PRoP: Personal Roving Presence

Eric Paulos **John Canny**

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA 94720-1776, USA
paulos@cs.berkeley.edu jfc@cs.berkeley.edu

ABSTRACT

Current internet applications leave our physical presence and our real-world environment behind. This paper describes the development of several simple, inexpensive, internet-controlled, untethered tele-robots or PRoPs (Personal Roving Presences) to provide the sensation of tele-embodiment in a remote real space. These devices support at least video and two-way audio as well as mobility through the remote space they inhabit. The physical tele-robot serves both as an extension of its operator and as a visible, mobile entity with which other people can interact. PRoPs enable their users to perform a wide gamut of human activities in the remote space, such as wandering around, conversing with people, hanging out, pointing, examining objects, reading, and making simple gestures.

Keywords

Tele-action, tele-presence, tele-conferencing, tele-embodiment, telecommunications, tele-robotics, gesturing, tele-work, robotics, computer-mediated human–human interaction.

INTRODUCTION

We already live in a society accustomed to ubiquitous telecommunications [4]. Telephones are in every office, cellular phones are in many automobiles, and many individuals are reachable at any time via a pager or personal phone. More recently, the internet has increased our tele-connectivity by allowing us to exchange text, images, sound, and video with anyone whose interests we share, professionally or socially.

But obviously something is missing from these tools compared to direct human contact. This difference is so important that we expend a great deal of time and money traveling to experience these direct contacts. There is something about the extended experience of "being there" and the gamut of activities that we can perform when physically present that makes us prefer to be there in person.

We do not believe that we can ever replace true human-human interactions, nor is it our goal to do so. However, we do feel that it is possible to identify and distill a number of human behavioral traits or skills that are inherent to human communication, understanding, and interaction. Employing computer networking and robotic technologies to implement these traits, our goal is to ultimately provide a compelling overall experience for both the remote and local users and more importantly to create a usable system for tele-embodiment.

Why Tele-embodiment?

PRoPs allow human beings to project their presence into a real remote space rather than a virtual space, using a robot instead of an avatar. This approach is sometimes called "strong telepresence" or "tele-embodiment" since there is a mobile physical proxy for the human at the end of the connection. We coined the term *tele-embodiment* to emphasize the importance of the physical mobile manifestation.

Our approach differs fundamentally from more traditional versions of strong telepresence which involve an anthropomorphic proxy or android. Instead, PRoPs attempt to achieve certain fundamental human skills without a human-like form. More importantly, our research is driven by the study and understanding of the social and psychological aspects of extended human–human interactions rather than the rush to implement current technological advances and attempt to re-create *exact* face-to-face remote human experiences.

Why Not Video Teleconferencing?

While standard internet-based video teleconferencing provides an arguably more realistic interface than many other forms of telecommunications, it is more of an enhancement to existing technology rather than a new form of communication. With video teleconferencing we find ourselves fixed, staring almost voyeuristically through the gaze of an immovable camera atop someone's computer monitor. As actions and people pass across the camera's field of view, we are helpless to pan and track them or follow them into another room. The result is a "one-sided" experience where the remote user feels immersed but there is no physical presence at the remote end with which people can interact. In essence we still lack mobility and autonomy. We cannot control what we see or hear. Even if we had cameras in every room and the ability to switch between them, the experience would still lack the spatial continuity of a walk around a building.

We claim users want a more realistic perception of physical remote embodiment. We realized the importance of immersing the PRoPs user in the remote space by providing continuity of motion and control of that motion. These elements provide the user the visual cues necessary to stitch together the entire visual experiences into a coherent picture of a building and its occupants. Our system also supports various levels of communications and interaction between the PRoP user and the remote world's inhabitants. Furthermore, our system is accessible to any user on the internet with standard software running on currently existing computer architectures.

Outline

This paper first reviews previous and related work in telepresence and personal, remote interaction devices. However, discussion of some related research is postponed until more appropriate sections of the paper. We then describe the implementation and design of airborne and later terrestrial PRoPs. We discuss the results of our use of these devices and identify several important elements in providing a compelling tele-embodied experience. Next we confront several relevant social issues involving PRoP development. Various future PRoP design plans are investigated. Finally, we complete the paper with a summary of PRoPs, their contributions, and importance.

