

Polycom[®] Using SoundStructure[®] to Compensate for Audio Delays in Displays Used in Conferencing Applications

Overview

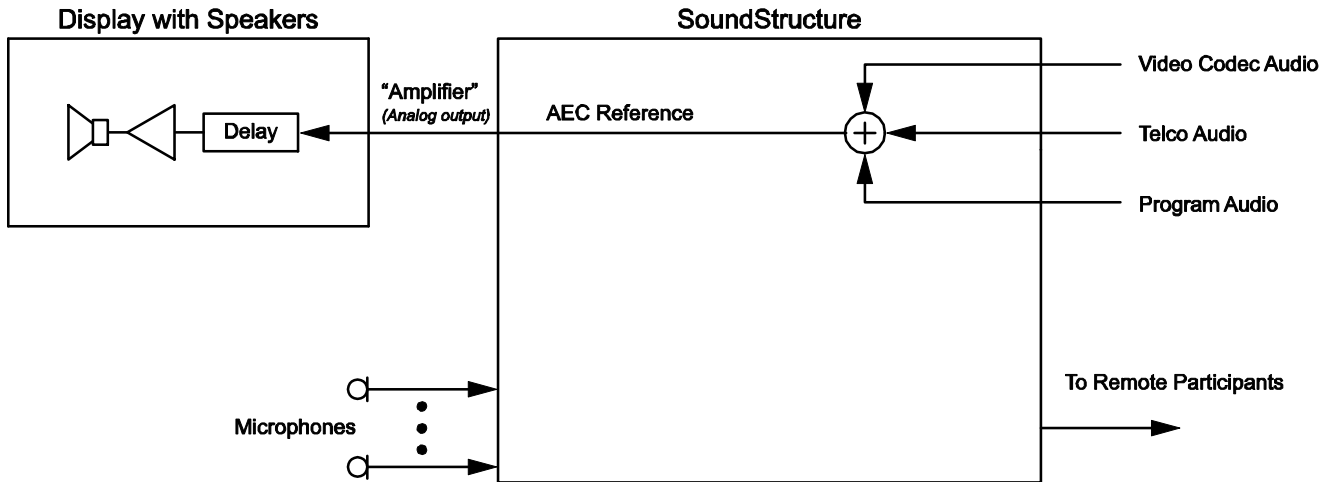
Many of the displays used for video conferencing applications include loudspeakers that serve as the audio playback system for room conferencing systems. If these displays introduce delay into the audio path due to inherent video processing delays, the performance of a room-based echo canceller will degrade, possibly causing acoustic echoes to be heard by the remote participants. This application note describes how to compensate for audio delay introduced by a display.

Display Audio Delay

Digital display technology has evolved over the years to include sophisticated video resolution scaling and other video enhancements to improve the resulting video image. In many applications, displays that are used for video conferencing applications contain loudspeakers that serve as the audio playback system for the room conferencing system. Sophisticated video processing within the display may require buffering multiple frames of video data which adds delays to the video playback on the order of 100msec. To maintain synchronization between the audio and video signals, the audio path will also be delayed by the same amount.

In a conferencing system, a delay of the audio signal in the display will reduce the performance of a conferencing system because the audio played into the room via the display's loudspeakers is delayed relative to the remote audio used in the echo canceller to remove acoustic echoes that are present at the microphones. The audio delay in the display causes a direct tail-length reduction of the echo canceller. For example, if an echo canceller has a 250msec tail time and a display adds 100 msec of delay before the audio is heard in the room, an echo canceller will have been waiting 100 msec to remove the acoustic echo before the audio is even present in the room, effectively reducing the echo cancelling time window by the same 100 msec to only 150msec of tail-time. This reduction of the echo canceller time window can result in acoustic echoes being sent back to the remote participants. Adding delay in the display is similar to the loudspeaker being moved 100ft away from the microphone yet still providing an audio signal that needs to be echo cancelled at the microphones.

