Effective Design of Audio/Video Conference Rooms

Gregory A. Coudriet, The Sextant Group, Inc., Pittsburgh, Pennsylvania
Jeffrey E. Babich, Babich Acoustics, Pittsburgh, Pennsylvania

Economic pressures, changing business models, and environmental concerns are increasing the use of video conferencing, telepresence, and distance-learning technologies in academic and corporate settings. A successful return on investment for these systems rests on their adoption among prospective users. Acceptance is maximized when the conferencing system provides a realistic virtual presence between users. Acoustics and speech intelligibility are key components of this effect. This article presents an overview of different conferencing technologies, their acoustical requirements, and built-in acoustical controls such as echo-cancellation, noise reduction, and voice lift.

The ultimate intention of any conference system is to facilitate communication between remote participants. This communication can include speech, program audio, video, presentation materials, computer files, and images of physical documents. Many systems are also capable of capturing this content for future viewing. Conference systems employ a variety of technologies in vastly different settings and for vastly different purposes. The ultimate goal is to provide users with the desired level of communication while balancing convenience, cost, flexibility, and fidelity.

Types of Conferencing Systems

Prior to discussing the acoustical requirements for audio and video conferencing facilities, it is necessary to understand the various functions of these rooms and the technologies that are employed. There is a continuum between each of the systems and technologies presented here. Each is scalable from a small single-user system to multiuser applications, including auditoriums and courtrooms. In general, the cost and complexity of each system is directly related to its functionality and the number of participants. Likewise, the need for good acoustics increases along these same lines.

Telephone Conferencing (Teleconferencing). A telephone conference between two endpoints equipped with standard telephone handsets or speakerphones is the simplest form of conferencing. There are also dedicated desktop conferencing systems containing specialized microphones and built-in speakers that can provide an improvement in pickup sensitivity and far-end audio fidelity over standard handsets. If multiline handsets or desktop conferencing systems are used, more than two endpoints can participate in a conference, though overall sensitivity and fidelity of the conference may suffer as more parties are added.

To overcome this, a remote (off site) audio bridging service may be employed. The teleconference host and invited participants typically call into a toll free number and enter a conference access number to join the audio conference. Audio quality is assured by the audio bridging service through its use of sophisticated audio conferencing electronics. Teleconferencing by the use of handsets, desktop conferencing systems, and third-party bridging services comprises a system that is simple to use and does not require the users to possess specialized equipment. A standard teleconference experience may be enhanced through the use of web conferencing software, such as Smart Technologies' Bridgit and Cisco's Webex.

Audio Conferencing. When conferencing audio fidelity is critical or when the number of participants at a conferencing site exceeds the performance capabilities of handset or tabletop conferencing systems, an audio conferencing system with a telephone system interface may be deployed. These high-end audio conferencing systems are similar in function to standard teleconferencing systems but utilize specialized audio conferencing equipment capable of handling many microphones and loudspeakers. This type of system can be quite complex and can incorporate a rack-mounted codec (coder-decoder) paired with digital signal processors and specialized microphones. A telephone-hybrid interface provides connectivity between the audio conference system and the telephone system. These systems are often permanently or semipermanently installed in purpose-built conference rooms. Like a standard teleconference system, an audio conference system can be enhanced through the use of web conferencing software, such as Smart Technologies' Bridgit and Cisco's Webex.

Desktop Video Conferencing. Desktop video conferencing is similar to basic telephone conferencing, but it is paired with a simple video feed of a participant's face. Audio and video from each participant are captured using a computer equipped with a simple web camera and possibly a headset microphone. Audio and video signals are typically transmitted over each user's internet connection in conjunction with desktop file-sharing and collaboration software. Other systems may use a video-capable telephone handset to transmit audio and video signals across a corporate local-area network (LAN) or an IP connection. Like telephone conferencing, this method of conferencing requires almost no special equipment. Unlike telephone conferencing, desktop conferencing systems offer more robust communication and almost limitless functionality.

Video Conferencing. Similar to high-end audio conferencing, high-end video conferencing takes place in a purpose-built conference room using a permanently or semipermanently installed system. These systems will generally use multiple microphones, loudspeakers, cameras, displays, and software to transmit audio and video communication between users. Application sharing allows users to share presentations, graphics, videos, electronic white boards, and document cameras. Some systems even allow virtual content creation and editing.

