Fact and Fiction Merge in Telepresence and Teleoperation

A Present and Future Perspective

Gordon M. Mair(✉)

Transparent Telepresence Research Group, Department of Design Manufacture
and Engineering Management, University of Strathclyde, 75 Montrose Street,
Glasgow G1 1XJ, Scotland, UK
`g.m.mair@strath.ac.uk`

Abstract. This paper examines current trends in commercially available
products and research related to telepresence and teleoperation systems. It also
compares them with some aspects of speculative fiction directly related to these
systems and their human interface. The presentation of the results in the form of
a parallel timeline highlights the research gaps that remain to be filled in order to
obtain the ideal telepresence system in which the technological mediation
becomes as transparent as in fictional representations.

Keywords: Teleoperation · Virtual reality · Augmented reality ·
Science-fiction · Human computer symbiosis

1 Introduction

Today the general population of the developed world is able to experience many
aspects of telepresence and mixed reality technology on a regular basis through cinema,
television, radio, and other media. With respect to the visual and aural senses the
experience is already of high fidelity. For example large immersive screens with sur-
round sound or head mounted displays and headphone delivered binaural audio are
available. Experiences can be further enhanced in many theme park rides with stim-
ulation of the vestibular, olfactory, and cutaneous senses. However the ability to
achieve a fully immersive multisensory experience in which the technological medi-
ation is transparent remains in the future.

2 Fictional Telepresence Concepts

The following are examples from fiction of such technologically transparent systems.
Firstly regarding simulated environments - as far back as 1950 Ray Bradbury wrote 'The
Veldt' in which there was a room providing a full multisensory simulation of an African
environment complete with lions, later the "holodeck" in the 1980s 'Star Trek -
The Next Generation' television series portrayed the same idea. Currently multi-screen
stereoscopic 'CAVE' (cave automatic virtual environment) type installations are the

closest we have to this concept but there is as yet no capability of providing high fidelity volumetric images in a natural environment, and certainly not images that could be touched as though solid. In fiction volumetric images have often been envisaged in the cinema, from the 'Forbidden Planet' in 1956, the first 'Star Wars' film in 1977, to the more recent 'Iron Man' films from 2008. True hi fidelity volumetric images are still out of reach although augmented reality can provide a substitute. A portrayal of this can be seen in a conference table situation as portrayed in the 2015 film adaptation of the graphic novel 'Kingsman'.

In 1965 Daniel F. Galouye published 'Counterfeit World' in which a simulated environment could be entered by donning a helmet that interacted with the brain. This concept was further developed in the 1980s and 90s by authors such as William Gibson in 'Neuromancer' and Neal Stephenson in 'Snow Crash''. However in the latter two novels there is no helmet involved and the virtual environment is experienced directly in the brain with no need to physically emulate the real world via screens, headphones, and the like.

Following the idea of full immersion in a simulated world even further, Greg Egan in his 1994 novel "Permutation City" uses the concept of mind uploading or copying into a simulated world in which the copied personas can live a separate life from the real world. Taking this idea to its conclusion and leaving the material universe completely is exemplified in the Iain M. Banks series of 'Culture' novels in which entire civilizations can decide to "sublime" to another plane of existence and explained probably most fully in his final science fiction novel 'The Hydrogen Sonata' published in 2012. But this is a step too far for the purposes of this paper and we will restrict ourselves to technological mediation.

With regard to telepresence and teleoperation and being able to use technology to be able to interact with the real world as though present at a distant location. This was shown in M.W. Wellman's story 'The Robot and the Lady' as long ago as 1938 in which a telepresence robot is used as a surrogate for the shy inventor. In 1942 there was Robert Heinlein's well known story 'Waldo' involving teleoperation from a geosynchronous orbit to the terrestrial surface. Later telepresence was merged with the concept of mind uploading in Frederick Pohl's 1955 novel 'The Tunnel Under the World' and much later in Venditti and Weldele's 2005 graphic novel series 'The Surrogates' made into a film in 2009. In the same year James Cameron's film 'Avatar' included a similar concept although in this case a biological surrogate was involved. In Alastair Reynolds book from 2012 'Blue Remembered Earth' we see many of these themes included such as telepresence "claybots", and mind uploading to animals in order to experience their perception of the world.

From the foregoing examples it is apparent that when a fictional virtual reality or telepresence environment is created we have the ultimate human computer interface exhibiting two common features. Firstly it is fully immersive in that the individual within the simulation is oblivious to the real world outside of the simulation. Secondly the fidelity of the experience makes the simulation seem completely real.