PREVIOUS AND RELATED WORK

Methods of achieving telepresence[1] are not new with one of the first electrically controlled mechanical teleoperational systems being developed by Goertz [6] in 1954. Since then a variety of applications for tele-operated robotics have been explored [16]. However, most of these system are designed for a single specific task and are quite complex. They also typically require special purpose dedicated hardware and a highly trained operator to control and interact with the mechanism in the remote environment. In our system we strived to constrain its development so that it would be accessible to a wide audience without additional, expensive, or extraordinary hardware. In essence, telepresence[2] for the masses.

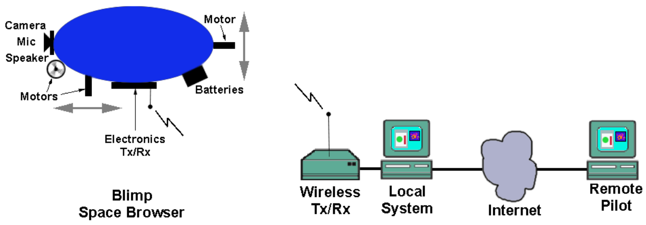

Figure 1: System overview of the basic space browser (or blimp PRoP) hardware configuration.

The exponential growth of the WWW over the past several years has resulted in a plethora of remote controlled mechanical devices which can be accessed via the WWW. Goldberg [7] developed a 3 DOF (Degree Of Freedom) telerobotic system where users were able to explore a remote world with buried objects and, more interestingly, alter it by blowing bursts of compressed air into its sand filled world. Soon afterwards, we developed Mechanical Gaze [13], a tele-robotic system where uses could control a camera's viewpoint and image resolution to observe various museum artifacts placed within the robot's workspace. By 1995, Goldberg had developed another telerobotic system called the TeleGarden [8] in which WWW users are able to observe, plant, and nurture life within a living remote garden. As of this writing, well over several hundred interesting mechanical devices are connected to the WWW with more spawning daily.

Social and psychological aspects of extended human-human interactions motivate the design of our PRoPs and we have identified a wide range of research in this area. Shared spaces and human interaction with video walls such as the VideoWhiteboard [17] designed at Xerox PARC and later Ishii's ClearBoard [9] are fundamental to designing usable PRoPs. We are also interested in the use of video in tele-connecting individuals which has been nicely explored by Kraut and Fish [10; 5] and others [3].

AIRBORNE PROPS: SPACE BROWSERS

The first PRoPs were simple airborne tele-robots we named *space browsers*. A space browser is a helium-filled blimp of human proportions or smaller with several lightweight motors directly connected to small propellers and no other moving parts. On board the blimp is a color video camera, microphone, speaker, simple electronics, and various radio links (see Figure 1). The entire payload is less than 600 grams (typically 400–500 grams). Our design choice was to use the smallest sized blimps that could carry the necessary cargo, thus making them easily maneuverable down narrow hallways, up stairwells, into elevators, and through doorways. At present we have iterated

[1] "To convey the idea of these remote-control tools, scientists often use the words *teleoperators* or *telefactors*. I prefer to call them *telepresences*, a name suggested by my futurist friend Pat Gunkel." [12]

[2] More specifically we are referring to tele-embodiment, tele-robotics, or tele-action. This is to avoid the ambiguity caused by the term *telepresence* which has grown in recent years to describe not only systems involving distant real spaces (*i.e.* tele-robotics) but also distant virtual spaces or VR.

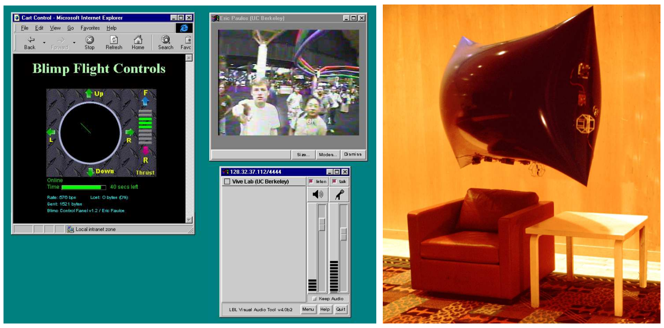

Figure 2: The remote pilot's Java interface with live audio and video (left) and a blimp PRoP in flight (right).

through several different configurations. Blimps ranging in size from 180x90 cm to 120x60 cm and shapes such as cylinders, spheres, and "pillow shaped" have all been flown. The smaller blimps consume about the same space as a standing person and are thus well-suited for moving into groups of people and engaging in conversation with minimal disruption. Even under full power a blimp moves at human walking pace. Figure 2 depicts one of the "pillow shaped" blimps in flight.

