Cisco TelePresence SX80, MX700, MX800
Using Extra Loudspeakers and Sound Reinforcement
A guide to providing a better auditory experience

D1536502
Produced: October 2016
All contents © 2016
Cisco Systems, Inc. All rights reserved.
Contents

What’s in this guide

Contents

Introduction.................................................................3
Codec SX80: Use with Extra Loudspeakers .......................4
Codec SX80: Expanding Scenario to Include Local Presenter ......6
MX700/800: Use with Extra Loudspeakers..........................8
MX700/800: Expanding Scenario to Include Local Presenter......10
Appendix: Relevant API Commands.................................12
Introduction

An MX 700/800 video system comes with built-in loudspeakers. You may, however, want to add external loudspeakers to the room to enhance the overall auditory experience.

Similarly, when using the Codec SX80, you may want to use more than one pair of loudspeakers in the room.

When using the SX80 with local sound reinforcement, extra loudspeakers are strongly recommended to ensure that people in the room get the good auditory experience.

With an MX700/800, extra loudspeakers are required to enable local sound reinforcement functionality.

To set up the system you will need to use the API of the codec. The relevant API commands are presented in the appendix.

We recommend delaying the sound from the extra set of loudspeakers to take advantage of the Haas effect.

The Haas effect (also known as the Precedence effect) causes humans to localize sound sources in the direction of the first arriving sound despite the presence of a single reflection from a different direction. A single reflection arriving within 5–30 ms can be up to 10 dB louder than the direct sound without being perceived as a secondary auditory event.

To a listener located in sweet spot, the two extra loudspeakers will combine into a single phantom source and thus cause no violation of the single reflection requirement.

About Microphones and Echo Cancelling

The Codec SX80 and the codec built into the MX700/800 are identical in performance. If no more than 8 microphones are required, we do recommend that you stick to the 8 independent built-in echo cancelers of the codec. These are high-quality cancelers closely integrated with the internal mixer, loudspeaker sound processing and the general signal flow inside the codec.

The use of external echo cancelers will normally introduce additional delay with reduced lip-sync performance as a result. In addition, you will experience increased cost of installation and maintenance.

Note We do not recommend using local reinforcement in combination with SpeakerTrack.

Note Please do not use CE Console to set up this behavior.
Codec SX80: Use with Extra Loudspeakers

The illustration at left shows the external connections needed. Note that, for clarity, the camera connections have been omitted.

The illustration shows a configuration using passive loudspeakers and a power amplifier. Active loudspeakers (with built-in power amplifiers) may, of course, be used instead.

You may use any number of local microphones. Shown here is the maximum number the codec supports, which is 8 microphones—see also “About Microphones and Echo Cancelling” on page 3.

Your presentation source may be connected using either HDMI or DVI. In the latter case, audio must be provided separately to the codec, as shown.

For clarity, the HDMI PC-input audio alternative has been omitted in the illustration of the internal signal paths shown on the following page.
Codec SX80: Use with Extra Loudspeakers

- For the initial configuration when setting this up, we recommend the use of a sound source connected to the video system as "local presentation outside a call". Make sure you use a signal with a representative content and signal level, as well as sufficient dynamic range. Recorded speech should be a part of this signal.

- We recommend that level setting on the audio output connectors is kept at its default value (-10). The extra loudspeakers’ level should then be set as suitable, using the loudspeakers’ own gain controls. This will leave a range of -14 dB to +10 dB available for further adjustments, which will be needed if compensation for many people present in the room calls for a different level.

- Define a sweet spot (normally the position where things should sound the best).

- Measure the distance between the sweet spot and all of the loudspeakers, as shown at right.

- You should now introduce a delay to the sound from each of the extra loudspeakers. This delay will correspond to the distances \( l_1 - l_2 \) and \( l_3 - l_4 \). As a rule of thumb, use 3 ms delay per meter.

- A further 5-30 ms of delay should be added to the calculated delay difference to take advantage of the Haas effect—see page 3 for more on this. The exact amount of delay is difficult to predict, try listening to the entire system while adjusting the delay. Use the sweet spot. Ideally, you should hear a rich sound without being very much aware of the presence of the extra loudspeakers, while at the same time it should be clearly noticeable when you switch off the extra loudspeakers.

- This delay adjustment may involve output level adjustment for the extra loudspeakers, but not necessarily. Use output level in the codec API to adjust the balance between the main and extra loudspeaker sets.

- The entire setup should then be tested with people in the room, alternatively with an amount of absorption material corresponding to such a situation. This may call for further level adjustments, as outlined above.

- Be sure to make a video call to test that the far end experience has the quality required.

Please do not use CE Console to set up this behavior.

