitive to the realities of a typical workplace. The social model embedded in the behavior of the tool is *universal reciprocity* augmented by the concept of *public* and *non-public images*. Universal reciprocity states that "I am capable of seeing you, if and only if you are capable of seeing me"; thus, any users that want to be able to view other members of the community must make their image available for distribution and if they choose not to do so, then they will be prevented from viewing the images of the community. Images, for which distribution is governed by universal reciprocity, are *non-public* images. If users (including common images sources, such as the window camera) wish their image to be available to anyone, without the enforcement of universal reciprocity, then they would declare their images to be *public*; public images can be viewed by anyone, including those outside of the community. Finally, to reduce both image traffic and the number of images on a user's screen, a user can select from the set of available image sources, and have only these images displayed in the palette.

Figure 4: *Postcards* Client Displaying a Palette of Images and Selection List

The re-engineering of the application was significant. A *Postcards* deployment consists of three major components: the clients (one per user), the servers (usually one per "site", or local area of potential clients, but possibly serving only an intermediary store-and-forward function), and the frame grabbers (one or more allocated to one server in a site). Permitting more that one frame grabber per site, allows common framegrabbers to be used, or frame grabbers in personal workstations to be used, or any combination; the OTP deployment uses only a common framegrabber per site. Servers communicate in a spanning tree in order to forward images to clients that are at different sites (since a user's community may be spread over a number of sites), thus image transmission is reduced (an image is transmitted only once between two servers), and an image is grabbed and shipped only if there is a client that is specifically requesting the image. Each server maintains a database of images available for subscription (to be viewed), images being requested for subscription, and which users are permitted (public/non-public) to subscribe to an image. Every change to the database information about a user at one site is propagated to all other sites through the servers; i.e., it is a distributed database with a copy maintained at each site.

2.3 Desk Area Network (DAN)

The *Desk Area Network* application allows local control of many A/V devices in a room; this would suit an office with several video surrogates and devices such as VCRs, or a typical conference room. The current version provides for local switching of the A/V sources and sinks so that the room can be configured easily for any particular social use. For example, an office could be set to video tape both ends of a meeting involving local and remote visitors, or be quickly reconfigured to a simple face-to-face meeting. Previously defined *presets* are used to specify known or standard configurations of device interconnections. The devices are wired to a small analog A/V switch (such as an 8 x 8 cross bar switch) that is controlled by the DAN application from the user's workstation. The A/V trunks linking the room to the outside world (typically the A/V PBX in a TMS) are attached to the A/V switch, and form part of its set of connection alternatives.

Further objectives for the DAN include: control of devices such as VCRs, and pan/tilt cameras; automated meeting establishment among several rooms (typically conference rooms); and context sensitive ("reactive") room and device configuration (e.g., placing a paper under a document camera would automatically cause the camera to be routed to the remote visitors). Several aspects of these have been explored as long term research projects by OTP; Karam, McLeod, and Boersma have developed software infrastructure for distributed conference control [3] and Buxton and Cooperstock have developed reactive room concepts and prototypes.

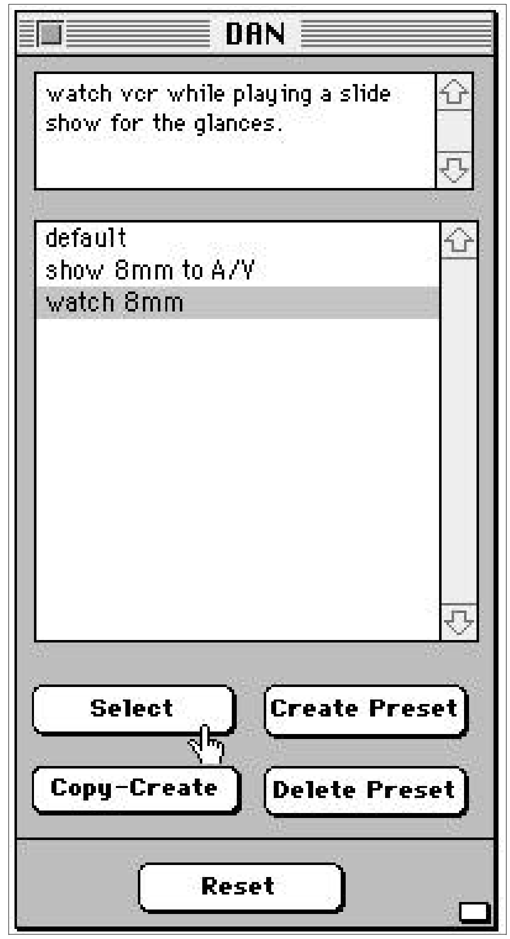


Figure 5: Desk Area Network Application

3. *Experiences on Route 66*

In the 18 month period from January 1993 to July 1994, the engineering team completed the following activities that represented most of our experience on *Route 66*. In early 1993, the basic single site TMS system was developed (starting with software from the *Cavecat* project) that included a prototype client user interface, basic connection services, and an outside line capability through a dial-up video codec. The TMS was deployed initially for OTP use, and then modified to support a field site with two remotely located offices. Even though the field site had two locations, it was managed as one TMS site, in which the analog switches and dial-up video codecs at each site were controlled from a central site (via the Internet) so that they formed a single logical switch fabric. Thus, all connection requests went to one site, from which control signals were issued. This was not very robust because it did not allow the remotely controlled site to function when failures occurred at the controlling site.

This basic system served the project for most of 1993. It was used for a variety of formal meetings: (1) the engineering team met weekly in a "room-to-room" meeting format; (2) the user interface group met weekly at times, and sporadically at others; (3) infrequent management meetings were held; (4) regular formal seminars were conducted during about a 12 month period (into 1994); (5) formal demonstrations and meetings with interested industrial partners and visitors were held; and (6) even an OTP Halloween party was held over the video link, to explore its use in a purely social situation. Informal use also existed, however such contacts, as with all formal contacts, were prearranged by other means (email or telephone). We recognized the need for a true multisite TMS that permitted group transparency; i.e., the ability to pick a name from the user list and automatically contact the individual, including all dial-up video codec signalling. The work on a multisite TMS began in late 1993.

In parallel, the Portholes software was ported to TMS, made operational in fall 1993, and made significant inroads to improving the sense of group awareness, but its limitations were apparent, and at the end of 1993, the designs for the new version, later called *Postcards*, were developed. From the formal seminars grew the demand for improved room control, and this resulted in a prototype of the DAN that came into reliable service in late 1993.

Starting in early 1994, full multisite servers were developed with plans to introduce them to the field site in June 1994. Simultaneously, all client applications (such as the TMS client and the DAN) which until then had been based on Supercard or Hypercard stacks, were re-implemented using a platform independent GUI builder, called XVT. This made an enormous impact on the performance of the interfaces, and the ability to put client applications on a variety of machines (different Unix platforms and the MAC platform). Also, a major reworking of the A/V PBX servers was performed to improve their flexibility, by providing them with an interpretive command language, TCL. By July, a fully operational multisite TMS had been deployed in OTP and the field site, including enhanced servers and new client applications. During this time numerous enhancements were made to the DAN and some maintenance to the existing the Portholes was handled.

The other major project in the first half of 1994 was the development of *Postcards* and the extensive debate on the appropriate social model that was to be supported. Much discussion was had among the OTP community while the infrastructure of servers and framegrabbers was developed. Later in 1994, a social model was fixed and the infrastructure and client were engineered.

On the whole, the end of June 1994 presented a much improved system over the first TMS deployed in mid-1993; people were happy with the performance and convenience of the new facility. The utility and limitations of 112 Kb/s dial-up video were well understood through our many hours of experience in a variety of situations. The blurry image and annoying time delay exacerbated larger meetings, and made many social interactions (jokes, interjection, banter) very awkward. Nonetheless, it allowed work to proceed much more smoothly, and inexpensively, than if we had been forced to do without. But when improved opportunities for collaboration presented themselves, we were keen to try them out. In July 1994, significant plans took shape that would lead OTP to bypass *Route 66* in favor of a brand new Interstate highway, an ATM network connecting the OTP sites in Toronto and Ottawa.

4. *Telepresence over ATM*

The deployment of the TMS and *Postcards* to function in an ATM environment was accomplished without significant re-engineering of the software; the DAN was not affected by this transition, but was used extensively for experiments in remote lecturing over the ATM network. This easy conversion was achieved in part, because of the flexibility of the A/V PBX server design and in part due to the choice of ATM equipment. The ATM network, as illustrated in Figure 6, provided three logical connections between the two OTP sites over a dedicated 45 Mb/s DS-3 line: (a) two full-duplex analog A/V circuits, and (b) an Ethernet link. The key elements of the network were the termination equipment --- two loaned Newbridge 36150 four-port ATM switches that were equipped with two A/V codecs, an Ethernet interface card and an outgoing circuit card. The A/V codecs provided JPEG compressed near NTSC quality video transmission (scalable according to the peak allocated cell rate) and stereo CD quality audio channels.