A user, anywhere on the internet, can use a simple Java applet running within a Java-enabled browser to pilot the blimp (see Figure 2). As they guide the blimp up and down or right and left the blimp delivers, via wireless communications, live video and audio to the pilot's machine through standard free tele-conferencing software that runs on a standard PC. The pilot observes the real world from the vantage of the blimp while listening to the sounds and conversations within close proximity to the blimp. The pilot converses with groups and individuals by simply speaking into the microphone connected to their desktop or laptop computer, the sound delivered via the internet and then a wireless link to the blimp's on-board speaker.

Problems with Blimps

Space browsers are far from perfect in that they are currently quite high maintenance. Stringent weight limitations allow for only a small amount of batteries to be carried on-board, yielding flights of about an hour before batteries need replacement. Although replacement is quick and rather straightforward, this process still prevents the blimp from operating continuously, as we would desire. As a result, remote conversations and explorations are often cut short.

Furthermore, piloting the blimp is often awkward. Typically the blimp exhibits erratic behavior and the sensation is more like swimming or floating than walking. Another problem is that unlike other robots, blimps are nearly impossible to bring to a complete halt.

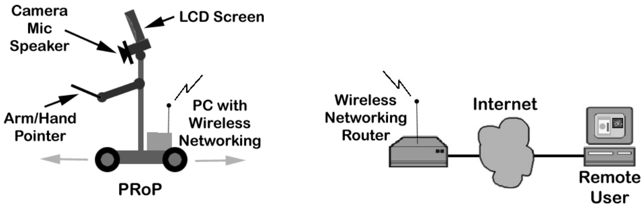

Figure 3: System overview of a Surface Cruiser (or cart) PRoP hardware configuration.

Some Solutions

Typically, a user wants to maintain a constant height while flying around, but instead must manually burst the lift motor at regular intervals to maintain that height. To solve this we have incorporated a simple lightweight sonar device onto the blimp to help maintain a constant height. Likewise, various simple aerial acrobatics such as nice ninety-degree turns and flying forward without rotating slightly left or right are difficult. In an attempt to solve this we have incorporated a simple, inexpensive electronic compass weighing less than 25 grams. Both the compass and sonar are carried on board and thus are not susceptible to the network delays experienced by the pilot attempting to correct these problems remotely. There is addition information [14] available for readers wishing to acquire more information about these airborne PRoPs.

TERRESTRIAL PROPS: SURFACE CRUISERS

Leveraging off of our previous research with airborne PRoPs, we developed terrestrial four-wheeled *surface cruisers* or *carts*. These carts are designed from simple remote-control vehicles with modifications to slow them to human walking pace and a 1.5 meter vertical pole to provide a realistic human vantage for the camera. On board the cart is a color video camera, microphone, speaker, color LCD screen, a few simple custom electronics, and various drive and servo motors. The basic system layout for the cart system is shown in Figure 3. Unlike the blimps, carts can travel outdoors, require less maintenance, and provide much longer battery life. Carts also carry a complete PC on-board with wireless networking hardware attached. Thus the multiple radios previously required to operate the blimp coalesce into a single wireless signal on carts. Furthermore, we leverage off of wireless communication infrastructures already in existence, greatly extending the inhabitable world of carts. A recently designed cart is shown in Figure 4.

RESULTS

The various PRoP design choices have been guided largely through trial and error experiments with the actual devices. Obvious methods of communication such as audio and video were part of our original design. However, after

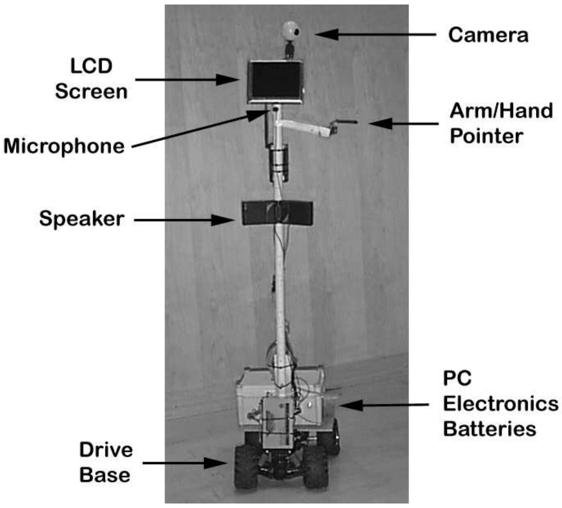

Figure 4: A cart PRoP with camera head, video LCD screen, controllable "arm/hand" pointer, microphone, speakers, and drive-able base.

extended use of these PRoPs it was clear that many important, and often subtle, elements were missing from the experience.