Useful API commands applicable to the scenarios described in this document can be found in "Appendix: Relevant API Commands" on page 12.
Codec SX80: Expanding Scenario to Include Local Presenter

Connecting to the Codec

The illustration at left shows the external connections needed when expanding the scenario to include a local presenter. The only difference from the original scenario is the introduction of a wireless microphone and its base station (shown in green). The use of a wireless microphone is not mandatory, wired microphones may also be used. In addition, you may also use any of the microphone inputs for the local presenter. If you need more than one for this, use the number needed, just observe the danger of feedback if care has not been taken (see the following page for more on this).

Note that, for clarity, the camera connections have been omitted.

The illustration shows a configuration using passive loudspeakers and a power amplifier. Active loudspeakers (with built-in power amplifiers) may, of course, be used instead.

You may use any number of local microphones. Shown here is the maximum number the codec supports, which is 8 microphones—see also “About Microphones and Echo Cancelling” on page 3.

In the illustration, the wireless microphone is connected to Input No.1, but you may connect it to any of the 8 inputs.

Your presentation source may be connected using either HDMI or DVI. In the latter case, audio must be provided separately to the codec, as shown.

For clarity, the HDMI PC-input audio alternative has been omitted in the illustration of the internal signal paths shown on the following page.
Codec SX80: Expanding Scenario to Include Local Presenter

- We assume that you have already set up your SX80 to work with a set of extra loudspeakers, as described on the previous page.
- Use a dedicated (and closely mounted) microphone for the presenter, preferably of the wearable type.

This will ensure high quality audio with low loudness variation from the presenter and low sensitivity towards other audio sources. It is important that the presenter microphone does not pick up much of the audio from the loudspeakers, when local reinforcement of the presenter audio is used.

This serves to avoid strong reverberation, coloration and in the worst case feedback, of the presenter audio.

Other microphones in the room, e.g. ceiling mounted microphones may still be operating, but not as a substitute for the personal microphone.
- The wearable microphone will typically be of the wireless type with a base station.

In order to avoid clipping of the codec’s microphone input, you may want to set the base station output to Mic out, rather than Line out.

If you do so, make sure the base station is not suffering damage from the phantom voltage supplied by the codec. See also Appendix for more.
- Another strategy will be to use Line out of the base station and reduce the input sensitivity of the codec accordingly.

The codec full scale level is 22dBu. If you have a base station supplying e.g. 18dBu, you should not set the codec input level higher than 4.

Each step of the codec input level setting corresponds to 1dB gain. For more on this, consult the SX80 Physical Interface Guide.

- Furthermore, we do recommend that any kind of processing in the base station is avoided. We are here thinking of such things as dynamic EQ, compression, noise reduction, AGC, de-essing etc. These types of time-varying and/or non-linear processing may be detrimental to the codec’s echo cancelling performance. Fixed EQ, however, should cause no problem.

The input signal level from the wearable microphone should be checked using the VU-meter feature of the codec. Green with occasional appearances of yellow is considered good.

Note: We do not recommend that you combine local reinforcement and SpeakerTrack.

- Adjust the gain settings so that the codec input levels are the same for all microphone inputs, for a given sound level.

Should you now need to adjust the sound level, use the following guidelines:
- Use MicrophoneReinforcement Gain to amplify the local presenter.
- Use Master Volume to amplify PC-input audio and far-end audio. The local presenter will not be affected by this.
- You may want to fine-tune the delay at this stage.
- The entire setup should then be tested with people in the room, alternatively with an amount of absorption material corresponding to such a situation.

This may call for further level adjustments, as outlined above.
- Be sure to make a video call to test that the far end experience has the quality required.

Useful API commands applicable to the scenarios described in this document can be found in “Appendix: Relevant API Commands” on page 12.

How to navigate to the VU-meters of the codec is explained on the same page.

In the below example two microphones are used for the local reinforcement, feel free to use any number of the 8 microphones.

Note: We do not recommend that you combine local reinforcement and SpeakerTrack.
MX700/800: Use with Extra Loudspeakers

The illustration at left shows the external connections needed. Note that, for clarity, the camera connections have been omitted.

The illustration shows a configuration using passive loudspeakers and a power amplifier. Active loudspeakers (with built-in power amplifiers) may, of course, be used instead.

You may use any number of local microphones. Shown here is the maximum number the MX700/800 supports, which is 8 microphones—see also “About Microphones and Echo Cancelling” on page 3.

Your presentation source may be connected using either HDMI or DVI. In the latter case, audio must be provided separately to the endpoint, as shown.