Group video-conferencing is highly desirable, because it allows face-to-face discussion between participants. These face-to-face communications allow participants to communicate using body-language as well as speech. The addition of video improves participant relationships and fosters collaboration. In today's global economy, businesses and educators seek to communicate with users throughout the world. Video-conferencing systems facilitate these communications, while eliminating the costs and carbon footprint associated with the travel required for an actual face-to-face meeting.

Currently, the largest impediment to video conferencing is universal adoption of technologies that are sometimes proprietary, meaning users at each endpoint must employ a system from a single manufacturer. Communication between different manufacturers' systems often has limited functionality and fidelity. Important meetings between different systems currently require a test call and may require adjustment of complex system settings. To this end, most video conferencing involves interorganization communications. Too often other communications are not user friendly. A second major barrier is user habit; however, with the advent of Skype, Google Chat, ubiquitous web-cameras, Apple's FaceTime, and cell phones with video conferencing capabilities, upcoming generations of digital natives will not only be comfortable with these technologies, they will demand them.

Realistic Telepresence. Face-to-face meetings are still highly desirable in many business settings. In any critical or sensitive

Based on a paper presented at Noise-Con 2011, INCE 25th Annual Conference, Portland, OR, July 2011.

meeting, it is imperative for participants to study body language and other nonverbal cues. While certain meetings are always likely to require face-to-face interaction, improved technologies are slowly chipping away at the need for actual face-to-face meetings. Telepresence systems use cutting-edge technologies to simulate the experience of a face-to-face, in-room meeting with a remote participant.

While there is a continuum between basic video conferencing system and a high-end telepresence system, a robust application will include the several common elements. The primary aim is to create the illusion that users are seated in the same room. All endpoints are outfitted with identical finishes, furniture, lighting, and technology. To users, the remote room looks, sounds, and feels the same as the local room. Furniture, displays, and cameras at each endpoint are arranged so that the video feed of remote participants is situated directly across the table from the local participants. Finally, the illusion of a shared presence is polished off with the capture and transmission of high-definition audio and video to limit aberrations that would draw ones attention to the presence of an artificial communication medium. In the same vein, negative acoustical and lighting effects are sure indicators of an artificial presence and must be tightly controlled.

Lecture Capture. A lecture capture system is another variant or hybrid of the systems described above. Typically employed in a university setting, a basic lecture capture system will capture audio, video, and computer content from an instructor. This feed is subsequently transmitted to students in remote locations or digitally stored for future viewing. High-end systems will use a multi-input conference paired with numerous automated microphones and cameras to capture student questions and discussions in addition to the lecture. Some installations attempt a form of telepresence, with the ultimate goal of facilitating discussions between students in both the local and remote classrooms. In this scenario, each student is outfitted with a microphone and loudspeaker.

Acoustical Design Considerations

Different conference systems provide users with a range of communication options, from simple utilitarian communications to the illusion of life-like, face-to-face meetings. Typically, more critical applications (and expensive installations) demand progressively higher levels of speech intelligibility and have a lower tolerance for intruding background noise. Far from a single set of criteria, the acoustical goals for any conference application begin with a firm understanding of the technology that will be employed and the system's programmatic requirements. Ultimately, the acoustical goal is to ensure that poor acoustics do not impair communication.

Acoustical Design Goals. Basic desktop audio and video conferencing often involves low-importance, impromptu communications undertaken at desks in open and private offices. Participants generally speak and listen through a telephone handset or a headset. The most common acoustical issue in this application is background noise, especially in an open office environment. Users may have difficulty hearing other participants in the presence of both local and remote noise. It is not uncommon to find that at least one participant is located in a noisy open office or calling into the conference from a mobile phone. The quality of the conference bridge is also a key factor to eliminate noise and ensure that all participants are heard at an equal volume.

All remaining types of conference systems employ microphones and speakers placed at some distance from participants. These can include microphones built-into conferencing equipment, goose-neck microphones, table-top microphones, ceiling-mounted microphones, and steerable microphone arrays. Depending on the conferencing application, each microphone may be intended to pick-up one or multiple participants. Remote participants are heard in the conference room via loudspeakers deployed in a similar range of installations. The coverage pattern and placement of these loudspeakers must be carefully coordinated with the microphone locations in the room to avoid creating an audio feedback loop.