Through this evolutionary development we have been able to identify several behavioral traits which we consider essential to providing the most compelling overall experience for both the remote and local PRoP user. In the following subsections we iterate these elements and discuss their role in creating convincing tele-embodiment. While some of them may seem obvious, there are others that we have found to be surprisingly important and far less evident.

Audio

Perhaps the most apparent element of communication, two-way audio allows users to engage in remote conversations. Audio is also the channel whose usefulness is most susceptible to quality degradation from reduced network bandwidth and/or network packet loss. An unexpected result was the importance of background "noise" near the PRoP. The experience of using the PRoP was noticeably more compelling when users were able to gauge the general mood of the remote location by receiving a variety of subtle aural cues such as doors opening, elevators arriving, people approaching, nearby conversations, music playing, automobile traffic, wind blowing, etc.

Video

A visual portal into the remote space is another obvious element of the immersive experience. Video is also the most demanding tele-embodiment trait in terms of network bandwidth consumption and processor usage. Despite the increased resource consumption, color video is far superior to grey-scale for distinguishing details and identifying objects and locations remotely as well as for providing a more realistic perception of the remote space. The effects and tradeoffs of video and image quality and its resulting perception by humans has been extensively studied by Reeves and Nass [15].

We learned that sophisticated video compression algorithms are essential to make video signals usable. Surprisingly, in many cases the overall quality of the resulting video is far less important than the ability of that video to provide subtle information about the motions, actions, and changes at the remote location. When the user is navigating or when significant amounts of activity are occurring within the camera field of view, the importance of high video frame rates dominates over the resulting video quality. We learned that during times of "high video activity" the resulting compressed video signal should convey to the remote user at least an approximate representation of the remote motion and/or activity. However, during periods of small temporal video activity such as when the user is conversing with an individual, examining an object, or reading a sign, it is clearly the overall quality of the video signal that dominates over frame rate.

On several occasions PRoP drivers lost their way in a familiar hallway. Users performed a two-step process to orient themselves. First, using the high-frame-rate low-quality video, users steered the PRoP towards a door name-tag or room number sign. Accurately positioning the PRoP in front of such a room marking is prohibitively difficult without the use of high-frame-rate video to provide visual feedback to the user. Second, when motion stopped, the user requests high-quality low-frame-rate video which is used to easily resolve the name on the door (or room number), thus identifying their location in the building.

This led us to identified the need for at least two levels of video resolution or "Telepresence with Extra Eyes" [18]. The system should provide a wide angle view similar to the human eye for navigating and recognizing people (and objects) and also a smaller field of view for reading text on paper, white-boards, doors, and computer screens.

We noticed that with only one-way video, PRoPs can be mistaken as tele-operated surveillance tools or autonomous reconnaissance drones. Both of these tasks are far from the intended application of PRoPs. We removed this video-asymmetry by adding a small (15 cm diameter) LCD screen with a video feed from the remote user. This two-way video is also an appropriate mechanism for transmitting a richer representation of the remote user through their facial gestures and expressions. When bandwidth is a problem and the screen is used only to display a still image of the remote user, we find that it still succeeds in conveying the identity and existence of the remote user.

Mobility

Mobility, and in fact all of the remaining behavioral traits, are notable PRoP enhancements to standard video teleconferencing. While the space browsing blimps possess more mobility (*i.e.* flying in the third dimension), the carts are actually much more effective PRoPs. The main reason is that blimps actually provide too much freedom and no mechanism to stop. That is, the blimps, despite several on-board sensors, are unable to hold a fixed position while the cart PRoPs can easily halt and interact with a group of people.

So how sophisticated should the mobility be? We found that simple car-like navigation of a PRoP on the ground was fairly straightforward for a user to understand and control though a relatively simple interface. It also provided enough freedom for users to maneuver within (and outside of) buildings. This was the simple design of our first PRoP.

However, since human interactions occur where humans can travel, PRoPs must be able to reach much of the world accessible to humans. Again, we are not attempting to create an android or anthropomorphic robot so we will not handle what we call dextrous human motions. In particular PRoPs do not need to climb fences, swing from ropes, leap over ditches, repel down cliffs, slide down poles, *etc*.

