

Development of an OTA (Over the Air) Mobile
Learning Telepresence Platform

BY

Kyle Galvin

HBS Sc COMPUTER SCIENCE, LAKEHEAD UNIVERSITY, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MSC COMPUTER SCIENCE
IN THE DEPARTMENT OF COMPUTER SCIENCE

COPYRIGHT KYLE GALVIN, 2013
LAKEHEAD UNIVERSITY

ALL RIGHTS RESERVED. THIS THESIS MAY NOT BE REPRODUCED IN
WHOLE OR IN PART, BY PHOTOCOPY OR OTHER MEANS, WITHOUT THE
PERMISSION OF THE AUTHOR.

SUPERVISORY COMMITTEE

Development of an OTA (Over the Air) Mobile
Learning Telepresence Platform

BY

Kyle Galvin

HBSC COMPUTER SCIENCE, LAKEHEAD UNIVERSITY, 2013

SUPERVISORY COMMITTEE

JINAN FIAIDHI AND SABAH MOHAMMED

Supervisor

Co-Supervisor or Departmental Member

Departmental Member

Outside Member

Contents

1	Lightweight Telepresence Technologies	4
1.1	Emerging Mobile Technologies	5
1.1.1	Microcontrollers & customized System on a Chip (SoaC) components	6
1.1.2	Mobile phones & Cellular devices	9
1.2	Telepresence & Real-Time communications	9
1.2.1	Audio/Video compression	9
1.2.2	Cellular network bandwidth flow & optimization	9
1.2.3	Privacy	9
1.2.4	Cloud Based Assisted Technologies	9
1.3	Digital identification and modeling	10
1.3.1	Bar codes & QR codes	10
1.3.2	RFID; NFC	11
1.3.3	real-time digital modeling	11
1.3.4	Image recognition and classification	12
1.4	Identifying the Major Elements for a Lightweight Mobile Telep- resence System	13
1.5	Summary	13

Abstract

Telepresence has been used in many forms in academia university for more than a decade by now as entities that help to maintain the relationship with learners and provide them with collaborative experiences without the expense of physical travel. However, the emerging technology has shifted its focus from the large class-room telepresence equipments to be scaled- down to mobile, wireless-networked telepresence products. With this technology shift, we are required to provide the learner with the ubiquitous ability to explore core learning contents deployed over the internet as well as to enable learners to interact with many other remote physical learning environments (e.g., Web environments, instructors, colleagues) through the use of mobile devices. This project aims at exploring this research area and to come with a solution for implementing a new type of learning objects that can be used over the air for telecollaboration and telepresence suitable for mobile platforms.

Chapter 1

Lightweight Telepresence Technologies

Microprocessors have shaped the world over the last century. Reducing in size over time exponentially, we are now able to achieve things that would have been unimaginable in the past. We can squeeze more bits per volume, transport more information and crunch more data each second than ever before. With this explosion of portability and connectivity comes a renaissance of technological growth that is unfolding before our eyes.

Density of information and computation as well as the speed of communication are at the core of modern digital technology, yet focusing on these features displays a very hands-on white box approach. There is also much to be learned with respect to the interaction between digital components and their environments, which could be considered more of a black box "I/O" style description. The interface a device supplies for others to interact with is just as important as the computational and communicative abilities the device has internally to process the environment around it.

By extrapolating on current computational growth trends, we can imagine many applications in which technology will soon improve our every day lives. By studying these applications both mundane and whimsical alike, we are likely to find many exciting ideas which are attainable much more immediately than they first appeared.

Arthur C. Clarke once wrote "Any sufficiently advanced technology is indistinguishable from magic". Indeed many amazing discoveries can find roots in sci-fi and futuristic predictions which push the boundaries of our collective knowledge and explore the potential and logical conclusions of current technological progress. The most recent ideas which are moving from science fiction to science fact are telepresence and augmented reality.

To introduce these ideas, I will borrow from the definitions others have

supplied:

Telepresence systems provide a human operator with the feeling of actual presence in a remote environment, the target environment. The feeling of presence is achieved by visual and acoustic sensory information recorded from the target environment and presented to the user on an immersive display.

[AH11]

Likewise, an accurate and compelling description of augmented reality:

AR is a variation of the more known concept of Virtual Reality Technology (VR), which is often defined as the use of real-time digital computers and other special hardware and software to generate a simulation of an alternate world or environment, which is believable as real or true by the users. VR technology creates an environment in which the user feels and seems to be moving inside a computer-created virtual world in the same way people move inside natural environment; while immersed in the virtual world, the user cannot perceive the real one which still surrounds him. On the contrary, AR allows the user to see the real world, augmenting it with superimposed virtual objects. In other words, while VR replaces reality, AR supplements it, creating an environment in which real and virtual objects harmonically coexist.