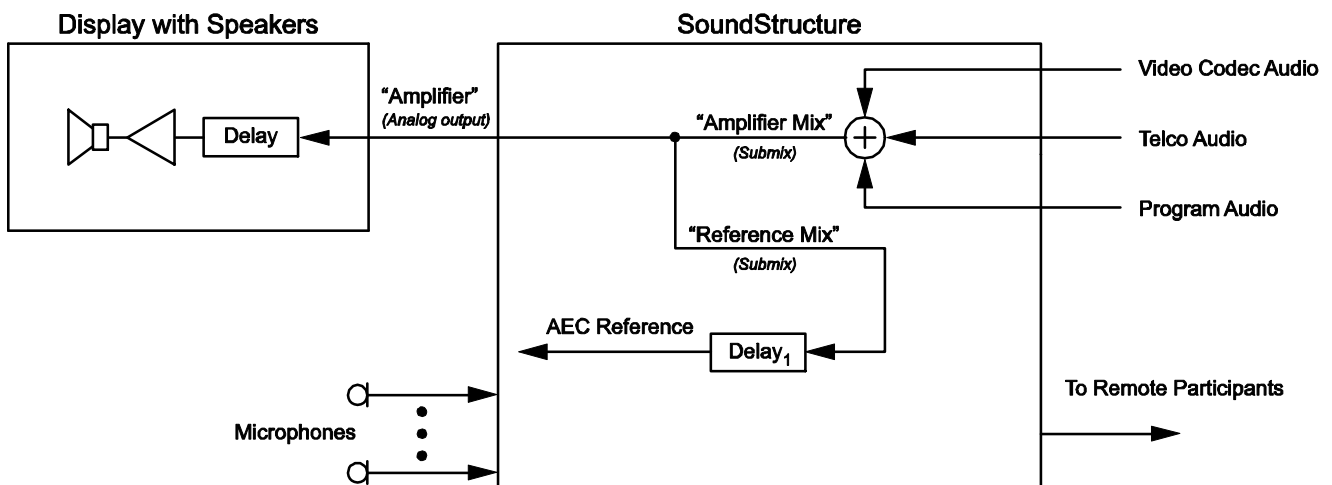
A typical conferencing system with a display adding delay (Delay) is shown in the following figure. The remote audio sources are connected to a SoundStructure system and are mixed together and sent to the loudspeakers in the display. Any remote audio picked up by the microphone will be delayed relative to the AEC reference in the system. If there is enough delay added by the display, this delay will reduce the effectiveness of the echo canceller possibly causing echoes to be heard by the remote participants.



Compensating for a Display Device's Audio Delay

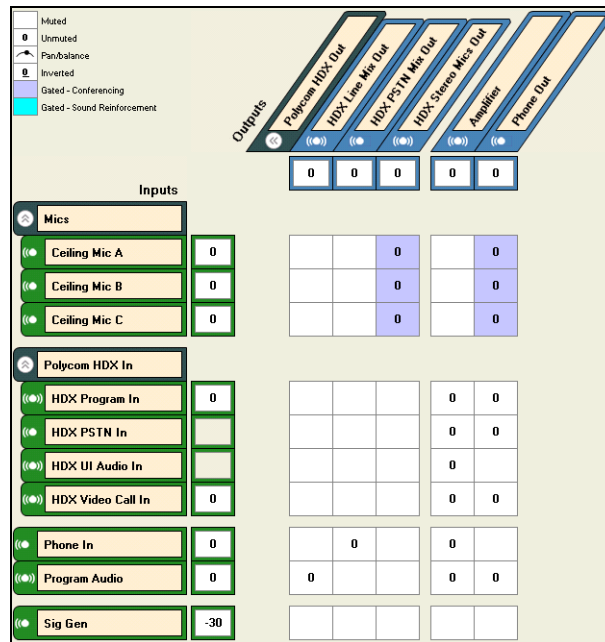
To compensate for the display's delay, it is necessary to delay the AEC reference inside SoundStructure by an equivalent delay as shown in the following figure. By adding Delay₁ to the AEC reference where Delay₁ ≤ Delay, this will tell the echo canceller that the echoes will be showing up later in the room.

It is important that Delay₁ be less than or equal to the delay added by the display. To determine the appropriate delay value, the installer should start with a small value of delay, for example 10msec, and gradually increase the value to ensure that less echo is heard by the remote participants. After each change in the delay value, allow the echo canceller to reconverge with up to ten seconds of only the remote participants talking. If Delay₁ is set larger than the delay in the display, persistent residual echoes will result because the echo will appear before the echo canceller knows to cancel it.

