NON-VERBAL INTERACTION IN THE DESIGN OF TELEPRESENCE ROBOTS FOR SOCIAL NOMADIC WORK

by

Jennifer S. Milne

M.Eng Product Design Engineering
University of Glasgow / Glasgow School of Art, 2010

Submitted to the Department of Mechanical Engineering
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Mechanical Engineering

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2012

Signature of Author: ______________________
Department of Mechanical Engineering
May 25, 2012

Certified by: ______________________
Federico Casalegno
Director, MIT Mobile Experience Lab
Thesis Supervisor

Certified by: ______________________
Maria C. Yang
Assistant Professor of Mechanical Engineering and Engineering Systems
MechE Faculty Reader

Accepted by: ______________________
David E. Hardt
Chairman
Committee on Graduate Students

NON-VERBAL INTERACTION IN THE DESIGN OF TELEPRESENCE ROBOTS FOR SOCIAL NOMADIC WORK

by

Jennifer S. Milne

M.Eng Product Design Engineering
University of Glasgow / Glasgow School of Art, 2010

Submitted to the Department of Mechanical Engineering
On May 25, 2012 in Partial Fulfillment of the
Requirements for the Degree of

Master of Science in Mechanical Engineering

ABSTRACT

Telepresence robots have emerged as a novel solution to meeting the social communication needs of nomadic workers. This thesis provides an overview of non-verbal communication cues for telepresence robot applications, and a snapshot of the competitive landscape for commercially available telepresence robots today. It then follows the design of a low-cost telepresence robot which can be remotely operated whilst running Skype, and discusses how further non-verbal communication cues could be incorporated to increase the feeling of social presence. Specifically, face tracking and the ability to communicate gaze is developed in the final prototype.

Thesis Supervisor: Federico Casalegno
Title: Director, MIT Mobile Experience Lab

MechE Faculty Reader: Maria C. Yang
Title: Assistant Professor of Mechanical Engineering and Engineering Systems

ACKNOWLEDGEMENTS

Thanks to Federico Casalegno who showed great enthusiasm in my work and provided the opportunity for me to join the MIT Mobile Experience Lab. I felt very at home in this truly interdisciplinary lab, enjoyed being surrounded by fellow Europeans, and found a fantastic project to work on. Thanks also to our collaborators, Michele Vianello and his colleagues at the Venice Gateway for Science and Technology (VEGA) who provided the inspiration for this project and the financial support to develop the prototype. Thanks also to Prof. Maria Yang of the Mechanical Engineering department for working with me and showing genuine interest in my project and my point of view, it was a pleasure work with and get to know her.

In developing the prototype the students of MAS.551J and members of the lab provided valuable feedback and encouragement. Prof. Dave Barrett also helped direct the development of the robot, providing valuable insights from his professional experience at iRobot and Disney. He helped focus the development offering practical advice and refreshing enthusiasm. Great thanks goes to Jeremy Scott, a graduate student in CSAIL who gave his time and expertise to develop the control script for the robot and to test drive the robot at various stages in its development.

I would like to mention those who have supported my time here at MIT, particularly the UK Kennedy Memorial Trust for my initial fellowship, and Federico Casalegno, Prof. Steven Eppinger, Prof. Warren Seering, and Prof. Matt Kressy who provided my two Teaching Assistant positions whilst at MIT. Both positions allowed me to indulge in my love of design and work with many talented and interesting graduate students.

Finally I would like to thank my family and friends back in Scotland, for having always encouraged me in work and play. My parents brought me up with a passion for making things. Between my Dad patiently restoring classic motorbikes, and my Mum teaching us arts and crafts, it seems unsurprising that I found my way into product design engineering. My success has all been down to a mixture of hard work, common sense, creativity, and empathy; traits I attribute to the two of them; and for that I am very grateful.

BIOGRAPHICAL NOTE

Jennifer Milne holds a First Class M.Eng degree in Product Design Engineering from the University of Glasgow department of Mechanical Engineering and the Glasgow School of Art School of Design. Upon graduation in June 2010 she was awarded the Glasgow University Engineering Society Medal and Prize as the most distinguished graduate from the School of Engineering. Jennifer came to MIT as a Kennedy Scholar, and in her final year was the Teaching Assistant for MAS.551J Design Without Boundaries, and 2.739 Product Design and Development. During her time at MIT Jennifer also undertook internships at two prominent global design consultancies, as a design strategist at Continuum, in West Newton, MA and as an engineer at IDEO, in Chicago, IL.

TABLE OF CONTENTS

3. DESIGN AND IMPLEMENTATION 47

LIST OF FIGURES

1. INTRODUCTION

1.1 Abstract

Telepresence robots have emerged as a novel solution to meeting the social communication needs of nomadic workers. This thesis follows the design of a low-cost telepresence robot which can be remotely operated whilst running Skype, and discusses how further non-verbal communication cues could be incorporated in future models to increase the feeling of presence experienced when operating such a device. Specifically, face tracking and the ability to communicate gaze is developed in the final prototype.

1.2 Introduction

Good relationships mean everything in business. Whilst an executive might be considering two impressive pitches from competing consulting firms, the fact that one pitch was led by an old buddy from college or that the other company took him out for a great round of golf, will influence his decision. Relationships, whether personal or professional rely on positive social interactions, and they all start by two people talking to each other. What does this mean for 20% of today's US work force who work remotely, at home or from their hotel? (Dieringer Research Group Inc. & WorldatWork, 2011)The remote workers who miss out on such face to face communication? Unsurprisingly perhaps, they tend to miss the company of others, feeling disconnected and isolated, and due to limited visibility in the work place, they have to work harder for a promotion (Mulk, Bardhi, Lassk, & Nanavaty-Dahl, 2009). There are a host of technologies which support remote collaboration, but these tend to focus on the information, and do not consider the social processes that support information exchange. (Hinds & Kiesler, 2002)

This is surprising at a time where people feel more connected than ever. An aversion to personal interaction has emerged from constant connectivity, in favor of asynchronous tools such as email or instant messaging. Some of this relates to politeness; it has become much easier to shoot off an email than risk bothering someone with a phone call. Another reason cited for avoiding real-time conversation is the lack of control; with an email, you can take the time to make your point much more eloquently than if someone was to put you on the spot (Turckle, 2011). This leads to several issues; firstly, that real-time communication becomes reserved for formal interactions only, and secondly, that asynchronous texting, emailing and posting online, does not provide a high level of information richness, which further inhibits social exchange (Marshall, Michaels, & Mulki, 2007). Face to face interaction is described as rich since it allows for rapid mutual feedback. A message can be adjusted, clarified, and reinterpreted instantly (Daft, Lengel, & Trevino, 1987), the interaction is dynamic and personal, where both members can influence and affect the other. It remains the 'golden standard' against which all other communication mediums are compared (Resnick, Levine, & Teasley, 1991).

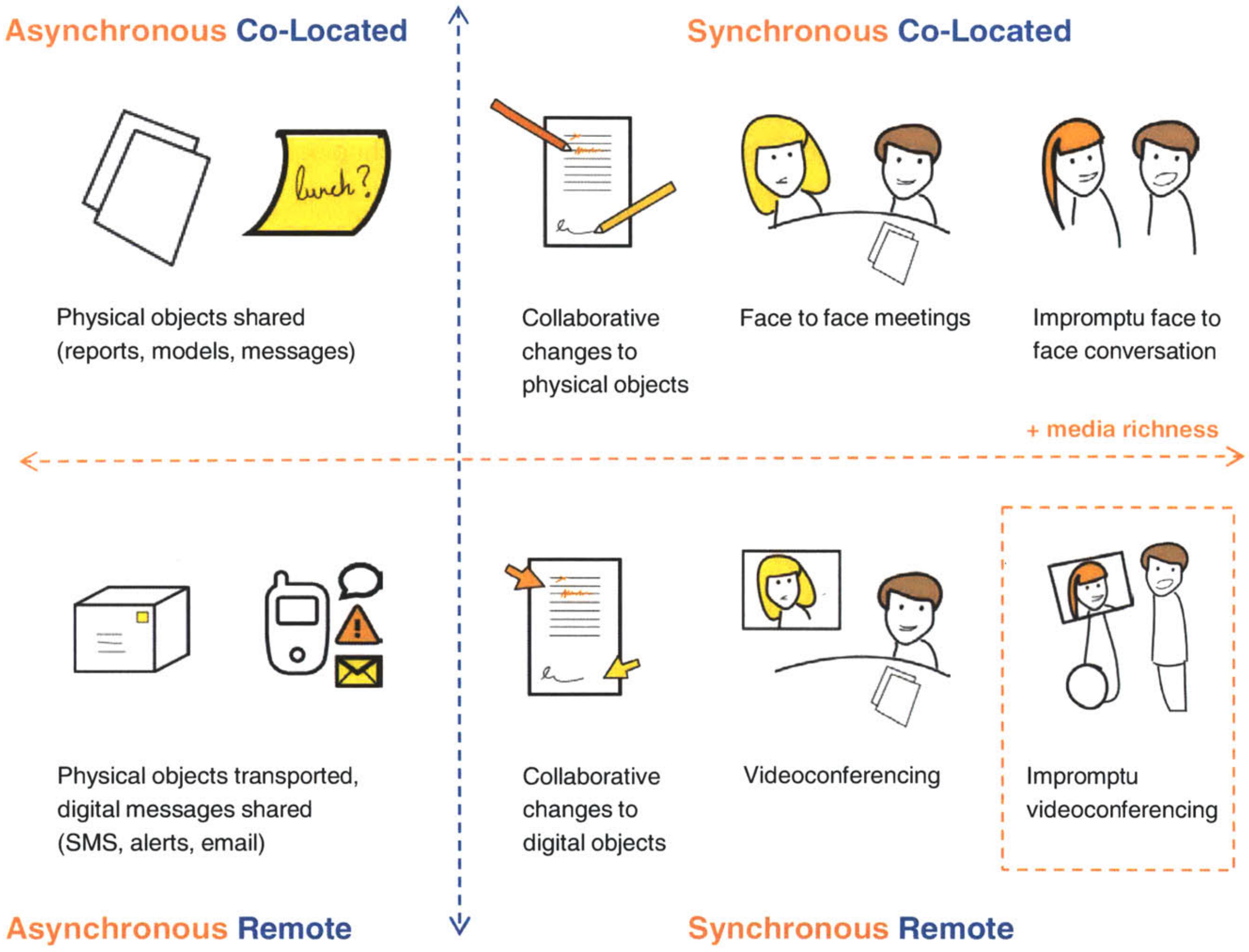


Figure 1: Matrix Comparison of Communication Channels for Co-Located and Remote Collaboration

Figure 1 provides an overview of the synchronous (real-time) and asynchronous communication channels which typically occur between co-located workers and distributed workers. Real-time communication is incredibly complex; language is used to deliver news and information, but in parallel, a subconscious stream of non-verbal cues such as tone of voice, facial expressions, gestures, touch and eye contact, tell a fuller story. Non-verbal cues help to communicate ideas, feelings and emotions, and are just as important as linguistic content in determining whether a social interaction has a positive or negative outcome. (Pentland, 2004) This non-verbal stream is much more important in human affairs than most people think, and specifically people tend to judge relationship quality through non-verbal cues (Remland, 2000).

Daft et al. defined ‘social presence’ as how well communication technologies convey physical presence, and nonverbal and social cues to participants (Daft, Lengel, & Trevino, 1987). If face to face interactions are the gold standard of communication, videoconferencing is the current gold standard of remote communication since it takes place in real-time and incorporates two very important cues, tone of voice and facial expressions. Telepresence robots take this one step further by additionally incorporating movement.

Telepresence refers to the sensation of being elsewhere, especially when remote controlling a physical object in that distant location. Telepresence robots take videoconferencing outside formal meeting rooms, and for the first time, give the remote participant the ability to engage in purely social interactions around the work place. Remote workers can have the same power as their colleagues to initiate or terminate a conversation, removing the hierarchy which previously existed, as outsider was invited in. Their communication zone is extended and their physical presence is dramatically increased by their robotic embodiment. These robots open up new usage scenarios, where the remote participant can step aside with someone during a meeting to have a private word, they can drive up to someone’s desk if they are ignoring an urgent email, and they can approach a group of people standing around a coffee maker to ask them how their day is going. Whilst several commercial solutions exist for the work environment, they have focused on the ability to be present in multiple locations, from the lab, to the warehouse, to the factory floor. Enough focus has not been given to the new social interactions which can occur; social interactions, which we know could dramatically improve the experience of working remotely.

Social presence can be improved further by adding more non-verbal cues. Whilst there is a great deal of work relating to non-verbal cues and social robots, the author found no thorough overview relating this research to telepresence robots. The scenario based approach taken in this thesis can help guide the development of social telepresence robots, and the overview of methods to communicate non-verbal cues aims to provide a framework for increasing social presence, recognizing that good social relationships breed good working relationships.

1.3 Thesis Statement

A new generation of robots is making it possible to be in two places at once. They allow remote workers to literally knock on their boss' door, and are the perfect solution for more informal and social interactions which remote workers are lacking. This thesis will document the design and development of a low-cost telepresence robot. The robot will be controlled remotely over the internet by the remote operator whilst they communicate over video; telepresence will be achieved as the remote worker can move autonomously through the environment and participate in distant events. Aesthetics were not considered in the scope, rather form followed function.

To frame the design, this thesis explores plausible strategies to enhance the interaction by adding further non-verbal communication cues. It will present an overview of previous relevant work in such HCI topics as telepresence, social robots, and computer supported collaborative work. The larger goal was to develop a low-cost telepresence robot as a test platform to experiment with multiple non-verbal communication cues such as gaze and physical gestures. This proposal will be developed in the further work section.

1.4 Thesis Overview

The thesis presents previous work which has influenced the design and development of commercially available telepresence robots used in the workplace. It then highlights key lifestyle trends which support the growing opportunity in this market, before finally drawing attention to how hacking and open-source communities have spurred innovation in the design and development of low-cost telepresence robots.

The bulk of the thesis relates to the physical design and development of one such robot. This was used to demonstrate the technology, and is intended to serve as a test platform for further work into non-verbal social cues. Specifically, face tracking was explored as an intuitive way to control the robot camera and manipulate the robot's point of view. This helps communicate gaze direction, indicating to those interacting with the robot whether the remote person is focused on them or distracted.

1.5 Methodology

This research project emerged as part of the MIT Media Class MAS.551J Design Without Boundaries Fall 2011 led by Federico Casalegno and the author as Teaching Assistant. The research problem was loosely proposed as 'the future of the workplace' in line with previous work and experience of the MIT Mobile Experience Lab. The problem was defined in detail during several days of workshops and collaborative discussion with the sponsors, VEGA (Venice Gateway for Science and Technology), where expectations, goals and a schedule was set prior to the start of the workshop.