[HCRM11]

1.1 Emerging Mobile Technologies

As microchip density increases, so does the mobility of computational and processing devices. While PDA and handheld gaming devices have been around for decades, the advance of cellular networks which allow for on-the-go personal telecommunications and widely dispersed access to internet services has really driven the shape and design of the current generation of mobile devices.

Smart phones and telecommunications aren't the only technology in this arena, but they are certainly the largest and most influential. Other devices to consider when discussing telepresence devices are lightweight microprocessors and system on a chip designs. These devices can allow industry and hobbyists alike to create a wide array of telepresence hardware that is capable of interacting with the environment around it on another's behalf. In

this case, we are now less bound by strict computational limits and are now merely bound by the sensors, motors, and analog/digital conversions available to read from (and interact with) the environment around us.

When we combine our new-found freedom to invent any sort of sensory device with our fully connected and always online 'internet of things', we can begin to explore and create all sorts of ideas that were inaccessible to the real-world and thus bound to the realm of fiction, futurism, and sci-fi.

1.1.1 Microcontrollers & customized System on a Chip (SoaC) components

Microcontrollers have become the de-facto platform for lightweight, mobile, and miniaturized devices. Much of the power of microcontrollers has been achieved by clever specializations and optimizations of the CPU unit. By diverging from the Desktop model where raw power takes precedence over power consumption, there are now a variety of architectures and specialized components that are well-suited to mobility. In fact there is now a wide spectrum of hardware ranging from high performance to low power consumption. With server and desktop hardware on one end of the spectrum, we have recently expanded the power efficient end with a range of miniaturized ARM general-purpose CPUs capable of supporting a general-purpose operating system and related peripheral components in an extremely small enclosure. To continue down the spectrum we depart from a traditional operating system and move towards programmable integrated circuits (PIC) and pure-hardware components which perform more specialized tasks using even less space and power.

Because of this balance between portability and power, it is important to keep in mind the practical limitations on the applications a device can support. While a computationally demanding task such as image processing, we are unlikely to get satisfactory results with a PIC controller, yet the current generation of ARM controllers which are recently emerging are just beginning to practically handle these tasks.

An interesting idea to consider is delegation of the heavy processing to a more capable device. However, this strategy simply moves the problem from the computational I/O boundaries of the device towards the networking I/O boundaries. To achieve practical results, we can use a hybrid approach by allowing the device to pre-process the data (assuming such a pre-processing can reduce the size of the raw data, preferably an order of magnitude or more) before sending the derived result over the network to a more capable device.

Moving from our internal capabilities towards our external interactivity, we are capable of hooking up a wide array of sensors, displays, and feedback devices. From motors to spectroscopy sensors to audio capture and processing (and many, many more) the possibilities are limited only by the fidelity/accuracy of the sensors around us and our own creativity.

Before we wander into fanciful and eclectic descriptions of advanced robots and artificial intelligence which could easily be the plot of the next sci-fi thriller, we should take note of the challenges and limitations of today's robotic and so-called 'smart' devices. For example, to apply the concept of tele-robotics (where the telepresence user is represented and mimicked as a robotic avatar on the opposite end) the following challenges have been identified and discussed:

In this telepresence domain, as remote manipulators become more sophisticated and the tasks they undertake become more complex. Three main sensory obstacles to effective widespread use have been identified [15]:

1. The detection of sensory signals
2. The feedback of real-time sensory information to the operator
3. The presentation of this information in a form that can be easily detected, processed by the brain as a reflex action and responded to, since an excessive need for thought would detract from performance of the primary task.

[CWKG96]

Aside from end-to-end synchronization and operator awareness, the question of usefulness and usability must also be addressed. Applicability has been formally described as follows:

The main objective of tele-robotics has been to develop methodologies for the control of robots at remote sites by human users at local sites. Tele-robots are suitable in certain situations such as:

1. The robot must operate in environments that are hazardous to human health.
2. The robot must operate at a scale that is much smaller or larger than the human size and scale.

3. The robot must operate in a location where it would be too costly for the human to be present (in terms of budget, timing requirements, and human safety).

