John Edwards

Telepresence: Virtual Reality in the Real World

Imagine a teleconferencing system that's so realistic, so natural, that you feel you could reach out and shake a participant's hand, even though the individual may be sitting in a room hundreds or even thousands of miles away. That's the idea behind telepresence, an emerging teleconferencing technology that's designed to not only connect people together but to make them feel like they're all collaborating together inside the same room.

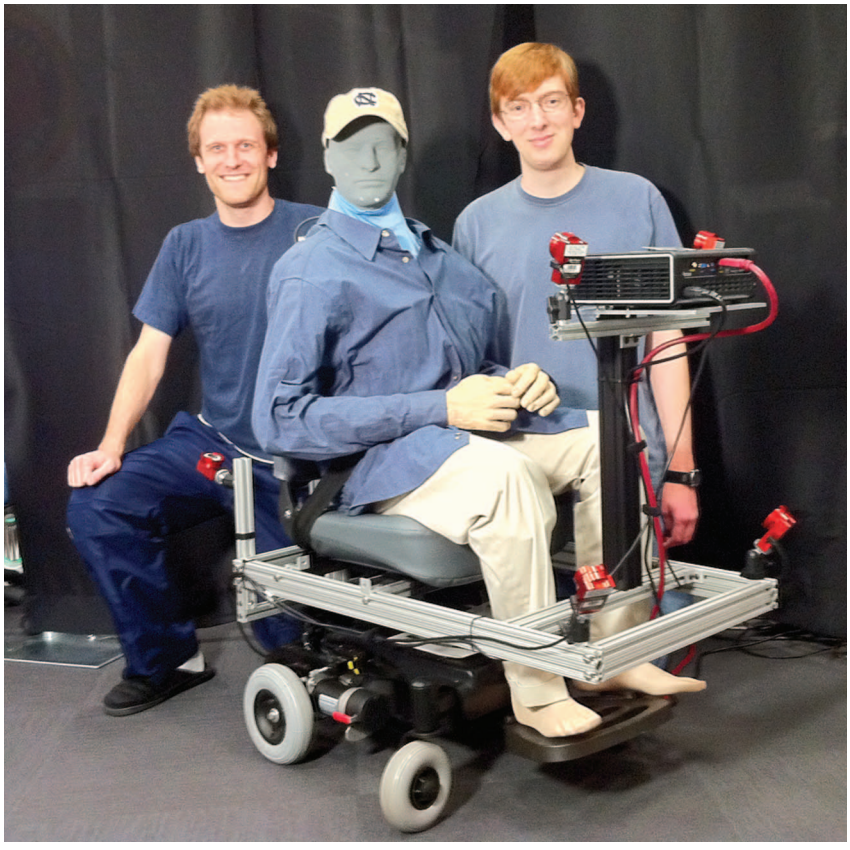
Telepresence provides an immersive meeting experience based on the highest possible levels of audio and video clarity. Ideally, participants are life-sized. Ideally, every sound, gesture, and facial expression are replicated in high-definition (HD) video and high-fidelity surround sound. The ultimate goal is to make telepresence and in-person meetings virtually indistinguishable from each other. "Telepresence has a great potential for getting people to work together collaboratively, efficiently, and naturally regardless of their physical location," observes Zhengyou Zhang, a principal researcher focusing on telepresence technologies at Microsoft Research in Redmond, Washington.

With business becoming increasingly global, and travel getting ever more expensive and cumbersome, telepresence's popularity is soaring. According to a December 2010 study by Wintergreen Research of Lexington, Massachusetts, worldwide telepresence sales are projected to reach US$6.7 billion by 2016. Yet the technology offers an important benefit that goes far beyond eliminating expensive and inconvenient travel itineraries—the ability to reach out to anyone, anywhere as if really there. "Our mind often has sparks, and those ideas need to be immediately debated and further brainstormed with colleagues and friends; otherwise, they will be just gone forever," Zhang observes.

Henry Fuchs of the University of North Carolina.