Our basic philosophy is that PRoPs should be able to access the *majority* of locations most humans inhabit daily. Aiming for simplicity, we feel that PRoPs should be able to perform simple locomotion through fairly benign terrains such as mild inclines, curbs, stairs, and small variations in ground surface (*i.e.* sidewalks, grass, dirt, *etc.*). This includes traveling outdoors and also means that PRoPs must be be untethered (*i.e.* wireless). It is also important to impede the overall speed of the PRoP, typically through various gear reductions, to roughly mimic human walking pace.

Directed Gaze

Although remote users can see, hear, and move around, navigating remains a tedious task and does not facilitate the ability to quickly glance around a room to get a feel for its size, occupants, *etc*. This problem is remedied by incorporating a small movable "head" (*i.e.* a camera on a controllable pan-tilt platform) onto the PRoP. Our device is similar to the GestureCam [11] which allows a remote participant in a conversation to have direct control of his or her visual field of view. This relatively simple PRoP "head" provides a vitally important element of human communication, direction of attention or gaze as discussed by several researchers [2; 9]. This allows PRoPs to perform human-like conversational gestures such as turning to face someone in order to see them, address them, or just give attention to them. These actions are also visible to people interacting locally with the PRoP and provide simple gestural cues to let individuals know when they are being addressed or looked at by the remote user.

Pointing and Simple Gesturing

We learned quickly that gestures are very important for human communication. Remote users immediately found the need to point out a person, object, or direction to the individual near the PRoP. Although the movable head could be used as a crude substitute, it lacked the correct visual gestural aesthetic of pointing and was often ambiguous to individuals watching the PRoP. We added a simple 2 DOF pointer so that remote users can point as well as make simple motion patterns. These motion patterns allow the PRoP user to express additional non-verbal communications gestures such as interest in a conversation, agreement with a speaker, or to gain attention for asking a question in a crowded room.

We found that adequate pointing does not require a mechanism as complex as a human hand, since it is gross motion and not dexterity that is needed for the social function of gesturing. We have also been exploring several optional "arm/hand" designs to accomplish basic gesturing functions. More complex gesture interfaces and mechanisms are an important piece of our long term research agenda.

There has been a significant amount of research into gesture *recognition*. These systems typically aim to identify a human motion, typically made with a mouse, and interpret it as a known gesture. For example, a quick up-down motion of the mouse may be recognized as the "scroll page" gesture. However, we are making a conscious choice to use such symbolic descriptions of gestures only as a last resort. Instead we prefer to use continuous input devices like mice and joysticks to provide direct gestural input from the user to the PRoP. For example, compare typing text to a speech synthesizer, with spoken text transmitted through a speech compression algorithm. The synthesis approach may provide clean-sounding speech at low bandwidth, but all nuance and emotional content is lost. Similarly, music which is generated by computer from an annotated musical score is lifeless compared to music played by a human from that score, even if the recording mechanism is identical (*i.e.* MIDI).

We believe that any human communication beyond the very simplest cannot be ascribed a unique or personal meaning. We are further motivated by Wittgenstein who succinctly expressed that communication is about *connotation* (what the communication could mean) and not about *denotation* (some particular meaning). Post-structural literary theory underlines the importance of social, political, and historical context in the understanding of any text. It illustrates just how complex the meaning of short "symbolic" text can be. It also explains why symbolic representations such as text are capable of such great richness.

In fact it is not really surprising that through these crude devices and narrow communication channels, that rich and complex communication is possible. Recall that actors transmit their gestures to audience members tens of meters away, dancers and mimes work without speech, and puppeteers work without a human body at all. All of us use the telephone without a visual image of our interlocutor. Our task in gesture transmission is to isolate the key aspects of gesture so as to preserve meaning as closely as possible. Some factors are clearly important, such as time-stamping to preserve synchronization and velocity. Others, such as mapping human degrees of freedom to robot "arm/hand" degrees of freedom are much less so.

Reflexivity

The ability of a user to experience their own existence and actions through the PRoP turns out to be an extremely important element in providing a compelling tele-visit. When users could point the camera downward and actually see the wheels and base of the PRoP there was a noticeable improvement in the quality of the immersive tele-experience. Likewise, the experience was enhanced when users could steer the wheels, move forward and backwards, or position the pointer while visually watching the actions that resulted from their commands. Imagine if you were able to give commands to your arms and legs but never sense the result? Clearly, the experience would be lacking a significant element which we call *reflexivity*.

