Being There

With apologies to Peter Sellers, a well designed telepresence room can only add to the effect of being where you aren't.

Source: ProAV MAGAZINE
Publication date: 2008-11-07

By Thom Mullins, CTS

With travel costs rising and companies growing ever more global, it isn't always possible for colleagues to meet in person. Video teleconferencing (VTC) has long been an option, but even that seems so last century compared to what AV pros can enable today.

Manufacturers such as Cisco, LifeSize, Polycom, Tandberg, and others have started pushing VTC and the design of teleconferencing spaces to the limits of their capabilities in an effort to reduce or eliminate as many of the communication filters inherent to VTC as possible. They've even given this VTC on steroids a name–telepresence.

Big or small, a telepresence room should be designed to enhance the illusion that the meeting participants on-screen are actually sitting in the same room.

Technology filters–factors that often render the VTC experience underwhelming–are well known: inadequate bandwidth, inconsistent audio, poor lighting, and lack of interoperability, to name a few. But the biggest filters relate to human ergonomic and personal communication issues–issues that merely underscore the fact that the person you're meeting with really is in another room, far away.

It turns out the design of a room and the placement of technology within the room can have more impact on the quality of a virtual meeting experience than the actual technology used. Just as the right conference room design will maximize the AV systems in that room, the right telepresence room design will create a truly immersive meeting experience. Here's how.

Defining the Experience

First, what exactly does the term "telepresence" mean? "Teleconferencing" literally means "to confer (or meet) at a distance." As you think through the that definition, you can form a good mental image of what is intended. Telepresence, however, should paint an entirely different picture. How do you achieve "presence at a distance?"

Pete Nutley, director of global product marketing at Tandberg, describes telepresence this way: "It's

a video meeting experience that creates the feeling of being across the table from the remote participants."

In order to achieve this experience, all participants are seated at a pre-determined distance from their respective cameras so as to appear life-size to the people at the far-end of the network connection. High-resolution, low-latency video, coupled with high quality directional audio help create this effect. And other visual cues–all telepresence rooms painted the same wall colors, equipped with the same lighting, and outfitted with the same furniture–round out the design.

"So telepresence really is about the experience," Nutley says. "When one or more of the criteria, such as background, distance from camera, desk color, etc., is not met, the experience of being in the same room deteriorates."

Jeff Machtig, co-founder of Digital Video Enterprises (DVE)*, which designs and sells telepresence solutions, agrees. "Psychology dictates that there are certain key elements that allow for effective communication and interaction between two or more individuals. How successfully we relate to one another depends greatly upon eye-to-eye contact."

The human factors of communication, therefore, are as or more important than the technology itself. The goal is to design a facility–technology and space working together–that fools the participant into thinking they are in the same space as the individuals they are meeting with. In other words, minimize the filters and distractions and make them as transparent as possible by paying close attention to several factors.

This story was edited to correct the name of Machtig's company, **Digital Video Enterprises (DVE)**

Factor 1: Acoustics

Although we're discussing video teleconferencing, let's start with the pieces of the puzzle that usually spawn the most complaints–the acoustical environment and audio system. It's relatively easy to demonstrate why these would be an issue in telepresence applications. Take a small handheld recorder into a conference room with you and record what you hear for five minutes. In fact, do it twice–once while you're in the room and once when you are out of the room. As you listen to the recordings, you will be very aware of all the background noise.

Setting up and coordinating camera, display, and furniture so that it looks like both parties are sitting at the same table helps facilitate better communication, especially in situations where telepresence meetings could go on for hours.

It is possible for people in a source room, i.e., the telepresence room they're sitting in, to ignore much of this. It's almost impossible for listeners at the remote site to do the same. The audio network is carrying both signal and noise, which means colleagues have to work harder to pull useful information out of the link.

The areas to address here are ambient noise levels, reverberation time or decay rates, microphone and loudspeaker placement, and system bandwidth.

One of the key contributors to clearly understanding the spoken word is the level of background noise (those sources that do not contribute positively) in the communication environment. Ideally, this level should be in the range of NC-30, which is roughly equivalent to a measured level of 36 to

38 dB(A) and is what the listener hears from a very quiet room fan at low speeds.