ROUTE 66

"By the mid-1950s, the wrecks, the breakdowns, and the congestion had made their point: even the people who needed 66 the most understood that it had to change. Many of them had known about (...laws and plans...) for a limited-access highway system. They may not have liked it, and they may not have known what to do about it, but few people on *Route 66* were surprised when the old road was replaced."
--- Susan Croce Kelly from
Route 66 --- the Highway and its People [5]

The remainder of the network provided the path between the terminal equipment. In Toronto, the terminal ATM switch was located a significant distance from the DS-3 access provided by the carrier. Another third loaned Newbridge ATM switch was used to convert the electrical DS-3 to a 140 Mb/s LATM fibre connection so that underground fibre at the University of Toronto campus could be used to reach the terminal ATM switch. From the University of Toronto campus, Bell Canada donated a DS-3 channel to the Ottawa area where it connected to OCRINet, an experimental ATM network connecting numerous Ottawa companies, Universities and Government labs. Carleton University, as one of the OCRINet nodes, provided the DS-3 connection to the Ottawa termination on the Carleton campus.

Permanent virtual circuits were created between the two terminal points for the duration of the trial; nominally, the maximum data rates were established as follows: full duplex video circuits - 18 Mb/s; full duplex, stereo audio circuits - 1 Mb/s; and Ethernet - 1 Mb/s, for a total capacity of 39 Mb/s (2 x 18 + 2 x 1 + 1). This represented most of the available payload capacity of the DS-3 connection. The control interfaces of the ATM switches were put on the respective LANs at the two Universities so that the ATM switches could be remotely controlled, and reconfigured as desired --- including changing bandwidth allocations.

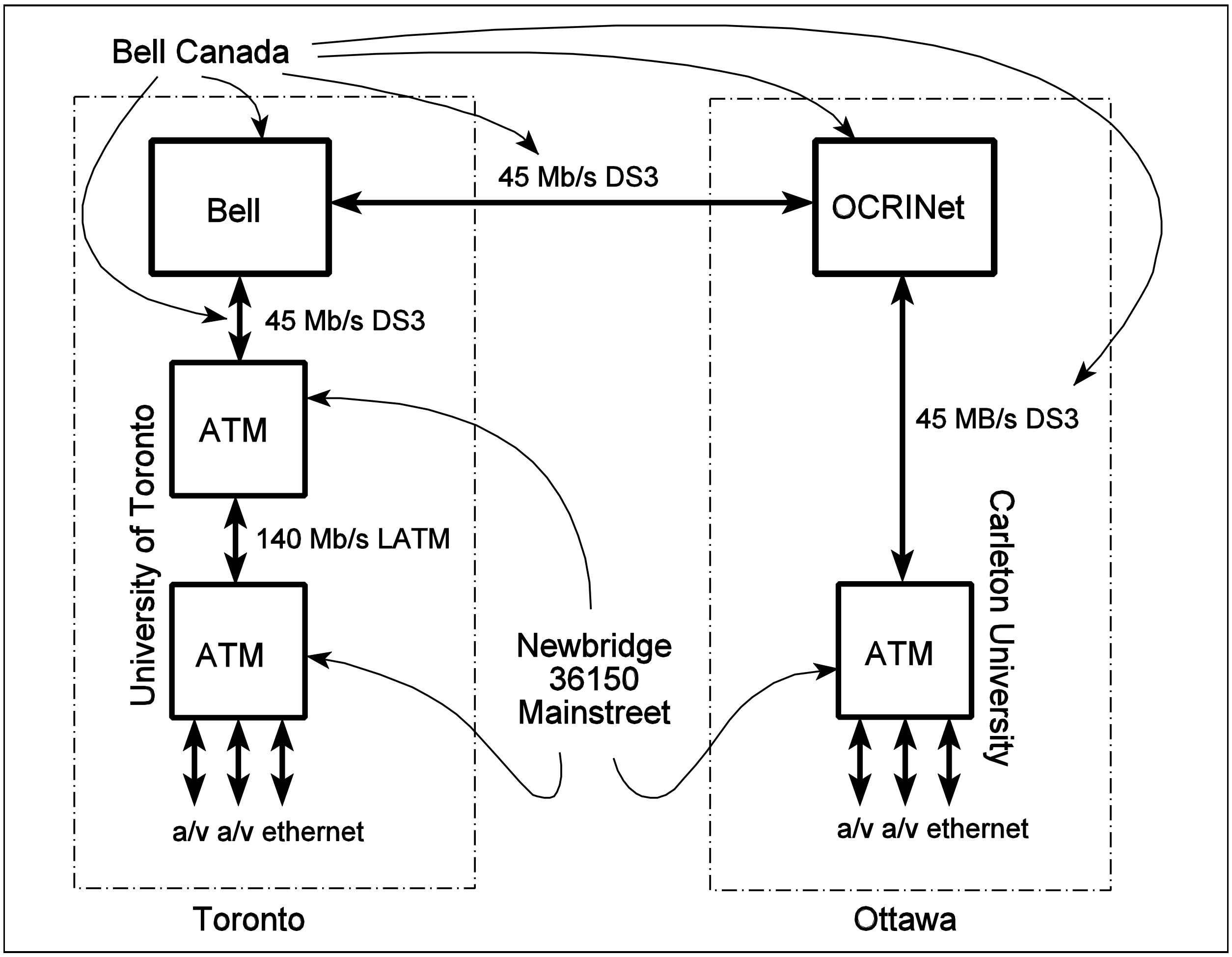


Figure 6: ATM Network Interconnecting OTP Sites in Toronto and Ottawa

There were several goals for the experimentation of telepresence over ATM, although time did not permit all of them to be achieved before the intercity DS-3 service was discontinued.

- Demonstrate single wire integration of OTP multimedia signals (audio, video, data, and control) over ATM.
- Demonstrate the effect of varying the bandwidth for video transmission for different applications; i.e., the quality of the image and the round trip delay.
- Demonstrate the use of multiple full-duplex channels of A/V in telepresence applications.
- Demonstrate the integration of current A/V technology (such as analog, and dial-up video codecs), and future technology (A/V transmission over ATM), and suggest models for how the two might be economically combined.

A number of specific experiments and uses of the telepresence applications were devised in order to illustrate as many of these points as possible. The active term for the trial was about 2-3 months, some of which was unfortunately over the Christmas holiday season. The intercity DS-3 service continued unexpectedly for an extra 2.5 weeks beyond the formal trial and allowed more information and experience to be collected.

4.1 Operation over the ATM Network

The TMS was configured to treat the A/V circuits as "very long", dedicated analog links interconnecting the two sites, as illustrated in Figure 7. Therefore, in order to establish A/V connections over the ATM network, only the configuration files of the A/V PBX servers needed modification so that they would search the direct links between sites before resorting to links provided by the dial-up video codecs. The Ethernet link between the two sites was somewhat more complex to establish as there was an explicit requirement not to permit conventional Internet traffic to be carried by the experimental ATM service. This forced a *gateway* concept to be used, where the host computers for the A/V PBXs at each site, acted as data transmission gateways between the public-access Internet on one side of the gateway, and the private LAN on the other side; i.e., each computer was physically attached to two LANs. Routing information for the Internet was configured so that only the gateway machines could transfer data over the private link.