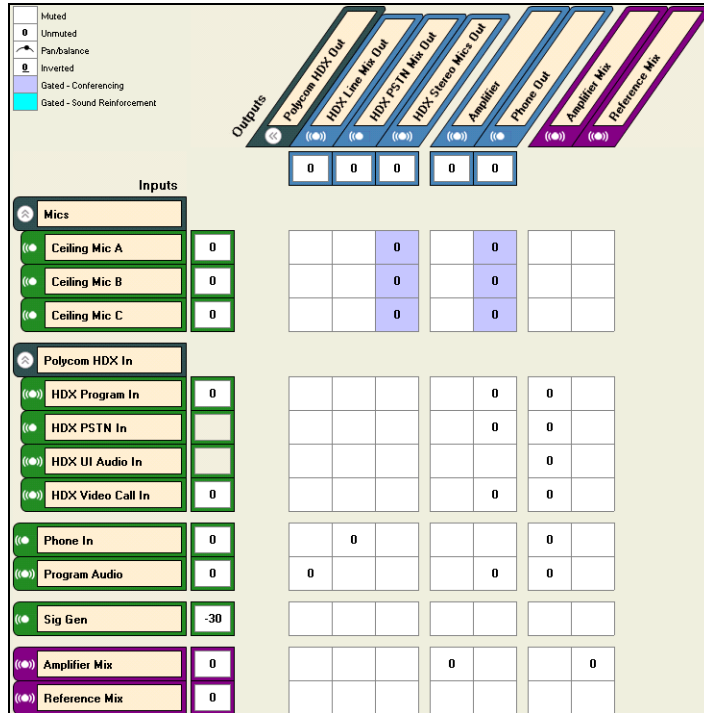


To implement this solution with SoundStructure, follow these steps:

1. Design the SoundStructure project as usual. In this example, the system has a Polycom HDX video codec, a SoundStructure telephony interface, one HDX ceiling microphone, and stereo program audio. Once the project has been designed, the matrix will look like the figure below.

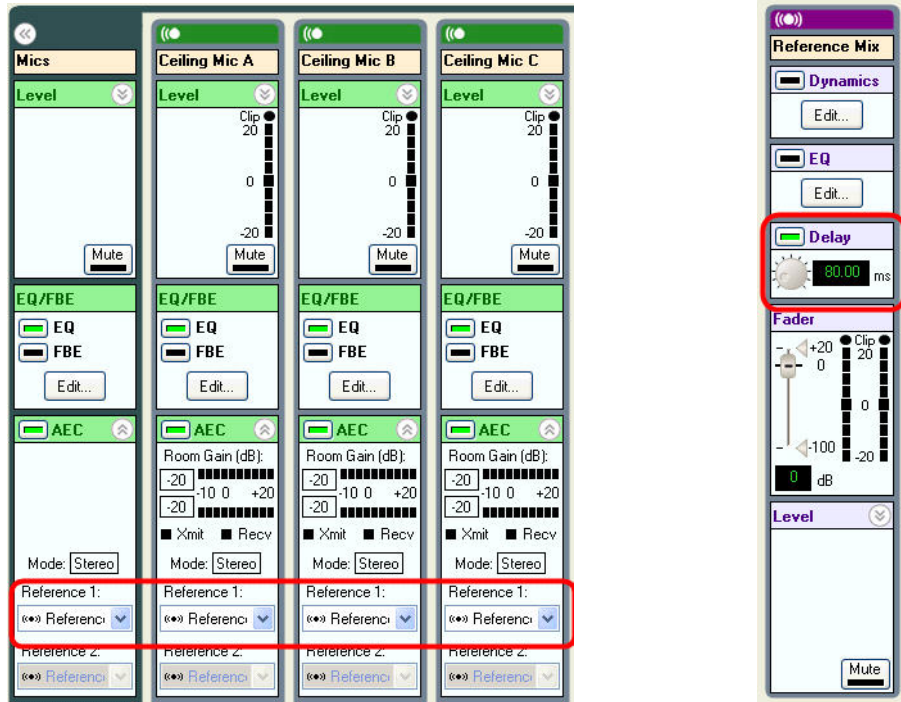


2. Use the Edit Channels control to add two submixes. If the audio to be played in the room is stereo, create Stereo Submixes. Rename the submixes to “Amplifier Mix” and “Reference Mix”.
3. On the matrix page, route all the remote audio sources to “Amplifier Mix” and route “Amplifier Mix” to the “Amplifier” output signal. In addition, route the “Amplifier Mix” to the “Reference Mix”. The resulting matrix should look like the matrix in the following figure.

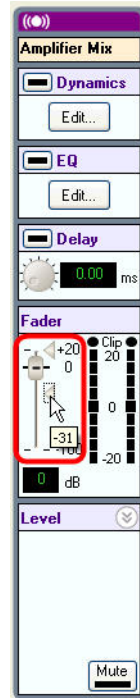


- On the channels page, set the reference for the “Mics” to be “Reference Mix”, and then enable the delay on the “Reference Mix” and set the delay to 80 msec.