For clarity, the HDMI PC-input audio alternative has been omitted in the illustration of the internal signal paths shown on the following page.
MX700/800: Use with Extra Loudspeakers

- For the initial configuration when setting this up, we recommend the use of a sound source connected to the video system as “local presentation outside a call”. Make sure you use a signal with a representative content and signal level, as well as sufficient dynamic range. Recorded speech should be a part of this signal.

- We recommend that level setting on the audio output connectors is kept at its default value (-10)

The extra loudspeakers’ level should then be set as suitable, using the loudspeakers’ own gain controls.

This will leave a range of \(-14\) dB to \(+10\) dB available for further adjustments, which will be needed if compensation for many people present in the room calls for a different level.

- Define a sweet spot (normally the position where things should sound the best).

- Measure the distance between the sweet spot and both of the extra loudspeakers, as well as the distance between the sweet spot and the video system, as shown at right.

- You should now introduce a delay to the sound from each of the external loudspeakers.

As a rule of thumb, use 3 ms delay per meter.

A further 5–30 ms of delay should be added to the calculated delay difference to take advantage of the Haas effect—see page 3 for more on this.

The exact amount of delay is difficult to predict, try listening to the entire system while adjusting the delay. Use the sweet spot. Ideally, you should hear a rich sound without being very much aware of the presence of the external loudspeakers, while at the same time it should be clearly noticeable when you switch off the external loudspeakers.

This delay adjustment may involve output level adjustments, but not necessarily.

- Use Output level in the codec API to adjust the balance between the integrated and extra loudspeakers.

- The entire setup should then be tested with people in the room, alternately with an amount of absorption material corresponding to such a situation.

This may call for further level adjustments, as outlined above.

- Be sure to make a video call to test that the far end experience has the quality required.

Useful API commands applicable to the scenarios described in this document can be found in “Appendix: Relevant API Commands” on page 12.

Please do not use CE Console to set up this behavior.
MX700/800: Expanding Scenario to Include Local Presenter

The illustration at left shows the external connections needed when expanding the scenario to include a local presenter. The only difference from the original scenario is the introduction of a wireless microphone and its base station (shown in green). The use of a wireless microphone is not mandatory, wired microphones may also be used. In addition, you may also use any of the microphone inputs for the local presenter. If you need more than one for this, use the number needed, just observe the danger of feedback if care has not been taken (see the following page for more on this).

Note that, for clarity, the camera connections have been omitted.

The illustration shows a configuration using passive loudspeakers and a power amplifier. Active loudspeakers (with built-in power amplifiers) may, of course, be used instead.

You may use any number of local microphones. Shown here is the maximum number the MX700/800 supports, which is 8 microphones—see also “About Microphones and Echo Cancelling” on page 3.

In the illustration, the wireless microphone is connected to Input No.1, but you may connect it to any of the 8 Inputs.

Your presentation source may be connected using either HDMI or DVI. In the latter case, audio must be provided separately to the endpoint, as shown.

For clarity, the HDMI PC-input audio alternative has been omitted in the illustration of the internal signal paths shown on the following page.
MX700/800: Expanding Scenario to Include Local Presenter

- We assume that you have already set up your MX700/800 to work with a set of extra loudspeakers, as described on the previous page.
- Use a dedicated (and closely mounted) microphone for the presenter, preferably of the wearable type.

This will ensure high quality audio with low loudness variation from the presenter and low sensitivity towards other audio sources. It is important that the presenter microphone does not pick up much of the audio from the built-in loudspeakers of the MX700/800 or the extra loudspeakers, when local reinforcement of the presenter audio is used.

This serves to avoid strong reverberation, coloration and in the worst case feedback, of the presenter audio.

- The input signal level from the wearable microphone should be checked using the VU-meter feature of the codec. Green with occasional appearances of yellow is considered good. Note that there is no peak hold, so testing by listening will be needed.

If you do so, make sure the base station output is set to Mic out, rather than Line out.

- The wearable microphone will typically be of the wireless type with a base station.

In order to avoid clipping of the codec’s microphone input, you may want to set the base station output to Mic out, rather than Line out.

If you do so, make sure the base station is not suffering damage from the phantom voltage supplied by the codec. See also Appendix for more.

- Another strategy will be to use Line out of the base station and reduce the input sensitivity of the codec accordingly.

The codec full scale level is 22 dBu. If you have a base station supplying e.g. 18 dBu, you should not set the codec input level higher than 4.

Each step of the codec input level setting corresponds to 1 dB gain. For more on this, consult the SX80 Physical Interface Guide.

- Furthermore, we do recommend that any kind of processing in the base station is avoided. We are here thinking of such things as dynamic EQ, compression, noise reduction, AGC, de-essing etc. These types of time-varying and/or non-linear processing may be detrimental to the codec’s echo cancelling performance. Fixed EQ, however, should cause no problem.