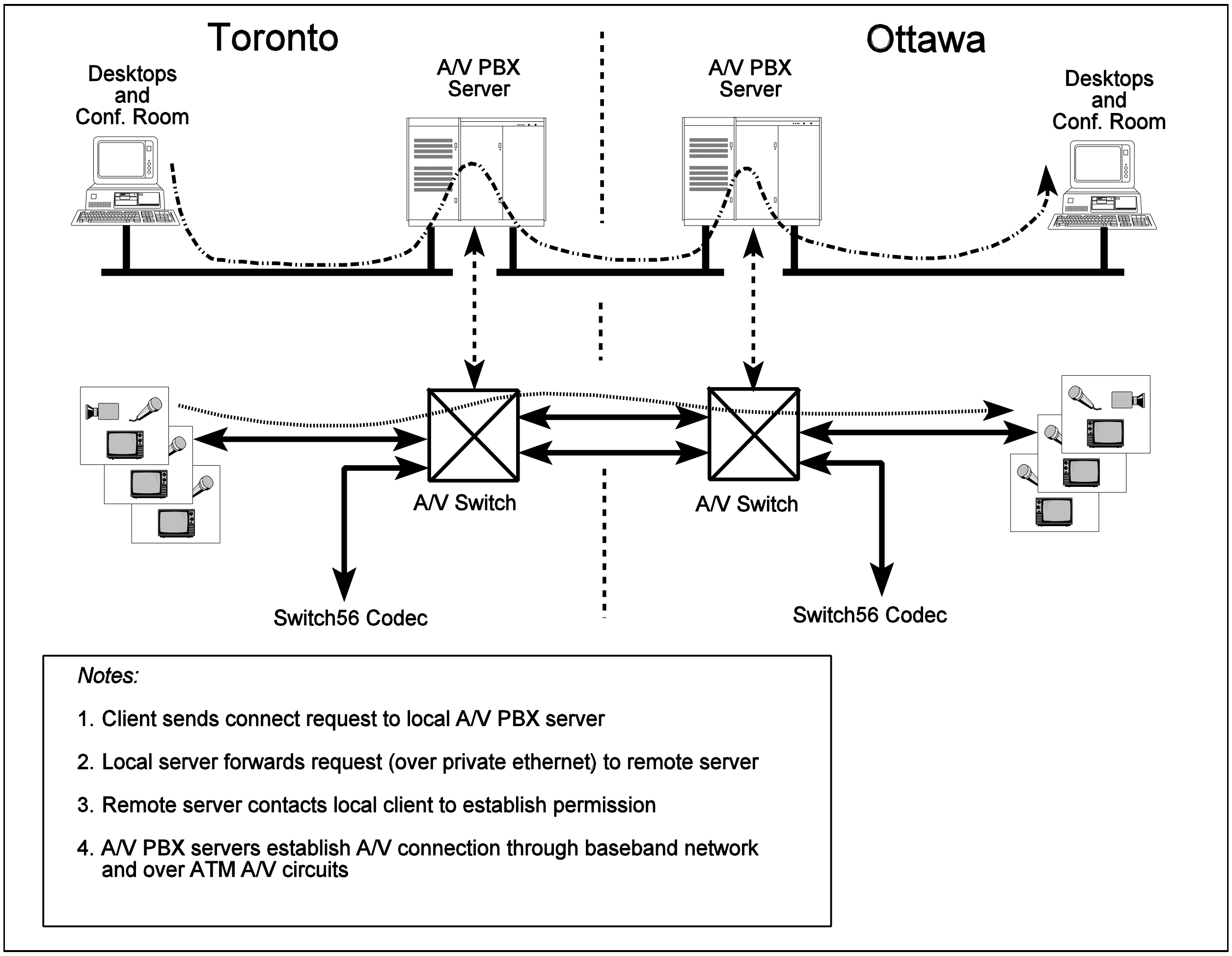


Figure 7: Connecting over the ATM network

Control messages exchanged between servers were routed over the private data link, therefore signalling times were expected to be only marginally longer than those experienced for local connections. The sequence of actions for connection establishment with a remote user, as shown in Figure 7, are as follows:

1. A user invokes a connection to a recipient; the user's client application sends a transaction to the local A/V PBX server over the public LAN.
2. If the A/V PBX server finds that: (a) there is an available intersite A/V link (over the ATM network), and (b) that the remote user is currently available for a video meeting, then it proceeds to establish a connection.
3. The local A/V PBX server allocates a physical path through the switch fabric from the user's node to the available intersite A/V link.
4. The A/V PBX server uses the private LAN to send a transaction to the peer A/V PBX server requesting the completion of the connection at the remote end.
5. The remote A/V PBX server sends a request to the client of the remote user. If the user's door is ajar or closed, then the user must provide an explicit confirmation of the connection

request. If the user's door was open, then the request causes an audible warning of an incoming video call.

6. Once the connection is to be accepted by the remote user, the remote A/V PBX server allocates a physical path through the switch fabric from the remote user's node to the intersite A/V link.
7. The remote A/V PBX server sends a message to the local A/V PBX server, confirming a successful connection; this information then permits the original client transaction to complete.

There is nothing unique about this configuration except for the use of the ATM network for A/V transmission and the private LAN. The same configuration is illustrated later using a fully analog link between sites located in two nearby buildings that share a campus-wide LAN.

The *Postcards* application functions in exactly the same manner as it would in a regular Internet environment, except that the Toronto and Ottawa *Postcards* servers would exchange images and information over the private LAN. Unfortunately, due to technical delays and the limited time and resources available for the ATM trial, *Postcards* was never made to operate over the private LAN.

4.2 Communications Technology Integration

Integration of current and future technologies was accomplished entirely through the flexibility of the TMS design. There are three technologies to be linked: (1) all-analog intersite A/V transmission; (2) digital intersite A/V transmission using conventional H.261 dial-up video codecs; and (3) digital intersite A/V transmission using ATM video codecs. Popular techniques in workstation based digital A/V codecs and data transmission could also be accommodated within the general design, but are not a significant focus of OTP research (with the exception of the *Postcards* application, that is at best a low resolution, very low frame rate variant). Figure 8 illustrates a network of TMS sites (and one non-TMS site) that interoperated during the ATM trial to demonstrate the integration: (1) the Carleton University OTP site; (2) the University of Toronto OTP site in the Sir Sandford Fleming building; (3) the *Information Technology Research Centre* (ITRC) site located on the University of Toronto campus in the D.L. Pratt building (adjacent to the Sir Sandford Fleming Building); (4) the Xerox Webster site, outside of Rochester, NY; and (5) the *Telecommunications Research Institute of Ontario* (TRIO) located in Ottawa. The A/V PBX server interconnections are shown if Figure 9. Users at sites that are directly connected appear in each others name lists, and may be contacted easily. Users at sites that are not directly connected could reach each other only through the outside line feature. Not shown in Figure 8, but appearing in Figure 9, is the ITRC site located at University of Waterloo, in Waterloo, Ontario. It has a dial-up video codec as an external A/V link and uses the Internet for communications with other TMS sites.

ROUTE 66

"The interstates came slowly, a piece here, a stretch of new concrete there, a bypass around a city. Where it had taken twelve years to pave U.S. Highway 66 the first time, it took two-and-a-half times that long to replace it. Besides finally taking the traffic away from the deteriorating two- or three-lane highway, the long gestation of those interstate highways had the effect of scaring many business people into years of inaction."
--- Susan Croce Kelly from
Route 66 --- the Highway and its People [5]

An additional notion in TMS is the *trusted site*; that is, a site has permission to use the outside lines of remote site with which the trusted site has a dedicated A/V connection. For example, the Ottawa OTP site is a trusted site of Toronto OTP, thus the Toronto OTP dial-up video codec could be accessed by Ottawa OTP users (it was shown in Figure 2 as "**codec-telep**"). This allows a site to pool dial-up video codecs even though the codecs are not physically resident at the site (indeed, the