Considered below is the present day technology available to approach such a technologically transparent environment. Its limitations are apparent and it is hoped that implicit in the text, and from the eventual comparison between what is available and fictional ideals, research gaps can be observed.

3 Contemporary Telepresence Technology

This paper is being written in 2015 and it is acknowledged that technological capabilities and the human computer symbiosis are increasing exponentially in varied areas. However it is possible to consider the current capabilities of technology as they highlight many of the limitations with existing methods with regard to a fully immersive realistic simulation. Figure 1 shows the elements necessary for a conventional telepresence system. A convention has been adopted of naming the location of the person experiencing telepresence as the "home" site. This person is called the operator, viewer, user, or driver depending on the context. The top half of the figure includes the interfaces, i.e. controls, displays, and computer, used today for virtual reality and, if required, the telecoms interface for remote multi-user participation. The lower half indicates the sensors such as cameras and microphones that are necessary when a telepresence and teleoperation system is required. It is also extended to include telepresence robots which are currently finding application in hospitals, care for the elderly, education, and commerce; in this case human interaction is necessary at both 'home' and 'remote' sites.

3.1 The Human Computer Interface and System Components

What follows is a snapshot of what is currently available at the time of writing.

Visual Displays. With respect to visual displays at the home site the resolution possible for large screens has reached a level that is adequate for most viewers, e.g. currently Ultra HD television offers a resolution of 3840 × 2160, the 4 K cinema equivalent is slightly higher at 4096 × 2160. With regard to the future 8 K, or Super Hi-Vision, with 33 million pixels will become available. The 2012 Olympics in London has already had events captured in 8 K and the Japanese company NHK plans to use the format to record and transmit the 2020 Olympics in Tokyo. [1] The newer display technologies are also providing greater color depth and dynamic range thus approaching a perceptual equivalence of our natural non-mediated viewing experience.

Head Mounted Displays (HMDs) with 4 K resolution are not yet available for the consumer market. Normally stereoscopic HMDs when used in telepresence fashion include head tracking. This allows the remote stereo camera platform, say on a telepresence robot, to be slaved to the head movements of the HMD wearer. Currently the most widely known low cost HMD is the Oculus Rift [2] which is available to content developers and researchers and should soon be able to be purchased as a consumer product.

Volumetric images, i.e. images which appear solid to the viewer are not yet available in high resolution or full color. There are a number of techniques available, one of which utilizes a high speed rotating screen and projection to produce voxels rather than pixels. The Perspecta volumetric display whose patents are now held by Optics For Hire [3] is a good example of this method. However the physics of producing an apparently solid image in air as seen in popular films is extremely difficult and not yet achievable. Also the computing power and speed to produce the necessary
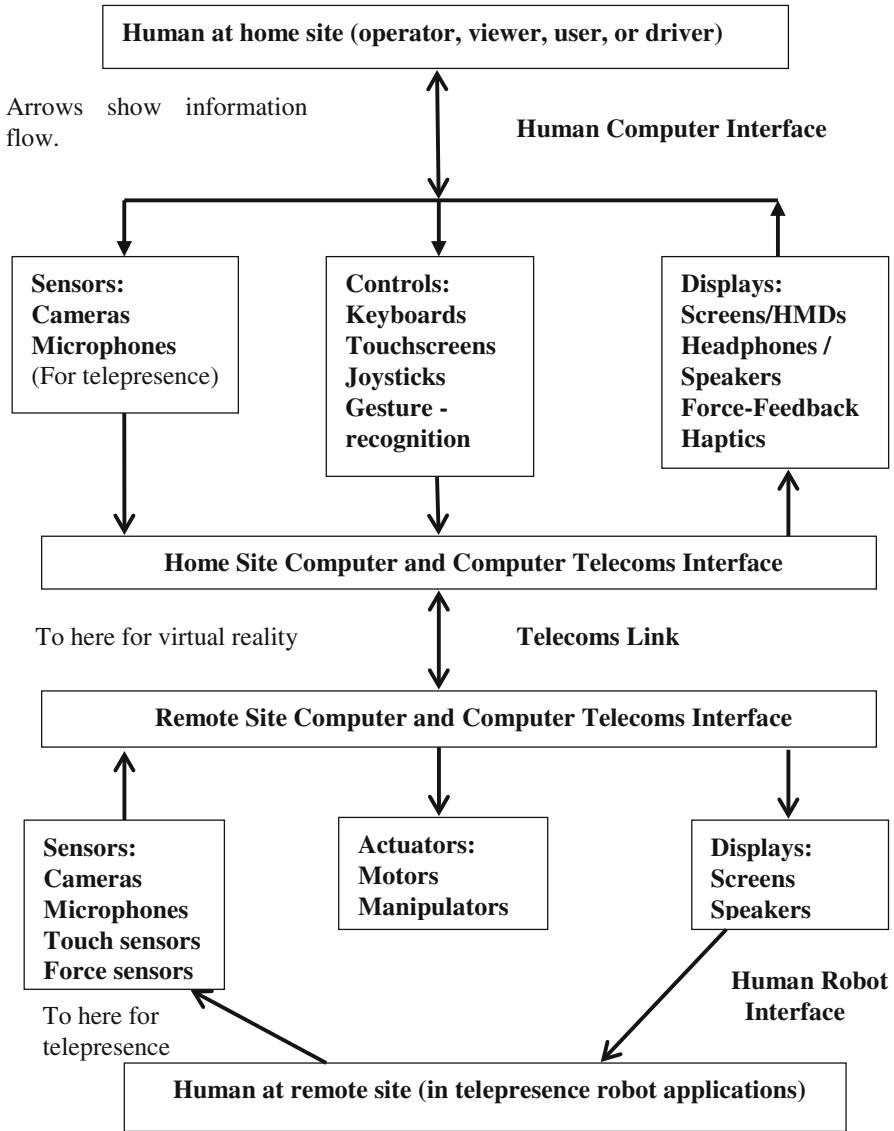
Human at home site (operator, viewer, user, or driver)

