

Polycom Video Communications

Advanced Audio Technology for Video Conferencing

Author:
Peter L. Chu

September 2004



 POLYCOM®

Connect. Any Way You Want.

If you ask frequent users of video conferencing technology what they consider to be the number one product feature for providing a high quality video conferencing experience, their response will most likely be the requirement for excellent audio. High quality audio is far more important to the perceived quality of a video conference as dialogue, discussion and decision making are usually the most important elements between meeting participants.

With an undisputed leadership role in the audio conferencing segment, Polycom also takes the leadership role with our outstanding audio quality in video conferencing systems. It is Polycom's goal to create and provide video conferencing solutions that deliver the highest quality audio capabilities for any environment. With Siren™ 14 and now patent pending Polycom StereoSurround™ audio technology, Polycom once again raises the bar for quality audio standards and delivers true performance to our customers.

Polycom's leading advantage: Siren 14

Polycom's Siren 14 technology brings 14 kHz audio fidelity for the most natural conversation. With twice the fidelity of other vendors, and four times the fidelity of a telephone call, Polycom is the only company in the industry to offer such "super-wideband" audio capabilities for video conferences.

With 14 kHz audio in a video conference call, participants can better hear and understand each other, while eliminating the fatigue often associated with longer duration calls under inadequate audio quality.

Why Siren 14 Audio is Important

Speech information is usually contained in the 100-11,000 Hz range, so having frequencies above 11,000 Hz is not important except for music. While a small minority of people can hear up to the limit of 20,000 Hz in audio frequency, the large majority of people cannot hear the difference between 14,000 Hz and 20,000 Hz high frequency limits for even the most demanding music material. The Siren14 codec in the iPower 9000 and VSX 3000, VSX 7000 and VSX 8000 provides audio capability up to 14,000 Hz, allowing for transparent reproduction of both speech and music.

The Siren 14 algorithm suite, developed and patented by Polycom is designed to provide a scalable wideband audio solution that will obviate any reason to use narrow band (3.4 kHz) audio coding. When combined with Polycom patented audio echo cancellation, the suite provides an environment that allows both the local and remote speaker to engage in conversation simultaneously, without cutting each other out. Siren 14 operates at 14 kHz audio at bit rates between 24 and 48 kbps, the usual bandwidth allocated to audio transmission in a video conference. Polycom's unique Siren 14 Super wide band audio capabilities provides incredible audio quality approaching compact disk and MP3 quality, while maintaining a modest network bandwidth requirement. In comparison to other audio algorithms, such as MPEG 3 and MPEG 4 at lower bit rate transmissions, Siren 14 actually delivers a smoother and more superior sound.

Polycom's Digital Tabletop Microphone: A natural companion for Siren14

Most video conference systems use simple directional microphones. Typically these microphones have a (+/-) 45 degree pickup pattern as measured from the axis pointing from the front of the microphone. The disadvantage of this type of microphone is that participants who are sitting away from the pickup pattern cannot be heard as well as those sitting closest to the microphone. Polycom digital tabletop microphones deploy technology that allows for voice pickup in a 360 degree radius covering distances up to 10 feet away from the microphones. Polycom digital tabletop microphones also use background noise suppression functions which eliminate unwanted background noise such as that generated by laptop or projector fans and vents. The benefit of automatic noise suppression creates a virtually noise-free communication for most meeting environments.

By using Polycom's digital tabletop microphone technology combined with the Siren 14 audio algorithm, conference attendees can sit anywhere in the room without having to worry whether their voices are being heard at the remote sites. An example of this capability is exemplified within the tabletop microphone shipped with the VSX™ 7000 product series. The VSX 7000 tabletop microphone deploys a three element microphone array that forms a narrow beam that automatically points to the person speaking. The narrow beam eliminates noise and reverberations from other directions. When several microphones are linked together, the microphone nearest the person speaking is automatically and seamlessly chosen while noise from the remaining microphones is eliminated.