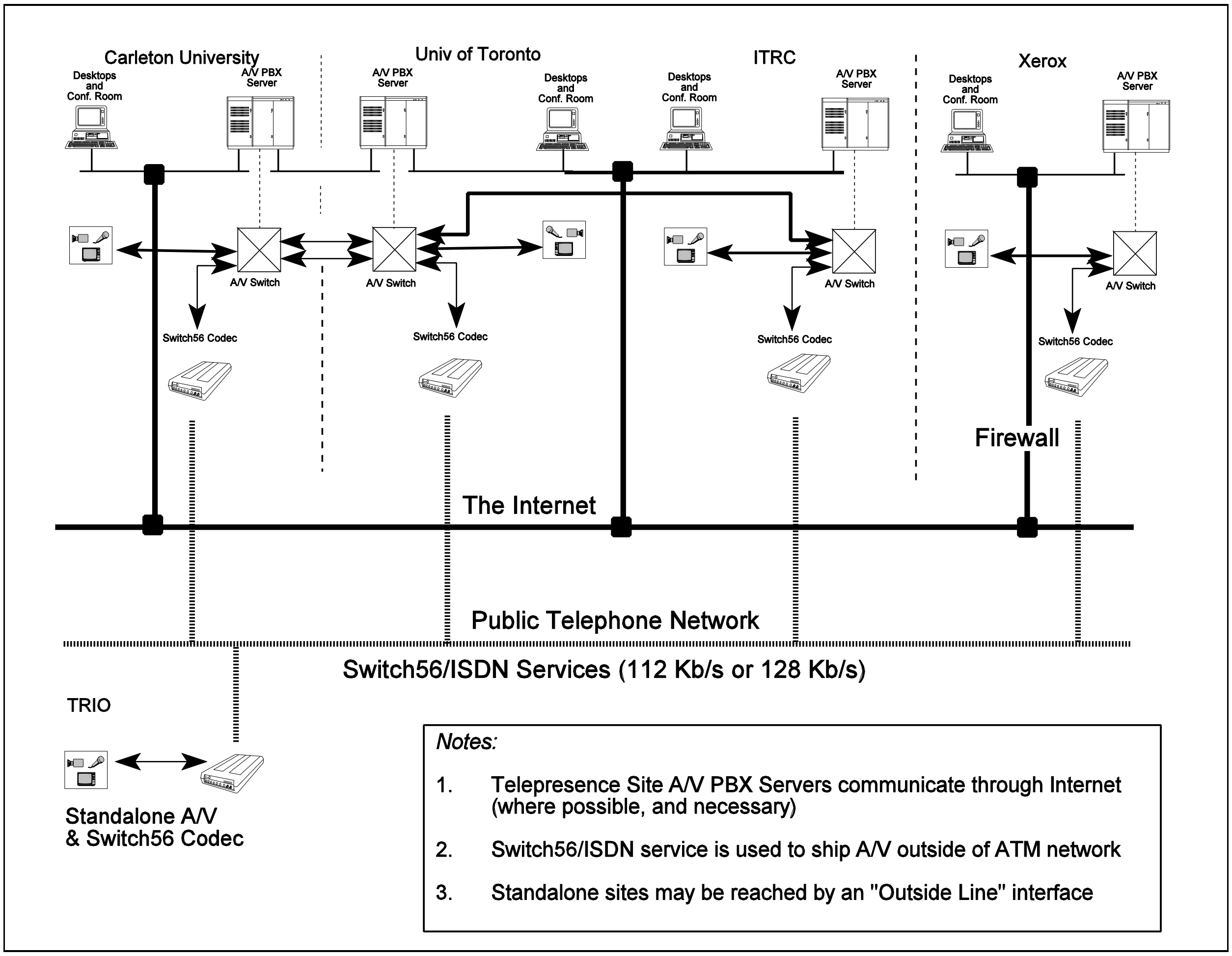


Figure 8: Multisite Integration using Current and Future Transmission Systems

Ottawa dial-up video codec was not in operation for the period of the ATM trial, therefore Ottawa OTP users regularly employed the Toronto OTP dial-up video codec for video meetings to non-TMS sites).

The multisite organization of Figure 8 tests the various combinations of connections as follows. The Toronto OTP site and Toronto ITRC site are interconnected by a dedicated full-duplex analog A/V link over copper unshielded twisted pairs. Their A/V PBX servers communicate over the campus LAN. The ITRC site is known only to the Toronto OTP site (and the Waterloo ITRC site), thus users at ITRC can contact users at Toronto OTP, and vice-versa, however, other sites connected to Toronto OTP cannot contact the ITRC users. The Toronto ITRC and Toronto OTP sites are *trusted sites* with respect to each other. Thus, they can access each others outside lines, effectively giving each a logical pool of two dial-up video codecs. The Waterloo ITRC site is connected to both the Toronto ITRC site and the Toronto OTP site, however the A/V link is only through dial-up video. The Xerox Webster site is connected to the Toronto OTP site through an Internet firewall machine used to provide security for the Xerox corporate network against unwanted access through the Internet. Several modifications were required in the A/V PBX server in order to cross the firewall in a standard and safe manner (similar to the way in which applications like News, FTP and TELNET are handled). The A/V link is through dial-up video. As described earlier, the Ottawa OTP site is connected to the Toronto OTP site through the dedicated A/V links over the ATM network, and communicates over the private LAN; the OTP sites each consider the other as *trusted sites*. Finally, the TRIO installation is a non-TMS site, that is reached by an outside line, dial-up video codec only.

With this arrangement, the Toronto OTP site has a TMS user community of people at the Ottawa OTP site, Xerox Webster site, Toronto ITRC site, and Waterloo ITRC site; furthermore it is a trusted site with Ottawa OTP and Toronto ITRC. This arrangement permits some interesting point-to-point and multipoint meetings to be established. For example, the Toronto OTP site could have a four-way multipoint meeting with Xerox Webster, Ottawa OTP, and TRIO by: (a) connecting to Ottawa OTP over the ATM A/V link; (b) connecting with a Xerox Webster user through its dial-up video codec; and (c) connecting to TRIO by using the outside line from the Toronto ITRC dial-up video codec. In another example, Ottawa OTP could establish a point-to-point meeting with Xerox Webster by using the outside line from the Toronto OTP dial-up video codec, and then manually arranging the meeting; that is, the A/V connection passes from a node through the Ottawa OTP switch fabric, travels over the ATM A/V link to the Toronto OTP switch fabric, where it is routed over the dial-up video codec, through the public switched telephone network, where it reaches the dial-up video codec at Xerox Webster, and then finally passes through the switch fabric at the destination site and terminates at a node. In one final example, Toronto OTP could have a four-way multipoint meeting with the two ITRC sites and the Ottawa OTP site, as follows: (a) Ottawa OTP connects over the ATM A/V link; (b) Toronto ITRC connects over its analog A/V link; and (c) Waterloo ITRC connects over the dial-up video codec.

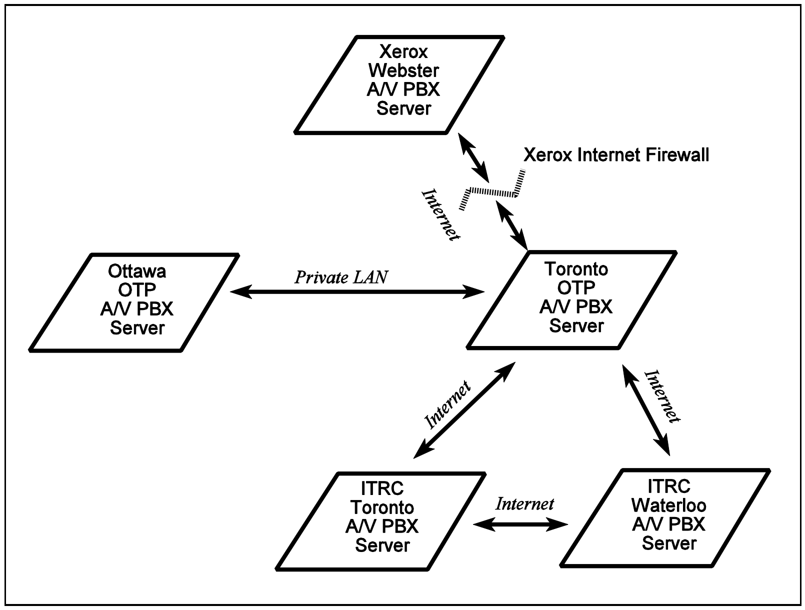


Figure 9: Network of A/V PBX Servers in Five TMS Sites

5. *Experiences on the Interstate*

OTP conducted a number of experiments and studies in order to gain as much insight as possible into the benefits of the ATM capability over our previous experiences with dial-up video. In most cases, the social science researchers are still evaluating results, but some early anecdotal comments are appropriate. Each of the evaluations, described below, was intended to explore one or more of our goals for telepresence over ATM.

Basic point-to-point connections. Point-to-point connections over the ATM network were an enormous improvement over the previous dial-up video transport mechanism. For comparison, the number of dial-up video codec connections from Toronto OTP to Ottawa OTP over a 128 day period was 26; unfortunately the number of connections in the reverse direction was not recorded due to technical problems. Even if the number of connections from Ottawa to Toronto was assumed to be about the same (this would not be unreasonable), this still represents fairly low usage. There are a number of reasons: (1) the time to set-up a call was anything from 35-80 seconds because of the dialling required to establish the video link usually required more than 20 seconds; (2) the quality of the image and annoying time delay were workable, but suppressed a real feeling of presence; (3) the 128 period included August, in which most staff were on holidays; and (4) there were staff fluctuations due to some competing projects. During some of this time period (about 30 days in October to November) the ATM capability was up, so there was little demand for the dial-up video codec.