Arrows show information flow.

Human Computer Interface

Sensors:
Cameras
Microphones
(For telepresence)

Controls:
Keyboards
Touchscreens
Joysticks
Gesture -
recognition

Displays:
Screens/HMDs
Headphones /
Speakers
Force-Feedback
Haptics

Home Site Computer and Computer Telecoms Interface

To here for virtual reality   **Telecoms Link**

Remote Site Computer and Computer Telecoms Interface

Sensors:
Cameras
Microphones
Touch sensors
Force sensors

Actuators:
Motors
Manipulators

Displays:
Screens
Speakers

Human Robot
Interface

To here for
telepresence

Human at remote site (in telepresence robot applications)

Fig. 1. Typical elements of a VR, telepresence, and teleoperation system

number of voxels is extremely high. For example to produce a telepresence image of only 1000 × 1000 × 1000 pixels at a minimum of 24 frames per second will require a transmission and projection rate of 24 billion pixels per second.

One way of achieving pseudo volumetric images is by using a stereoscopic display coupled with head tracking and gesture control. This can provide the system user with the impression that a solid object is being viewed since the stereoscopic image alters its

parallax appropriately as the viewer moves their head. Polarising screens and glasses can be used as in the zSpace system [4].

Another method is to use augmented reality. Here the system user can wear see through glasses onto which can be projected images whose perspective can be tied to the position and direction of the viewer's head. Six degrees of freedom head tracking will be required here. It can be seen that a number of users wearing the glasses could observe the object from their own perspectives thus creating an apparent volumetric image that can be viewed by a number of participants simultaneously. One of the latest potential offerings of this type is that of the Microsoft HoloLens [5].

All the volumetric images comments above apply to computer generated objects, none of them are as yet capable of displaying live real world objects such as people. To allow a volumetric image of a person to be shown in a manner that is able to be correctly observed by a number of viewers standing or sitting around the image, multiple cameras would have to be employed around the remote real world object or person of interest. This gathering of three-dimensional information would be transmitted as live video and even with image compression this will require very high bandwidth transmission. Subsequent image processing in order to achieve smooth viewing at all angles will also require significant real-time computing power.

At the remote site, in the case of a telepresence robot being used, the opportunity arises for the display of an image of the remote driver. This is currently normally done with a simple two dimensional video display. However an advanced method of providing a more three dimensional impression was demonstrated by Tachi et al. [6]. Here the home site operator's face could be seen on the telepresence robot's 'head' through the use of retro-reflective material and a projection system.

An alternative method of obtaining an apparent three dimensional volumetric image of the home site operator could be through the use of augmented reality. The person at the location of the telepresence robot could wear augmented reality glasses onto which a live representation of the home site operator is projected. This would be tied to the robot through a registration point on the robot to ensure appropriate perspective for any number of people at the remote site.