As soon as distant microphones and speakers are employed, room acoustics begin to impact the fidelity of the system. Specifically, background noise and unwanted acoustical reflections can degrade speech intelligibility. Unlike the human ear, a microphone cannot use spatial cues to perceptually separate speech from unwanted background noise; both speech and noise are transmitted to remote users at an equal level. Conference systems employ various techniques to control this noise; however, it is always prudent to employ passive acoustical controls to reduce reliance on these systems.

It is important to remember that participants are presumably listening for half of any conference. Nearly all of the sophisticated noise control technologies that are employed today can only treat the effects of poor acoustical conditions from the room where speech originates. These technologies cannot eliminate the effects of poor acoustics in the receiving room. At a minimum, each conference room should maintain acoustical conditions that are appropriate for face-to-face communication; however, use of loudspeakers and microphones typically demands slightly better conditions.

Finish Treatment. In most conference rooms, an absorptive ceiling is more than sufficient to control reverberation time. To this end, additional acoustical treatments in conference rooms are rarely aimed at reverberation control. Rather, they are selected to control specular reflections and the build-up of low-frequency modal energy. In practice, these aims often result in the application of functional, tackable acoustical panels around the perimeter of the room and possibly some type of low-frequency absorber.

Sound Isolation. Because microphones are more sensitive to interference from noise (see above), it is important to provide a slightly higher level of sound isolation than what might be typically needed for a passive conference room. In practice, a mid-level video conference room may be surrounded by full-height walls, constructed with multiple layers of drywall and batt insulation. MEP penetrations in these walls, such as transfer air ducts and receptacles, must be identified and treated. While an entry vestibule is preferred, entry doors are typically a solid, fire-core type outfitted with acoustical seals.

Video conference rooms are often featured spaces in a building, surrounded by storefront glazing or frameless glass and doors. Where these conditions can't be avoided, a laminated insulated glazing unit is often set in a heavy-gauge, store-front frame that is outfitted with neoprene gaskets and filled with mineral fiber insulation.

Desktop video and audio conference applications, which use close micing techniques, rarely require additional sound isolation beyond what is considered good practice for private and open office environments. On the other end of the spectrum, the envelope of high-end telepresence rooms may begin to approach that of a low-grade audio or video studio.

Mechanical System Noise Control. It is generally advisable to seek a background noise level that is 5-10 NC points below standard guidelines for similar passive conference rooms. This is particularly true for conference rooms that use ceiling-mounted microphones placed close to HVAC diffusers. Similar to sound isolation requirements, desktop conferencing doesn't necessarily demand a quieter HVAC system, while high-end teleconferencing installations warrant an extremely quiet HVAC system.

The duct system serving a basic audio/video conference room must maintain low airflow velocities and avoid turbulent layouts and transitions. Terminal units must be located outside of the room and powered units may require duct silencers. Finally, diffusers and grilles must generate low noise levels and be located far from ceiling microphones.

Built-in Conferencing System Enhancements. As mentioned above, audio conferencing systems employ a variety of sophisticated noise control techniques to limit the effects of poor conference room acoustics. The manufacturers' intentions are to make their systems work in the maximum number of environments, not just expensive purpose-built conference rooms.

Note that as the noise reduction algorithms apply filters, gates, and cancellation techniques to eliminate noise, the speech signal is inevitably affected. As acoustical conditions in the conference room worsen, the speech signal is subjected to greater and greater levels of noise reduction until speech intelligibility is eventually degraded. At their extreme settings, noise reduction algorithms may break down with unpredictable results.

Finally, it must be restated that these electronic control measures can only affect the speech signal until it enters the receiver room. They cannot eliminate the effects of poor acoustics in the receiving room.

Noise Reduction. Noise reduction algorithms use software to identify and eliminate background noise. These algorithms typically identify wanted speech by its patterns, envelope, and frequency. They are very effective at identifying a steady-state hum or hiss from a mechanical system as unwanted noise; however, they are unable to identify occupant noises (shuffling papers and chairs) or speech from an adjacent room as noise.

Automatic Gain Control. Most audio conferencing devices use automatic gain control to keep the speech at a constant loudness, despite level fluctuations that may occur while a talker is speaking. To function properly, the talker's voice must have a suitable signal-to-noise ratio. If the background noise in a room becomes too loud, the automatic gain control will respond to the background noise instead of a quiet talker.

Echo Cancellation. Echo cancellation refers to algorithms that are employed to eliminate two forms of "echo." The first form is the familiar acoustical type that affects a talker's voice in the source room as reverberation, slap echo, and unwanted specular reflections. Echo-cancellation algorithms may emit and record a test signal in each endpoint at set-up to identify the acoustical signature of each room so that it may later be removed from the transmitted speech signal.