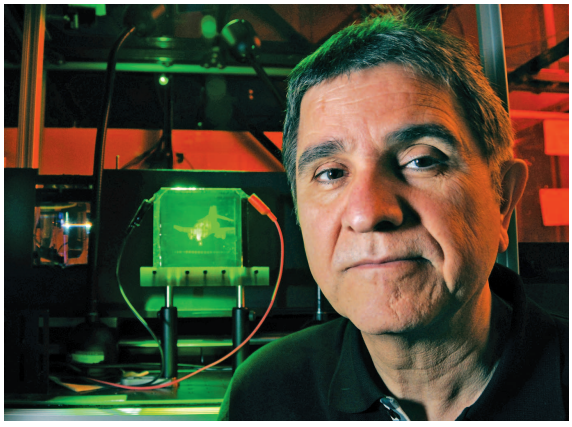
READY FOR THE REAL WORLD

Once the fodder of science fiction writers and filmmakers, telepresence is already being adopted by enterprises in sectors ranging from government to retail to manufacturing to finance to utilities. Currently available, telepresence systems are usually designed to join an enterprise's existing communications portfolio—including Web conferencing, Internet telephony, instant messaging, and social media communication tools—to help employees boost performance, increase sales, protect investments, and create rich collaborative



Ryan Schubert and Peter Lincoln (right) are grad students working on the University of North Carolina mobile avatar project.

Digital Object Identifier 10.1109/MSP.2011.941853
Date of publication: 1 November 2011



Nasser Peyghambarian, a professor of optical sciences at the University of Arizona in Tucson is shown with the hologram.

experiences with colleagues and customers. Communications integration means that participants can work face to face in a virtual meeting room while also inviting in colleagues who are limited to phone- or Web-based conferencing systems.

Telepresence systems, including cameras, displays, speakers, and other hardware and software components, are currently available from several vendors, including Cisco Systems, Tandberg, and Polycom. Prices can run anywhere from several thousand dollars to upwards of a quarter of a million dollars. Basic components in a high-end system typically include at least three HD 65-in or larger displays, several HD cameras, multiple microphones and speakers, special lighting arrays, two or more 1-Gb Ethernet ports in the room, and approximately 2–3 Mb/s of bandwidth per active screen.

While telepresence is already meeting the needs of a rapidly growing number of organizations, researchers worldwide are developing new technologies with the aim of making electronic meetings even more like "being there." Among the innovations being investigated are robot avatars that will allow people to create a virtual physical presence almost anywhere. Another important research area is three-dimensional (3-D) technology, which promises to allow colleagues, friends, and family members to pop an electronic representation of themselves into places as easily as dialing a phone. Researchers are also working on transporting realistic representations of entire meeting rooms and offices to distant locations. Zhang says telepresence's potential uses are far-reaching and limited only people's imagination. "Applications include effective team collaboration, immersive entertainment experiences, and improved customer satisfaction in marketing," he says.

ANYONE FOR ANYBOTS?

While most telepresence technologies are developed for incorporation into dedicated conference rooms, Mountain View, California-based Anybots has created a system that's designed to bring telepresence sessions directly to people at their desks or in virtually any other indoor location. The company's mobile Anybots robot is designed to function as a seeing and hearing personal avatar that can freely roll around an office, production site, sales floor, convention center, or lobby. The system is operated via a Web browser using a wireless fidelity, third generation, or fourth generation wireless connection. A video camera and speaker mounted inside the robot's head allows operators to converse with remote colleagues from a remote location while also being able to hear and see them.

Trevor Blackwell, Anybot's chief executive officer, says that ongoing advancements in video, wireless broadband, and robotics technology are making it easier and more practical for people to project their presence to remote sites. "It's now becoming possible for computers to control robot bodies in much more interesting ways than in the 1980s or 1990s," he says. "So it was time to try to build a general-purpose robot that would be useful in the office."

Blackwell states that Anybots can be used to support a variety of training, sales and customer service activities. He believes that the technology is particularly well suited for technology support services, such as helping an employee master a new software application or fix a minor computer glitch.

"Most small offices these days need a tech support person maybe one or two hours a week," he explains. He notes that small organizations often can't afford the expense of hiring a full-time support professional. Meanwhile, a part-time hire may not be available on site when needed. Blackwell says that using a telepresence robot allows a business to combine the cost and availability benefits of phone support with the interactivity attributes of a personal visit. "Being able to support people through a robot means that instead of being on the phone with someone, you can send your robot to stand behind the person, look at what they're doing on their computer, talk to them about it, look at where the wires are going, and do whatever else is necessary to troubleshoot the problem," he says.

An Anybot telepresence robot lets users project their presence to a work site.

The technology can also be put to use as an office greeter, allowing a human receptionist to remotely welcome guests without leaving the front desk vacant. "Around our office, our robot greets visitors at the door and shows them where the diet Cokes are and leads them to a seat in the conference room,"

Blackwell says. "And then it goes and gets me."