There is also reflexivity in the audio channel. In fact the importance of full-duplex audio, that is the ability to hear and speak at the same time, allows the remote user to hear their own voice when speaking. Users also use this mechanism to regulate the tone or volume of their voice to suit the acoustical or aural mood of the remote space the PRoP is inhabiting.

Limited resources such as bandwidth can inhibit reflexivity and distract from the immersive experience. This occurs when network delays cause the user to feel detached from the PRoP. The lag between moving the control joystick or mouse and seeing the results can sometimes be several seconds. By then the PRoP may have wandered far from where the user intended. The impression that the user gets its that the PRoP has "a mind of its own" which is exactly the opposite of an immersive experience. Expected technological advances will eventually solve this problem. However, substantial transmission delays are going to be a fact of life on the internet for at least a few more years. That means that real-time control of PRoPs over the internet will continue to be cumbersome.

Physical Appearance and Viewpoint

Although not anthropomorphic, we observed that PRoP design is loosely coupled to a few human-like traits which are important visual cues for successful communication and interaction. Clearly, a small ground-based robot conveys a rodent-like perspective of the world. However, a large robot is typically unable to navigate down narrow hallways, pass through doors, and impedes normal human traffic flow in a building. Furthermore, larger more industrial-type mobile robots are also more likely to frighten people, detracting from their use in human communication and interaction.

Since they stand in as a physical proxy for a remote user, it makes sense that PRoPs should be roughly the same size as a human. We attached a 1.5 meter vertical pole at the center of the PRoP to provide a realistic human vantage for the camera. In general we have found that the positioning of various attachments on the PRoP (*i.e* head, pointer, arm, *etc*.) should have some correspondence to the location of an actual human body part that provides the equivalent functionality. Also, all of the communication channels should be from the point of view of the PRoP (*i.e.* from on-board the tele-robot). It does not suffice to simply have a camera someplace in the room where the PRoP is currently located.

We have also experimented with the overall height of the PRoP after discovering that it was intimidating to people shorter than the camera head on the PRoP. It is still to early in our research to determine the useful range of heights or overall torso designs that facilitate human interaction with PRoPs.

Browsing and Exploring

When designing PRoPs we found that even with all of the previously discussed traits, it is essential to allow a remote user to wander, explore, and travel throughout buildings and spaces in much the same manner as humans normally do. It is this higher level browsing and exploring behavior that is perhaps the most important element of tele-embodiment.

PRoPs should allow users to search for an individual or a particular location such as a laboratory, office, or conference room. They should support exploration and wandering, where the user has no specific target. This behavior is intended to mimic the action of walking around a location noting the names and functions of rooms, wandering around looking for people they want to visit, or checking out the progress of experiments in a laboratory.

When exploring and browsing, the user can automatically generate a spatial time-line or "visual scrapbook" of the visit. By simply recording high-quality still images as the user travels with the PRoP, a rich context of the visit can be generated. For example, a user may record that they met their friend X in the south hallway of a building, then went to Professor Y's office, then went for coffee at the cafe down the street, *etc*. The visual time-line that is created assists the user in remembering the visit and in acquiring an overall feel for the remote space.

More importantly, browsing and exploring are autonomous operations performed by a remote user and do not require any support from individuals in the remote space, beyond the availability of the PRoP itself. The benefit of this is that PRoPs can be installed and used in a location with little overhead and disruption to the inhabitants of the remote space. A remote PRoP user can also be given a tour of the remote space by one of its local occupants. In fact either or both the visitor and guide may be PRoPs. Overall, a wide gamut of human activities can be performed without any local assistance such as attending meetings, seminars, conferences, discussions, *etc*.

Hanging Out

A surprising but important social function is the ability to simply hang out. We know that physical presence often improves the length, quality, and priority of one-on-one group interactions. This is a purely social phenomenon. In many work situations, individuals are willing to talk to you for a much longer period in person than they would be willing to do over the phone. When phone and in-person communications collide, most people try to terminate the phone call so they can return to their live interlocutor. We would like to better understand the factors that influence this preference and see where a PRoP presence fits into the priority ordering.

A visit to a remote place is an extended activity in which a person shares space with others. During this time, the task being performed by each person changes, and there may or may not be tasks in common at any given moment. These tasks serve as additional stimuli for communication between the individuals, leading to multiple communication episodes with different subjects.