If the noise level exceeds 40 dB(A), it can be distracting to the listener trying to concentrate or stay focused on studies. Contributors to ambient noise include traffic (automobiles and aircraft), airflow from HVAC systems, elevator machinery, hallway conversations, lighting ballasts, and footfalls from the floor above, to name just a few.

People usually compensate for high ambient noise levels by either moving closer to one another or by raising their voices. Neither of these options is really viable in the telepresence environment, so the problems have to be addressed at the source. Careful attention to the room envelope (walls, ceilings, floors, doors, and windows), HVAC system design, and type of lighting used will help mitigate these problems.

Of equal importance is the reverberation time or decay rate of the room. Most people are familiar with reverberation; it's based on a uniform distribution of acoustic energy and random directions of propagation that can usually be achieved in larger rooms. There are a number of formulae that have been developed to help calculate reverberation time. In general, reverberation time increases with room size or volume and varies inversely with the amount of absorption in the room.

However the available formulae aren't very accurate for the smaller rooms in which telepresence systems might be deployed. For that reason, we should look at the acoustic decay rate of the room (the reduction of acoustic energy measured in dB/second). This is a measure of how quickly the acoustical energy in a space decays after the room is excited.

A "fast" room decays more rapidly (120-240 dB/s), which contributes to higher speech intelligibility and a greater sense of acoustical intimacy for the listener. Most telepresence rooms are quite small, or should be, and should hold no more than 10 to 12 participants at the most. At roughly 30 square feet per person, this translates to a room of roughly 300 to 400 square feet and an overall volume between 2,400 to 3,200 cubic feet. At this size, careful attention should be paid to room shaping and reducing resonant nodes.

Factor 2: Speakers and Mics

In a well-designed room, loudspeaker placement is less critical than microphone placement, but only slightly less. The listener on one end should be able to localize on who at the other end of the teleconference is speaking. This will typically mean placing a loudspeaker above the image of the person speaking.

Some manufacturers have included this ability in their audio signal processing equipment, and can recreate a reasonable sound field using two loudspeakers (see "Advancing the Art of Positional Audio," page 34). It isn't necessary to use a lot of loudspeakers, or that they be capable of extremely high sound pressure levels. After all, the goal is to recreate speech levels at the listener's ears in a conference room. This means that measured sound pressure levels should be between 66 and 69 dB(A) at the listener's ears.

Proper microphone placement will contribute to that sense of intimacy and increased intelligibility. Ideally, each participant should have a microphone, and have it as close to their mouth as possible. This results in a higher ratio of direct sound to reverberant sound at the microphone capsule. It also means less noise from shuffled papers, pagers, or other distractions.

If cost prohibits the use of personal wireless systems, there should be no less than one wired microphone for every two meeting participants, preferably mounted 18 to 24 inches from the front

edge of the conference table. It is possible to use ceiling-mounted microphones, if proper attention has been paid to HVAC noise, vibrations, and the other ambient noise and decay rate issues. This is usually a choice of last resort, however.

All that said, carefully placed hypercardioid microphones can be placed further from meeting participants. It is possible to retain the sense of aural intimacy and also reduce noise from shuffling papers while "hiding" the microphones a bit further away. Some manufacturers have begun producing phased arrays of microphones to achieve similar results.

Factor 3: Audio Bandwidth

When designing the audio side of a telepresence room, don't overlook system bandwidth. With today's modern microphones, mixing equipment, signal processing, and amplification, it's often not even thought of, but that would be a mistake.

Remember that the audio bandwidth of the signal sent to the far side often does not exceed 7 kHz. Don't use microphones with a bandwidth equal to or less than the bandwidth of the signal link; the system will sound worse than it would if you used microphones typically found in either recording or live sound reinforcement applications.

Factor 4: Lighting

Obviously, perhaps the most important part of telepresence room design has to do with the visual experience: the cameras, display size and resolution, and lighting.

Room lighting for telepresence systems is no different than what is recommended for a typical video teleconferencing system. The Illuminating Engineering Society of North America (www.iesna.org) offers handbooks on the subject and Pro AV editorial advisor Tim Cape, CTS-D, has written on the topic before. (Visit www.proavmagazine.com and search on "lighting for videoconferencing.") In a nutshell, you're balancing comfortable lighting for local participants with proper lighting to ensure a camera, codec, and display system can reproduce an optimal experience.