Blackwell notes that besides enabling efficient video and audio streaming, signal processing played an important role in developing the robot's physical attributes. "Our focus on signal processing has been mainly in the mobility and movement of the robot," he says. "The new technology is all about controlling the way the robot moves, making sure it doesn't crash into things, making it move smoothly and as quietly as possible, and making it seem very responsive to the remote driver even though there is some small delay over the network as they're driving it."

Such attention is vital since physical movements that are trivial for most people can be major challenges for a robot. "There are all kinds of things that can happen as you're rolling around an office—like one wheel hits a bump and then all of a sudden it's off the ground," Blackwell says. "Then, if you try to control the robot by spinning that wheel, it's going to do the wrong thing." Signal processing algorithms help the robot know what to do in specific situations.

Blackwell says that he and his team gain inspiration by observing the real world. "We spend a lot of time looking at what people do in offices and how much of that could be done remotely," he says. "We're pretty optimistic that we can replace a significant fraction of office jobs with people controlling a robot from home."

JUST LIKE BEING THERE

Robot avatar technology also figures into a telepresence system currently being developed by the BeingThere Centre, a joint project of the University of North Carolina (UNC) in Chapel Hill, North Carolina, Nanyang Technological University (NTU) in Singapore, and the Swiss Federal Institute of Technology (ETH) in Zurich. The project unites 32 leading telepresence research leaders working across three continents

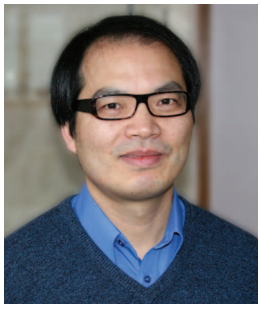
A concept developed by Gregory Welch, a UNC research professor, is to create a technology that gives people the option of sending a mobile mannequin to represent themselves at a conference table. Such a unit would act as an avatar that could freely navigate to a distant environment and take on the appearance and gestures of its far-away human host. The project's ultimate goal is creating an autonomous virtual human with memory and awareness capabilities that can take the place of its host whenever he or she is absent.

An Anybot telepresence robot allows remote workers to collaborate with on-site personnel directly and informally.

Another mobile-oriented telepresence application being developed by BeingThere Centre researchers is a portable display that can be used to bring a 3-D graphical representation of a friend, relative, or colleague to a meeting room, office, home, or other location. Like the mobile mannequin, the semitransparent display is designed to allow a user to create a virtual physical presence in a place perhaps hundreds or even thousands of miles away.

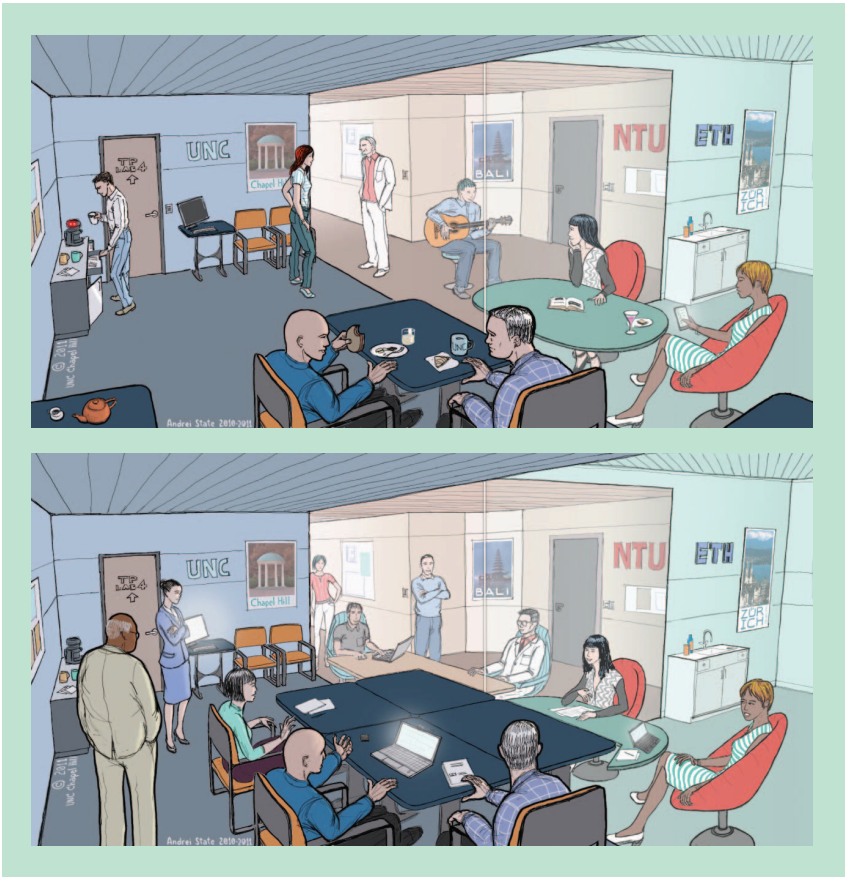
Yet another telepresence model being investigated by BeingThere Centre researchers is a conference room that connects with other similarly equipped rooms to form an entire virtual office suite. Each facility would feature wall-sized displays designed to create the illusion that the rooms are adjacent to each other and separated only by glass walls, even though they may be located in different countries. Such "glass wall" displays will provide each person in the room with a correct, personalized stereo view into the remote rooms, giving the illusion that all the of participants, both local and distant, share one common space. The idea is to allow enterprises with offices located around the world to function as if everyone were located inside the same building.