The number of ATM video connections in a 97 day period of operation from Toronto OTP to Ottawa OTP was 310. There was most likely a similar number of connections in the reverse direction, although there is no log data to corroborate this. This number was subject to several caveats as well. In support of high traffic: (1) there was a spur of learning, experimentation, and demonstration early on, that accounted for some of the additional traffic; (2) communications were increased because planning ATM experiments necessitated it; (3) connections were very easy to establish as there was no dialling time (connection set-up time was typically 5-15 seconds, depending on the network load); and (4) the high quality of the image and (virtually) delay-free communications conveyed a much greater sense of presence and made both social and business communications effortless. On the other hand, the traffic level may have been reduced for some of the following reasons: (1) OTP staff were declining so there were fewer people with whom to interact; (2) about three weeks were lower traffic due to the Christmas holiday season; and (3) the last 18 days of traffic occurred after OTP had officially ended and there was reduced need and opportunity to communicate. In summary, the point-to-point experiences were very positive, and the higher connection traffic largely validates this conclusion.

Some measurements were taken to show the connection times using the complete ATM system (ATM video links, and the private LAN for control signalling): (1) complete ATM --- about 5.3 seconds was required for A/V PBX server overhead and LAN communication required about 2.6 seconds, thus the user saw a delay of about 8 seconds, and this was constant as OTP did not load the data channel; (2) ATM video only --- again with an A/V PBX server overhead of 5.3 seconds, the communications over the Internet consumed about 6.6 seconds under average conditions for a total of about 12 seconds, although this could vary upwards considerably due to Internet load. Thus, while sending the control signals over the ATM was 2.5 times faster than the Internet, the end user only perceived a 1.5 times improvement because of the basic server overhead.

The meetings using ATM A/V links were often very short, as they were to some degree more convenient and interesting than phone calls or email (telephone traffic declined considerably between Ottawa and Toronto OTP sites, although email traffic was still used a lot). Some connections were also much longer, because users did not feel constrained by the telecommunications cost and simply let the length of the meetings fit the workpractice. This differs from the dial-up video links that were simply not as convenient to use, did not provide the same quality of interaction, and calls were typically focussed on a specific purpose. Cost was not really an issue for the dial-up video link, even though regular phone rates (the cost of two long distance phone calls) were incurred, but usually was the line was not used to excess as we knew it was not free (we did have 2 hour meetings, but not every day, and probably would avoid a 4 hour open connection). Also, the tiring nature of the dial-up video codec image quality and time delays, discouraged lengthy meetings.

Finally, point-to-point connections were invoked from Ottawa OTP to Toronto OTP to access the dial-up video codec in Toronto as an outside line. In the 97 day period some 50 uses were recorded, although about 70% of these were likely experimental in nature. Of particular interest was the interaction between the Newbridge JPEG video codec, that provided the analog-to-digital-to-analog conversion over the ATM network, and the PictureTel H.261 dial-up video codec, that provided the analog-to-digital-to-analog conversion over the telephone network. Several sites were contacted, and overall, the recipients did not notice much degradation over their usual experiences with dial-up video at 112 Kb/s (or 128 Kb/s) rates. Thus the JPEG codec worked very well in this situation. We found however, that the JPEG codec was rather sensitive to the quality of the analog video source; a noisy analog source, such as bad cabling and some computer-to-NTSC composite video scan converters, caused the JPEG video codec to lose synchronization with the video signal and generate a freeze frame effect that we nicknamed, "the Max Headroom Effect", after the jittery computer generated character from late 1980's television.

Multipoint meetings. A number of multipoint meetings were held; in some cases they involved users connected by analog and ATM video intersite links, and in other cases also involved dial-up video intersite links. We found that in the first case, there was no difference between an all-local user multisite meeting, and a meeting involving ATM A/V link users (at times we had two ATM A/V link participants since two A/V links were available). As there was no significant time delay in the ATM A/V transmission, interactions proceeded smoothly. Multipoint meetings involving dial-up video participants, proved to be a little more difficult, as the majority of the participants had no time delay, and the dial-up video participants had poorer picture quality and suffered the usual annoying communications delay. In this case, the meeting tended to revert to a protocol of waiting for answers and stilted conversation when the discussion was directed at the dial-up video participants, and tended to be more freewheeling when they were not the main focus. One such example was the OTP Final Review Meeting for industry collaborators that included a multipoint meeting of: (1) two people in a Xerox Webster office (by dial-up video); (2) more than a dozen people in the Toronto OTP conference room (the site of the *multipoint bridge* --- where the meeting was constructed); (3) four people from the Ottawa OTP conference room (by ATM A/V link); and (4) one person from an Ottawa OTP private office (also by ATM A/V link).

Virtual open office collaboration. In a virtual open office model, a point-to-point or multipoint connection is left on for an extended period of time (e.g., for half or the entire business day), in order allow distant collaborators the opportunity to share a *virtual open office*. Two offices (one in Ottawa and one in Toronto) used by 3 of our collaborating engineers were connected in this way for about 30 days. They reported 29 positive interactions and 6 distractions resulting from the open office. This data is being analyzed by our social science researchers, however, overall, the engineers were impressed by the convenience and were satisfied with the experience. The engineers reported that since the virtual open office has been discontinued, their productivity has been impaired as the typical "quick question, quick answer" that was afforded by the shared office, is now handled by many exchanges of short email messages. During the study period, email usage was more limited to comprehensive messages that were a more effective means of communications than spoken word or video.

The Telepresence Tunnel. The *Telepresence Tunnel* was a background full duplex A/V connection between a public space at the Ottawa OTP site (in the engineering building) and a public space at the Toronto OTP site (also in a building occupied primarily by engineering students). Students, faculty, and indeed the general public were able to walk up to their end of the *Telepresence Tunnel* and readily communicate with peers at the other end. Its purpose was to understand how people in physically widely separated buildings, but with some common background (they were mostly engineering students) would socialize using audio and video, once given the opportunity. It would also illustrate how video conferencing novices would react to interactive video, and perhaps to future publicly available video communication services.

The A/V PBX servers were slightly modified so that the *Telepresence Tunnel* used one of the two ATM A/V links whenever one was not in use for normal meetings. Typically the *Telepresence Tunnel* was in operation 24 hours a day for eight weeks (subject to some interruptions for our normal video meetings during the business day), although for about 3 weeks there were final exams at both sites, followed by the Christmas break. Images were sampled at five minute intervals at both ends of the *Telepresence Tunnel* and formed into MPEG movies. With this sample rate, we gained some sense of usage through a 30 day period and generated about 17,280 images. Questionnaires were also available to be completed by tunnel users at both ends. Both the movies and the questionnaires are being analyzed. Anecdotal observations suggest that the *Telepresence Tunnel* was widely enjoyed: some people "hung out at the tunnel" for extended lengths of time to engage in conversation; people would arrange (by email or telephone) to meet one another; even OTP staff passed each other in the tunnel on occasion. When there were lots of people in the building there were often periods of significant interaction (during class change and at lunch time), followed by periods of quiescence. One major problem was the low audio level, which proved to be very challenging because of the environment (hallways) and the lack of an echo canceller on the ATM A/V link (this is discussed later).

Remote teaching/lecturing. We ran three evaluations of a remote teaching experience in a three week period. In all cases, the lecturer was at the Toronto OTP site and the audience was at the Ottawa OTP site in the conference room. Also the DAN was present at both ends to easily reconfigure the rooms for different phases of the lecture. The goal was to assess the impact of image size, image quality, sense of presence, and the effect of having two full duplex A/V channels rather than the traditional one. In all cases there were questionnaires completed and interviews or comments by the lecturer --- as these will be assessed formally at a later time, the remainder of this discussion will focus on the configurations we used and some anecdotal information.

Lecture 1: A lecture for an undergraduate class (about 60 seniors in electrical and computer engineering) was given by a lecturer with no video conferencing experience. Initially, it was planned that one A/V link would be used for the lecturer's image, and one would be used for a document camera in which the overheads would be placed, and that would capture the lecturer's hand gestures. The lecture room had the lecturer's image displayed on a large screen using a projection television, and other monitors would be used to display the document camera. Unfortunately, technical difficulties prevented us from using two links for most of the lecture, thus, the lecturer's image was alternated with the document camera by an assistant. The lecturer saw two static images of the class on 19 inch monitors at a distance of about 20 feet from the lectern, thus the individual members of the class had little presence to the lecturer. During question period, an assistant in the lecture room used a remotely controlled camera to zoom-in on students asking questions and therefore significantly improved their presence to the lecturer, however, poor audio hampered the experience. Nonetheless the discussion flowed easily because there was no time delay and the lecturer's image and overheads were easily seen.