After this alignment stage, several weeks of collaborative research into design precedents and academic topics led to a shortlist of opportunities for the class to explore. The topic of telepresence robots was raised and selected by the author to pursue as a good opportunity to demonstrate both engineering ingenuity and user-centered design. The class comprised of formal 3-hour sessions each week, supplemented by one on one time with the Professor. The class setting shifted between lectures with group work, student presentations, reviews and guest lectures. The VEGA team was constantly updated during all the workshop phases and was available to address questions as they arose. Collaboration happened via file sharing using Dropbox (www.dropbox.com) and email, but also via face-to-face meetings in Boston and in Venice. A blog was maintained (http://newmediabuilding.wordpress.com/) to keep track of related research topics and maintained by the author. It served as a useful tool to provide context to guest lecturers in advance, and to keep VEGA informed of the path the research was taking. As of today (05/25/11) the blog contains 58 posts and has received 1,079 all-time views from 7 different countries. Whilst this required some work to maintain, students preferred the reduction in email 'spam' as links were posted to the blog and not emailed to everyone. Another benefit to the blog was that the interdisciplinary research was captured for longevity and not lost at the end of the class.

Invited guests included Andrea L. Kavanaugh (Associate Director of the Center for Human-Computer Interaction, at the Department of Computer Science, Virginia Tech) who advised the students, discussed projects and gave feedback on the work. David Bartlett (Vice President of Industry, Solutions, Energy, & the Environment at IBM) also lectured and responded to the students' work whilst Mick Clare (design strategist at Continuum) provided feedback during the prototyping phase. The invited guests provided a useful industry perspective.

The majority of class sessions involved intensive collaborative design sessions amongst students. The goal was to tackle a particular problem in a multidisciplinary way. It was a very intensive focused innovation process, where the students presented their weekly progress, shared their ideas and raised any issues. Students and instructors provided thoughts and insights helping them see things from a different perspective.

Throughout the semester, students passed through formal stage gate reviews to critique their concept, prototype and final deliverable. Half-way through the workshop, students travelled to Venice to present their concepts to VEGA and to see the site of the new building. On Wednesday, November 9th, Zoe Schladow of the MIT Mobile Experience Lab attended Ecomondo, the 15th International Trade Fair of Material and Energy Recovery and Sustainable Development in Rimini, Italy. Schladow presented the work of all students during a break-out conference session in the afternoon. Led by Director Michele Vianello and his talk “Pandora, a Living Organism in Marghera”, this session was part of the Città Sostenibile (Sustainable City) section of Ecomondo.

After the conclusion of the class the project continued throughout the Spring Semester. Weekly class sessions were replaced by weekly lab meetings where MIT Mobile Experience Lab members shared their progress and presented their work. This was valuable in quickly resolving issues and seeking advice. The project was also presented at a Mechanical Engineering seminar and discussed with Prof. Dave Barrett, Visiting Associate Professor at MIT and former VP of Engineering at iRobot Corp to illicit further feedback.

The prototypes were developed in an intensive test and iteration cycle with sequential improvements to control software, stability, and social presence. 5 distinct iterations were made beyond the first functioning robot demonstrated to VEGA at the final presentations of MAS.551J. The prototype was tested by a remote operator and observed by the author whilst moving through a space and interacting 'in the wild' with anyone or anything they came across. This fits into the notion of user-centered design (Gabbard, Hix, & Swan, 1999), where usability is incorporated into the product throughout the development process. The remote operator, a CSAIL graduate student at MIT, was able to provide feedback relating to the control software or interactions, whilst the author as observer could understand how the physical robot might be improved. This informal testing process allowed for rapid iterations to occur. The final prototype was demonstrated at the MIT Mechanical Engineering de Florez Award Competition and was awarded 5th place in the Graduate Design track.

2. PREVIOUS WORK

2.1 Nomadic Work

Nomadic workers are an extreme form of remote workers. Strictly speaking they travel to where the work is, and are location-independent (Su & Mark, 2008). These people represent a growing segment of the workforce, due to recent changes in place and technology. The internet has altered the perception of place in two ways; with wireless ubiquitous computing allowing for place-independent interactions, and with mobile and location services providing access to place-specific information (Foth, Jaz Hee-jeong, & Satchell, 2011). It is now possible to be anywhere in the world and yet simultaneously immersed in what is occurring at a specific distant location.

Remote work typically relates to working from home; an independent study found that 63% of remote workers conducted work at home. A significant number of those surveyed (between 33-40%) also reported working in the car, on vacation, from a hotel or motel, from a café or restaurant, or from a customer/client's place of business (Dieringer Research Group Inc. & WorldatWork, 2011). This is why the term nomadic worker is preferred since it emphasizes the variance in location and flexibility of workers. The study also found that in 2010, most nomadic workers were male, college graduates, and around 40 years old. The typical user profile then is that of a knowledge worker at a relatively late stage in his career, likely considered an expert in his field.

The rise in demand for freelance work has also spurred the rise of the nomadic worker, particularly in the creative industries; in 2009 28% of all designers in the UK worked freelance (UK Design Council, 2010) . Websites such as Dribble (dribble.com) make it easy to promote and connect talented designers with companies who require assistance with user experience or graphic design challenges.

2.1.1 BENEFITS AND CHALLENGES

There are a great number of benefits to nomadic work. The workers enjoy the flexibility, experience reduced anxiety and better health, and are more productive, satisfied in their work, and loyal to their employer (Kurkland & Bailey, 1999). For the employer too there are benefits; they can access a more diverse talent pool (Morgan, 2004), and experience reduced absenteeism (Lesonsky, 2011). It also lowers their overhead costs substantially and their workers are more likely to work longer hours, potentially increasing productivity. This benefits of this lifestyle are reinforced in Microsoft's research, where the majority of workers surveyed, expressed a desire to work remotely twice as frequently as they currently did. (Ipsos Public Affairs, 2011) Such flexibility is no longer viewed as a perk, but has become a crucial part of today's ideal job (Dieringer Research Group Inc. & WorldatWork, 2011).

However, there are some downsides to this working lifestyle, particularly relating reduced social interaction (Mulk, Bardhi, Lassk, & Nanavaty-Dahl, 2009). Workers may come to miss the company of others (Marshall, Michaels, & Mulki, 2007), and the cohesiveness of groups may suffer; telepresence is no substitute for the way that face-to-face meetings build and strengthen personal relationships. From a career perspective, the nomadic worker is far less involved with the companies work, they may not receive the same level of training and mentoring that other employees do (Nelson, 2003), and their limited visibility in the organization and interaction with supervisors can also limit their chance of promotion (Olson, 1982).

Whilst popular with employees, remote work is less supported by employers who require a different management style to accommodate such flexibility. Due to the reduced visibility of their subordinates, there is a perception that it will difficult to know if they are working hard and performing well. It is clear that in order to support remote workers, and address the key concerns of reduced face-to-face discussion and difficulty in getting quick responses from colleagues, that future business tools and services should aim to fit in with richer ways of collaborating (Liu, Stappers, Pasman, & Taal-Fokker, 2011).

2.1.2 THE FUTURE OF THE OFFICE

The concept of a 'non-building,' or 'work without walls' (Lesonsky, 2011) is becoming more popular; if employees no longer need dedicated desk space, what does the office of the future look like? How might we redefine what it means to go to the 'office'? Dixon & Rose provide examples of how companies have benefited from following this trend. Yell, a UK based directory service, realized their sales staff was working out of 35 nationwide offices, but was constantly on the road, leaving these offices underused. They closed 15 offices and equipped their sales team with portable technology and access to independent telework centers. In doing so they reduced their property overhead by $2.5m a year (Dixon & Ross, 2011). Microsoft also dramatically rethought its office space; refurbishing all its offices around the world, and converting them to activity-based working, where no one has a permanent desk and they connect to work through laptops and mobile phones. Employees each have a locker for their belongings and the office is designed to be almost paperless (Smith, 2012). Even the highly conservative British government has recognized the change in working styles. A book written in 2008 looks at how government offices can downsize, to create better workplaces and more intelligent distribution. In doing so, they hope to also vastly increase how often these spaces are used (Bridget, et al., 2008)

The virtual office is an extreme interpretation of the 'non-building'. Many businesses experimented in reducing their carbon footprint and unnecessary downtime due to employee travel by hosting meetings and conferences online using virtual spaces. Cisco, a leader in telepresence solutions used Media Net to run virtual meetings. This allowed their annual GSX sales conference traditionally held in Las Vegas for 13,000 people at $4,037 a head, to go down to only $437 a head when held virtually in 2010, and an additional 3,000 people 'attended'. In 2009, they realized over $299m of cost savings on telecommuting utilizing over 700 of their telepresence rooms globally (Dixon & Ross, 2011). Research in this area continues to explore how interactions in virtual worlds such as Second Life (www.secondlife.com) compare to face to face interactions (Friedman, Karniel, & Lavie Dinur, 2009), and independent virtual collaborations spaces continue to be developed which try to draw from old social dynamics. Communico is an example of this, an online space which tries to recreate the experience of overhearing discussions in a physical workplace (Dullemond & van Gameren, 2011). Another benefit to using virtual spaces is that in comparison to video, which does provide synchronous interaction, but in passing, creates very little shared memory, interactions in virtual spaces automatically produce a digital history; a common shared knowledge which everyone can refer back to (Casalegno & McAra McWilliam, 2004). From this asynchronous history, data mining and social analytics can be used to understand who is talking about what. Researches at IBM China created a program called Pharos to internally help IBM BlogCentral users in navigating the blog site and locating their interests. Pharos provides a top level summary and can auto-generate social communities based on "hottest" activities at that time (Chi, Liao, Pan, & Zhao, 2011).

A distributed network of independently operated, physical office space is also emerging to provide all the resources of an office environment at a local level. Often referred to as co-working space or previously, telework centers, these spaces bring nomadic workers together, addressing some of the social interaction lost by working alone. Typically co-working spaces offer a good internet connection, desk and chair, printing, faxing, scanning and sometimes free coffee. French company Neo-Nomade (www.neo-nomade.com) provides such office space around Paris. Workers using the service can easily find the nearest desk to get going on their project using a smart phone app. The co-working movement believes that independent professionals and those with workplace flexibility work better together than they do alone and there is a database of free and affordable places to work all over the world (wiki.coworking.info).

2.1.3 TELECONFERENCING AND VOIP

Telepresence is an essential communication tool for any growing organization operating from distributed locations. Cisco is the world leader in corporate telepresence solutions with systems such as the Tandberg-T3, available from $300K, providing immersive board room environments for those in different locations to come together virtually (www.cisco.com). These systems use life-size ultra-high definition transparent displays to create the illusion that everyone is physically together around the same table. This is extremely impressive technology, but the limited nature of the interactions which can take place in the room where it is installed, do not accommodate for informal, creative or spontaneous discussion.

A benefit to telepresence over face to face conversation is that it is very easy to collaborate around digital objects, since both participants are already connected to the internet and a display screen they can quite easily pull up images, reports or digital models whilst maintaining the flow of conversation. In such creative scenarios this ability to talk whilst sharing a digital point of view is almost superior to face to face interaction. From a multi-disciplinary educational workshop, Casalegno provides anecdotal evidence where students working across a language barrier found the ability to share objects to be more important than the social cues provided by video, "as long as we were able to show them a picture of what we had, we didn't have to see them: as long as we were able to share what we were looking at." (Casalegno & McAra McWilliam, 2004)

On the other end of the spectrum, Skype (www.skype.com) is a software application offering free video calls over the internet using a webcam or integrated camera that now comes with most laptops, netbooks or smart phones. Skype is used from one monitor to another as participants sit in front of their computer screens, and it does have some issues with signal quality and delays. Skype does not offer the same reliability as Cisco's integrated solutions by any stretch of the imagination; nevertheless, users happily tolerate such flaws given the ease with which they can meaningfully connect with friends, family and colleagues all over the world. ooVoo emerged in 2007 as a competing service offering free 6-way video conversation, thus supporting group video conferencing, and providing real-time performance diagnostics for an early indication of latency in the call (www.oovoo.com). In comparison, Skype users can only run group video calls by upgrading to their premium service for a monthly fee. Both services profit from offering such premium features as screen sharing, free landline calls and large file sharing. Free video calling services are now mobile, with Apple's CEO announcing FaceTime for the iPhone at his keynote speech in June 2010 (www.apple.com/iphone/built-in-apps/facetime). Skype and ooVoo also offer iPhone/Android applications, as do new competitors such as Fring and Tango (www.fring.com) (www.tango.me).

2.1.4 DISPLAY TECHNOLOGY

Improving image resolution is an obvious thread in the development of telepresence systems, and with 3D televisions entering the mainstream, 3D telepresence is in pursuit. Cisco demonstrated holographic telepresence on-stage at the launch of their Globalization Center East in Bangalore, India in 2007 using an ultra-high definition camera and codec overlaid onto Musion's patented transparent eye-liner foil stretched across the stage (www.musion.co.uk/Cisco_TelePresence). Truly holographic 3D telepresence is different to this example of 3D images now available, in their ability to display parallax; the object looks different when viewed from different angles Researchers have developed a screen that displays dynamic color holographic images which could eliminate the problem of viewing a 3D world from a fixed camera's perspective, however, the refresh rates are a long way from rivaling live video, with resolution at one frame every 2 seconds to achieve half parallax and 3 minutes for full parallax. (Blanche, et al., 2010).

2.1.5 INCREASING PRESENCE

The term presence comes from the Latin *praesentia,* 'being at hand'. This presence is apparent in the power exuded by our smart devices; we have at our fingertips, the ability to connect with our entire social network, 24 hours a day, 7 days a week. This connection we feel to others through our technology, is comforting, even if we do not interact, the value comes from knowing you can reach out if you want to. Presence is often described as a subjective sense of "being there" (Heeter, 1992), and our phones make us feel that others are indeed there for us. The true weight of this presence is realized by the overwhelming loss which is felt when a hard drive crashes or a phone goes missing. It interrupts our feeling of social security when the presence such an object commands disappears.

Increasing the feeling of presence is a common goal of research projects in the area of telepresence robots. Chertoff comments on the varying heuristics which have been used in Presence Theory, believing that a better understanding of the elements that contribute to presence is essential to design effective telepresence experiences (Chertoff, 2008). Chertoff references the work of Lee (Lee M. K., 2004), who before him, listed 30 Presence articles which described 25 psychological and physical factors which were found to cause a feeling of presence. These ranged from personal moods and experience with technology to the inclusion of sound and image resolution.

It is proposed that an increased sense of social presence is related to a decreased perception of workplace isolation. Another approach to creating a feeling of presence might be, to instead, focus on lessening the feeling of isolation. Marshall et al. found workplace isolation to be two-dimensional, relating to individuals' perceptions of isolation and perceived isolation from colleagues and company support networks. Their insights informed the development of a self-reporting workplace isolation scale (Marshall, Michaels, & Mulki, 2007).