- The input signal level from the wearable microphone should be checked using the VU-meter feature of the codec. Green with occasional appearances of yellow is considered good. Note that there is no peak hold, so testing by listening will be needed.

If you do so, make sure the base station output is set to Mic out, rather than Line out.

If you do so, make sure the base station is not suffering damage from the phantom voltage supplied by the codec. See also Appendix for more.

- Another strategy will be to use Line out of the base station and reduce the input sensitivity of the codec accordingly.

The codec full scale level is 22 dBu. If you have a base station supplying e.g. 18 dBu, you should not set the codec input level higher than 4.

Each step of the codec input level setting corresponds to 1 dB gain. For more on this, consult the SX80 Physical Interface Guide.

- Furthermore, we do recommend that any kind of processing in the base station is avoided. We are here thinking of such things as dynamic EQ, compression, noise reduction, AGC, de-essing etc. These types of time-varying and/or non-linear processing may be detrimental to the codec’s echo cancelling performance. Fixed EQ, however, should cause no problem.

- The input signal level from the wearable microphone should be checked using the VU-meter feature of the codec. Green with occasional appearances of yellow is considered good. Note that there is no peak hold, so testing by listening will be needed.

If you do so, make sure the base station output is set to Mic out, rather than Line out.

- The wearable microphone will typically be of the wireless type with a base station.

In order to avoid clipping of the codec’s microphone input, you may want to set the base station output to Mic out, rather than Line out.

If you do so, make sure the base station output is set to Mic out, rather than Line out.

- The wearable microphone will typically be of the wireless type with a base station.

In order to avoid clipping of the codec’s microphone input, you may want to set the base station output to Mic out, rather than Line out.

If you do so, make sure the base station output is set to Mic out, rather than Line out.

Should you now need to adjust the sound level, use the following guidelines:

- Use MicrophoneReinforcementGain to amplify the local presenter.

- Use Master volume to amplify PC-input audio and far-end audio. This will not influence the level of the local presenter.

- You may want to fine-tune the delay at this stage.

- Finally, the entire setup should be tested with people in the room, alternatively with an amount of absorption material corresponding to such a situation. This may call for further level adjustments, as outlined above.

- Be sure to make a video call to test that the far end experience has the quality required.

Please do not use CE Console to set up this behavior.

Useful API commands applicable to the scenarios described in this document can be found in “Appendix: Relevant API Commands” on page 12.

How to navigate to the VU-meters of the codec is explained on the same page.

Note We do not recommend that you combine local reinforcement and SpeakerTrack.

In the above example two microphones are used for the local reinforcement. This is just an example, feel free to use any number of the 8 microphones.
Appendix: Relevant API Commands

Sign in to the web interface and navigate to Setup > Configuration. For more API commands, see the API guide for your system.

Audio MicrophoneReinforcement Input Microphone [1..8] Mode
**Description:** Set the microphone reinforcement mode individually on each microphone. The signals from all microphones with Mode = On will be mixed and fed to the selected MicrophoneReinforcement outputs. Also consult the Audio MicrophoneReinforcement Output Line Mode setting.

**Valuespace:** On/Off (default)
- **On:** The microphone signal will be fed to the selected MicrophoneReinforcement outputs, as well as to the far end.
- **Off:** The microphone signal will be sent only to the far end. It will not be fed to the selected MicrophoneReinforcement outputs.

Audio MicrophoneReinforcement Output Line [1..4] Mode
**Description:** Set the microphone reinforcement mode individually on each line output. If Mode = On, the line output will add the microphone reinforcement mix to its ordinary output signal. Also consult the Audio MicrophoneReinforcement Input Microphone Mode setting.

**Valuespace:** On/Off (default)
- **On:** This output will deliver far-end audio, local presentation audio and the microphone reinforcement mix.
- **Off:** This output will deliver far-end audio and local presentation audio.

Audio MicrophoneReinforcement Gain
**Description:** The gain (in dB) that will be applied to the mixed microphone signal that is fed to the selected MicrophoneReinforcement outputs.

**Valuespace:** Integer (-54..15); default: -54
Use this gain to adjust the level of microphone signal that should be fed to the local output. The value -54 means “Off”, and no signal will be fed from the microphone to the output.

Audio Input Microphone [1..8] Type
**Description:** Set the specified input as Line or Microphone. When set to Line the phantom voltage feeding will be turned off.

**Valuespace:** Line/Microphone

Audio Output Line [1..4] Delay DelayMs
**Description:** Set the delay to a specified value (milliseconds)

**Valuespace:** Integer (0..290)

Audio Output Line [1..4] Level
**Description:** Output gain in decibels.

**Valuespace:** Integer (-24..0)