The second type of echo occurs when a talker's voice is emitted from loudspeakers at a remote site, recaptured by the microphones in that room, and retransmitted back to the original talker. This condition results in a considerably pronounced and delayed echo. Echo-cancellation algorithms will employ a variety of gates and noise cancellation techniques to eliminate this form of echo.

Voice Lift. Employing multiple microphones in a conference room inevitably increases the amount of room and occupant noise that is captured by the system. An extreme example would be a distance-learning classroom where every student or pair of students is outfitted with a table-top microphone and loudspeaker. The combined noise from all of these microphones is obviously undesirable and should not be transmitted to remote users. A voice-lift system mutes the signal from each microphone until it receives a signal from a physical push-to-talk button or perceives speech above a preset threshold. A truly intelligent voice-lift system will not only transmit a student's voice to a remote classroom, it will also send the signal to every other student's personal loudspeaker within the local classroom. The level of speech sent to each loudspeaker can even vary based on its physical distance from the talker. These same devices can also be connected to automated pan-tilt-zoom cameras and microphone arrays, which focus the systems attention to a specific talker.

Microphone Techniques. Designers of conferencing systems will use a number of different microphone techniques depending on the application, functionality, and budget for each system. Acoustical designers must understand the ratio of microphones to talkers in each conferencing application. Some applications will use a single, distant, omnidirectional microphone to pick-up the voices of all talkers in a room. This configuration is simple and cost effective but introduces the room's acoustical environment into the transmitted signal. Other applications will use a single directional microphone for each talker, often in the form of a table-top or lapel microphone. This configuration passively rejects much of the room's acoustical environment from the transmitted conference signal.

Other Design Considerations. Conceptual acoustic designs for video and audio conference rooms and even telepresence rooms are fairly simple. However, one of the biggest challenges is to deploy and coordinate acoustical treatments with other functional and architectural requirements.

Furniture and Conferencing Equipment. Most conference rooms are filled with fixed or flexible seating, either at a conference table or, in the case of a distance-learning classroom, in rows. Telepresence rooms will often include "half" of a conference table terminating into a display showing remote participants. It is important for the acoustical designer to understand the orientation of each talker (and loudspeaker) in relation to the room's microphones. For instance, talkers may be speaking into fixed table-top microphones, or the placement may be flexible and change from day to day. Voices may be captured by a distant ceiling-mounted, omnidirectional microphone susceptible to noise from the HVAC system or footfalls on the floor above.

Further, the room's wall surfaces are quite often taken up by large displays, electronic white boards, and credenzas. If acoustical wall treatments are deemed necessary, they must be incorporated into these elements or deployed between them.

Aesthetics. Perceptually, the conference system needs to remain transparent to facilitate effortless communication between remote participants. Like acoustical disruptions, visual distractions can quickly degrade this transparency. Any acoustical finish treatments in view of the room's cameras must have a neutral finish. Bright colors, complex patterns, and bold textures are to be avoided. Horizontal and vertical slats or grooves should be avoided or covered with a neutral acoustical fabric to avoid moiré patterns on video images.

Lighting. Conference room designers will go to great lengths to control lighting conditions in video conferencing facilities. Whenever practical, black-out drapes are used to eliminate outdoor light. Participants are instead lit using dedicated video lighting with a specific color temperature to give subjects a bright and natural appearance.

Conference rooms are frequently designed with one long wall consisting entirely of an exterior glazing and an opposing corridor wall made up entirely of interior store-front glazing. To address lighting concerns, the glass surfaces are simply outfitted with black-out drapes that are automatically deployed when the conferencing system is in use. Unfortunately, these black-out drapes are acoustically transparent. They do nothing to address the acoustical issues posed by two opposing glass walls or the poor sound isolation capabilities of glass.

A secondary set of heavy fabric drapes can offer some acoustical control for these surfaces; however, it is important to remember that they are usually deployed only when the conferencing system is in use. Even a well used video conferencing room will serve as a basic, passive conference room from much of its life, and the drapes will not be deployed in this scenario. If used, these drapes should be automatically deployed by the room's control system to avoid reliance on users who may not understand their purpose. Another promising treatment is the permanent installation of clear polycarbonate absorbers over the room's glass surfaces.

This situation is more common in mid-level video conferencing facilities. High-end telepresence rooms are generally seen as purpose-built environments, and it is easier to make the case to eliminate glass walls entirely.