Aural Displays. For telepresence an audio display of sound from a remote site needs not only hi-fidelity reproduction of the sound quality but also directional and distal information. How this is done depends on whether the sound information is being directly related to the head orientation of the system user, as in a Head Mounted Display, or is independent of the user position, as in a large screen or CAVE environment. In the latter case there are many available systems providing surround sound using multiple speakers for sound spatialisation and localization. However for the operator using an HMD with head tracking capabilities, binaural sound through headphones is necessary in order to provide accurate directional and distal information. Demonstration of binaural sound can be heard at the websites of QSound Labs [7] and 3DIO [8]. How the sound is acquired is noted in the sensor section below.

Other Displays. Somatosensory displays include those which can provide the system user with a sense of touch, e.g. the ability to feel the texture of a remote surface, and cutaneous e.g. feel the movement and temperature of air over the skin. Haptic sensing occurs when proprioception is combined with touch, this allows shapes to be discerned

by movement of the hands and joints. The ability of the system user to sense force is also important in teleoperation. Finally in some cases the ability to sense the orientation, acceleration or deceleration, of a telepresence robot or drone may be useful in some situations hence vestibular displays could be employed, however these are currently only found in flight simulators and theme park rides.

Today force feedback is used mainly in mechanical handling. Haptics are used commercially for providing vibrations when receiving data or calls, or simulating button clicks, on mobile phones. Displays of this type are also used in computer gaming and virtual reality. For example there are displays produced by the Immersion company [9] and there is the Phantom Omni haptic feedback device by Sensable which provides the illusion of touch for virtual objects [10]. Companies such as Cyberglove Systems [11] market gloves that provide both tactile and force feedback mainly for virtual reality applications. Also for flight simulation and gaming force feedback joysticks can be used such as those produced by Thrustmaster [12]. Despite these examples, and despite research in the area for more than two decades, displays to satisfy the somatosensory senses are not currently widely used in telepresence and teleoperation systems.

Controls. Conventional controls such as keyboards, mice, and joysticks are widely used commercially for both telepresence and virtual reality. However if combined with position sensors it is possible to use the previously mentioned gloves as controls as well as displays. There are also gesture recognition systems available today such as the LEAP system that can track all ten fingers with an accuracy 0.01 mm. It can be used for gaming, computer aided design, and potentially for telepresence. It has a 150 degree field of view and an update rate of 200 frames per second [13].

Remote Visual Sensing. For decades the conventional method for achieving live stereoscopic panorama images for interactive immersive telepresence has been to use an HMD with headphones slaved to a remote anthropometric 'head' containing a stereo camera system and binaural microphones as in the author's research group's early telepresence systems of almost twenty years ago [14]. Today the possibility is emerging of obtaining stereoscopic panoramic images from static camera clusters. This removes the need for electro-mechanical actuators and associated problems with power requirements, maintenance, and mechanical delay times due to inertia and resistance. An example of such a cluster is the Panocam 3D system [15] that can record and playback panoramic stereoscopic images. Due to the time required for image processing however this is not yet suitable for live telepresence.

Remote Aural Sensing. In order to provide the surround sound or binaural sound for the home site operator suitable microphones are required. For a mechatronic sensor platform capable of pan and tilt movement a pair of binaural microphones situated in an anthropomorphic relationship to a pair of stereoscopic cameras can be used. These microphones should be inserted in artificial pinnae separated by anthropometric dimensions and ideally with a mass between them similar to that of the human head in order to create a head related transfer function (HRTF).

For a static camera cluster a microphone cluster can be used such as can be found in the Sound City Project which uses four equispaced microphones with pinnae [16]. This system can record panoramic sound which can be listened to with headphones. As you pan

around an image of the location the recording was made in order to hear the sound relative to your gaze point. However just as with the panoramic stereo camera cluster the signal processing is not done in real time so the system cannot yet be used for telepresence.

Other Remote Sensing. Force and touch sensors can be mounted on remote grippers or anthropomorphic 'hands' and these are widely available in many forms [17, 18]. The gathering of information for vestibular sensing is not commercially available or as yet necessary but could be provided through accelerometers and gyroscopes.

4   Fact and Fiction Timeline Comparison

The following table shows in a parallel fashion relevant developments in technology and concepts presented in fiction. It is proposed that this is evidence of cyclical feedback in operation. Scientific insights and technological advances provide ideas for science fiction writers to extrapolate. These extrapolations then act as input to scientists and engineers to encourage further investigation and development.

5   Full Teleoperation and Telepresence Robot Systems

In conclusion there have been a number of research projects in the latter half of the 20th and the beginning of the 21st Century on creating full telepresence systems. Some of the later ones are now briefly mentioned.