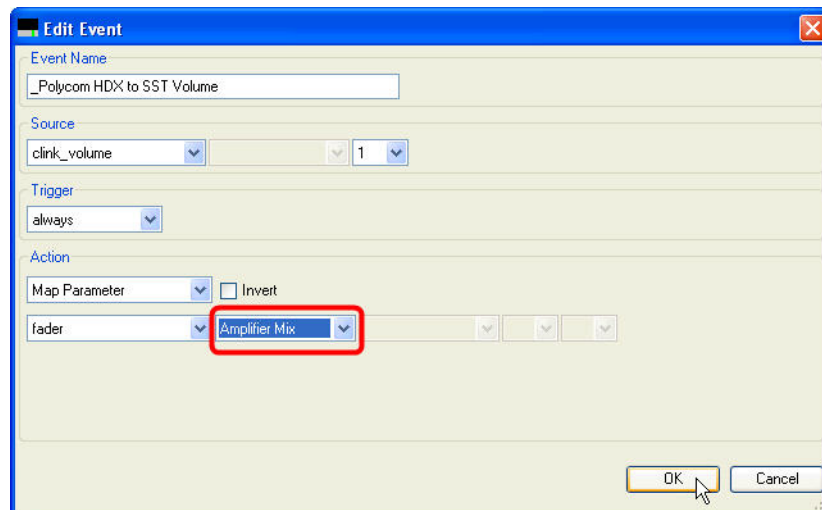
Note: The delay value 80 msec may not be sufficient for your display. As described previously, if the delay in the display is unknown, start with a small delay value, for example 10msec, and gradually increase the delay until the echo is removed. After each change in the delay value, allow the echo canceller up to ten seconds with the remote participants talking to reconverge.

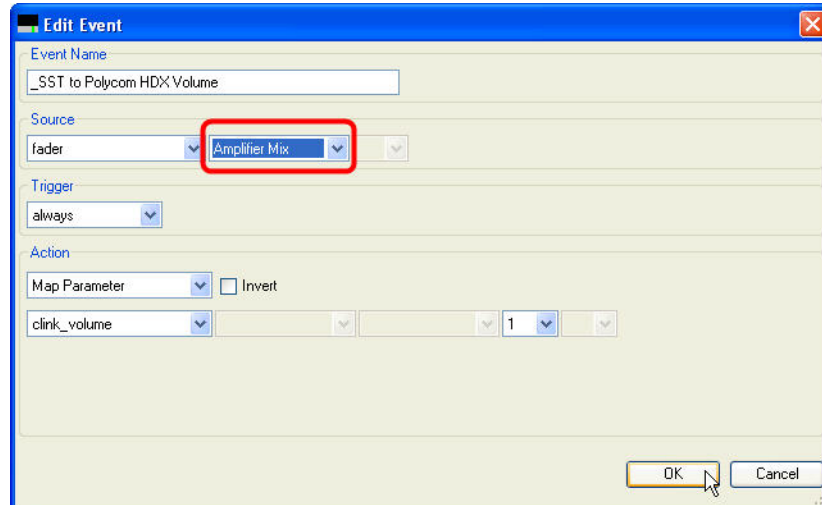


5. If the project has 'events' for volume control that use the "Amplifier" channel, drag the fader user min value on the "Amplifier Mix" channel to -31 on the Channels page as shown in the following figure. The user maximum will be at +20 by default.



On the events page, modify the volume control events, `_Polycom HDX to SST Volume` and `_SST to Polycom HDX Volume` to use the “Amplifier Mix” channel instead of “Amplifier”. This will ensure that volume adjustments to the system will adjust the “Amplifier Mix” signal allowing the echo canceller to also see the volume adjustments. To change the events, double click the event and make the changes as shown in the following figures.





6. Save the project to ensure the settings are stored to a preset.

Adjusting Volume

Volume adjustments to make the room louder or quieter should either be applied to the input sources (recommended) or to the room fader of the channel “Amplifier Mix”. Either approach will ensure that as the remote audio sources get louder; the echo canceller will also see the louder audio and continue to perform well. In step 5 above, the volume control was adjusted to ensure the “Amplifier Mix” fader was adjusted instead of the “Amplifier” fader.

Summary

To compensate for audio delays in displays that use built-in loudspeakers for conferencing audio, it is important to delay the AEC reference used by the echo canceller of the microphones within a SoundStructure system. This can be accomplished simply by creating several submixes and delaying the AEC Reference to compensate for the delay of the display. It is important that the delay on the reference be no more than the delay through the display.