2.2 Robots for Nomadic Work

2.2.1 INTERNET OF THINGS

Urban densification across the globe is leading to cities becoming a complex mesh of tangible and intangible parts, creating a 'hybrid space' between the physical and digital (de Souza e Silva, 2006). It is in this hybrid environment that there is an emergence of the 'Internet of Things', a term first coined by Kevin Ashton (Ashton, 1999). Whilst the definition is still evolving with technology, it generally refers to the organization of objects when each individual object is uniquely identifiable and connected via the internet. Neo-Nomade (www.neonomade.com), the co-working service mentioned previously, encompasses this concept. By providing real time information about available office space, a nomadic worker can find a desk and chair on an as-needed basis. Just as Zipcar (www.zipcar.com) changed the way people rent cars, and airbnb (www.airbnb.com) provides a platform for people to sublet their spare rooms, there are likely to be big changes in the way we share and rent work space.

When considering the future of the workplace, it is also apparent that a worker only needs a desk during the parts of the day when they are out of meetings. In 'Pandora', the new office building being developed by VEGA, sponsors of this work, the architects propose having robotic walls, so that the size and number of office rooms can be controlled and varied to meet the requirements of the occupants. In co-working spaces, such as the Cambridge Innovation Centre in Cambridge Massachusetts, it is common to find all the furniture on wheels so that it can be redistributed and re-organized on an as needed basis.

Robotic furniture is a new research field where using the internet of things, the furniture can move and re-organize itself. Roombots is one such system from the University of Lausanne where individual components can move autonomously and come together to form chairs, stools and desks (Sprowitz, 2010). Currently, many workplaces use a central booking system to reserve meeting rooms. In the truly connected office, people would also reserve furniture, and program the setup and lighting requirements in the same way.

2.2.2 TELEPRESENCE: FROM WAR ZONES TO WEDDING PARTIES

In the 1980s, Marvin Minsky published an influential manifesto about the age of telepresence, citing the biggest challenge to be 'achieving that sense of "being there",' and questioning if telepresence could ever be a substitute for the real thing (Minsky, 1980). Whilst we have moved closer towards this vision, the question remains relevant today, 30 years on. Telepresence robots are the closest we have come to creating that sense of 'being there.' They are used in a wide range of scenarios, including allowing an 82 year old mother, too frail to travel from her home in Las Vegas, to attend her son's wedding in Paris and even dance with her family at the reception. (Glazer, 2012)

Initially, the motivating scenario for developing telepresence robots was to allow humans to explore environments too dangerous or remote, to access themselves (Clark, 1961). Examples of hostile environments include space, nuclear labs, the ocean depths, and enemy territory. Remote manipulators could substitute for humans in tasks which previously posed risk to human life, such as with the Canadarm (Aikenhead, Daniell, & Davis, 1993), a 50ft mechanical arm used to maneuver payload on the Space Shuttle. It was first used on STS-2 in 1981, and subsequently used in over 50 shuttle missions. Initial research projects typically had military or space applications, but the technology soon influenced other areas. Robots started being used to protect maintenance staff from radiation, by allowing remote inspection of nuclear plants (Kim, Johng, & Kim, 1999). Robots were used to keep military personnel informed, by performing covert reconnaissance missions in enemy territory (Chemel, Mutschler, & Schempf, 1999). Robots were used to protect civilians, by remotely detonating land mines (Tcherneshoff, 1997), and robots were used to replace firefighters, in handling highly explosive gas cylinders during a blaze (Hisanori, 2002).

As the cost of processing power decreased, autonomous robots went from assisting in dangerous pursuits, to slightly duller pursuits, that people would rather avoid themselves, such as vacuum cleaning (Andersen, Medaglia, Bimpel, Sjølin, & Mikkelsen, 2010) and repetitive factory work (www.kivasystems.com). Whilst the autonomous tasks robots were being programmed for got duller, telerobotic surgery (Ottensmeyer, Thompson, & Sheridan, 1996) provided an intricate and life-critical remote operation challenge. The incredible DaVinci system from Intuitive Medical (www.intuitivesurgical.com) represents the state of the art, allowing a surgeon in one part of the world to remotely perform minimally invasive surgery on a patient located somewhere else. Remote surgery is a pertinent example of telepresence, where the surgeon is completely immersed in an activity occurring in a distant location. Telepresence robots continued to help people in other ways, such as PEBBLES, an assistive robot from Ryerson University (Yeung & Fels, 2005) that allows sick children to attend school remotely, raise their hand in class and scan and fax their homework to physically hand it in, just as they normally would. In Boston, MA Boston Children's Hospital has begun sending telepresence robots home with post-operation patients so medical staff can check up on them remotely, without the patient needing to return to the hospital several times during their recovery (Dillow, 2011).

The past decade has seen telepresence robots filter down into more everyday scenarios. IRobot (www.irobot.com), a leader in the consumer robotics space, were one of the first companies to attempt to break the mass market with their two robots, the iRobot LE and the iRobot coworker. The iRobot LE was intended for residential use with its ability to climb stairs, and the coworker was an industrial model, intended for use on the factory floor. Whilst these robots performed well, the price point was wrong, and it was not until recently that telepresence robots became affordable enough to gain traction and mass appeal.

2.2.3 COMMERCIALLY AVAILABLE TELEPRESENCE ROBOTS

The past couple of years have seen a surge in newly founded companies offering telepresence robots with mass appeal such as Anybots (www.anybots.com), founded in 2001 in Mountain View California. Their QA model, demoed at the 2009 Consumer Electronics Show (CES), was well received but at $30,000 it was still too expensive, and the company voiced their desire to get this down to $5000, a goal they are still pursuing. In summer 2011 they released the more affordable QB model available today for $10,000. With its Segway-like mobility and comical appearance (often pictured wearing a bowtie), it certainly provides a talking point. Their QB model is a frontrunner in the space having received great interest from the mass media, and valuable exposure from an article in the New York Times (Markoff, 2010). Also featured in that article was the VGo (www.vgocom.com) from VGo Communications, more accessible at just $5,000. Other competitors listed were robots from inTouch Health, Willow Garage and RoboDynamics (www.intouchhealth.com) (www.willowgarage.com) (www.robodynamics.com). IRobot, a relative veteran in the consumer space, recently demoed a robotic development platform called Ava at CES 2012 and indicated an agreement with inTouch Health, the main provider of telepresence robots for the healthcare industry. Ava can be used as a telepresence robot, but is also available to developers to create their own unique applications. Outside the US there are several International competitors both in Europe and Asia; companies such as Gostai (www.gostai.com), who released their Jazz robot in 2011 (€7,900), and Tmsuk co. (www.tmsuk.co.jp/english/robots) a Japanese company who offer a wide spectrum of robots, including the tmsuk-4, a shopping robot allowing anyone to enjoy a remote retail experience via their mobile phone.

The basic functionality of the robots is the same, there is a mobile base which can be controlled remotely, and a structure supporting a screen, microphone and camera such that live video and audio can be streamed and received over an internet connection. These robots differ physically in terms of height, screen size, additional features, and form, and experientially in terms of how easy they are to operate, set up and the video/audio/connection quality. The QB from Anybots for example can be extended to a height of 6ft, offering a good experience for standing conversations, whereas the VGo is fixed at 4ft. They allow the screen to rotate in order to provide a better viewing angle. The VGo screen is also larger than the QB's thus can be viewed from further away. A compelling feature of the QB is that it has web based controls and is therefore very accessible; anyone can be sent a web link and drive the robot immediately, without needing to install special control software.

2.2.4 DIY TELEPRESENCE ROBOTS

In parallel to these commercial systems, there are a host of research robots and robotic platforms which can be used and programmed to operate as telepresence robots. These projects are supported in huge part by the amount of open-source software available in this area. Both Texai from Willow Garage and the commercially available Jazz from Gostai leveraged the open-source ROS framework to develop their applications (www.ros.org). There is also a growing DIY/hacker/maker culture creating and sharing well-documented projects online. With a little bit of know-how and support from these communities, anyone with spare money, time and enthusiasm should be able to build their own telepresence robot.

IvanAnywhere built by Ivan Bowman from found parts for around $5000 was the first DIY project to gain online coverage in 2007 (Gaylord, 2008). Ivan built the robot to allow him to work from home for his company iAnywhere in Ontario whilst still attending important meetings. Since then, projects such as Sparky and Sparky Jr have risen to fame on hobbyist website instructables.com (Gomi Style Crew, 2012), a DIY telepresence robot was featured in IEEE Spectrum (Schneider, 2010), and Google engineer Johnny Lee demonstrated on his 'procrastineering' blog, how he built a Low-cost video chat robot to keep in touch with his long distance girlfriend for just $450 with an iRobot Create and a netbook (Lee J. , 2011).

2.2.5 NON-VERBAL COMMUNICATION IN SOCIAL ROBOTS

Social presence was defined earlier as how well the robot communicates physical presence, and non-verbal and social cues to participants (Daft, Lengel, & Trevino, 1987) . All the telepresence robots listed previously have a strong presence from tone of voice, facial expressions, and movement. In this section, the entire spectrum of non-verbal and social cues will be examined, drawing from research in related fields and applying this to the design of social telepresence robots. The author was unable to find a high level overview of this nature in the literature and believes it provides a very useful framework for developing robots which will enable rich social interaction.

Facial Expressions

Eminent social psychologist Michael Argyle, states that the face is the most important non-verbal channel for expressing attitudes and emotions (Argyle, 1988) Video-calling is more expressive than phone conversation for this reason, you can communicate both news and information, and, more clearly express your feelings and emotions at the same time.

Since facial expressions communicate so much, it was curious at first to consider that several commercially available products actually have a very small video screen, to the extent that this valuable information is being lost from the interaction. The Anybots QB and VGO robots have 3.5" and 6" diagonal screens respectively, much smaller than what most people currently use in monitor to monitor Skype video-calls. An initial hypothesis was that this was a deliberate design decision to force those interacting with the robot to relate more to the robot's 'face' rather than the live video feed, and that creating a character, as done with the QB, Jazz, and Luna robot from RoboDynamics, was somehow more engaging than a large image of the remote person's face.

However, anecdotal evidence from expert interviews suggests this is not the case. Prof. David Barrett, former vice president of engineering at iRobot, worked on iRobot Coworker (2007) and iRobot LE (2000). He confirmed that it was not a conscious decision to use a small screen, but one dictated by hardware and delivering an affordable product. Several companies such as iRobot, Anybots, and Gostai might have been compensating for a small screen by creating a sense of robotic character. This is supported by the fact that both iRobot and Gostai have made a larger screen a feature of their latest product releases.

It is however, also possible to convey emotions through a robotic face, rather than relying on the live video feed. Much work has been done in the field of social robots. AIDA, a project from MIT's Personal Robotics Group (MIT Personal Robotics Group, 2012) lives in your car dashboard, and communicates warnings to drivers via a cartoon digital expression. The Philips iCat research robot can generate a range of complex expressions by manipulating its eyes, eyelids, eyebrows, and mouth (Philips Research, 2012). On the far end of the spectrum, the Geminoid DK humanoid robot can generate the widest range of lifelike facial expressions (Ackerman, 2011).

It is also possible to understand facial expressions generated by a remote operator through computer vision techniques (Aðalgeirsson, 2009). Avatar Kinect from Xbox uses facial recognition to generate a believable digital avatar to allow people to have expressive virtual conversations (Xbox, 2012). Such techniques could be incorporated into telepresence robots in order to compliment or exaggerate the real-time facial expressions of the remote operator. Another path would be to forgo the live video stream and create a charming robotic face which conveyed virtual expressions in real time using facial recognition.

Deliberate Gestures

In most research examples, gesture is remotely controlled through deliberate user input or less consciously through direct translation of the remote user's gestures. Deliberate input refers to the user actively specifying gestures, such as a child clicking a virtual button to cause the robot to raise its hand as with PEBBLES (Yeung & Fels, 2005) or using virtual sliders to individually position arm joints, or even physically manipulating a sympathetic controller to determine the remote orientation of an arm as with MeBot (Aðalgeirsson, 2009). Whilst this kind of control adds a new element of interaction it is somewhat limited in scope, and the remote operator must consciously tend to this control. This is very unlike natural gesturing and becomes more burdensome in a scenario where that person is also holding a conversation and also driving the robot.

Deliberate gestures are relatively easy to incorporate as new features, but their inclusion might limit the use scenarios for which the robot is appropriate. PEBBLES was designed specifically to allow sick children to attend school. The student is able to perform deliberate gestures appropriate to that context, such as raise their hand in class. PEBBLES also includes a scanner and fax so that the student can deliver paper homework assignments. This is a very specific use case, and the additional hardware to achieve this is already outdated. Some deliberate gestures however are universally understood. These include counting with fingers, pointing to things, and waving. Pointing to things could add the most value when having a discussion around places or objects. Anybots' QB robot is able to do this using a laser pointer.

Non-deliberate Gestures

Non-deliberate gestures are trickier to convey by their very nature; they are delivered sub-consciously. Studies which have required remote operators to input general posture positions tend to agree that "direct manipulation is only good for delayed telecommunication scenarios" (Yim & Shaw, 2011). They found that deliberately inputting non-deliberate positions overloaded the remote operator, "it is difficult to manage the boundary between operating the robot and doing something locally with your arms that is unrelated to the interaction." (Aðalgeirsson, 2009)

Robotic gestures could be used to exaggerate the emotional state of the remote person. A study from Stanford (Sirkin & Ju, 2011) used facial recognition to infer the intent of the remote operator, and then used a robotic arm to exaggerate this with pre-determined gestures. For example, if the remote person leaned in, the robot would gesture forward in tandem. Generally it is risky to infer things from the user in this way because when there is room for error, the user might become frustrated if their robotic presence is wrongly representing their intent. The study tried to reduce the potential for miscommunication by recording a series of head movements coordinated with an arm gesture. They then used Mechanical Turk (www.mturk.com) to ask people to indicate how they interpreted the combination. They were then able to only move forward with a range of gestures which were well understood. The study found that additional gestures increased the 'presence' of the remote operator, and both members of the conversation felt equals in the discussion.

Direct translation of real-time gestures avoids any inference confusion. This has been done successfully using additional hardware such as gloves, the Xbox Kinect, or additional cameras. In gesture tracking examples where only one camera is used, the remote individual is required to stand directly in front of the camera; a scenario which is not perfect for nomadic workers in public places where they may want to be more discrete.

Subtle, non-deliberate gestures such as leg crossing, arm crossing or trembling hands, which communicate a great deal about an individual's mental state, are often trying to be hidden. Undesirable gestures such as these are referred to as non-verbal leakage (Multa, Yamaoka, Kanda, Ishiguro, & Hagita, 2009). Whilst being able to convey such gestures through the robot would share honest signals and be beneficial for interpersonal communication, the desirability is debatable. Whilst this is an interesting research question, it is probably not a valuable function for a commercial robot. Would people want to buy or use a robot that is clearly nervous when giving a presentation to their boss? Chertoff also suggests that "we may find that increased processing power and more realistic sensations offer diminishing results in terms of presence," (Chertoff, 2008) and so it is important to assess whether the pursuit of a particular non-verbal cue is truly going to add to the experience.