Factor 5: Camera Positioning

From there, visual design is about recreating that in-person experience. People prefer to speak eye-to-eye and face-to-face with one another. Think of what it's like when someone doesn't look directly at you when they respond to a question or a request.

The face-to-face dynamic can be replicated through the proper placement of cameras and displays. Ideally, the center line of a camera's lens should be at seated ear height (+44 inches above finished floor, or AFF) or slightly above (up to 48 inches AFF). Note: In telepresence it isn't necessary to provide a camera with pan/tilt/zoom capabilities because the room and where people sit are meant to be tightly controlled, unlike in a typical VTC environment.

Lens focal length should be determined by the distance the two (or more) parties would be seated from each other in a real meeting. For instance, if the client would meet an in-person colleague across a 5-foot-wide table, the focal distance of the lens should take into account the virtual size of the person you're meeting with across the void and the number of people each camera will see.

So where is the camera placed relative to the display? As DVE's* Machtig points out, this is the determining factor in the success of an immersive telepresence environment because so much of non-verbal communications is expressed with our eyes.

Again, if we were building the ideal room, we would align the center axis of the camera with the center axis of the displayed image, so you had the illusion you were talking directly to the people at the far site. This would place the camera behind a piece of two-way glass while the display would be mounted to reflect its image onto the glass.

Most solution providers choose not to implement this model, however, so compromises are made in this vital area. Choices include placing the camera above the display (which can be unnatural for the viewer as they look down on the top of your head, especially in a shorter throw application); or placing the camera below the display, which forces the participants to look up; or placing the camera between displays.

Although alignment with both axes of the display is ideal, I would opt for placing the camera either between the displays (first choice) or above the display. But reasonable designers can disagree. The more your client is willing to spend on high-end, complete VTC solutions, the more likely you are to find a system that fits the camera in small, unobtrusive positions, like between blended, bezel-free displays, or built into the system just above the screen.

This story was edited to correct the name of Machtig's company, **Digital Video Enterprises (DVE)**

Factor 6: Display Positioning

Determining the image size you need requires you to follow the usual formulae that tie image size, target distance, and focal length together. Most imaging chips are either 1/3- or 1/4-inch CCD devices, with either 720p or 1080p resolutions.

This then drives the size of the display you want to use. Since participants are meeting across a virtual table, you want them to be able to see a portion of that table and up to a reasonable height above the table. Because seated ear height is +44 inches AFF (on average), add an allowance of 10 to 12 inches (including forehead) to ensure you don't give someone an unwanted crew cut on the monitor. This puts the top of the image at +56 inches AFF.

The displays you choose should be large enough to accompany life-size images of remote meeting participants–as many as two per display–and enough of their surroundings to underscore the in-person experience.

Considering the surface of a conference table is normally at +29 inches AFF, you'd be looking at an image height of roughly 27 inches. If you allowed for including some of the conference table (ideal for telepresence but usually only if the local and remote rooms are well controlled and include the same conference table), this would place the bottom of the image at +26 inches AFF, requiring an image size of roughly 31 inches.

At a minimum, you should choose a 54- to 60-inch diagonal display device. With this size display, you may be able to fit two people onto one display. (It seems counterintuitive, but the wider your virtual conference table, the smaller the display needs to be.) High-end systems often come with larger displays (see "Telepresence in a Box, page 32").

Factor 7: The Displays

Finally, the resolution of the display should match the resolution of the cameras being used to

capture images, or the minimum resolution of the codecs being used. In most cases, this means a 1280x720 display, because the minimum HD resolution used in most telepresence systems is 720p. At the image sizes discussed above, this resolution will actually be higher (typically 1365x765). This is the lowest common denominator for most manufacturers, and typically acceptable by most users. The larger the displays you choose to create the immersive experience, the more likely you are to need higher definition displays.

Clearly, there are many factors that go into telepresence design that are beyond room and system setup, such as bandwidth (a 1080p telepresence meeting could consume up to 18 Mbps) and frame rates. After you have a handle on the room design, these technology factors will require analysis of the client's application requirements and expectations.