SOCIAL ISSUES

Although PRoPs provide interesting new methods of remote interaction, there are also numerous social issues that must be carefully addressed before PRoPs become ubiquitous tele-embodiment tools. In this section we discuss a few immediate and obvious concerns in this research area.

Safety

When control of a physical mechanical system is accessible by anonymous individuals, great precautions must be taken to insure the safety of people and objects sharing the space with the PRoP. We are all aware of the interest in hacking into computers and manipulating, stealing, or destroying digital data. One can easily image the fascination of taking control of a potentially dangerous device to use to one's own ends.

Unlike many other robots, it is vital that safety be a primary concern when designing PRoPs. We propose a teleoperational variation on Asimov's first law of robotics[3] which stipulates that at no time should a PRoP ever be capable of injuring a human being, regardless of the action or inaction of the remote tele-operator.

The tele-robot's abilities, physical attachments, and even the basic construction must all be considered. Even an out of control PRoP must safely interact with humans and property. Since we desire these tele-robots to co-habitate with humans, this constraint is of the utmost importance. For example, even a seemingly un-threatening and safe, but slightly heavy, tele-robot may accidentally be knocked over near a stairwell, causing it to tumble recklessly down the stairs, and impact an innocent human in the stairwell. The importance of the safe co-habitation requirement cannot be overstated as it relates fundamentally to the acceptance, approachability, friendliness, and interactivity of PRoPs and humans.

Privacy and Security

As fascinating as a floating blimp or cruising cart is, it is somewhat disconcerting when you wonder who may actually be at the other end of the connection. Should that person be allowed to access to the space occupied by the PRoP? We envision a system to easily provide secure access to PRoPs using techniques not to different from those currently used to limit access to your individual files and computer hardware. People would invite individuals into their local space by issuing them a digital token that would authenticate and validate individual access to the local PRoP and hence the space. Also, although they are designed to be small and agile so that they can navigate within a building, PRoPs have no super human abilities that enable them to pass through walls, push elevator buttons, unlock doors, or for that matter even open a closed door.

CURRENT AND FUTURE RESEARCH

Current and future research plans center around enhancing the overall quality of the experience delivered by the PRoP and its functionality as a communication tool. We have had only limited experience with pointers on the PRoPs and would like to experiment with richer gesturing devices. We expect these to be more mechanically complex than the simple 2 DOF pointer, but we are still unsure of the actual design, methods of use, or quality of gestures we can produce.

We are also studying human physiology to better understand the dynamics of the human hand and arm. Similar research is being conducted using simulated dynamical systems to prototype simple gestural mechanisms. We plan to design dynamic regularizers (perhaps using simple PID controllers on the motors) that translate simple mouse gestures into resulting "arm/hand" dynamics on the the PRoP that mimic human arm/hand motions. The goal is to provide realistic and thus recognizable human gestures on the PRoP.

[3] "A robot may not injure a human being or, through inaction, allow a human being to come to harm." Handbook of Robotics, 56th Edition, 2058 A.D., as quoted in *I, Robot* by Asimov [1]

Certain common tasks are far too cumbersome using the current system. For example, navigating down to the end of a long hallway requires constant user attention and subtle command tweaking to avoid colliding with the walls. Our plan is to design a simple system that allows for a more intuitive point and move-to mode. A user would simply click on the video image causing the tele-robot to move to the corresponding location on the ground plane.

We expect that our simple click-and-move system will lack the accuracy of a high performance (and expensive) tele-robotic system. In fact it will likely require a few refinements in the positioning as the robot moves down the hallway. However, the benefit of moving through large areas using only a few mouse clicks would relieve the user of a tremendous burden and increase the overall functionality of PRoPs.

We would also like to be able to record higher quality images at certain locations, save them, and use them to mosaic the larger room or space. Mosaicing may also occur automatically when the PRoP is idle or when it detects an interesting event such as an extended conversation. These higher quality images server as a "scrapbook" of the interesting episodes during the user's PRoP tele-visit.

As described in the results section, there are distinct situations in which a PRoP should trade off video frame rate for resolution. We would like to be able to automatically detect these events and control the level of trade-off in the video compression algorithm. Currently, controlling this trade-off requires user intervention.

So far our cart PRoPs have not been able to negotiate stairs. We currently have several simple tread-based robot designs that we hope can be constructed to allow PRoPs to use stairs and travel over curbs.