Movement

The key differentiator of telepresence robots vs. other communication technology is the ability to move through the distant space which the person you are conversing with, is occupying. This makes the interaction more of a shared experience. Movement is particularly useful in initiating and terminating a conversation. Body language experts say that if you want to know if someone is really interested in a conversation, look at where their feet are pointing. If they are pointed towards the door they are probably looking for a way to get out of the conversation. This is crucial in a work environment where productivity is important and too much talk can waste time.

Through movement, telepresence robots change the conversation dynamic with remote workers. Instead of only being able to communicate with familiar contacts, they are now able to approach and converse with anyone they wish, even if they have never met before. A study of a workplace trialing telepresence robots found that around 50% of conversations the robot engaged with were initiated by the remote operator, reflecting the shift in power. (Tsui, Desai, Yanco, & Uhlik, 2011) Compare this to a typical video conference where remote workers usually have to be invited to the conversation by the main location hosting the meeting.

The MeBot project conducted user trials comparing an interaction with a stationary 'non-expressive' telepresence robot and the same robot with expressions, such as movement and arm gestures. Their study concluded that these expressive aspects improved psychological involvement and behavioral engagement, and led to better cooperation and overall general engagement. One interesting finding was that participants felt more competitive with each other when the robot was stationary, and more comfortable with each other when the robot was expressive, suggesting that these non-verbal cues help establish a flat hierarchy (Aðalgeirsson, 2009). This is further supported by the work of Sirkin & Ju who found that gesturing helped both participants feel like equals during the interaction. (Sirkin & Ju, 2011)

Movement also plays a role in supporting phatic communication, or small talk, discussed in more detail in the next chapter. If a remote worker is driving down a hall to attend a meeting, they might pass someone and be able to ask "hey, how's it going?" but keep on moving. They appear friendly and sociable, and the other person feels their presence, but they do not need to start a full conversation. Alternatively they can choose to slow down and stop, indicating that they want to converse with that person. This ability to more seamlessly transition in and out of discussions through physical movement mimics how we naturally conduct face to face conversations.

The telepresence robot developed in this thesis moves in response to remote keyboard control. Commercial products have worked to improve the control interface and minimize how much time must be spent actively driving the robot. Previously mentioned was the Jazz Connect from Gostai. With their interface, the user can click to a space in the distance where it will autonomously drive to, or select a person for the robot to follow. These kinds of features make it easier for the operator to focus the majority of their concentration on the conversation.

Paralinguistics

Paralinguistics, or tone of voice, help to communicate the emotional intent behind a message. Largely a sub-conscious social cue, tone of voice is among the most powerful of all social signals (Nass & Brave, 2004) Paralinguistics are already conveyed through the audio stream associated with video-calling, however, interesting work from Kyoto University attempts to exaggerate this signal, inferring emotional delivery by analyzing the audio stream and syncing the robotic motion appropriately. A happy voice will cause the robot to gesture excitedly, and a sad voice will cause the robot to move slowly (Lim, Ogata, & Okuno, 2011). The Jerk-o-meter, a project from MIT's Human Dynamics lab also uses speech analysis during telephone calls to infer whether or not the person talking is being a jerk (Madan, 2005). Perhaps there are ways speech analysis can be used in telepresence to encourage better social etiquette by making the robot look bored if someone is hogging the conversation, or raising the volume of the robot's voice if others in the room are talking over one another.

Eye Contact and Gaze

Eye contact is an extremely powerful non-verbal communication cue that helps regulate the conversation, and signals the exchange of speaker roles. It occurs during 10-30% of natural conversation and can be used to acknowledge or avoid the presence of others in passing (Webbink, 1986). A key complaint about web cameras is that whilst you might be looking directly at the other person on screen, direct eye contact is not correctly transmitted. Some commercial products such as SeeEye2Eye, $49.99 (www.bodelin.com/se2e) or Iris 2 Iris $2,449 (www.iris2iris.com) have tried to address this, but they remain bulky, non-integrated solutions. In the MeBot project, (Aðalgeirsson, 2009) they addressed the problem of eye contact by building a display with the camera embedded at the center. The remote window was then projected on to the display from a video projector mounted behind the remote person, giving the effect for the person looking at the robot, that the robot is looking straight ahead.

In the realm of Computer Supported Collaborative Work, screen share features in Skype and simultaneous document editing features in Google Documents (docs.google.com) can be extremely useful for conversations around digital objects. Gaze is crucial in conversations around objects because you are constantly sending cues about what point on the object, or in a document, you are referring to. Eye tracking is used as a research tool to understand what someone is looking at and dual eye-tracking can be used to augment the shared screen experience to communicate what both people are looking at whilst having a discussion around that document. (Gergle & Clark, 2011) This should help reduce miscommunication and misunderstanding in remote discussion.

In relation to telepresence robots, face tracking, which encompasses eye tracking, is a useful tool in understanding gaze direction, and in the MeBot project, it was used as an intuitive way to control the robot's head without requiring operation of an additional device. The remote operator simply had to turn their head and the robot would also move its head in tandem (Aðalgeirsson, 2009). Whilst in projects such as MITRO (Alers, Bloembergen, Bügler, Hennes, & Tuyls, 2012), head tracking was used on the opposite end of the interaction by the robot's camera in order to retain focus on an individual automatically, reducing the amount of repositioning the remote operator must do.

Haptics

Haptics, or the sensation of touch, could lead to completely new interactions in personal communication or discussions around objects. InTouch (Brave & Dahley, 1997), a project from MIT's Media Lab, presented a prototype system providing haptic feedback to interpersonal communication through three wooden rollers which could be manipulated in synchronicity from both locations. Brave & Dahley reinforce that physical contact is a basic means to create a sense of connection, and indicate intention and express emotion. The physical object is crucial in increasing the social presence experienced during an interaction; it is possible to tangibly feel that the other person is there.

Touch is more appropriate for personal communication scenarios, since it is often culturally regulated in organizations (Harris, 2002). In a work setting, tactile communication may be reduced to interactions such as handshakes.

Researchers at the University of Salford are working on improving remote discussions about physical objects, having created a glove with built in sensors that creates the sensation of an object pressing against the finger as it moves around a 3D virtual object (Counsell, 2003). Haptic feedback could allow a remote operator to move through an environment and feel a realistic sensation of touching and interacting with the physical objects they encounter.

Appearance

It is common knowledge that the first impressions made about you when you meet someone dramatically shapes their impression of you. Appearance is generally the first non-verbal message someone receives about you. It influences if they perceive you as similar to themselves and is used to evaluate credibility and general attractiveness (Hickson & Stacks, 1993). A general complaint with telepresence robots is that they do not convey the personality and individuality of the remote person they embody, who now essentially looks the same as anyone else who might drive the robot other than the video displayed on screen. Anybots' QB robot provides some level of personalization with the screen located on the center of the chest below a robotic face. Typically this displays the live video of the remote person, but it also allows for some customization; you can create the effect that the robot is wearing a tuxedo or some other image could be used that reflects your personal style.

The importance of projecting personal appearance was raised in popular TV show the Big Bang Theory, when Willow Garage's Texai telepresence robot makes a cameo appearance as "ShelBot" a robot operated remotely by main character Sheldon. Sheldon chooses to attach a hanger with the outfit he is wearing below the video display of his face so that the robot is instantly recognizable as him and not someone else (Willow Garage, 2010). This has also been seen with inTouch Health robots used in hospitals, where the robot is wrapped in a white coat and stethoscope so that the remote doctor can approach patients in the orthodox fashion.

3. DESIGN AND IMPLEMENTATION

3.1 Initial Motivation

Telepresence robots were discussed in Federico Casalegno's MAS.551J Fall 2011 Class, Design Without Boundaries at MIT. The class was sponsored by the Venice Gateway for Science and Technology (VEGA) to deliver proposals relating to the future of the office. The goal was not to design a building, but to design for connectivity. In this class, students considered how the workplace should facilitate effective communication as an essential tool for productive, successful teamwork. Cisco's systems offer an intimate telepresence experience for the traditional board room discussion, but the nature of work is changing, and telepresence should also support informal social discussions.

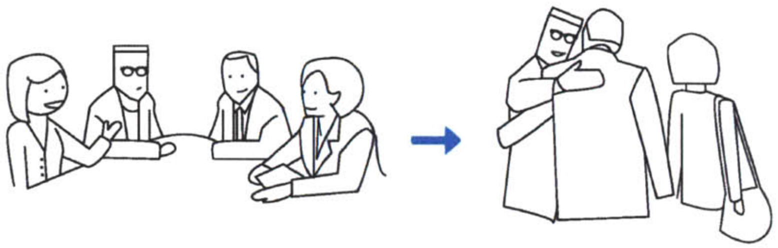
A nomadic workforce represents a population of people who are very non-traditional in their definition. They might be individuals working in the creative sector, knowledge workers, entrepreneurs or researchers. Their culture demands flexibility and the organizations they work for might not have the significant investment required to provide the infrastructure Cisco delivers. Telepresence robots offer increased functionality in comparison to free services such as Skype, but with relatively low initial investment, and have therefore proved a novel and attractive option for start-ups and other small-medium enterprises. Where telepresence robots significantly differ, is that in addition to supporting formal information exchange through video conversation, they just as easily enable small talk. The remote worker can engage in formal board room discussions and then just as easily engage in water cooler talk, or hang around for the after-work beer meeting.

3.1.1 PHATIC COMMUNICATION

Phatic communication, or small talk, is an essential part of natural conversation. It typically frames the meat of the conversation and its value should not be underestimated. It is used to open a conversation, such as asking, "Hey, how's it going," as a space filler during quiet stretches or moments of awkwardness, and to wrap things up. Also referred to as grooming talk, these interactions help everyone feel comfortable and equal in a conversation.

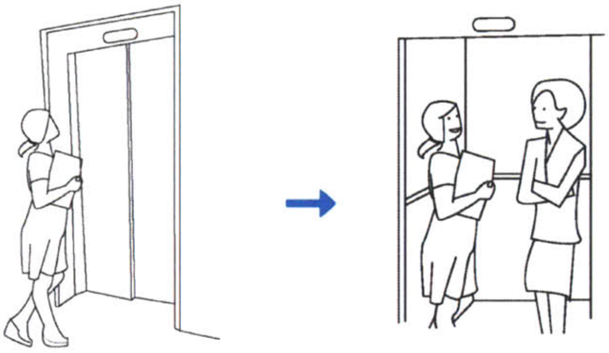
Phatic communication is lacking from current video conversation. Instead of easing into the interaction, the experience is binary; the call is live, or terminated. The issue is apparent to anyone who has experienced awkwardness surrounding the finality of hanging up on someone. Anecdotal evidence suggests that people feel self-conscious about video conversation and often go to freshen up or get their 'game-face' on before initiating or accepting an incoming call. Participants in video conversations can also spend a lot of time monitoring how they are appearing on the camera feed, rather than fully focusing their attention on the other person.

Insight was gained by understanding the current experience of nomadic workers. As previously mentioned, the Microsoft study highlighted particular downsides for individuals such as feeling less involved in everyday office life, being less likely to be offered a promotion, and missing the company of others (Lesonsky, 2011). These scenarios are summarized in figure 2 on the following page. By creating a communication channel which encourages phatic communication, more social interaction will naturally occur, addressing some of these downsides and improving the experience for nomadic workers. Telepresence robots offer such a communication channel. Figure 3 demonstrates that a telepresence robot could just as easily participate in any of the phatic communication scenarios from figure 2, which remote workers are currently excluded from.



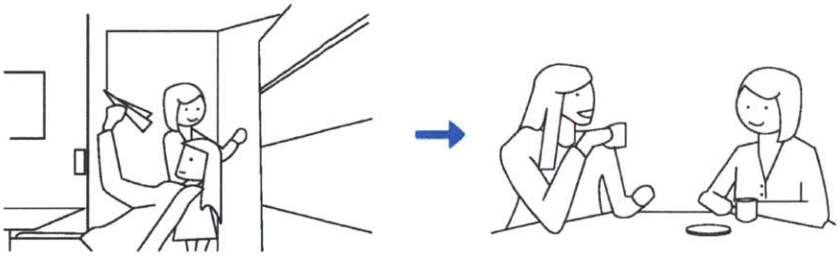
Afterthoughts:

Even after the meeting, sometimes the most important discussions happen as everyone leaves the room.



Networking:

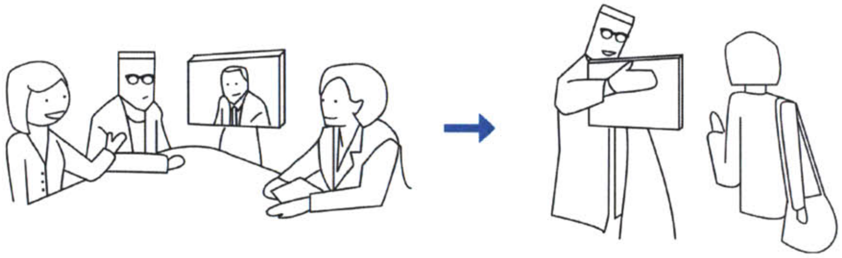
Simply by being in the office, you become known by those who might help further your career.



Socializing:

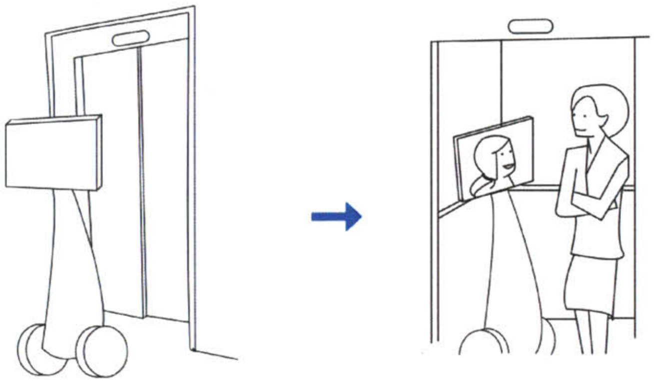
These fun, impromptu interactions are what people miss about coming into work.

Figure 2: Scenarios Highlighting the Key Interactions Missed by Remote Workers



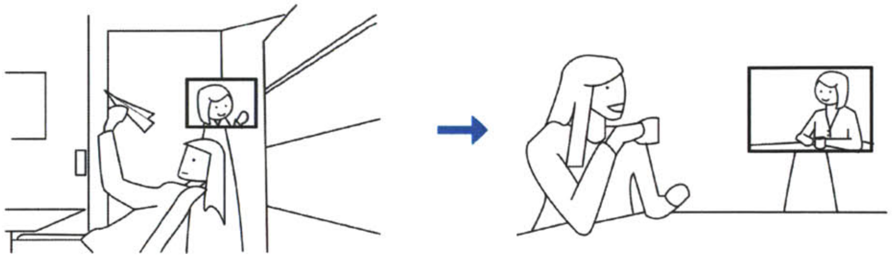
Afterthoughts:

Even after the meeting, sometimes the most important discussions happen as everyone leaves the room.