Because telepresence can vary in scope, and because so many acoustical, design, human, and technology factors can affect the total experience, there are obviously compromises along the way that might result in a less than ideal telepresence experience for the viewer. We already see some of these compromises introduced by manufacturers seeking to find price points for their pre-integrated equipment that can meet customers' various needs and budgets.

But if you as an AV pro begin with the end result in mind–marrying design, technology, and physics to make the users feel like they're face to face with distant colleagues–then you've laid the proper groundwork and likely optimized the experience regardless of the scope and complexity of the system itself. Even if the client doesn't opt for a quarter-million-dollar system, they can still enjoy an in-person discussion with people on the other side of the table–or the world.

Pro AV contributing editor Thom Mullins is a senior consultant with BRC Acoustics & Technology Consulting in Seattle.

Telepresence Systems

Telepresence In A Box

If you or your client is in need of a serious, high-end telepresence system; you need it set up quickly and optimally; and you want all the design criteria thought out in advance, there are all-in-one telepresence rooms available that come with everything from the codecs to acoustically enhanced walls. Among the latest is Tandberg's Telepresence T3, which comes with a trio of 65-inch LCD displays, Tandberg's own C90 codec, and built-in 1080p cameras. It also comes with LED-lit blue glass backgrounds, a conference table with built-in touch-screen control and presentation systems, customizable walls, specialized room lighting, and carpet (yes, carpet).

Rick Snyder, president of Tandberg Americas, showed Pro AV the company's own T3 installation in the Washington, D.C., suburbs and said that Tandberg worked with design experts and social anthropologists to ensure the T3 package delivered an ideal experience.

"For instance, the blue backgrounds we found were very comforting to telepresence meeting participants," says Snyder. While the walls, which hang on a specially designed metal frame, are acoustically treated, they're not overdone. During its research, Tandberg found

some of a room's natural resonances are often preferable during long meetings. The system uses full-duplex spatial audio with super wideband stereo sound to ensure the right participants are heard clearly when they speak.

The T3 was introduced last month and is scheduled to ship in January, though Snyder says the company already has a backlog. It's not a one-size-fits-all solution, however. Snyder says Tandberg works through its integrator partners to provide guidelines for this type of room-within-a-room installation. The room itself where the T3 will be installed should be the right dimensions, with provisions for the proper electrical and HVAC systems (T3 plus lighting consumes 3.8 kW, on average). Tandberg publishes a specification document with all the requirements.
The telepresence system itself goes for $299,000. The room, including the wood paneling, glass, LED light bar, carpeting, etc., costs an additional $39,000. Snyder says customers could purchase the system separate from the room.

Photos: Brad Grimes

Tandberg's new Telepresence T3 comes with everything for an optimal room design, including the total system, plus specially designed glass walls and lights.

–Brad Grimes

Telepresence Products

Advancing the Art of Positional Audio

When Polycom began the process of updating its Vortex conferencing system, the company had two goals in mind. First was to create a teleconferencing-specific audio device that would incorporate changes recommended by contractors and integrators. Key among those was creating a seamless, scalable audio matrix across a number of boxes that would look to a control system like a single device. Achieving the goal would result in a system that would be easier to install, setup, and adjust.

The second goal was to create a box that would respond to the market as it moved toward telepresence. Achieving that goal came down to supporting what Greg Suchomel, AV applications engineer with Polycom, refers to as "positional audio." The results was the company's new SoundStructure series.

Most audio codecs utilize only a single reference channel for acoustic echo cancellation (AEC). Positional audio, which would allow more accurate alignment of the audio signal with the video displays, uses a two-channel reference for AEC. This allows the AV professional to set up a sound stage of adequate width that would support accurate positioning of people within the sound stage. And it requires an AEC algorithm that can easily respond to multiple microphones and differing signal levels in each loudspeaker.

While these changes to Vortex appear to address the issues adequately, Polycom is already receiving requests to push the technology further. Are eight channels likely? It might be possible, but not practical. The practical limit is directly related to the number of displays that can be

deployed. In most telepresence rooms, this would be four or five displays. Any more than this would result in a room that's too long and reduces effective communication from one end of the room to the other.