Lecture 2: A seminar was given by an OTP staff member to an audience of about 40. In this case a single ATM A/V channel was used by choice. The lecture was given from a private office and used an inexpensive picture-in-picture unit to project both the lecturer's image and the slides (from PowerPoint on a MAC using a scan converter to NTSC video); i.e., for most of the talk, the lecturer's image was in one corner of the main picture --- the presentation materials. The video in the lecture room was again projected to a large screen. The poor quality of the picture-in-picture unit significantly degraded the quality of the lecturer's image, and failed to convey a sense of presence. Worse still, the lecturer really did not watch the remote audience, who were projected on a very small monitor in the periphery of the lecturer's view. In essence, the lecturer just spoke to the camera leaving the audience with a perception that they could have been watching a video tape. When the formal presentation was concluded and open discussion ensued, the lecturer's image was the primary image and the use of the picture-in-picture was discontinued. Again, the discussion was easy and free flowing as the remote controlled camera panned and zoomed around the room to the current speaker in the audience, and the lack of any noticeable delay permitted interruption, debate and natural social situations.

Lecture 3: Another seminar was given by another OTP staff member to an audience of about 25. In this case both ATM A/V links were used. This lecture was done from a room that allowed the audience to be displayed on a large projection screen. A document camera was used for paper overheads and was transmitted over one ATM A/V link; the lecturer's hand gestures could be seen in the document camera. In the lecture room, the document camera was displayed on multiple monitors around the room, and the lecturer's image, transmitted over the second ATM A/V link, was displayed using the projection television. In the lecturer's room, a still image of the room was sent over one ATM A/V link and displayed on the projection television, and a roaming image (guided manually by an assistant in the lecture room) was displayed on a nearby monitor (less than 8 feet directly in front of the lecturer). During the formal presentation, the assistant used the remote control on the roaming camera to pan and zoom around the room as a lecturer might do naturally while giving a talk. Again, once the question period was reached, a fluid discussion happened, as in the other cases.

This last configuration gave the most satisfying result to both the audience and the lecturer because: (1) the audience saw the full size image of the lecturer at all times, and saw good interaction with the presentation devices --- the two were not viewed as separate; (2) the lecturer saw a large (albeit grainy) image of the audience and thus always had a feeling of their presence in the lecturing room; and (3) the lecturer got a more intimate view of the audience members *during the talk*, due to the constantly roaming camera, and thus could get a better feeling for how the material was being received. In short, both audience and lecturer had a sense of telepresence. All lectures suffered audio problems, that are described next.

General comments. Overall, the ATM network performed beyond everyone's expectations, however we were constantly plagued by audio problems because the Newbridge JPEG codecs did not have any echo cancellation, and we had no external echo cancellation. This resulted in a slightly annoying echo at times (about 0.1 seconds), as the remote microphone picked up the sounds from the remote speaker, and sent it back to the originating site. But worse, was the audio feedback that would result if the audio gain was raised too high. This forced us to keep audio levels lower than would normally be used and interfered particularly in the conference rooms and the *Telepresence Tunnel*. For the tunnel, the large physical spaces with poor acoustics and much reflection (concrete hallways) exacerbated the audio feedback problem. The conference rooms were also a major problem as they were large, and their acoustic characteristics changed widely (especially in the Ottawa OTP conference room) when an audience was present. In particular, the audience voices were much harder to pick up in the microphones, thus more gain would be needed, but more gain would lead to greater ambient room noise in the audio system, thus leading to feedback. If we had echo cancellation, our audio quality would have been significantly improved, thus conveying an overall more positive experience (people often perceive that a video experience is poor, even if the problem is actually the audio). The acoustic characteristics of private offices could usually be managed sufficiently well to avoid both feedback and echo problems.

6. Conclusions

Before the ATM experience, OTP had demonstrated the utility and characteristics of the synchronous A/V media space in supporting workplace collaboration. The advent of telepresence applications over ATM showed us the future could be much better than we had imagined and yet left us realizing that there was so much more to be explored, both in the situations that we studied, and those that we omitted for lack of time and resources. The experience gained in combining current and future communications technologies, validated much of our design approach, as well as showed the degree to which they could interact today. It also suggested some possible models for deployment in the near term as ATM transmission services become available in metropolitan and long haul networks. The example of Figure 10, shows a large company with several buildings on a campus or spread across a city could use ATM and analog technology to reduce much travel between sites and increase the quality of interactions. However, as long haul ATM services are bound to be expensive, it may be more practical to connect to sites located in other cities (perhaps branches of the same company, business partners, or clients) using conventional dial-up video services. Again our approach not only easily permits this model, it allows sites that are directly linked to share these codecs very easily. In the example, site A directly controls three dial-up video codecs, and site D, across the city, controls two dial-up video codecs. Since sites A through E are all connected by dedicated links (ATM or analog) and are trusted sites to one another, then the pool of codecs available to users in the city is five. Note: that this configuration assumes that a TMS site could forward A/V connections (e.g., from site D to E to B to A); this is not currently supported by the TMS as all intersite links must be point-to-point. However, this would not be a difficult modification to the design.

ROUTE 66

"All told, it took almost thirty years to completely bypass *Route 66* after Congress passed in the Interstate Highway Act, in 1956. Nonetheless, that law and the first bits of new concrete that were poured, cast a cloud over the Great Diagonal Highway. People were not so likely to seek their fortune on the edge of a doomed road, and of those who were already there, fewer and fewer saw any value in upgrading or expanding or --- sometimes --- doing basic maintenance. After 1956, *Route 66* remained important, but its importance was slowly moving away from the concrete toward the glorification of what the highway had been."
--- Susan Croce Kelly from
Route 66 --- the Highway and its People [5]

In this suggested configuration users in the city could: (1) maintain virtual open offices, (2) have sophisticated, multichannel conferences, (3) have point-to-point or multipoint meetings, (4) have shared public spaces (like the *Telepresence Tunnel*) in common areas, and (5) have a wide range of window cameras to view (even if only through short term connections, or as *Postcards* images). Finally, users company-wide could maintain convenient and close visual ties with one another, all in a cost-effective manner.

Many experiments had to be left unfinished due to the limited time and resources. Most regrettable was the lack of time to evaluate the effect of different video data rates (and hence lower video quality and increased delay) on the different applications. The main obstacle was the complexity involved in changing the peak cell rates on the permanent virtual circuits for the JPEG video cards. Late in the ATM trial we received control hardware that allowed us to use TELNET scripts to alter the ATM switch parameters from A/V PBX server host computer. We made some progress in developing the control software but never had an opportunity to finish it. The scripts would perform VT-220 emulation to disconnect circuits, and then reconnect them with different parameters; this would have to be done for the three ATM switches in the network.