Networking:

Simply by being in the office, you become known by those who might help further your career.



Socializing:

These fun, impromptu interactions are what people miss about coming into work.

Figure 3: Telepresence Robots Involved in Phatic Communication Scenarios

3.1.2 ORIGINAL PROPOSAL

A new generation of robots is making it possible to be, in effect, in two places at once. They allow remote workers to literally knock on their boss' door, and are the perfect solution for more informal conversation and social interaction. There are many solutions on the market; however, I propose to develop a telepresence robot which also features non-verbal gestural cues. A human-centered design methodology will be used with the nomadic worker in the forefront of design decisions.

The robot should allow a remote worker to be present in the following scenarios located in the workplace. 1) Board room conversations with multiple participants, 2) walk and talk conversations as part of a small group while moving down a corridor and 3) informal location independent scenarios such as in an elevator, a kitchen or other social area, or someone else's desk. The robot should also be able to transition between these scenarios without interruption. Any of these group scenarios should also consider multiple remote workers interacting via multiple robots.

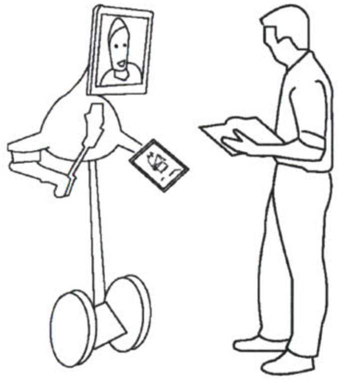


Figure 4: Drawing Generated for Initial Proposal

3.2 Process

The design process followed a double diamond approach, with 2 phases of expansion and contraction as presented by the UK Design Council. The initial phases will be described here, whilst the development will be described in detail in the following section on robot design.

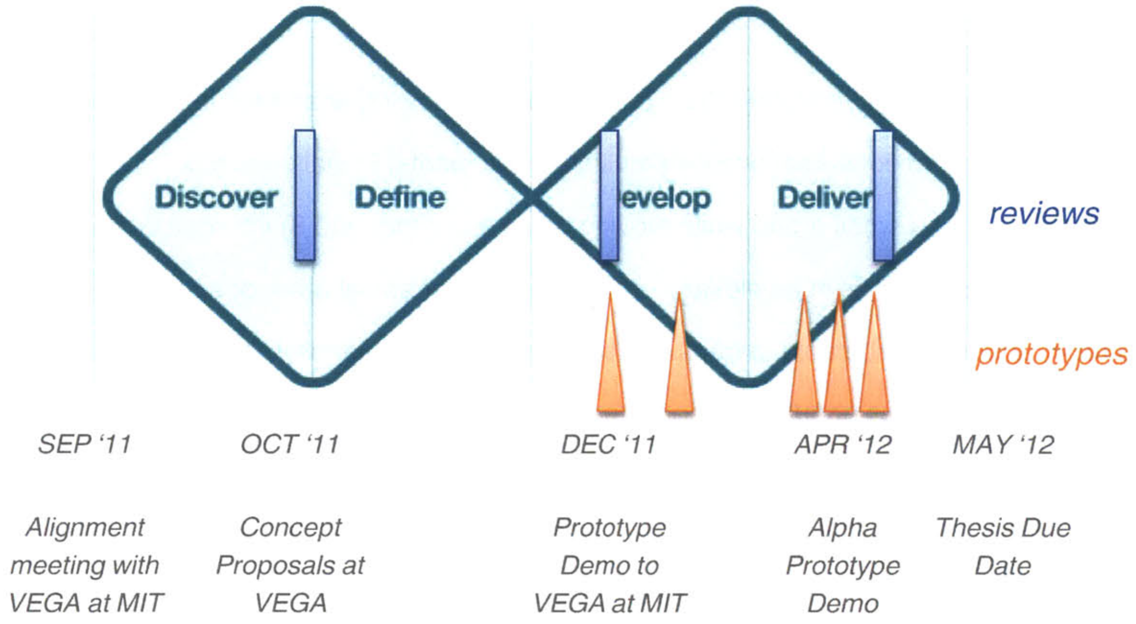


Figure 5: Double Diamond Design Process

Initial research was conducted into parallel and competitive products to understand the design space. A design strategy was chosen to find the shortest critical path to creating a robot with the two basic functions of remote drive control and 2-way live video streaming. The chosen design was selected due to its construction from a small number of affordable parts, and that there was a good community of online support to assist with any issues which might arise during the development. After the design review the budget was confirmed and it was possible to build the initial prototype. After testing it was apparent there were three main areas for improvement, software reliability, stability, and non-verbal communication cues. At this stage, expert consultation was used to learn from those with experience in the telepresence robot industry and HCI. The software was improved by writing a new control script for the second prototype, stability was improved in the third prototype, and non-verbal communication was developed in the final two prototypes in the form of gaze direction.

3.2.1 INITIAL RESEARCH

During the Discovery Phase several research themes were explored as identified with VEGA. Students in MAS.551J researched and discussed interesting design precedents relating to the future of the workplace. Robotic furniture was initially explored in line with the building having robotic walls; the space can be transformed to meet the requirements of those occupying the building at that point in time. This then prompted a discussion about how to offer some level of personality to such objects if they never truly belong to one individual. Telepresence robots then emerged as an initial step in embodying a personality in a pre-defined vessel. It was proposed that current products on the market are a step in the right direction, but there is still a lot of room for improvement by focusing on using the robots as a tool for social interaction.

3.2.2 STRATEGY

Given their rise in popularity, DIY telepresence robots (essentially a laptop on wheels), have been well documented on the internet. For this project it was not appropriate to purchase a telepresence robot and adapt it; instead, advantage was taken from previous work documented online. Extensive research and analysis was performed on various designs, from small desktop robots, to ground based vehicles, to more structured robots which come closer to interacting at eye height.

Tippy the Telepresence robot by researchers at the University of British Columbia is a desktop design which interfaces with an iPhone (Wang, Tsao, Fels, & Chan, 2011). To control the robot, direction commands are sent as an overlaid image on the video stream and understood by light sensors mounted to the screen. Sparky and Sparky Jr are ground based robots documented on instructables.com (Gomi Style Crew, 2012). For the application of the nomadic worker however, it is essential that the design be ground based and operate at eye level. Two designs were identified which best met these criteria. One robot documented in an IEEE spectrum magazine which had a very robust custom built drive train (Schneider, 2010), and the low-cost video chat robot from Lee (Lee J. , 2011). Given the simplicity of Lee's system; an iRobot create, a frame, and a netbook, it stood out as the easiest platform to work with and came in at $400; an impressive jump from Ivan's $5000 robot from 2007 (Gaylord, 2008). Lee also makes software available on his website making his project a great starting point.

3.3 Robot Design

3.3.1 SPECIFICATIONS

The robot represents the remote worker through mobility and video presence. The robot must allow an individual to connect to it from a remote location, receive and transmit live video and audio, and move through a space.

Nomadic workers should be able to control this movement with a standard laptop. It is assumed that this includes an integrated webcam and microphone. This means that ideally no additional hardware should be needed to remotely control the robot. Su & Mark refer to this design criteria as 'Easing the assemblage of actants' in reference to the need to design for the ergonomic transport and easy integration of different devices. (Su & Mark, 2008)

The remaining specifications relate to the limitations of this project. Firstly, that the robot should be low-cost, and secondly that it can be built and controlled with entry-level skills in software and hardware development. Skills will be developed on an as needed basis to complete the project. This means that the design will proceed along the path of least resistance; learning from and using open-source projects as appropriate in order to achieve functionality relatively quickly.

3.3.2 STRATEGY

The project was executed in two stages. The first stage developed the prototype to a level where a video conversation could be held whilst also remotely moving the robot. The robot was demoed at this stage. Iterative improvements were made to the software, the stability and the functionality. Additional functions were added in a modular fashion, and the software developed to accommodate this. Specifically, a feature was developed using computer vision, whereby the remote worker could also control the camera angle by simply turning their head. This presence of gaze direction helps the remote worker navigate their surroundings, but also communicates their focus of attention to those interacting with them.

3.3.3 FIRST PROTOTYPE

The first prototype was built as per Lee's instructions and running his software made available via his blog (Lee J. , 2011). The robot consisted of an iRobot Create connected via a serial cable to a netbook mounted atop a podium set at eye level. The robot (server) could be controlled remotely by a laptop (client), running Lee's software written in C#. Movement was controlled by the 'wasd' keys and the software also allowed control from a virtual joystick. The laptop sends the drive instructions to the netbook wirelessly over the internet where the commands are sent via the serial cable to the iCreate in order to drive the robot.

The netbook was mounted on a podium made of gatorboard foam at an approximate height of 30" from the ground, allowing good face to face video conversations to be held at sitting height. The single podium design (seen in Figure 6) had a small footprint making it easy to maneuver, but the center of gravity was very high. This made it tip precariously when accelerating and decelerating. To increase the stability, weight was added at the base of the podium (seen in Figure 6 as two large water bottles), and the robot could be driven at 150-200rpm without tipping. Water bottles provided a quick solution but stability could be much improved. Further consideration was given to the center of gravity in the design of the third prototype.

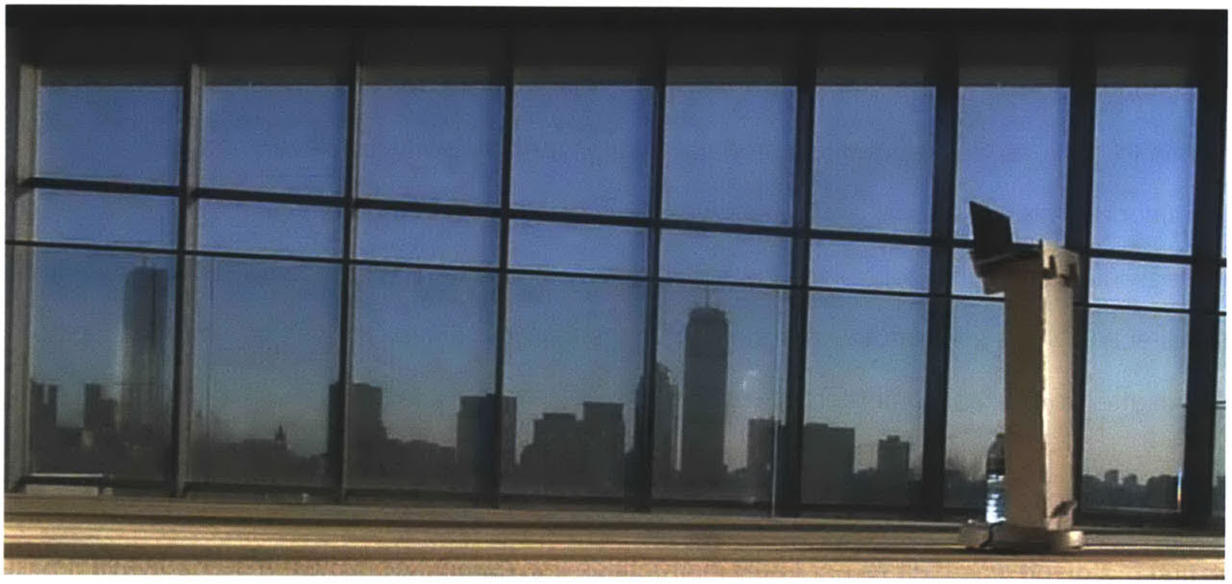


Figure 6: First Prototype Exploring MIT Media Lab, water bottles added for extra weight

Through using the robot, it was clear, that the key area for improvement was the control software which was frequently causing blue screen errors. A successful demonstration occurred for the sponsors at the final presentation of MAS.551J, but this was unprecedented. Both the remote laptop running the client software and the robot's netbook running the server software were prone to crashing and the cause was unclear. This was frustrating and could be dangerous. If the error occurred whilst the robot was moving, the robot would keep moving in that direction, requiring someone to chase after it and turn off the iRobot Create in order to stop it. In a re-design, it was important that if the system failed it should fail safely and elegantly. This was achieved in the second prototype. Lee's software was written in C# and makes use of the iRobot Create library. Whilst it was unclear what the main issue was that was causing the frequent crashes, Lee states himself that "The UPnP port forwarding is far from perfect and is not well tested at all. If it works for you, consider yourself lucky." In order to improve the software, new code was written. It was believed that writing new code from scratch would allow for easier debugging and increase the understanding of how the system worked.

3.3.4 SECOND PROTOTYPE

The second prototype was physically the same (Figure 6), but the new software eliminated the disruption from blue screen errors, and improved the signal drop failure mode. This time, when the signal dropped, the robot failed safely. Now when the signal dropped, the robot would freeze up rather than continue moving off in some direction. Blue screen errors were a huge problem as they require the system to be completely restarted, meaning that it takes several minutes to get the system back up and running again. The new software constantly checks the signal connection, and alerts the person interacting with the robot when it has dropped. This helps to reduce confusion and frustration, as that person can confidently reset the software and have the robot up and running again in around 30 seconds.

The new software was written in Python. Functionally, it worked in a similar way to Lee's program, but was an improvement in that it did not lead to any blue screen errors. The software was now reliable enough for longer tests to occur and the robot was let loose to explore the MIT campus. The robot was controlled by someone located ¼ mile away in a separate building using the 'wasd' keys and their laptop webcam. The robot completed two test runs around the Media Lab travelling 350ft and 480ft respectively before problems arose. In both cases, this was due to the connection being lost; something the software alerted us to. Whilst the robot was moving through the environment, the signal was prone to drop from time to time. The software tried to automatically reconnect but this was not always possible without resetting the system. When the signal did drop, the Skype call would run into difficulty before the drive control was lost. This gave the person interacting with the robot an indication that there was a signal problem. They could then confirm this by checking the connection status provided by the control software, and were able to quickly intervene by restarting the program and re-establishing the connection. Alternatively if they were unable to intervene the signal drop did not cause the robot to careen off in one direction, rather, it froze up and so was not a safety hazard to others.

Whilst it was disappointing that the robot was not able to maintain a connection over a long distance on the MIT campus, these distances are comparable to those observed in a study looking at alpha and beta prototypes of commercially available robots, the Anybots QB and VGO respectively. These robots were tested by traveling down a 500ft path and maintaining a conversation with someone. Out of 24 runs, only 16 runs were counted as having completed whilst the rest were terminated due to signal problems or other technical difficulties. (Tsui, Desai, Yanco, & Uhlik, 2011)

It was first proposed that the signal drop might be made worse by operating on an open campus network such as MIT. The control system acquires the IP address of the remote robot and by moving around, the IP address might be continually reset. However, when the signal dropped and the program was restarted the IP address was found to have remained the same so this was not the issue.