[AWZ98]

It begins to become clear that tele-robotics is not an idea that would lend itself to any situation. The reduction in fidelity we introduce by using a proxy or avatar robot is severe enough that we have only found successful use-cases in well-defined environments such as laboratories and factories where there are little or no unknown variables. Situations where there are people in close proximity to the robots are often too dangerous for the technology to be applied, and furthermore the decrease in the awareness a user would have with their remote surroundings makes the potential for interaction very limited. This barrier to adoption can be seen echoed in the following passage:

Traditionally most robotic applications have involved the use of single static (non-mobile) manipulator platforms, with this technique being particularly suited to applications where the actual task is relatively well defined, the work volume is limited and safety considerations make even slightly unexpected motions totally unacceptable.

[CWKG96]

...the BCI collects the EEG brain activity and decodes the user's intentions, which are transferred to the robot via the Internet. The robot autonomously executes the orders using the navigation system (implemented with a combination of dynamic online grid mapping with scan matching, dynamic path planning, and obstacle avoidance) or the camera orientation system. Thus, the shared-control strategy is built by means of the mental selection of robot navigation or active visual exploration task-related orders, which can be autonomously executed by the robot.

[EAM12]

The idea of presence fidelity can be considered a continuum. We can consider verbal descriptions and printed material on the low-fidelity end of the spectrum, while actual presence would be on the highest end. Our goal is to bring virtual telepresence farther along this continuum until it is as close to actual presence as possible [AWZ98]

1.1.2 Mobile phones & Cellular devices

When discussing mobile telepresence, the recent emergence of smart phones is undeniably the most important event to occur to date. In a learning environment they are indispensable for personal organization, time management, and real-time updates. It has been stated by others that they can now allow students to "...be informed of all necessary notices, assignment deadlines and supervisor advices during their busy schedule". In fact, "Mobile education is defined ... as any service or facility that supplies a learner with general electronic information and educational content that aids in acquisition of knowledge regardless of location and time." [Hos07]

[Bin11]

1.2 Telepresence & Real-Time communications

1.2.1 Audio/Video compression

[SlZSSyX07] [LLD09] [AKK09]

1.2.2 Cellular network bandwidth flow & optimization

[Kim11] [KV04] [LYC04]

1.2.3 Privacy

[TVH08] [TW08] [LL12] [PHE02]

1.2.4 Cloud Based Assisted Technologies

With the advance of internet connectivity, it is rare that modern mobile cellular devices are offline. If we are constantly in communication by means of a global (universal) IP address, It follows that we can achieve two things which we previously could not.

First, we can synchronize our local data with the data of others in real-time. This lends itself to instant news aggregation, social media, e-mail, instant messaging, and even VoIP technologies. This is not in itself extremely surprising as these features have existed in the scope of the desktop application since the dawn of always on high-speed broadband connections, but the ability to bring this to a widely distributed array of mobile devices brings the connectivity of our society (and the speed of information travel as a result) to an all new level.

Second, we can now outsource services which are not desired or capable of running on the mobile device to another computer. While this is typically (as of yet) a cloud service provider's machine, it is reasonable to consider that over time software will develop which allows users to host their own content from a simple always-on home computer which serves as a personal hub for content including but not limited to public social media, geosecure proxy access, private home surveillance, and data storage. By using strict private/public tags, and 'group' authentication on a server's data as well as RSS-style content aggregators, it should be possible to design a decentralized 'home cloud' service which can serve many useful purposes to a mobile user in the field.

1.3 Digital identification and modeling

As our current trend of minniaturized mobile networked devices continues, the ability to stream high definition real-time media from many simple cellular devices is beginning to unfold. With higher quality cameras emerging in consumer devices as well as faster network connectivity emerging in the forms of 3G and 4G telecommunication services it is completely feasible (even today) to use two phones to achieve a long distance video conversation.

It is interesting to consider that if we can stream enough information between two points to effectively allow a user to 'see' and 'hear' what is in another location, then theoretically speaking we must have processed and moved (via the VoIP phone system) enough information about the two locations to effectively understand (to the same degree the two communicating people were capable) what is happening at each of the locations. We as humans don't think about the image processing we do on a daily basis. We take for granted the fact that, somewhere between the rods and cones of our eyes and the high-level understanding we have of our surroundings, a lot of information was processed, stored, and acted upon.

While the topic of Artificial Intelligence and Computer Vision is far from bridging this gap, impressive results have been found by re-thinking the ways we can facilitate enhanced modelling and identification techniques.