"Teleconferencing is not taking a camera image and sending it through a computer network and putting it onto a TV screen," says Henry Fuchs, a UNC computer science professor and an adjunct professor of biomedical engineering. Fuchs, who is also one of the BeingThere Centre's three codirectors (along with Prof. Markus Gross of ETH and Prof. Nadia Thalmann of NTU), feels that to provide a truly realistic immersive experience, cameras



Zhengyou Zhang with Microsoft.

John Apostolopoulos with HP.



At the BeingThere Centre, researchers are looking into the possibility of using telepresence to link geographically isolated offices into collaborative virtual suites. (Images used with permission from the BeingThere Centre.)

and other sensors need to be able to capture the 3-D geometry of everything inside a room and to project that representation as accurately as possible to the remote destination. "Not just the people, but the furniture, books, writing on the white board, waste basket, telephone, coffee mugs—everything," he says.

THE 3-D FUTURE

Three-dimensional technology has become one of the most active areas of telepresence research. Microsoft's Zhang is one of many researchers focusing on creating a 3-D virtual audio-visual environment that's both life-sized and photorealistic. "Such as that when you meet remote people in that environment, you feel as if they were sitting with you at the same table," he says. "This means you have the correct mutual gaze, accurate motion and stereoscopic parallax, and immersive spatial audio."

Zhang notes that his research is in a preliminary stage and is encountering the same road blocks that most advanced telepresence projects are bumping into: restricted Internet connection speeds and display resolution limits. "The current Internet bandwidth is not high enough, so we cannot send the full resolution of our entire 360° panoramic video," Zhang explains. "And even if we did send this in full resolution, a normal screen does not have enough resolution to display all of this."

Telepresence researchers also face a subtler challenge, one that most end users would have trouble defining, yet which gives many virtual conference participants the feeling that something just isn't "quite right." Telepresence experts call the problem gaze awareness: the ability to tell what someone is looking at by watching the direction of their eyes. Zhang says the challenge boils down to a matter of user perception. "People who are sitting at the same table see things differently: If I look at you, you see my face and your partner may see the side view, so your partner knows I am looking at you," he says. "If I turn my head and speak to your partner, then he sees my face, and you see the side view, and you know I'm not looking at you any more."

Conquering gaze awareness would go a long way toward improving telepresence realism and making sessions more like real world get-togethers. "In a virtual meeting, if you misinterpret a cue, you make the whole collaboration less effective," Zhang explains. "That's why, many times, people ask again and again for someone to clarify what's going on."

Nasser Peyghambarian, a professor of optical sciences at the University of Arizona in Tucson, believes that holography could provide the key to creating a truly immersive 3-D telepresence experience. "Holographic is closest to the way humans see their surroundings," he says. "It's also an approach that doesn't require any eyeglasses or other special eye wear, unlike when you go to see a 3-D movie or watch 3-D TV."

Peyghambarian notes that holographic presentation differs from today's movie and TV 3-D offerings in another important way. "If you go to *Avatar*, you see 3-D, but it has a very limited number of perspectives; something like two perspectives for one eye, and one for the other eye."

Holography, on the other hand, promises a vastly expanded number of perspectives, which would be handy for addressing challenges like gaze awareness. "Let's say it's a live object and there are 100 cameras taking 100 pictures from different angles," Peyghambarian says, "so there are 100 different perspectives that are coming in to provide detail." Using Peyghambarian's approach, the data generated by the camera array is

(continued on page 142)

important diagnostic information contained in medical images. It is therefore important to provide specific objective IQA measures that can help maximize the level of compression, but without affecting the diagnostic value of medical images. Moreover, many modern medical imaging devices acquire images with much higher dynamic range of intensity levels than what can be appropriately shown on standard dynamic range displays. Therefore, it is desirable to employ those IQA measures that can provide meaningful quality evaluations of the images after dynamic range compression. Furthermore, both data rate and dynamic range compression of medical images should be optimized for the IQA measures specifically designed for medical applications.