Maintaining signal appears to remain a significant problem for commercially available telepresence robots even today. Some robots such as Jazz from Gostai have taken steps to minimize the impact of signal drop via a clever interface. When driving the Jazz robot, if a person hovers the mouse over the video screen it becomes a 3D pointer, they can then click in that location and the robot will navigate to that point. Similarly, they can click on a person and the robot will autonomously follow them. This means that even if the signal drops, the robot will continue to move through the space where it will perhaps re-establish the connection en route.

Figure 7 shows images from testing the robot around the MIT Media Lab. The social presence of the physical robot was apparent in the way people felt compelled to approach the robot and find out where the person on the screen was controlling it from.

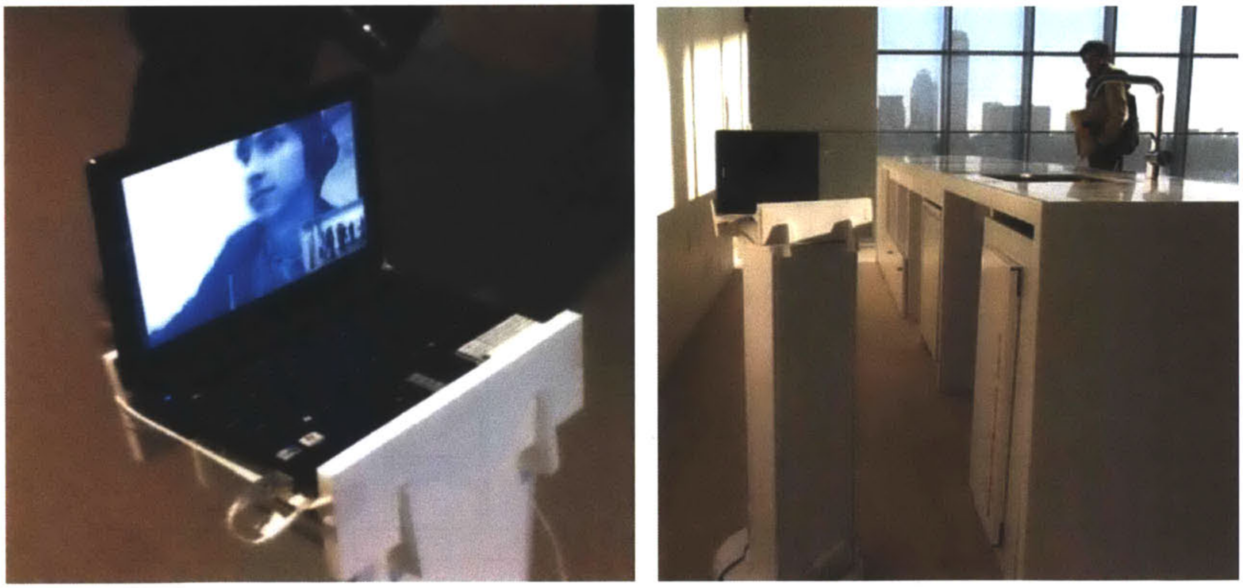


Figure 7: Second Prototype, much more reliable. Testing led to impromptu interactions with passersby.

3.3.5 THIRD PROTOTYPE

Having made some progress with software reliability, the third iteration aimed to address the stability issue. The center of gravity was determined and the structure re-designed to lower and center this. A design was proposed (Figure 8), where the podium also featured a platform which could be manually raised and lowered. This was based on feedback from the person operating prototype 2. Whilst the original height of 30" worked well when interacting with seated individuals, driving around the MIT Media Lab and approaching others who were walking around gave the driver a strange perspective of staring at the midriff of anyone he approached too closely. Trying to increase the height of the structure however with the single podium design would only worsen the stability. Therefore a new podium design was proposed which extended the wheelbase by removing the central wheel on the iRobot Create and projecting two caster wheels out to establish an effective 4 wheel base (Figure 9).

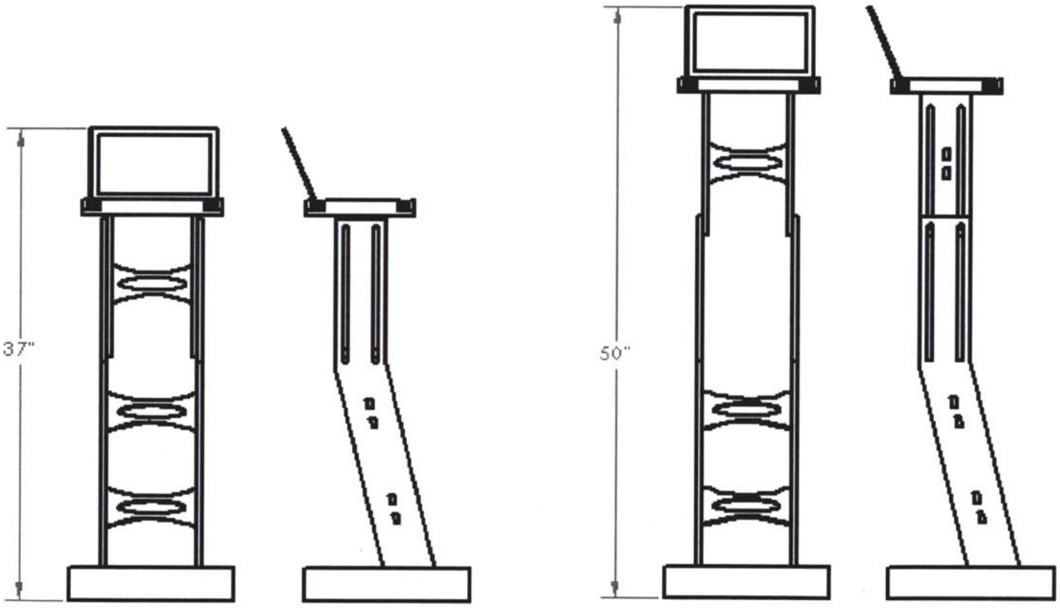


Figure 8: Height adjustment Stand Design

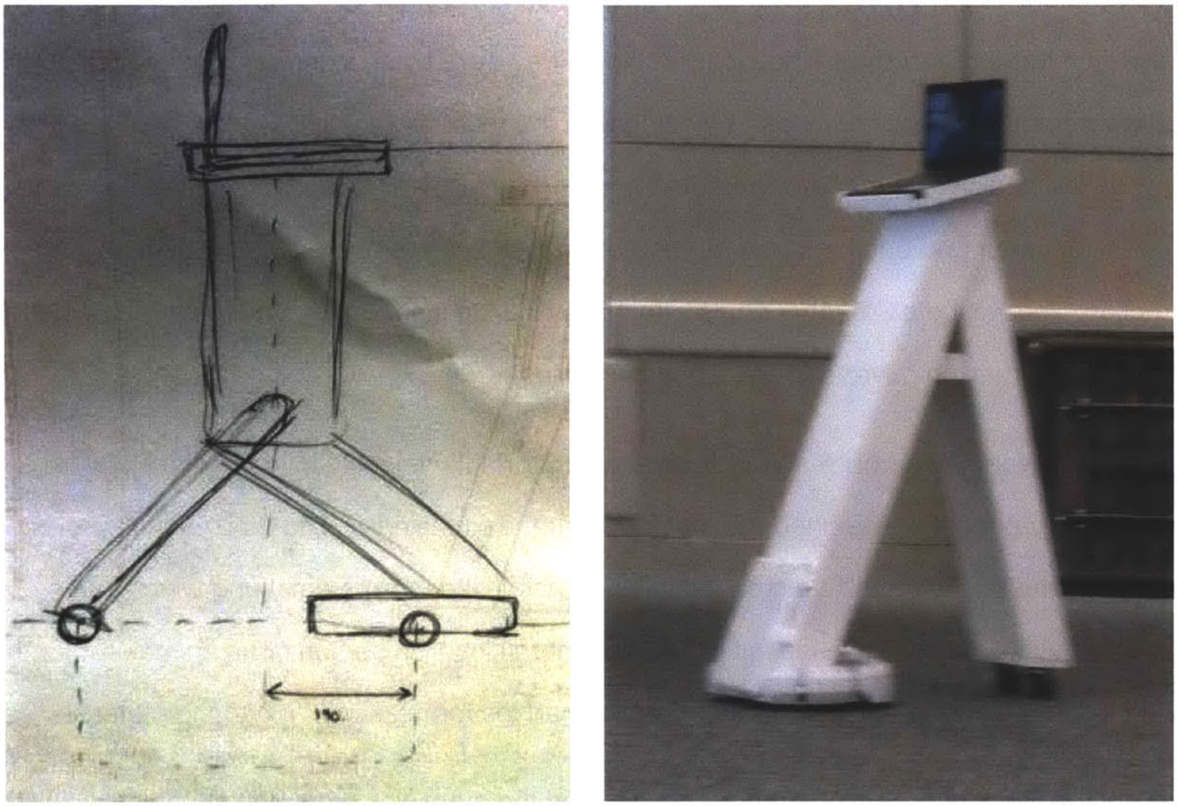


Figure 9: Stable Design with 4-Wheel Base

The third prototype (Figure 9) was significantly taller, at 45”, and more stable being easily driven at around 350rpm. This was considered too fast however for moving through crowded spaces and was reduced to 250rpm. The prototype was demonstrated at an event where it navigated a room full of standing guests where the increased height worked well. A height of 45” is competitive with robots from VGo and RoboDynamics.

As was somewhat anticipated, the larger wheel base of the third prototype was not an ideal solution since it reduced maneuverability and created blind spots. When the driver turned the robot, the back legs were liable to knock against the legs of people standing nearby. This was lessened by some sensitivity from the driver to accommodate these wider proportions. The netbook camera was also tilted towards the ground to improve navigation for the remote operator, but this interfered with their view of those they were trying to interact with. Minimal collisions were achieved by someone standing with the robot giving the driver a hand signal to stop when the back legs were turning into someone. To improve this, sensors could also be used to alert the driver or prevent them from turning when the back legs were approaching an obstruction. The Anybots QB robot deals with this by having a separate camera stream looking down to the floor so the driver can see if anything is in the way.

With this prototype, the netbook screen was at a comfortable height to interact with standing guests and that benefit outweighed the problem of some clumsiness. The robot visually seemed more robust and its greater size increased its presence in the room. Something which came out from the form was that with two legs, it resembled more closely a person or an animal, rather than the single podium from previous prototypes which had the form factor of a parking meter.

The event involved 38 prototype demonstrations and the room was crowded and noisy. Whilst the driver, again, located around ¼ mile away in a separate building, was able to hear what was being said by those standing close to the robot, the netbook speakers were insufficient to project their voice. To further improve the robot it would be beneficial to have additional speakers which could regulate the volume of the remote operator in relation to background noise. In this demonstration the robot was confined to a large room and the software performed perfectly, with no signal interruptions. The remote operator successfully controlled the robot non-stop for 1.5 hours. This suggests that the signal drops experienced with prototype 2 were related to lower Wi-Fi strength in the atrium of the building where the failures occurred.

3.3.6 STRATEGY

Now that the robot functioned with satisfactory reliability, development turned to incorporating more non-verbal communication. Based on the previously stated specifications, face tracking seemed a promising gesture input tool to work with since it requires no additional hardware; the nomadic worker could use this for robotic control with just their laptop and integrated webcam. It was decided to use face tracking to communicate gaze direction through the robot. In order to achieve this, the entire netbook on the robot would need to be rotated, or an additional webcam could be used and rotated separately. By making the robot's gaze direction dynamic, it was believed this would help communicate the remote worker's current focus of attention. Were they focused on the person in front of them, or more interested in what was going on around them? For example, in Figure 7, the robot operator is visible but looking to the left because he is working on another computer. He is not paying attention to the people around him in the remote environment. It would be useful to communicate this distraction. As a second use case, in prototype 3 it was clear that it would have been useful for the camera to be able to tilt up and down, allowing the remote operator to focus on the face of the person they were talking to, or look to the ground to see if there was an obstruction in front of them.

Face tracking was pursued to provide intuitive control of the camera to manipulate the robot's webcam and thus their point of view. It was hoped this would indicate to those interacting with the robot whether or not they had the remote operator's attention. Finally, it was believed that this additional functionality would make the robot more useful as a form of presence when stationary. A study exploring use cases for telepresence robots (Tsui, Desai, Yanco, & Uhlik, 2011), tested two commercially available products at Google offices in California. They generally found that the robots were useful for moving hallway conversations, but when used in conference meeting rooms, participants often reverted back to using traditional video conferencing technology due to the perception of increased reliability. They felt that once the remote person had driven in and "parked" up by the desk, the robot had no benefit over traditional video conferencing. It is proposed then, that gaze control could add value in this conference room scenario. With this feature, even when the robot is stationary in a parked position, the camera would swivel around, indicating who the remote operator is focused on and listening to, or who they are addressing with a comment.

3.3.7 FOURTH PROTOTYPE

In the board room scenario, it would be useful if gaze was communicated by rotating the entire netbook, so that the remote person could change the camera's point of view to look at someone beside them, and the person they are looking at could look back at them face to face. For ease, it was decided to demonstrate this interaction by only changing the position of a webcam attached to the netbook. The netbook would remain stationary, facing forward, but the webcam would rotate. This allowed rapid prototyping since low power servo motors and standard servo brackets could be used. This feature was developed separately to the functioning robot described in the Third Prototype.

The hardware consisted of a webcam mounted to a servo via a bracket (Figure 10). The bracket could then be attached to the lid of the robot's netbook so that it was positioned centrally above the screen. This already improved the experience since the separate webcam provided better video quality than the notebook's integrated webcam used in previous iterations. The servo motor was controlled by an Arduino board attached to the netbook and required its own power source which was provided by four AA batteries. This setup required that the netbook had at least two USB ports to support the Arduino and the additional webcam.

For the software, face tracking was achieved using FaceAPI (Figure 11); open-source software available under a research license (Seeing Machines, 2012). Using their libraries, a face tracking application was written in C++. The user calibrated their head position upon launching the application and then position data was sent to the Arduino, determining whether they moved significantly to the left or right. If the head position was above a certain threshold (used to determine if it was a deliberate head movement), the camera would move to a pre-determined position of + or – 45 degrees. This initial prototype worked with extremes of location, with the camera turning to the right, facing straight on, or to the left.

The setup was also rotated so that head movement in the vertical direction caused the camera to tilt up and down. The bracket used could accommodate two servo motors, and thus, a full pan/tilt mechanism was possible. However, in moving forward it was decided to focus on panning from left to right, to assist with the conference room use case described previously.