CONCLUSION

Our claim is that PRoPs provide an extremely useful, functional, powerful new tool for supporting human communication and interaction at a distance. They enable a variety of important work and social tele-activities far beyond what we perform currently with our computers and networks.

PRoPs are also an ideal platform for studying computer-mediated human interaction because they operate in existing social spaces and can interact with groups of humans. Despite our limited experience using PRoPs, we have been able to identify several factors that we consider vital to providing the most compelling overall experience for both the remote and local users. This is why our research draws as much on the sociology of group interactions as on sensing and actuation techniques. In fact we need the former to drive our choices for the latter.

Interestingly, we found that the absolute performance of the PRoP hardware is less important that the human activities that it enables or impedes. In fact the technologies needed to construct compelling and highly immersive tele-embodiment PRoPs already exist, and they are surprisingly inexpensive. The space browser prototype cost under two-thousand US dollars and the carts only slightly more. With their low cost and enormous commercial potential, we feel that it is quite possible that such devices will someday become ubiquitous home computer accessories.

ACKNOWLEDGMENTS

This research was supported by BMDO, grant number N00014-96-1-1200 administered by ONR. We would also like to thank the following individuals for numerous enlightening personal discussions surrounding this work: Ken Goldberg, Mark Pauline, Scott Fisher, James Landay, Benjamin Zorn, and Allison Woodruff.

References

[1] ASIMOV, I. *I, Robot*. Garden City, N.Y. : Doubleday, 1950.

[2] DONATH, J. The illustrated conversation. In *Multimedia Tools and Applications* (March 1995), vol. 1.

[3] FINN, K., SELLEN, A., AND WILBUR, S., Eds. *Video-Mediated Communication*. L. Erlbaum Associates, 1997.

[4] FISCHER, C. S. *America Calling: A Social History of the Telephone to 1940*. University of California Press, 1992.

[5] FISH, R., KRAUT, R., ROOT, R., AND RICE, R. Evaluating video as a technology for informal communication. *Communications of the ACM 36*, 1 (1993), 48–61.

[6] GOERTZ, R., AND THOMPSON, R. Electronically controlled manipulator. *Nucleonics* (1954).

[7] GOLDBERG, K., MASCHA, M., GENTNER, S., ROTHENBERG, N., SUTTER, C., AND WIEGLEY, J. Robot teleoperation via www. In *International Conference on Robotics and Automation* (May 1995), IEEE.

[8] GOLDBERG, K., SANTARROMANA, J., BEKEY, G., GENTNER, S., MORRIS, R., SUTTER, C., WIEGLEY, J., AND BERGER, E. The telegarden. In *SIGGRAPH* (1995).

[9] ISHII, H., AND KOBAYASHI, M. Clearboard: A seamless medium for shared drawing and conversation with eye contact. In *ACM SIGCHI* (1992).

[10] KRAUT, R., FISH, R., ROOT, R., AND CHALFONTE, B. Informal communication in organizations: Form, function, and technology. In *ACM Conference on Computer-Supported Cooperative Work* (1988).

[11] KUZUOKA, H., KOSUGE, T., AND TANAKA, M. Gesturecam: a video communication system for sympathetic remote collaboration. In *ACM Conference on Computer-Supported Cooperative Work* (October 1994).

[12] MINSKY, M. Telepresence. *Omni 2*, 9 (June 1980).

[13] PAULOS, E., AND CANNY, J. Delivering real reality to the world wide web via telerobotics. In *IEEE International Conference on Robotics and Automation* (1996).

[14] PAULOS, E., AND CANNY, J. Ubiquitous tele-embodiment: Applications and implications. *International Journal of Human-Computer Studies/Knowledge Acquisition* (1997). Special Issue on Innovative Applications of the World Wide Web.

[15] REEVES, B., AND NASS, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge Univ Press, 1996.

[16] SHERIDAN, T. B. *Telerobotics, Automation and Human Supervisory Control*. MIT Press, 1992.

[17] TANG, J., AND MINEMAN, S. Videowhiteboard: Video shadows to support remote colaboration. In *ACM SIGCHI* (1991).

[18] YAMAASHI, K., COOPERSTOCK, J. R., NARINE, T., AND BUXTON, W. Beating the limitations of camera-monitor mediated telepresence with extra eyes. In *ACM SIGCHI* (1996).