Events in fact and fiction leading to VR, telepresence, and teleoperation	
Fact	Fiction
1850 to 1899	
1873 Photoconductivity of Selenium discovered by Willoughby Smith	1878 'Punch' magazine publishes an imaginative sketch of a "Telephonoscope" similar to a telepresence display
1876 Patent for the telephone filed by Alexander Graham Bell titled "An Improvement in Telegraphy"	1895 In *The Remarkable Case of Davidson's Eyes* HG Wells describes a telepresence experience
1884 Scanning Disk invented by Paul Gottleib Nipkow	1897 *The Crystal Egg* short story by HG Wells includes the concept of receiving visual images from Mars
1893 Radio control of a submersible boat by Tesla	
1900 to 1949	
1900 Constantin Perskyi coins the word "Television" in a paper presented to the International Electricity Congress at the World Fair, Paris, on August the 25th	1917 JRR Tolkien begins work on his mythopoeia part of which would become *The Lord of the Rings* written between 1937 and 1949. This contained the "palantir" stones used for viewing remote places and communication

Events in fact and fiction leading to VR, telepresence, and teleoperation

Fact	Fiction
1926 John Logie Baird demonstrates world's first televised moving images to Royal Institution in London January 26th	1932 *Brave New World* by Aldous Huxley includes "feelie" multisensory stereoscopic cinema
1928 Baird makes the first transatlantic television transmission between London and New York.	1938 *The Robot and the Lady* by M.W. Wellman published, includes full concept of immersive telepresence in a surrogate robotic body.
1939 Filing of patent application Pat No 2,344,108 by H.A. Roselund for paint spraying machine – an early 'robot'	1942 *Waldo* by Robert Heinlein published, the story incorporates teleoperation from a geosynchronous satellite
1943 An HMD patent filed titled "Stereoscopic Television Apparatus" by Henry McCollum incorporating miniature CRTs	1949 George Orwell's *1984* is published, it includes the concept of ubiquitous "Big Brother" two way television thus removing privacy in the home
1949 For the USA Atomic Energy Commission Goertz produces technical Report No. AECD-2635 "Master–slave manipulator	
1950 to 1999	
1952 "A Force-Reflecting Positional Servomechanism", R.C. Goertz and F. Bevilacqua. Nucleonics 1952; 14:43-55	1950 *The Veldt* by Aldous Huxley includes a fully immersive multisensory environment with solid simulacra.
1954 "Handyman", Ralph Mosher's master arm exoskeleton and slave arm, 12 years after Waldo published	1952 *Bridge* by James Blish, this short story includes the concept of immersive telepresence via a head mounted display
1958 to 1961 Mort Heilig build three "Sensorama Simulators"	1964 *Counterfeit World* by Daniel F. Galouye based on concept of immersion in a computer generated world
1960 Mort Heilig presents "Telesphere Mask" HMD patent design as a proposal to the RCA Research Centre	1966 "Teleoperation" word first used by E.G. Thomson
1961 "Headsight" HMD developed by Comeau and Bryan	1967 *Lord of Light* by Roger Zelazny includes electronic mind transfer
1967 Surveyor III land on Moon with extendable arm for sample gathering	1984 "Cyberspace" coined by William Gibson in *Neuromancer*
1968 "A Computer with Hands, Eyes, and Ears", J. McCarthy et al. Fall Joint Computer Conference AFIPS Proceedings, pp 329-328	1992 "Metaverse" coined by Neal Stephenson in *Snowcrash*
1968 Ivan Sutherland builds HMD	1995 *Permutation City* by Greg Egan published, includes mind transfer into virtual worlds

(*Continued*)