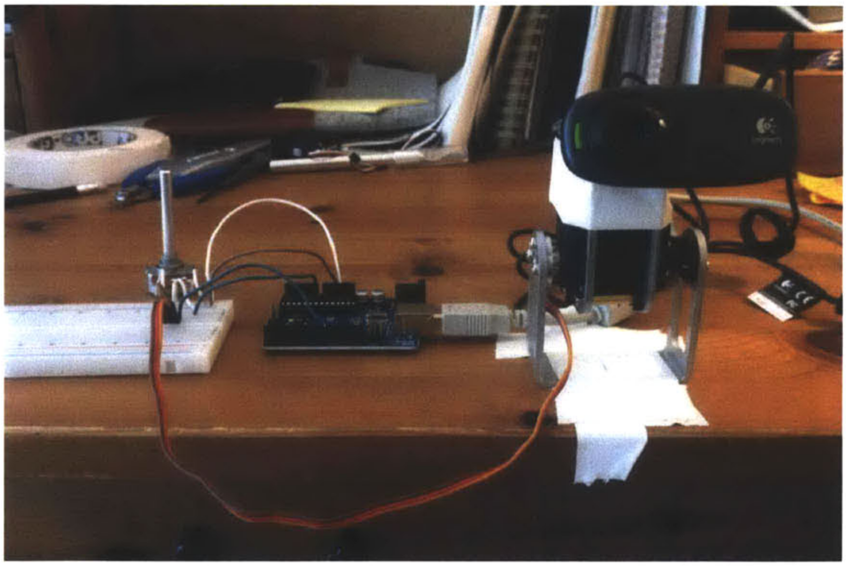


Figure 10: Arduino Controlled Webcam Tilt Mechanism

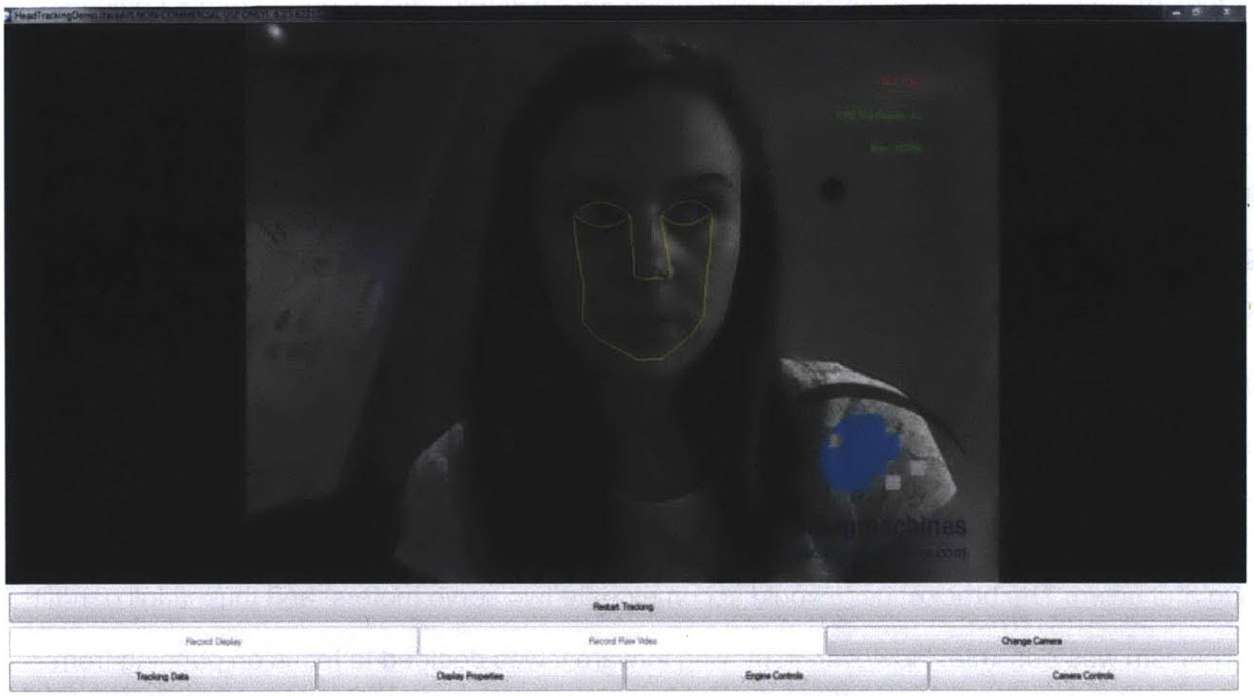


Figure 11: Screenshot from FaceAPI Demo

3.3.8 FIFTH PROTOTYPE

In the final iteration the head control was made more sensitive. The relative head position data was mapped directly to the servo position to provide tighter control of camera position. The previous prototype used a continuous servo motor, however direct mapping in this fashion was easier when using a traditional position servo. Smoothing was used to reduce the noise in the signal by taking the average of the position data over 8 readings. Again, a threshold was applied so that only large, more deliberate gestures caused the camera to rotate.

It was not possible to run the face tracking application whilst simultaneously making a Skype call since both applications need to access the webcam. This feature was therefore developed separately and never fully integrated with the robot from Prototype 3. It is apparent that there are ways to split the webcam signal for use with multiple applications but this could not be implemented in time. In further work, by splitting the webcam signal it might be possible to run our Python drive control program and the C++ head tracking control program simultaneously. It might also be beneficial to combine these into one program. This would allow us to test if head tracking was in fact an intuitive way to control the robot's camera angle, and determine how easy it is to drive the robot when the point of view can change. For example, if the robot is looking to the left but you drive forward the robot will not move 'straight ahead'.

3.5 Evaluation

3.5.1 ASSESSMENT CRITERIA

Testing and assessment of the prototype occurred frequently and informally. It was used to inform the development of the design and is responsible for the five iterations discussed previously. Signal drop remained the predominant failure mode when testing the prototype on campus, and so when tests were exploratory in nature the success was determined by how far the robot travelled before issues arose. For demonstrations confined to a room, there were no system failures, and success was determined based on observations relating to the quality of the interactions which occurred. The approachability of the robot was important, how did it engage with others and what was the quality of that interaction; generally this was judged by observation. The intuitiveness of the control was also important; this was gauged anecdotally since typically the person controlling the robot was in a remote location. Finally, when issues did arise, frustration was an important assessment criterion too. Presence requires the remote operator to truly become immersed in the environment they are projecting into, and delays or other connection issues will lead to a poor experience of presence for the operator.

3.5.2 SOCIAL INTERACTION

The main hypothesis presented was that increasing non-verbal communication cues leads to less more social interaction between an individual and a remote worker. This was apparent in the way people interacted with the robot. During the first demonstration of prototype 1, the robot was controlled by someone interacting with a group of colleagues they had met before. Due to the unreliable software the operator simply stepped outside of the room to demonstrate the remote control. As the robot approached individuals they would wave to the robot, bring others over to the robot, and took pictures with the robot. Whilst some of these actions relate to the novelty and excitement of interacting with such a system for the first time, it is rare to think of two colleagues waving to one another during a video conference. In this instance it was used as a friendly way to get the robot's attention and start a conversation.

With prototype 2, the software was improved to the extent that testing occurred with a remote operator driving the robot from another building and an observer exploring the campus environment alongside the robot. This was necessary to observe the interaction, intervene if the signal dropped, and also to prevent theft of the hardware. During these tests, people who did not know either the observer or the operator approached the robot to ask who they were and where they were controlling the robot from. Interactions with strangers are not currently supported with current videoconferencing technology. In another test run, the robot happened upon a social event where students were singing karaoke. Again, people came up to the robot, made jokes as they offered the robot drinks, and even asked the operator if they would sing karaoke. (The operator politely declined).

With prototype 3, the robot was tested in a more formal scenario during a poster session at an MIT competition. Again, the majority of conversations were initiated by those approaching the robot and one person who came to talk about the project towards the end of the session commented that they had seen the robot earlier and wanted to have a word. This remark speaks to the point that the robot did provide a significant physical presence in the room. Even those who were not interacting with the robot directly were aware that there was someone, located in another place, available to talk to should they wish. Whilst this same person may have had Skype running in the background on their phone, they were less aware of that in comparison to the physical presence of the moving robot, which they kept seeing in their peripheral vision, interacting with others.

To conclude these observational insights it is believed that:

1) The robot elicited more non-verbal interaction from others.
2) People were more comfortable to initiate a conversation with a stranger.
3) People were more frequently aware of the remote operator's presence and availability than they would be with other remote communication systems.

4. CONCLUSION

4.1 Future Work

The robot is now available to be used as a low-cost telepresence robot. It presents a fantastic platform for further work into increasing social presence through non-verbal social cues.

4.1.1 FULL INTEGRATION

The current gaze control feature should be integrated with the drive control after developing the camera splitting application to allow Skype and the face tracking application to access the webcam simultaneously. User testing can then determine how easy it is to control the camera point of view and drive the robot. It is anticipated that the head control will need to be refined, as it might be confusing if the camera is looking to the right whilst the robot is being driven forward, potentially into an unseen obstacle. To avoid this it would be possible to only implement gaze control whilst the robot is stationary, or have the camera revert to a forward position whilst the robot is in motion.

Once an intuitive control system was established it would be interesting to perform larger user trials in order to assess if gaze translation actually benefits those involved in the interaction. It is believed that the robot gaze feature will increase the presence of the remote worker whilst the robot is positioned stationary during a board room meeting, and provide a better interaction whilst the robot is in a conversation with a group of people.

4.1.2 RESPONSIVE AUDIO

Whilst testing prototype 3 in a noisy environment, it was apparent the integrated netbook speakers were not sufficient for the remote operator to be heard even at full volume. Additional speakers should be used to increase the volume further and a control system to adjust the volume to an appropriate level based on the background noise. This could be overridden by the remote operator who might want the ability to shout, or whisper their message.

4.1.3 USER TRIALS

Collaboration

The 'Afterthoughts' scenario in Figure 2, depicts the robot transitioning from a formal meeting to an informal corridor discussion. The robot was only ever tested in informal scenarios and so this was never well tested. In order to facilitate the robot as a tool for formal collaboration, the control software has been made available on the internal server such that anyone in the MIT Mobile Experience Lab will be able to use the robot in meetings with a remote participant. The qualitative feedback gained from these interactions will further inform the robot's development.

Non-Verbal cues

It would be beneficial to understand which non-verbal cues most significantly improve the feeling of presence and in what scenarios. To do this, a benchmark task might be performed such as requiring one individual and a remote operator using the robot to come to a decision. This could be done using Skype, then using the stationary robot, then allowing the robot to move, and then using the gaze feature also. The robot was intended as a platform for non-verbal cues to be tested so it is anticipated that other features could be developed and experimented with. For example, there is room for a gestural arm to be attached to the main supporting structure, or to allow for customization of appearance.

4.2 Conclusion

This thesis documents the design and development of a low-cost telepresence robot for the workplace, which provides a more expressive and engaging alternative to Skype and other teleconferencing solutions. Over the internet using keyboard commands and webcam face-tracking, the remote worker can control the robot's movement and visual focus in the workplace. They can then participate in previously inaccessible types of events: ad hoc informal meetings around the coffee maker, or physically checking in at a co-worker's desk.

An overview of prior work in HCI relating to telepresence, social robots and computer supported collaborative work was presented. First, due to the many workplace situations in which remote interaction lags behind co-located interaction, a scenario-based and human-centered design approach was used to analyze the issues. Furthermore, it became clear that while commercial telepresence systems have been increasingly technology-driven and developments in display technology have led to more lifelike 3D visuals, many social issues remain unresolved.

Social presence was then more clearly defined to include subtle communication channels which are generally overlooked, but are highly important in improving the perceived presence of a remote worker. As such, this thesis' telepresence robot became a test platform to experiment with both movement and multiple non-verbal communication cues such as gaze and physical gestures.

An application to communicate gaze direction was developed, where slight head rotations to the right or left caused the robot's webcam to move in tandem. Face tracking was a viable input method for remote telepresence operation since it captures sub-conscious non-verbal cues and does not place a cognitive burden on the operator. It was also appropriate for nomadic workers, since it requires no additional hardware. For those in the work place, this application further increased social presence since it communicated in more detail where the remote user was looking, and also emphasized head gestures such as the remote user nodding or shaking of the head.

Insights from the design, development and testing of the robot are presented below.

4.2.1 INSIGHTS

The ability to control the robot's movement had clear implications for both the quality of the remote user's interaction, and the presence of the remote user. Through qualitative tests, two of the three initial scenarios depicted in Figure 2 were examined. Firstly, the networking scenario was encountered – that simply by being present in a situation, you might meet influential people who become valuable additions to your professional network. This occurred when testing prototype 2 in the atrium of the MIT Media Lab, where another student interested in the field of telepresence approached the robot, and suggested interesting precedents which should be researched. Secondly, the robot happened upon the socializing scenario – a karaoke party in the atrium. This allowed the remote operator to have a fun break from his work and joke around with people. The third scenario about afterthoughts was not fully experienced since it presents the transition from a formal meeting to an informal discussion outside. However, this did occur in a sense when, after testing the robot on the general public in the atrium space, the observer and operator seamlessly returned together to a private office where they were able to debrief and discuss the experiment.

In practical terms of the robot's design, Wi-Fi signal drops and reliability are a real and constant issue. During testing however, it did not prove an issue when the robot was operated in a limited space with good signal. This is interesting since commercial robots are being marketed on the fact that they can let you access labs, factories or store rooms across a company site. This could set consumers up for a disappointment. Instead it might be beneficial to focus on making the robots more expressive for stationary interaction, allowing the remote worker to be more involved during meetings than they could be with current videoconferencing. It is reasonable to anticipate that these signal issues will be resolved in time.

When increasing presence through non-verbal communication cues there are opportunities to automate some behaviors, for example, by automatically turning the robot to face the person who is talking or maintaining an acceptable distance to the person the robot is following. Whilst either of these features could be desirable, too much automation might detract from truly communicating the operator's personality and it might be preferable for the operator's input alone to be responsible for the robot's behavior. In this case, the focus can shift from improving the robot's behavior to making the operator's environment as realistic as possible. What does the operator see and hear in various situations? This could mean fitting the robot with stereo microphones to determine the direction of sounds in the environment, which could then be relayed to the operator through their right or left speaker so they can choose to follow a voice if they want to.

4.2.2 OTHER APPLICATIONS

Whilst this thesis opened by discussing communication in the workplace, the scenario based approach and focus on the social aspects of telepresence are widely applicable. Already, telepresence robots are used extensively in healthcare, and any way to provide a doctor with a more sociable bedside manner would further improve the patient's experience. Telepresence robots are also entering retail stores where the perceived quality of presence between salespeople and customers will directly influence sales and the service experience. Finally, telepresence robots will become more popular for personal use, especially for those in long distance relationships, or for workers to interact with their children at home. The more these robots support informal, social activity, the more they will enhance these relationships. Already there are examples where robots have been used at weddings, where the ability to dance, express and share emotions with family and friends is unparalleled by any other technologies

5. BIBLIOGRAPHY

Ackerman, E. (2011, 03 05). *Latest Geminoid is Incredibly Realistic*. Retrieved 5 24, 2012, from IEEE Spectrum: spectrum.ieee.org/automaton/robotics/humanoids/latest-geminoid-is-disturbingly-realistic

Aðalgeirsson, S. Ö. (2009). *MeBot: A Robotic Platform for Socially Embodied Telepresence.* Cambridge, MA: MIT.