Audio Reproduction from Polycom

Equally important to participants in a video conference is sound fidelity of the speakers that produce the audio signals. Clear speech, after all, is noted as one of the the most important elements of a video conference. Often, end users rely on the quality of television speakers for video conference audio reproduction. These speakers, however, are not well suited for reproducing speech, the most commonly transmitted audio signal in video conferencing applications. Television speakers are usually designed to offer

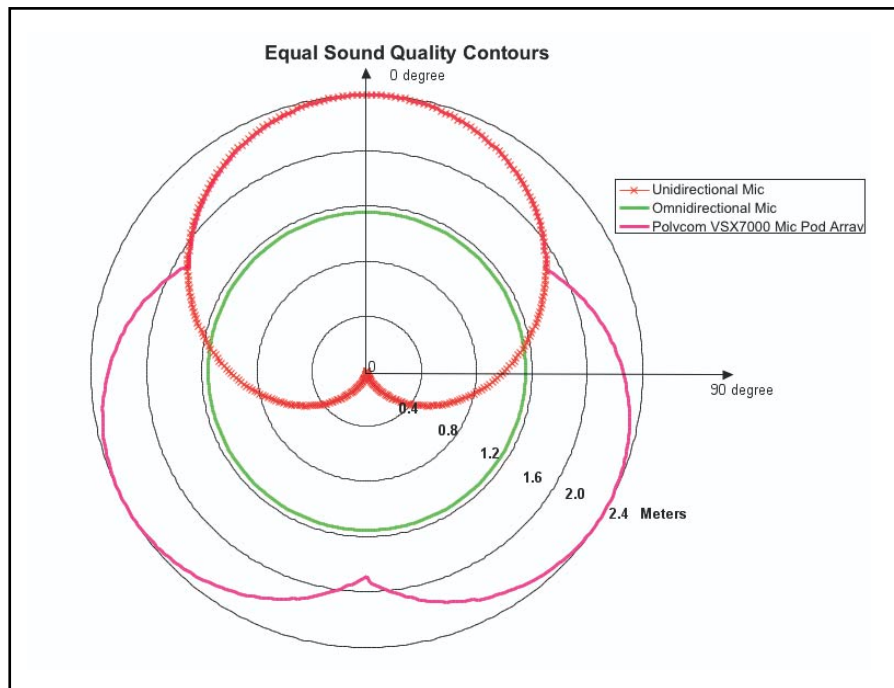


Diagram 1: Area pickup of Polycom Digital Tabletop microphone compared to other microphones.

merely adequate audio reproduction for a low price. As a result, these television speakers often sound similar to a portable stereo system.

Polycom offers custom high fidelity audio speaker systems to provide the highest quality speech reproduction for video conferencing. The VSX 7000 is the world's only set-top video conferencing system that includes an integrated speaker and companion subwoofer to provide near CD-quality audio in any setting. The Polycom Executive Collection offers a 270 Watt speaker subsystem with each Executive Collection configuration; ideal systems for this configuration include the Polycom iPower™ 9000, VS4000™ and VSX 8000. Polycom also provides an auxiliary Speaker Kit, ideal for adding Polycom StereoSurround capabilities to the VSX 7000 or VSX 8000. This Speaker Kit consists of a subwoofer with 150 watts output power with two 6-inch drivers with bass response down to 35 Hz and dual speakers that include a 3/4" silk fabric dome tweeter combined with two 3 1/2" composite midrange drivers for superb sound reproduction at frequency ranges of 100 Hz to 22 kHz.

The Total Polycom Audio Solution

With all of these great audio technologies, it is easy to see why Polycom's audio is the best in the industry and why Polycom continues to be the market leader in video conferencing solutions.

Setting a new standard: Polycom Stereo Surround

Consumers are used to enjoying stereo audio when listening to music, watching DVD movies, or even experiencing impressive multi-channel audio at movie theatres. Stereo makes the overall multimedia experience more realistic and less fatiguing, providing us the required spatial cues to separate out different sounds. On a video screen, when someone on the left hand side of the screen talks, we expect that sound to come out the left side, and stereo allows this to happen. In our everyday life, we are used to having sounds come from all directions. In mono conferencing, when every sound comes from a single point in space, it is very artificial and fatiguing.

To make stereo audio conferencing a reality, Polycom has developed breakthrough patent pending technology for the VSX 8000, VSX 7000, VSX 3000, and other products. The technology includes stereo echo cancellation, Automatic Gain Control (AGC), noise suppression and audio compression technology.