1.3.1 Bar codes & QR codes

QR codes (or two-dimensional bar codes) can reference nearly anything. Ranging in size from 25x25 to 177x177, they are most often used to redirect a user to a URL containing anything from videos to product info to social media content [ALYY11] Basic compression is done using run-length

coding, where sequences of identical values are replaced with a single instance of that value followed by the repetition count. [ALYY11]

1.3.2 RFID; NFC

[PJ09]

1.3.3 real-time digital modeling

Triangulation & Accelerometer based Orientated Positioning

If many radio-enabled devices (whether via wifi, bluetooth, gps, or other means) are within communication range, it is possible to use time-synchronized signals to triangulate the positions of the devices with respect to each other. If any of these devices were equipped with a GPS, it could communicate this information and allow neighbors to take their local position data (relative to each other) and place them globally.

If any of these devices were equipped with an accelerometer and a camera, it is theoretically possible to calculate a position and orientation vector for the camera, effectively letting the device give extremely precise descriptions of what region of space is being recorded. This information (spatial meta-data, as well as the raw audio/video data) can be combined with other similar information from the region in order to provide high-fidelity reconstructions of recorded events. This information can be further augmented by other sensory input such as spectroscopes, depth-sensors, and environment monitors.

Stitching multiple images/videos together

In cases where images are not aligned, unwanted artifacts can be produced. Color and lighting inconsistencies can also be introduced which would create an unbalanced effect. A lack of references and identifiable control points can also make it difficult to correctly position image fragments. [JT08]

[XP09]

Depth matrices/maps for 3D imaging

Depth imaging techniques have only previously existed in costly special-purpose applications. With the spread of large-scale production on Time-Of-Flight sensors (specifically, the Microsoft Kinect) the accessibility and spread of these devices has grown considerably. Time-Of-Flight technology involves measuring the delay between sending and receiving an infrared signal.

With this information, triangulation techniques can be used to measure the depth between the device and the target.

These devices have been designed with object recognition in mind and are not particularly suited for 3d scanning applications. Low resolution and a large amount of noise are certainly factors when re-purposing these technologies for scanning[CST⁺13]

1.3.4 Image recognition and classification

Image recognition is a complex problem often approached with a neural network model or similar fuzzy categorical organizer. If RGB information is augmented with depth information we can achieve much better results than if we were to rely on RGB information alone.

Vector Quantization

Uncertainty / Fuzzy Logic

Improved accuracy through domain-specific environments/contexts

telepresence surgery model - live data can be collected, then used to simulate the procedure virtually. This effectively allows us to generate training programs which are extremely accurate within the domain of the live collected data.

Imagine a doctor in front of device operates instrumentation which performs surgery on a remote patient. instrumentation includes a wide variety of I/O (controls and sensory output via microphone, video, and even tactile feedback) Assume access to highly detailed descriptions of our I/O over the duration of many operations (live experience captures) The challenge is to make a virtual model of the operating procedure in which the doctor can interact with a virtual patient in a way which is synonymous with the standard interactions they would encounter with a live patient.

The challenge is considered 'solved' when the doctor cannot differentiate between a live patient telepresence experience and a simulated patient telepresence experience I call this challenge the "Telepresence Turing Test", and it can be applied to any activity or domain in which telepresence can augment.

This challenge has a few interesting unknowns.

How 'synonymous' with live data can we realistically make the experience?

What are the most important factors we need to capture in our training data? What instrumentation can best capture those factors?

From the training data (live experience captures), how can we best create and improve upon a simulation model?[GHJS95]

1.4 Identifying the Major Elements for a Lightweight Mobile Telepresence System

1.5 Summary

Bibliography

- [AH11] A.P. Arias and U.D. Hanebeck. Motion control of a semi-mobile haptic interface for extended range telepresence. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 3053–3059, 2011.
- [AKK09] J.J. Ahmad, H.A. Khan, and S.A. Khayam. Energy efficient video compression for wireless sensor networks. In *Information Sciences and Systems, 2009. CISS 2009. 43rd Annual Conference on*, pages 629–634, 2009.
- [ALYY11] Hou A-Lin, Feng Yuan, and Geng Ying. QR code image detection using run-length coding. In *Computer Science and Network Technology (ICCSNT), 2011 International Conference on*, volume 4, pages 2130–2134, 2011.
- [AWZ98] A. Agah, R. Walker, and R. Ziemer. A mobile camera robotic system controlled via a head mounted display for telepresence. In *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on*, volume 4, pages 3526–3531 vol.4, 1998.
- [Bin11] Huang Bin. The study of Mobile Education development based on 3G technique and Cloud Computing. In *Uncertainty Reasoning and Knowledge Engineering (URKE), 2011 International Conference on*, volume 1, pages 86–89, 2011.
- [CST⁺13] Yan Cui, S. Schuon, S. Thrun, D. Stricker, and C. Theobalt. Algorithms for 3d shape scanning with a depth camera. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(5):1039–1050, 2013.
- [CWKG96] D. Caldwell, A. Wardle, O. Kocak, and M. Goodwin. Telepresence feedback and input systems for a twin armed mobile robot. *Robotics Automation Magazine, IEEE*, 3(3):29–38, 1996.