Aikenhead, B. A., Daniell, R. G., & Davis, F. M. (1993). Canadarm and the Space Shuttle. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 126-132.

Alers, S., Bloembergen, D., Bügler, M., Hennes, D., & Tuyls, K. (2012). MITRO: an Augmented Mobile Telepresence Robot with Assisted Control. *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012 Demo Track)*, (p. to appear). Valencia, Spain.

Andersen, K. N., Medaglia, R., Bimpel, G., Sjølin, P., & Mikkelsen, L. S. (2010). Robots on the Move from the Production Line to the Service Sector: The Grand Challenges for Contractors, Workers, and Managers. *eChallenges*, (pp. 1-7).

Argyle, M. (1988). *Bodily Communication.* New York, NY: Methuen.

Ashton, K. (1999, July 22). *That 'Internet of Things' Thing.* Retrieved May 5, 2012, from RFID Journal: http://www.rfidjournal.com/article/view/4986

Blanche, P. -A., Bablumlan, A., Boorakaranam, R., Christenson, C., Lin, W., Gu, T., . . . Peyghambarian, N. (2010, November 4). Holographic Three-Dimensional Telepresence Using Large-Area Photorefractive Polymer. *Nature, 468*(7320), pp. 80-83.

Brave, S., & Dahley, A. (1997). inTouch: A Medium for Haptic Interpersonal Communication. *CHI '97 extended abstracts on Human Factors in Computing Systems Looking to the Future*, (pp. 363-364). New York.

Bridget, H., Graham, R., Stansall, P., White, A., Harrison, A., Bell, A., & Hutton, L. (2008). *Working Beyond Walls: The Government Workplace as an Agent of Change.* London: OGC.

Casalegno, F., & McAra McWilliam, I. (2004, Nov). Communication Dynamics in Technological Mediated Learning Environments. *International Journal of Instructional Technology and Distance Learning, 1*(11), 15-33.

Chemel, B., Mutschler, E., & Schempf, H. (1999). Cyclops: Miniature Robotic Reconnaissance System. *Proceedings of the 1999 IEEE International Conference on Robotics & Automation. 3,* pp. 2298-2302. Detroit, MI: IEEE.

Chertoff, D. (2008, August). Improving Presence Theory Through Experiential Design. *Presence, 17*(4), pp. 405-413.

Chi, C., Liao, Q., Pan, Y., & Zhao, S. (2011). Smarter Collaboration at IBM Research. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)* (pp. 159-166). New York, NY: ACM.

Clark, J. W. (1961). MOBOTRY: The New Art of Remote Handling. *IRE Transactions on Vehicular Communications, 10*(2), 12-24.

Counsell, M. (2003). *Haptic Communication for Remote Mobile and Manipulator Robot Operations in Hazardous Environments.* Manchester, UK: University of Salford.

Daft, L. R., Lengel, H. R., & Trevino, K. L. (1987, September 1). Message Equivocality, Media Selection, and Manager Performance: Implications for Information Systems. *MIS Quarterly*, pp. 355-366.

de Crescenzio, F., Miranda, G., Persiani, F., & Bombardi, T. (2009). A First Implementation of an Advanced 3D Interface to Control and Supervise UAV (Unihabited Aerial Vehicles) Missions. *Presence, 18*(3), 171-184.

de Souza e Silva, A. (2006, August). From Cyber to Hybride: Mobile Technologies as Interfaces of Hybrid Spaces. *Space and Culture, 9*(3), pp. 261-278.

Dieringer Research Group Inc. & WorldatWork. (2011). *Telework 2011 A WorldatWork Special Report.* WorldatWork.

Dillow, C. (2011, Dec 24). *Children's Hospital Boston Sends Telepresence Robots Home with Post-Op Patients.* Retrieved 5 24, 2012, from POPSCI:

http://www.popsci.com/technology/article/2011-12/childrens-hospital-boston-turns-telepresence-robots-post-op-patient-care

Dixon, M., & Ross, P. (2011). *VWork: Measuring the Benefits of Agility at Work.* Regus.

Dullemond, K., & van Gameren, B. (2011). Communico: Overhearing Conversations in a Virtual Office. *ACM 2011 Conference on Computer Supported Collaborative Work (CWCW '11)* (pp. 577-578). Hangzhou, China: ACM.

Foth, M., Jaz Hee-jeong, C., & Satchell, C. (2011). Urban Infomatics. *ACM Conference on Computer Supported Collaborative Work (CSCW 2011)* (pp. 1-8). New York, NY: ACM.

Friedman, D., Karniel, Y., & Lavie Dinur, A. (2009). Comparing Group Discussion in Virtual and Physical Environments. *Presence, 18*(4), 286-293.

Fugitt, R. B. (1975). Design and Operation of Two Remotely Manned Undersea Vehicles. *OCEAN 75 Conference*, (pp. 870-876).

Gabbard, J., Hix, D., & Swan, J. E. (1999). User-centered design and evaluation of virtual environments. *IEEE Computer Graphics and Applications, 19*(6), 51-59.

Gaylord, C. (2008, May 23). *Send Your Robot to Work.* Retrieved 5 18, 2012, from The Christian Science Monitor: http://www.csmonitor.com/Innovation/Tech-Culture/2008/0523/send-your-robot-to-work

Gergle, D., & Clark, A. (2011). See What I'm Saying? Using Dyadic Mobile Eye Tracking to Study Collaborative Reference. *Proceedings of the ACM 2011 conference on Computer supported Cooperative Work (CSCW '11)*, (pp. 435-444). New York.

Glazer, E. (2012, March 10). When a Robot Becomes the Life of the Party. *The Wall Street Journal*, p. A1.

Gomi Style Crew. (2012). *Sparky.* Retrieved May 6, 2012, from Gomi Style: http://gomistyle.wordpress.com/sparky/

Harris, T. E. (2002). *Applied Organizational Communication: Principles and Pragmatics for Future Practice.* Mahwah, NJ: Lawrence Eribaum Association.

Heeter, C. (1992). Being There: the Subjective Experience of Presence. *Presence: Teleoperators and Virtual Environments, 1*(2), 262-271.

Hickson, M. L., & Stacks, D. W. (1993). *NVC Nonverbal Communication Studies and Applications.* Dubuque, IA: WM C. Borwn Communications.

Hinds, P. J., & Kiesler, S. (2002). *Distributed Work.* The MIT Press.

Hisanori, A. (2002). Present Status and Problems of Fire Fighting Robots. *Proceedings of Society of Instrument and Control Engineers Annual Conference.* Osaka, Japan.

Ipsos Public Affairs. (2011). *Microsoft 'Work Without Walls' Report: U.S. Telework Trends 2011.* Powerpoint Deck.

Kiddy, J. S., Chen, P. C., & Niemczuk, J. B. (2002). Low-Cost Inflatable Lighter-than-Air Surveillance System for Civilian Applications. *SPIE, 4708*, p. 304.

Kim, S., Johng, S., & Kim, C. H. (1999). Preventitive Maintenanceand Remote Inspection of Nuclear Power Plants using Tele-Robotics. *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems, 1*, 603-608.

Kurkland, N. B., & Bailey, D. E. (1999). Telework: The Advantagesand Challenges of Working Here, There, Anywhere, and Anytime. *PsycINFO Organizational Dynamics, 28*(2), 53-67.

Lee, J. (2011, Feb 9). *Low-Cost Video Chat Robot.* Retrieved 05 16, 2012, from Procrastineering Blog: http://procrastineering.blogspot.com/2011/02/low-cost-video-chat-robot.html

Lee, M. K. (2004). Why Presence Occurs: Evlutionary Psychology, Media Equation, and Presence. *Presence, 13*(4), 494-505.

Lesonsky, R. (2011). *Work Without Walls: Best Business Practices to Enable Remote Working.* Microsoft.

Lim, A., Ogata, T., & Okuno, H. G. (2011). Converting Emotional Voice to Robot Gesture. *11th IEEE-RAS International Conference on Humanoid Robots* (pp. 472-479). Bled, Slovenia: IEEE Xplore.

Liu, W., Stappers, P. J., Pasman, G., & Taal-Fokker, J. (2011). Supporting Generation Y Interactions. *ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)* (pp. 669-672). New York: ACM Press.

Lord, C. G. (1980). Schemas and Images as Memory Aids: Two Modes of Processing Social Information. *Journal of Personality and Social Psychology, 38*(2), 257-269.

Madan, A. (2005). *Jerk-O-Meter: Speech-Feature Analysis Provides Feedback on Your Phone Interactions*. Retrieved from MIT Media Lab: http://www.media.mit.edu/press/jerk-o-meter/

Makato Su, N., & Mark, G. (2008). Designing for Nomadic Work. *Proceedings of the 7th ACM conference on Designing interactive* (pp. 305-314). New York, NY: ACM.

Markoff, J. (2010, September 5). The Boss Is Robotic, and Rolling Up Behind You. *The New York Times*, p. A1.

Marshall, G. W., Michaels, C. E., & Mulki, J. P. (2007, March 1). Workplace Isolation: Exploring the Construct and Its Measurement. *Psychology & Marketing, 24*(3), pp. 195-224.

Minsky, M. (1980, June). Telepresence: A Manifesto. *Omni*, pp. 45-52.

MIT Personal Robotics Group. (2012, 5 24). *MIT Researchers Develop Affective Intelligent Driving Agent (AIDA).* Retrieved from MIT Personal Robotics Group: http://robotic.media.mit.edu/projects/robots/aida/overview/overview.html

Morgan, R. E. (2004). Teleworking: an Assessment of the Benefits and Challenges. *Euopean Business Review, 16*(4), 344-357.

Mulk, J., Bardhi, F., Lassk, F., & Nanavaty-Dahl, J. (2009, October 1). Set Up Remote Workers to Thrive. *MIT Sloan Managmenet Review, 51*(1), pp. 63-69.

Multa, B., Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Nonverbal Leakage in Robots: Communication of Intentions through Seemingly Unintentional Behavior. *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction (HRI '09)* (pp. 69-76). New York, NY: ACM.

Nass, C., & Brave, S. (2004). *Voice Activated: How People Are Wired for Speech and How Computers Will Speak with Us.* Cambridge, MA: MIT Press.

Nelson, P. (2003, March 25). *Homeworkers Miss Out on Training and Development.* Retrieved from Personell Today: http://www.personneltoday.com/articles/2003/03/25/18112/homeworkers-miss-out-on-training-and-development.html

Newcome, L. R. (2004). *Unmanned Aviation: A Brief History of Unmanned Aerial Vehicles.* AIAA.

Olson, M. H. (1982). New information technology and organizational culture. *MIS Quarterly, 6*(4), 71-92.

Ottensmeyer, M. P., Thompson, J. M., & Sheridan, T. B. (1996). Telerobotic Surgery: Experiments and Demonstration of Telesurgeon/Assistant Cooperation Under Different Time Delays and Tool Assignments. *SPIE 2901*, (p. 156).

Pentland, A. (2004). Social Dynamics: Signals and Behavior. *International Conference on Development and Learning (ICDL '04).* Cambridge, MA.

Philips Research. (2012). *Robotics.* Retrieved 5 24, 2012, from Philips Research: http://www.research.philips.com/technologies/robotics.html

Remland, M. S. (2000). *Nonverbal Communication in Everyday Life.* Boston: Houghton Mifflin.

Resnick, L. B., Levine, J. M., & Teasley, S. D. (1991). *Perspectives on Socially Shared Cognition.* Washington, DC: American Psychological Association.

Schneider, D. (2010, October). *A DIY Telepresence Robot.* Retrieved from IEEE Spectrum: http://spectrum.ieee.org/geek-life/hands-on/a-diy-telepresence-robot/1

Seeing Machines. (2012, May 21). *faceAPI.* Retrieved May 17, 2012, from seeingmachines: http://www.seeingmachines.com/product/faceapi/

Sirkin, D., & Ju, W. (2011). Communicating Meaning and Team Role Through Gesturing Telepresence Robots. *RSS 2011 Workshop on Human-Robot Interaction.* Los Angeles, CA: Stanford University.

Smith, F. (2012, Feb 28). Next Evolutionary Stage: the Office Vanishes. *Financial Review*, p. [Online].

Sprowitz, A. (2010, August). Roombots: Reconfigurable Robots for Adaptive Furniture. *IEEE Computational Intelligence Magazine, 5*(3), 20-32.

Su, N. M., & Mark, G. (2008). Designing for Nomadic Work. *Proceedings of the 7th ACM conference on Designing interactive* (pp. 305-314). New York, NY: ACM.

Tcherneshoff, M. (1997). Hazardous-Duty Applications for Unmanned Ground Vehicles. *Proceedings of the ANS Seventh Topical Meeting on Robotics and Remote Systems* (pp. 217-221). August, GA: ANS.

Tsui, K., Desai, M., Yanco, H., & Uhlik, C. (2011, August). Exploring Use Cases for Telepresence Robots. *Proceedings of the 6th International Conference on Human-Robot Interaction* (pp. 11-18). Lausanne, Switzerland: ACM.

Turckle, S. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other.* Basic Books.

UK Design Council. (2010). *UK Design Industry.* Retrieved 5 24, 2012, from designcouncil.org.uk: http://www.designcouncil.org.uk/Documents/Documents/Publications/Research/DesignIndustryResearch2010/DesignIndustryResearch2010_UKoverview.pdf

Wang, J., Tsao, V., Fels, S., & Chan, B. (2011). Tippy the Telepresence Robot. *Entertainment Computing (ICEC 2011). 6972*, pp. 358-361. Vancouver, BC, Canada: Lecture Notes in Computer Science.

Webbink, P. (1986). *The Power of Eyes.* New York, NY: Springer Publishing.

Willow Garage. (2010, 10 04). *Texai on Big Bang Theory*. Retrieved from Willow Garage: http://www.willowgarage.com/blog/2010/10/04/texai-big-bang-theory

Xbox. (2012). *Avatar Kinect*. Retrieved 5 24, 2012, from Xbox: http://www.xbox.com/en-US/Kinect/Avatar-Kinect

Yeung, J., & Fels, D. I. (2005). A Remote Telepresence System for High School Classrooms. *Canadian Conference on Electrical and Computer Engineering* (pp. 1465-1568). Toronto, On: IEEE Xplore.

Yim, J.-D., & Shaw, C. D. (2011). Design Considerations of Expressive Bidirectional Telepresence Robots. *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*, (pp. 781-790). Vancouver, Canada.