Echo Cancellation: The Enabler for Polycom StereoSurround

If stereo audio is so ubiquitous in the consumer industry, why is it absent in the video conferencing industry? The problem has to do with echo cancellation. Echo cancellation is required to eliminate the pickup of the loudspeaker by the microphone. Without echo cancellation, users would hear feedback which would make the conference very annoying. Conventional echo cancellers identify only one path between the loudspeaker and the microphone. In stereo, there are a multiplicity of paths between the two loudspeakers and microphone, which confuses the echo canceller. To solve this problem, many researchers have resorted to distorting the left and right loudspeaker signals, which enables correct estimation of the multiple paths. Unfortunately, this distortion can often be heard, especially for music.

After a decade of research, Polycom has developed patent pending stereo echo canceller technology which doesn't distort the signal, yet quickly identifies the multiple paths of the stereo loudspeakers and eliminates echo, allowing for full-duplex conversations. There is no "white noise" training sequence, as found in competitor's mono echo cancellers. Positioning of the loudspeakers and microphones is not critical. If microphones are moved, the stereo echo canceller quickly adapts to the new position within one or two words of speech. For the frequency range from 50 Hz to 14,000 Hz, the echo canceller adaptive filter length is 160 milliseconds which is adequate for even large rooms of 30 by 50 by 9 feet in dimension.

The echo canceller works as well with stereo music as it does with speech.

Other Polycom stereo acoustic signal processing: AGC and Noise Suppression

Polycom's artifact-free stereo noise suppression is applied to the microphone signals to eliminate hiss and fan noise. The noise suppression adapts within several seconds to new noise sources which are introduced, like someone turning on a projector. Additionally, a voice-activated stereo AGC is included to automatically increase microphone gain for people who are further away from the microphone.

The Basis for Polycom StereoSurround

A stereo version of Polycom's Siren14 codec has been developed to give high quality 14 kHz bandwidth stereo audio at lower bit rates than MPEG 4, thereby allowing video to have more bits for sharper images. Polycom StereoSurround works at 48, 64, and 96 kilobits per second. Polycom hopes that offering stereo audio capabilities as part of the standard feature set on video conferencing system will popularize the use of high-quality audio in video conferencing. In the international standards body for video conferencing, the ITU-T, Polycom is sponsoring the adoption of Siren14 as a standard extension to the existing G.722.1, which is also a Polycom developed codec.

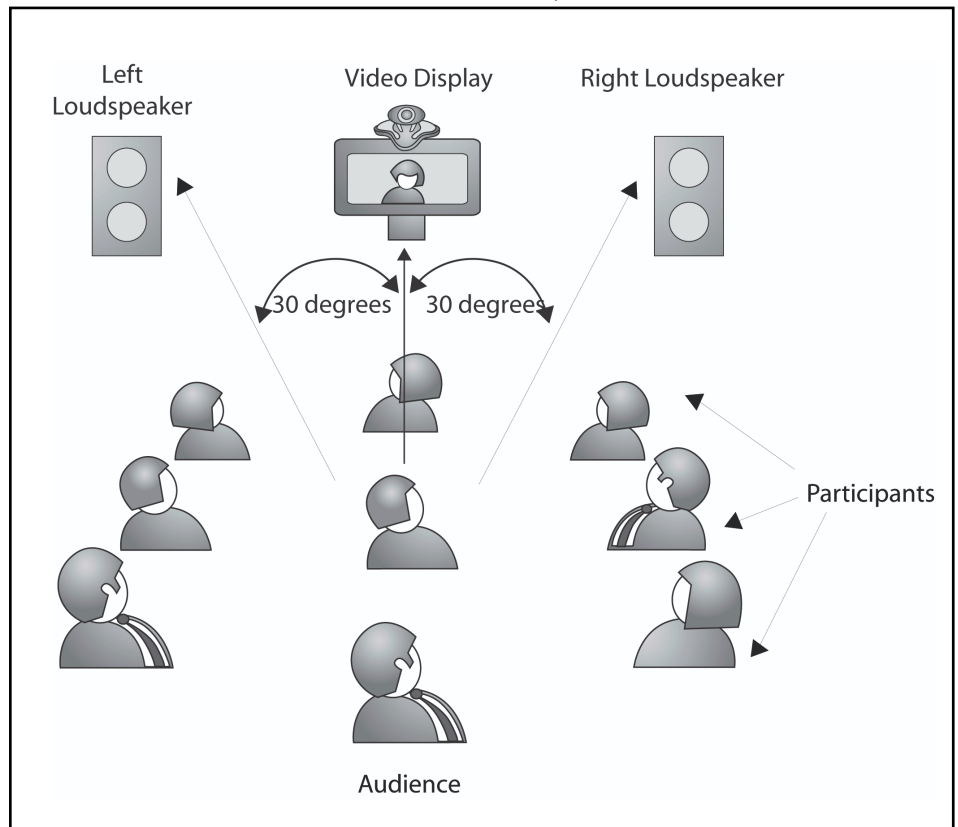
Speaker and Microphone Positioning on the VSX 3000