OUTLOOK

We have discussed the application aspects of modern objective IQA methods. Rather than providing an exhaustive survey of all applications, we have emphasized on the great potentials of IQA applications, provided instructive examples, and also discussed the main challenges. In the future, it is expected that the development and application sides of objective IQA measures will mutually benefit each other. On one hand, more accurate and more efficient IQA measures will certainly enhance their applicability in real-world applications. On the other hand, new challenges arising from real applications (e.g., desired mathematical properties for optimization purposes) will impact the new development of future IQA measures.

AUTHOR

Zhou Wang (zhouwang@ieee.org) is an associate professor in the Department of Electrical and Computer Engineering, University of Waterloo, Canada.

REFERENCES

[1] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Processing Mag.,* vol. 26, pp. 98–117, Jan. 2009.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing,* vol. 13, pp. 600–612, Apr. 2004.

[3] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Processing,* vol. 15, pp. 430–444, Feb. 2006.

[4] H. R. Wu and K. R. Rao, Ed., *Digital Video Image Quality and Perceptual Coding.* Boca Raton: CRC, 2005.

[5] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Processing,* vol. 15, pp. 3736–3745, Dec. 2006.

[6] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. Image Processing,* vol. 19, pp. 3116–3132, Dec. 2010.

[7] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E.-H. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Trans. Image Processing*, vol. 15, pp. 1680–1689, June 2006.

[8] M. C. Q. Farias, S. K. Mitra, M. Carli, and A. Neri, "A comparison between an objective quality measure and the mean annoyance values of watermarked videos," in *Proc. IEEE Int. Conf. Image Processing,* Rochester, MN, Sept. 2002, pp. 469–472.

[9] A. Rehman and Z. Wang, "Reduced-reference SSIM estimation," in *Proc. IEEE Int. Conf. Image Processing,* Hong Kong, China, Sept. 2010, pp. 289–292.

[10] D. Brunet, E. R. Vrscay, and Z. Wang, "A class of image metrics based on the structural similarity quality index," in *Proc. Int. Conf. Image Analysis and Recognition (Lect. Notes Comput. Sci.),* vol. 6753, Burnaby, BC, Canada, June 2011, pp. 100–110.

[11] J. Fierrez-Aguilar, Y. Chen, J. Ortega-Garcia, and A. K. Jain, "Incorporating image quality in multi-algorithm fingerprint verification," *Lect. Notes Comput. Sci.,* vol. 3832, pp. 213–220, 2005.

[12] Video Quality Experts Group Web site [Online]. Available: www.vqeg.org

[SP]

fed into a spatial light modulator, a device that can modulate light spatially in amplitude and phase. "You could get all that information and display it in 3-D and it can actually be in real time," he says.

While most existing telepresence systems are geared toward conferencing applications, Peyghambarian feels that real-time holography has the potential to drive the technology into a wider range of fields. "Benefits include 3-D social networking, 3-D remote surgery, and 3-D collaborative research," he says. "The advantage of our technology is that it can continuously read and replace data, so you can use it in an magnetic resonance imaging or computer assisted tomography scan system that would provide the information it gathered in 3-D to doctors." The technology could, for example, help surgeons performing brain surgery or other types of delicate operations. "They could use that [technology] to see the information as they do the operation," Peyghambarian says.

John Apostolopoulos, director of the Mobile and Immersive Experience Lab at Hewlett-Packard (HP) Laboratories in Palo Alto, California, believes that signal processing will be vital to overcoming many of the challenges telepresence researchers currently face. "This includes video and audio capture, noise reduction, compression, transmission over a packet network, packet-loss concealment, multichannel echo cancellation, efficient signal-processing algorithms for multicore and GPU systems and so on," he says. "I believe that advances in signal processing will continue to be central to improving telepresence in the future."

None of these improvements will come too soon for Microsoft's Zhang, who admits that he has a personal interest in seeing sophisticated telepresence systems becoming commonplace. "I have frequent phone calls with my parents and family members in China as well as my research collaborators at Microsoft Research Asia in Beijing," he says. "Telephony is a great invention, but leaves much more to be desired compared with a face-to-face meeting."

[SP]