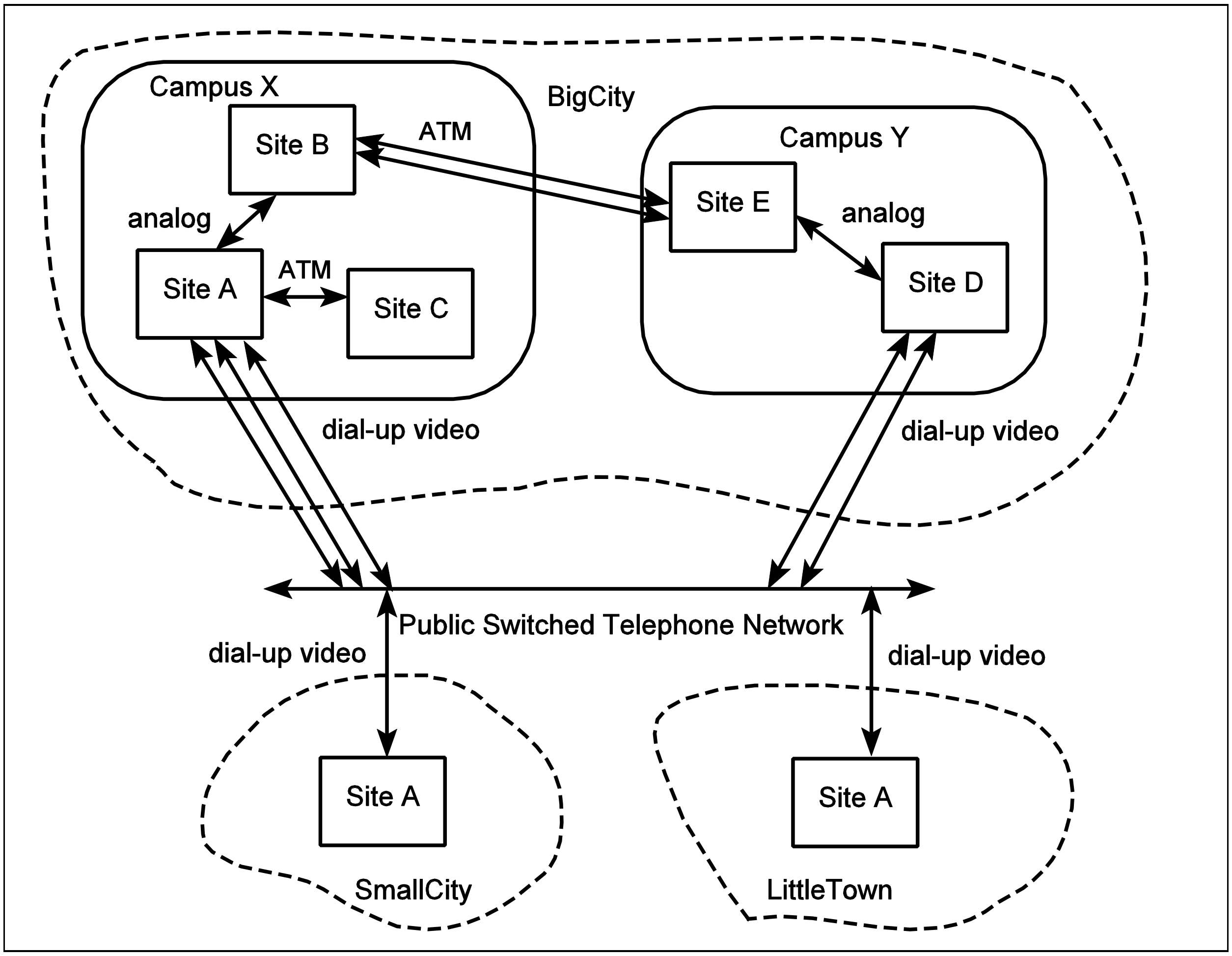


Figure 10: Example Integrating Analog, ATM, and Dial-up Video Services

The ATM experience showed the power and pitfalls of the remote lecturing application. More work is need to understand the impact and needs of multichannel video and lecturing styles. Furthermore, we never had the opportunity to understand and evaluate the A/V requirements of a *shared classroom*, in which the class exists in two or more rooms, one of which acts as the broadcast center for the lecturer, In this case, there is a desire to let the audience members have a sense of each others presence, and not just that of the lecturer.

The small stretches of the ATM Interstate on which we drove just made us beg for more --- it was fast, smooth, and let us accomplish our travels much more conveniently than the old road. *Route 66* served us well, and continues to serve us now since the ATM Interstate is still in its infancy and exists only in small segments. While the completion of the ATM Interstate is, to some degree, inevitable, it will probably take many years due to some of the same challenges that plagued the US Interstate Highway System. To justify the roadway, we will need services and traffic, and our experience suggests that these are not easily defined nor understood without considerable experimentation and study with the real user community.

Acknowledgements

This work represents many hours of hard work and commitment by the staff and researchers of the *Ontario Telepresence Project* in particular Bill Buxton, Gale Moore, Ron Riesenbach, Dick Dillon, Ross Jacobs, Tom Milligan, Tracy Narine, Chris Passier, Dominic Richens, Jingping Xie, Garry Beirne, Maciek Kozlowski, and Tabatha Holtz. The ATM phase of the project was a success only because of the gracious help from our industrial collaborators: *Bell Canada* seconded Michelle Valliere to us for a year, provided funding and the intercity DS-3 service; *Newbridge Networks Ltd.* provided the ATM switch gear loan, through the kind support of Jim Mackie, Richard Zwiep and others who made this a success. Also, *OCRINet*, *Carleton University*, and *University of Toronto* provided much assistance in this phase. Finally we are indebted to the Province of Ontario and out industrial partners and collaborators for funding and in-kind support throughout the project.

References

1 W. Buxton, "Telepresence: Integrating Shared Task and Person Spaces," *Proceedings of Graphics Interface '92*, Vancouver, B.C., May 1992.

2 P. Dourish and S. Bly, "Portholes: Supporting Awareness in a Distributed Group", *Proc. of ACM Conf. on Human Factors in Computer Systems CHI'92*, Monterey, CA, May 1992.

3 G.M. Karam, B. McLeod, and G. Boersma, "A Network Services Interface for Telepresence Applications," to appear in *IEEE Intl. Symp. on Autonomous Decentralized Systems* (8 pages), April 1995.

4 R. Riesenbach, "The Ontario Telepresence Project," *CHI'94 Conf. Companion, ACM Conf. on Human Factors in Computing Systems*, pp. 173-174, Boston, MA, Apr. 1994

5 Q. Scott and S.C. Kelly, *Route 66 --- The Highway and its People*, Norman OK: University of Oklahoma Press, 1988.

6 J. Steinbeck, *The Grapes of Wraith*, New York, NY: Penguin Books, 1939.

Section 4: Report of the Chief Sociologist

4. REPORT OF THE HEAD OF SOCIAL SCIENCES[1]

A basic premise of the Ontario Telepresence Project was that information and communication technologies are inherently social in nature. Consequently, it was essential in the design of these technologies to take advantage of the social skills and the knowledge of social protocols which future users already possess, building on existing skills rather than demanding or forcing new behaviours. Similarly, we recognized the need to incorporate an understanding of work and workplace variables in the design. The specification and elaboration of the relationships among work variables is a primary goal of our research. One size is unlikely to fit all and the successful development, adoption and diffusion of the Telepresence system must take into account existing work practices and organizational culture.

The Context for the Research Studies

While the Ontario Telepresence Project built on the work of the earlier CAVECAT Project and took advantage of research findings, particularly in the areas of cognitive psychology and human computer interface, the goal of the current research was on moving the design from the laboratory into the workplace. Taking the project beyond the laboratory and the local workplace into external sites broaden the research plan, both substantively and methodologically, but as the continuous development and testing of new technical applications was also central to the Project it was evident that it would be necessary to continue to do research internally. There was a great deal to be gained by using the first prototypes of new applications by users familiar with the system and tolerant of both of continuous change and software that was not sufficiently robust for external users. OTP had in fact been structured to encourage the team to work collaboratively across distance. Both the engineering and the social science group were divided between Ottawa and Toronto, and, in these locations members of the group were distributed within buildings. This was a natural laboratory for testing new ideas, applications, and technology, and encouraged the development of group processes so essential in any team, but particularly in one where a number of disciplines were represented and the majority of members had not worked together previously.

In the case of the external sites two approaches were taken. Our primary goal was to deploy the most robust version of TMS; Telepresence Media Space into workplaces outside the laboratory. A research agenda was established and a set of criteria were developed for the identification of field sites. The negotiation processes involved in gaining access to potential field sites required considerable effort and we had only limited success in achieving this goal. While the field studies would be conducted without financial cost to the client, the client would have to invest in hardware which depending on the specific installation could run from $100,000 - $250,000. The Project was thus dependent on the client's ability and willingness to incur these costs. For example, one field site was close to closure when the organization was bought out by an American firm.

---

[1]The first year of Ontario Telepresence Project ended on Dec. 31, 1992. It is important, however, to note that the social sciences component only became institutionalized in the project in July 1992 when a full-time social scientist was appointed

The second strategy adopted was to seize opportunities that were available and to study the use of commercial videoconferencing systems that were already in use in organizations, or were being introduced into workplaces to which we had access. This provided the opportunity to gain an understanding of the use and usefulness of these technologies in actual workplaces, insights that could be fed back into the design of the TMS.

The Field Studies

1. "Indigo"

A robust version of TMS technology was deployed in this site - a small research administration organization in which headquarters and the branch office are approximately 100 km apart - in March 1993. Research was carried out at this site for the next 18 months. This site provided the opportunity to put into practice user-centered design principles and to collect longitudinal data that are being used in the preparation of scientific papers in the areas of diffusion of innovation, organizational usability, and control of communication spaces and privacy. Preliminary results suggest the need to install systems that are tailored to support existing communication networks, both formal and informal. An unanticipated consequence of the deployment was the visibility the installation brought the organization within its environment and the symbolic meanings that emerged around this which furthered the organization's mission to be an information technology leader. One early report on this research in this site entitled The Creation and Control of Media Space was presented by in a workshop at the 1994 ACM Computer Supported Cooperative Work Conference.

2. "Provitel" Site

In the fall of 1992, Provitel Inc., a telecommunications organization who was deploying VISIT technology to selected groups within the organization agreed to deploy TMS linking two of the participants in the VISIT trial. This provided a unique opportunity to evaluate these two systems. Full details of the study can be found in an OTP Report - *Videoconferencing 1990s Style: Sharing Faces, Places and Spaces*. The results of this study demonstrated a clear need for an ecological approach to understanding videoconferencing technologies. Successful deployment, adoption, continued use and increasing demand requires a fit between the organizational culture and the work practices of the employees, and lent clear support to our hypothesis that technologies often fail for social rather than technical reasons. Valuable insights into the potential uses of video in the workplace were also gained.