Since the VSX 3000 ships with two integrated speakers (left and right) and two integrated microphones (left and right), it is ready for Polycom StereoSurround right out of the box without any additional hardware. Simply enable the stereo feature in the user interface and begin making video calls. The difference will be immediately noticeable!

Loudspeaker Placement for Group Video Systems

For the optimal realization of Polycom StereoSurround, it is necessary to place the left and right loudspeakers in the correct positions. The optimal angles for the left and right loudspeakers should be +/- 30 degrees with respect to an average audience. To ensure proper placement of loudspeakers, go to System/Diagnostics/Speaker Test. When the left speaker icon is enabled, the tone should come out of the left loudspeaker. Do the same test for the right loudspeaker. Loudspeaker placement for use with the VSX 7000 or VSX 8000 is illustrated in the diagram below.

Another way to view proper speaker placement is that the distance between left and right loudspeakers on either side of the TV monitor should equal the distance between the TV monitor and the midpoint position of the audience. So, for example, if the midpoint of the audience is 12 feet away from the TV monitor, the left loudspeaker should be 12 feet away from the right loudspeaker. Separating the loudspeakers by this amount allows the average listener to perceive sound coming from left loudspeaker, right loudspeaker, and points in-between the two loudspeakers.



Optimal placement of left and right loudspeakers for Polycom StereoSurround

Microphone Placement for Group Video Systems

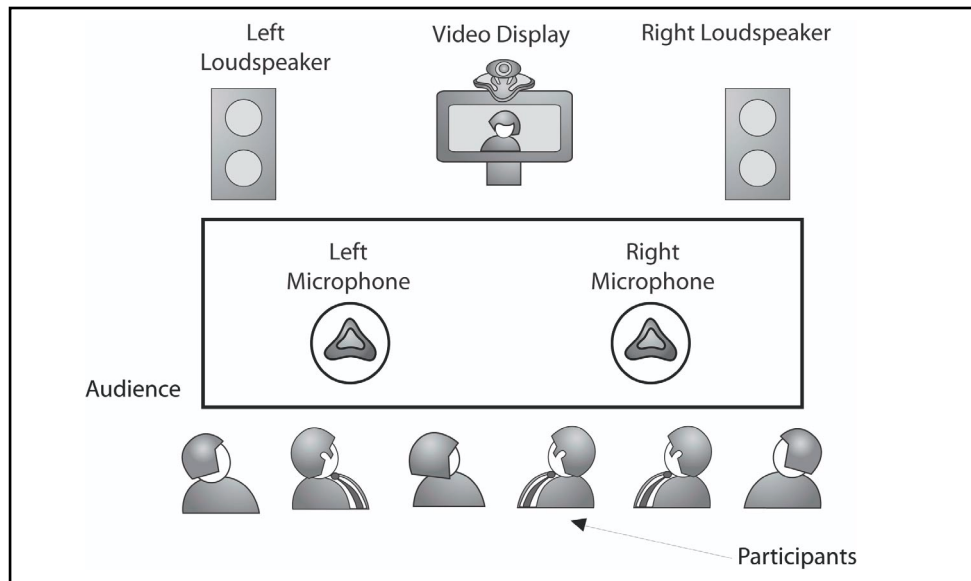
Stereo microphone pickup is accomplished in the VSX 7000 and VSX 8000 by utilizing two microphones. One microphone picks up all audio for the left audio channel while the other microphone picks up all audio for the right audio channel. Knowing which microphone is left and which is right can be determined by going to the VSX 7000 or VSX 8000 menu, Admin Settings/Audio/Stereo Settings. Lightly tap each microphone to identify their positioning by watching the Left and Right microphone meters in the UI. If necessary, swap the left and right microphones to correctly correspond to the meters.