- [EAM12] C. Escolano, J.M. Antelis, and J. Minguez. A Telepresence Mobile Robot Controlled With a Noninvasive Brain #x2013;Computer Interface. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 42(3):793–804, 2012.
- [GHJS95] P.S. Green, J.W. Hill, J.F. Jensen, and A. Shah. Telepresence surgery. *Engineering in Medicine and Biology Magazine, IEEE*, 14(3):324–329, 1995.
- [HCRM11] M. Hincapie, A. Caponio, H. Rios, and E.G. Mendivil. An introduction to augmented reality with applications in aeronautical maintenance. In *Transparent Optical Networks (ICTON), 2011 13th International Conference on*, pages 1–4, 2011.
- [Hos07] W. Hosny. Power engineering mobile education technology. In *Universities Power Engineering Conference, 2007. UPEC 2007. 42nd International*, pages 971–974, 2007.
- [JT08] Jiaya Jia and Chi-Keung Tang. Image Stitching Using Structure Deformation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(4):617–631, 2008.
- [Kim11] S. Kim. Cellular network bandwidth management scheme by using nash bargaining solution. *Communications, IET*, 5(3):371–380, 2011.
- [KV04] Sungwook Kim and P.K. Varshney. An integrated adaptive bandwidth-management framework for QoS-sensitive multimedia cellular networks. *Vehicular Technology, IEEE Transactions on*, 53(3):835–846, 2004.
- [LL12] Xinxin Liu and Xiaolin Li. Privacy Preserving Techniques for Location Based Services in Mobile Networks. In *Parallel and Distributed Processing Symposium Workshops PhD Forum (IPDPSW), 2012 IEEE 26th International*, pages 2474–2477, 2012.
- [LLD09] Limin Liu, Zhen Li, and E.J. Delp. Efficient and Low-Complexity Surveillance Video Compression Using Backward-Channel Aware Wyner-Ziv Video Coding. *Circuits and Systems for Video Technology, IEEE Transactions on*, 19(4):453–465, 2009.

- [LYC04] Kam-Yiu Lam, Joe Yuen, and E. Chan. On using buffered bandwidth to support real-time mobile video playback in cellular networks. In *Multimedia Software Engineering, 2004. Proceedings. IEEE Sixth International Symposium on*, pages 466–473, 2004.
- [PHE02] Sang Yun Park, Moon Seog Han, and Young Ik Eom. An efficient authentication protocol supporting privacy in mobile computing environments. In *High Speed Networks and Multimedia Communications 5th IEEE International Conference on*, pages 332–334, 2002.
- [PJ09] R. Pathak and S. Joshi. Recent trends in RFID and a java based software framework for its integration in mobile phones. In *Internet, 2009. AH-ICI 2009. First Asian Himalayas International Conference on*, pages 1–5, 2009.
- [SlZSSyX07] Zhao Shu-long, You Zhi-Sheng, Lan Shi-yong, and Zhou Xin. An Improved Video Compression Algorithm for Lane Surveillance. In *Image and Graphics, 2007. ICIG 2007. Fourth International Conference on*, pages 224–229, 2007.
- [TVH08] Lei Tang, S. Vrbsky, and Xiaoyan Hong. Collaborated Camouflaging Mobility for Mobile Privacy. In *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pages 1–5, 2008.
- [TW08] Caimu Tang and D.O. Wu. Mobile Privacy in Wireless Networks-Revisited. *Wireless Communications, IEEE Transactions on*, 7(3):1035–1042, 2008.
- [XP09] Yingen Xiong and K. Pulli. Sequential image stitching for mobile panoramas. In *Information, Communications and Signal Processing, 2009. ICICSP 2009. 7th International Conference on*, pages 1–5, 2009.