Events in fact and fiction leading to VR, telepresence, and teleoperation	
Fact	Fiction
1976 Viking lands on Mars	1999 *The Matrix* film introduces the concept of full immersion in a computer generated world to popular culture
1980 "Telepresence" word first used by Marvin Minsky.	
2000 to 2015	
2008 TELEsarPHONE humanoid telepresence robot with retroleflective material to provide realistic facial images of remote driver.	2009 The film *Avatar* directed by James Cameron introduces relatively high quality stereoscopic content via polarizing glasses. The story includes the concept of telepresence through mind transfer into a biological alien body.
2011 NASA's 'Robonaut 2' launched to the International Space Station (ISS) February 24th capable of being teleoperated via an HMD and force and tactile feedback gloves	In science-fiction aspects of telepresence, mixed reality, and human-computer symbiosis appear as integral parts of story lines as exemplified in the example below.
2011 ESA Eurobot demonstrated	2012 *Blue Remembered Earth* by Alastair Reynolds. In this story humans in the near future have their brains and eyes augmented by artifacts so that they can perform various acts that are logical extensions to what is shown here in the parallel the 'fact' column. For example they can "aug" without the need of special glasses in which they can access information by just looking at an object or person or actively enquiring access to particular data. They can also make themselves telepresent in "proxy bodies" called "claybots" that can take the features of whoever is driving them. This is a form of mind transfer. Also because all humans are equally augmented they can "voke" a mutually observable "aug" image which is seen as a volumetric image suitable for round table collaborative work.
2012 "Exploration Telerobotics Symposium" held at NASA Goddart Space Flight Center, May 2–3	
2012 'Curiosity' rover lands on mars on August the 6th at the Bradbury Landing site named after Ray Bradbury	
2012 Oculus Rift raises Kickstarter funding for low cost but good image quality HMD.	

<center>(*Continued*)</center>

Events in fact and fiction leading to VR, telepresence, and teleoperation	
Fact	Fiction
March 2014 Facebook purchases company for 2 billion dollars	
2013 First Google Glass augmented reality glasses prototypes April production to be ceased announced January 2015	
2014 Robonaut receives legs for mobility at the ISS	
2014 Paper from VERE project shows how a humanoid robot was controlled by thought using fMRI	
2015 Microsoft announce HoloLens augmented reality headset.	
2015 Sony's SmartEyeglass augmented reality developers kit goes on sale	

The Japanese Humanoid Robotics Project (HRP) Super Cockpit [19] included telepresence control of a robot from a cockpit incorporating an immersive display and master arms with force feedback [19]. The Robonaut program in the USA [20] has developed an anthropomorphic and anthropometric robot. Developed by NASA in collaboration with industry the robot has the capability of being controlled through telepresence as well as operating auton-omously. Finally the five year European VERE project due to finish this year has had the aim of "…dissolving the boundary between the human body and surrogate representations in immersive virtual reality and physical reality. Dissolving the boundary means that people have the illusion that their surrogate representation is their own body, and act and have thoughts that correspond to this" [21]. The project has successfully shown that a form of mind control, by using fMRI, can be used to control a humanoid robot through telepresence [22].

References

1. http://www.techradar.com/news/television/ultra-hd-everything-you-need-to-know-about-4k-tv-1048954. Accessed 18 February 2015
2. https://www.oculus.com/dk2/. Accessed 18 February 2015
3. http://www.opticsforhire.com/patents/. Accessed 18 February 2015
4. http://zspace.com/product. Accessed 18 February 2015
5. www.microsoft.com/microsoft-hololens/en-us
6. Tachi, S., et al.: TELEsarPHONE: mutual telexistence master-slave communication system based on retroreflective projection Technology. SICE J. Control Measur. Syst. Integr. **1**(5), 335–344 (2008)
7. http://www.qsound.com/demos/binaural-audio.htm. Accessed 19 February 2015
8. http://3diosound.com/examples.php. Accessed 19 February 2015
9. http://www.immersion.com/. Accessed 19 February 2015

10. http://www.dentsable.com/haptic-phantom-omni.htm. Accessed 19 February 2015
11. http://www.cyberglovesystems.com/all-products. Accessed 19 February 2015
12. http://www.thrustmaster.com/products/force-feedback-joystick. Accessed 19 February 2015
13. https://www.leapmotion.com/. Accessed 19 February 2015
14. Mair, G.: Transparent telepresence research. Industrial Robot **26**(3), 209–215 (1999). MCB University Press
15. http://www.panocam3d.com/index.html. Accessed 19 February 2015
16. http://soundcityproject.com/#/about. Accessed 19 February 2015
17. http://www.trossenrobotics.com/c/robot-force-sensor-fsr.aspx. Accessed 19 February 2015
18. http://www.meas-spec.com/product/t_product.aspx?id=2442. Accessed 19 February 2015
19. Tachi, S., et al.: Telexistence cockpit for humanoid robot control. Adv. Robot. **17**(3), 199–217 (2003)
20. http://robonaut.jsc.nasa.gov/. Accessed 19 February 2015
21. http://www.vereproject.eu/. Accessed 19 February 2015
22. Cohen, O., et al.: fMRI-based robotic embodiment: Controlling a humanoid robot by thought using real-time fMRI. Presence: Teleoperators Virtual Environ. **23**(3), 229–241 (2014)