The left microphone's purpose is to pick up sound in the left half of the audience. As such, it should be placed in the center of the left half of the audience, towards the front, as people tend to speak facing the video display. For the right microphone, a similar situation applies. The diagrams to the right illustrate proper placement of the microphones for both horizontal and perpendicular orientation of the table relative to the video display.

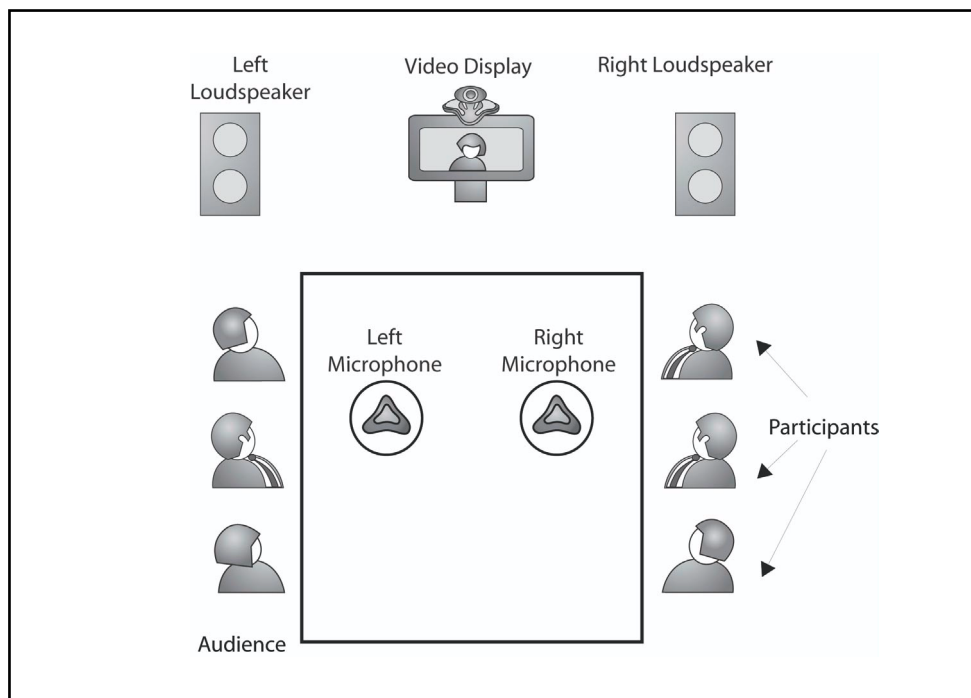
The Polycom Office™

With integrated video, audio, data, and Web capabilities, The Polycom Office is the only solution that offers an easy way to connect, conference, and collaborate any way you want. Work faster, smarter, and better with The Polycom Office.

Polycom, Inc. develops, manufactures and markets a full range of high-quality, easy-to-use and affordable voice and video communication endpoints, video management software, web collaboration software, multi-network gateways, and multi-point conferencing and network access solutions. Its fully integrated end-to-end solution, The Polycom Office, is supported by the Polycom accelerated communications architecture and enables business users to immediately realize the benefits of integrated video, voice data and web collaboration over rapidly growing converged networks.



Optimal placement of left and right microphones for Polycom StereoSurround on a horizontal table.



Optimal placement of left and right microphones for Polycom StereoSurround on a perpendicular table.



Polycom Headquarters:

4750 Willow Road, Pleasanton, CA 94588 (T) 1.800.POLYCOM (765.9266) for North America only.
For North America, Latin America and Caribbean (T) +1.925.924.6000, (F) +1.925.924.6100

Polycom EMEA:

270 Bath Road, Slough, Berkshire SL1 4DX, (T) +44 (0)1753 723000, (F) +44 (0)1753 723010

Polycom Asia Pacific:

Polycom Hong Kong Ltd., Rm 1101 MassMutual Tower, 38 Gloucester Road, Wanchai, Hong Kong, (T) +852.2861.3113, (F)+852.2866.8028

©2004 Polycom, Inc. All rights reserved.

Polycom and the Polycom logo design are registered trademarks and Siren, Polycom StereoSurround, VSX, iPower, VS4000 and the Polycom Office are trademarks of Polycom in the U.S. and various countries. All other trademarks are the property of their respective owners. Specifications subject to change without notice.

Rev 9/04