Equipment Noise. Depending on its complexity, the electronic equipment used in a conference room may generate quite a bit of noise. While a simple table top teleconference device will generate almost no noise, more complex systems may include video projectors, computers, amplifiers, and other equipment outfitted with noisy cooling fans. The ideal approach is to place this noisy equipment in an isolated equipment closet or enclosure and run the equipment remotely using a control system. Rear-projection rooms, where used, offer an excellent place to house noisy equipment; however, the acoustical designer should study the construction of the rear-projection screen and any loudspeaker alcoves to gauge how much equipment noise these elements will allow into the conference room.

Outgoing Sound Isolation. Video conference rooms are typically designed to provide a rich, multi-media playback from multiple audio and video sources. It is not uncommon to use these systems to view a multimedia presentation that includes loud music, effects, and speech. Without proper treatment, the noise produced in a typical video-conferencing room has the potential to seriously disrupt occupants of nearby spaces.

Material Durability. Conference rooms, by their definition, are public spaces where groups of people gather to converse and collaborate. To this end, they may see high levels of traffic, especially in distance-learning classrooms. Acoustical treatments, from finishes to door seals, must maintain an appropriate durability for public use.

Case Study

The Sextant Group was recently involved in the design of a corporate conference room that was outfitted with HP's Halo telepresence video conferencing system. The room was located on the 29th floor of a high-rise building in an urban area. Despite using telepresence-grade technology, this application did not seek to provide life-like, across-the-table video conferencing. Instead, the room contained a traditional conference table in the center of the room and a rear-projection screen on one wall to display remote participants and content.

To optimize the performance of their telepresence system, HP recommends specific acoustical criteria for conferencing rooms that will house their system. Specifically, HP recommends a sound transmission class (STC) of 50-55 for walls surrounding the room, an ambient noise level below 32-34 dBA, and reverberation time below 0.8 seconds at 1,000 Hz. These criteria were considered in conjunction with the acoustical design goals described previously.

The walls surrounding the conference room were constructed using two layers of 5/8-inch-thick gypsum wallboard on both sides of metal studs and batt insulation in the stud cavity. The walls extended from the floor slab to the deck above and were carefully sealed with a nonhardening sound caulk. A separate corridor with a waiting room was built outside of the conference room to provide an acoustical buffer zone. In addition, the conference room used solid-core wood doors outfitted with full-perimeter seals and a threshold.

A major sound isolation challenge was the presence of a glass curtain wall along the room's exterior wall. In general, this condition had an adverse affect on the quality of the video image and provided limited acoustical separation from exterior noise sources. In this particular application, however, traffic noise was less of a concern, because the conference room was located on the building's 29th floor. A more significant concern was obtaining a good seal between the conference room's interior partitions and the mullions in the curtain wall. This intersection detail was carefully studied and enhanced to maintain an appropriate level of isolation between the conference room and adjacent spaces.

The HVAC (heating, ventilation and air conditioning) system presented the primary source of background noise in the room. To maintain an appropriately quiet background noise level, VAV (variable air volume) boxes and large main ducts passing over the space were relocated to an adjacent space. In the remaining duct work, airflow velocities were kept to a minimum, and smooth transitions were provided to reduce turbulence and air flow noise.

Noise from video projectors, amplifiers, and equipment racks in the adjacent rear projection room also presented a significant noise concern. Equipment noise levels were analyzed to determine how much noise reduction had to be achieved between the two spaces. A half-inch-thick, glass, rear-projection screen was ultimately the limiting factor for noise reduction between the spaces. Seals between the rear-projection screen and the surrounding wall were selected to minimize sound transmission.

Controlling reverberation within the space was also a challenge. Three of the walls were scheduled to contain mostly hard, acoustically reflective surfaces. The south wall was comprised entirely of the exterior curtain wall; the west wall contained the room's rear projection screen; and the north wall was outfitted with two large writing surfaces. As such, the ceiling provided the best opportunity to introduce acoustical absorption into the room, so acoustical ceiling tiles were used for the majority of the ceiling. Fabric-wrapped acoustical panels were also placed on the available wall surfaces in the room.

Reviewing architectural documentation ensured that the conference room would be capable of meeting HP's acoustical requirements. Achieving these requirements in the field took careful coordination between the design and construction teams. S/V

The authors can be reached at: gcoudriet@TheSextantGroup.com, jeff@babichacoustics.com.