A second part of this work at this site focused specifically on the deployment of the VISIT system in the organization. Management had used what we have characterized as the Velcro Model of Deployment - i.e., toss it out and see where it sticks. The organization had designated a group of workers to receive desktop videoconferencing without concern for the existing work practices, communication networks or organizational subculture. As predicted, adoption rate was poor and the deployment problematic. These results and observations were relayed to management for use in future deployment activities.

3. "ResCorp" Site

ResCorp, a research division of a large multinational corporation, had invested heavily in a conference room videoconferencing facility, but found that it received little use. This mode of deployment we have characterized as The Field of Dreams Model, or, if you build it they will come.

In Sept. 1993 OTP undertook a study of a group of early adopters to identify a range of social, organizational and technical issues that affecting adoption. The results of this study have been reported in *A Tale of Two Cities: A Study of Conference Room Videoconferencing*. The technical issues were relatively easy to identify and to diagnose but the social and organizational issues were less visible. While less attention is generally paid to these aspects of videoconferencing, we concluded that it was the resolution of this class of problem that would in the long run determine whether the organization not only recovered the cost of its investment, but maximized the long term benefits to both the organization and the employees.

Lessons from Organizational Reality

The following insights were gained from our field studies. The research to support these brief statements can be found in the OTP published literature.

1. The three most important things in understanding the interaction between organizations and technology are .... context, context, context.
2. Technology fails often for social, not technical reasons
3. Technology is an enable, not a driver
4. Technology transfer is more than just software and hardware - *methodologies* are technologies too
5. Communication technologies have *at least* two ends. There is a need to understand the ways in which users and organizations are not autonomous and independent with regard to communication media
6. A new productivity metric is need for the new economy
7. Technology makes explicit what has previously been implicit. Negotiated social order and associated behaviours often become visible only when disrupted and technology can provide just this kind of disruption. Well designed technologies support dynamic accommodation to a variety of 'understood' social arrangements - i.e., they score high on a *social/technical accommodation index.*
8. It is important to understand how media work as *environments*

Other Activities

ATM Studies

In October of 1994 the Ontario Telepresence Project had the opportunity to incorporate an ATM environment into the existing TMS; Telepresence Media Space.[2] At the time the

---

[2] Newbridge Networks of Ottawa loaned the OTP three 36150 four-port ATM switches that were equipped with two audio/video codecs, an Ethernet interface card and an outgoing circuit card. The codecs provided JPEG compressed near NTSC quality video transmission (scaleable according to peak allocated cell rate) and stereo CD quality audio channels. Bell Canada donated a DS-3 channel between University of Toronto and the Ottawa area where it was connected into OCRINet, an experimental ATM network linking a number of Ottawa universities, government institutions and companies. The technical details of this installation are elaborated in a paper by Gerald Karam (1995).

ATM network was made available, members of the Project had over 2 years of experience with a variety of audio/video services[3] and were highly sensitive to the ways in which audio and video quality, and the ease of making connections - both in terms of the action required to make the call and the time required to make the connection - affect how the system might be used. The superiority of the ATM service was recognized immediately in terms of the television quality image (and stereo sound) and the speed at which connections could be made between the two sites, but it was three years of living in a media space that helped us 'see' beyond the obvious technical advantages and led to the design a series of studies exploring innovative ways in which the technology could be used to improve social relations.

Social scientists from the Ottawa and Toronto ends of the OTP in cooperation with the engineering group used the ATM technology to design a series of studies which are described briefly below. A set of working papers describing these studies in detail are currently in preparation.

1. Telepresence Tunnel

The Telepresence Tunnel was a live audio video link between a corridor at Carleton University and a corridor at the University of Toronto. The goal was to see how people in two widely separated buildings, with some common background- primarily engineering and computer science students and faculty- would socialize using an open audio/video channel if given the opportunity. This also gave us the opportunity to see how people without any experience of videoconferencing would react to interactive video, and allowed us to ask about their potential use of and concerns about this type of service if it were commercially available. The sample was self-selecting as the Tunnel was in operation 24 hours a day and questionnaires were simply left at the site. Additional observations were gathered from OTP staff who were encouraged to visit the Tunnel, speak to the users and report findings back to the social scientists. Overall, response was highly positive. It was used for serendipitous encounters as well as arranged meetings. Of interest to the telecos is the fact that about 70% of the respondents reported they would use a home video system or would go to a specific facility to use this type of service for long distance calls. Approx. 60% suggested they would use such a facility for local calls.

2. TeleLearning

Three remote teaching studies were conducted. The Conference Room at Carleton University was the location for all the Ottawa participants, but in Toronto, where the lectures originated, the venue was changed in each case - a large conference room, a mid-size demo suite and an office. Evaluations were completed by the Ottawa participants at the end of each lecture. We were particularly interested in what advantage, if any, would be gained from the use of 2 audio/video channels instead of the single channel that is traditionally available. Evaluations of the sessions showed that the use of the second channel made a difference in terms of the attendees feeling of 'presence' and in their ability to interact. These studies raised more questions than they answered, but certainly pointed to problematic areas in the current commercial installations of the electronic classroom.

---

[3] Analog video, PictureTel, CLI high-end codecs,

3. Virtual Office

Two of the OTP engineers working on a joint project , one in Toronto and one in Ottawa, shared a virtual office for a period of approximately one month. Overall, both reported extreme satisfaction with the experience and while they had worked together for over 2 years they developed a social relationship and sense of community during this period which they felt enhanced their productivity. In interviews after the ATM connection had been taken down, both resented the amount of time that they now had to spend on email and missed the spontaneity that the ATM connection had provide, and could be said to be experiencing 'video withdrawal'. While they had developed a social relationship during this period they felt that this was impossible to maintain, even though both desired it, when the available media were email and telephone. They are currently planning to get together in Ottawa.

Lessons learned

1. Bandwidth on demand has the potential to revolutionize work across time and space
2. The relationship between bandwidth and cost is unlikely to be linear. Pricing bandwidth is a crucial issue.

Research Bibliography

As research on the social impacts of information technology is a developing field and one which crosses a number of disciplines access to this literature is difficult. A student was employed in the summer of 1993 to carry out a exhaustive search of the literature, to make copies of the significant documents and produce a machine readable and searchable index to this file. This resource was used by members of the Project, students and visitors. Some copies have been distributed.

Methodology for the Introduction of New Technology in the Workplace

One of the goals of the social scientists was to develop a methodology for the introduction of TMS technology into organizations. The method consists of a series of modules that are specifically designed to be responsive to the multiple contexts of the participating organization - e.g., organizational culture, work practices. The strength of this approach lies in the fact that the methodology is informed by a philosophy which incorporates and respects the end user, it was developed with input from a number of disciplinary experts, and modified in light of experience in actual workplaces. Finally, the model has been tested under strict conditions as it was developed for deployment of a technology which brings video into the workplace. The use of video is still unfamiliar in most workplaces and raises many issues that could thwart adoption.

Sponsorship of a Workshop: Towards the Virtual Workplace: Implications for Social and Organizational Research

In Nov. 1994 OTP and CITI: Centre for Information Technology Innovation co-sponsored this workshop to bring together members of the Canadian academic community and researchers from labour, government, public policy and technology organizations concerned with the interplay between workplace organization/re-organisation and information technology. The goals were to assess the existing state of research in the field, identify future research agendas and areas for cooperation and collaboration, and in general to begin building a network for a group who are currently working in isolation from one another. A number of initiatives are now underway as a result of this meeting.

Continuing Activities

A number of academic papers based on field study research, in particular in Indigo, should be presented or published over the next year.

The senior managers of the Project have developed an outline and proposal for a book telling the story of the Project. It is tentatively *entitled Living the Future Today: The Story of the Ontario Telepresence Project*

While members of the project team have dispersed there are proposals currently being developed to utilize the expertise and relationships fostered by the Project for new initiatives. One local example involves the University of Toronto Library.

Finally, the social scientists have an ongoing relationships with partners from the OTP (e.g., Corel) and proposals are currently being developed to continue these relations into the future.

APPENDIX 1: PRODUCT INFORMATION

(Not available in on-line version)

For more information on the
Ontario Telepresence Project, contact:

Ron Riesenbach, President
Telepresence Systems, Inc.
ron@itrc.on.ca
ph. (416) 482 6698