audio-technica

ATND1061LK
ATND1061DAN

Beamforming Array Microphone

Voice Lift Function

Ver. 1.0

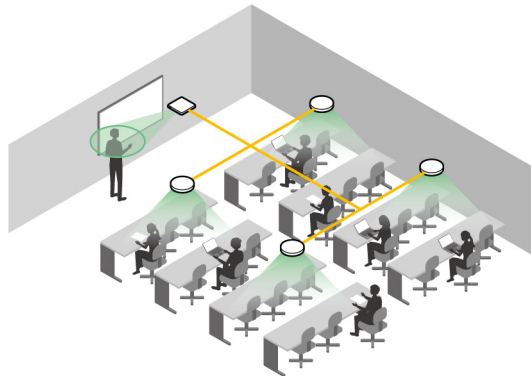# What is "voice lift"?

Introduction

Voice lift is a general term for solutions that enable people in the same room to hear voices uniformly and clearly, and is necessary for enabling natural conversations among multiple people.

Although the ATND1061 is designed to be used for web conferencing, it has also been increasingly used in hybrid conferencing, which has become more commonplace in recent years with meetings split between in-person and remote attendees.. Generally speaking, it is difficult to hear natural voices when approximately 7 m (23') away from the talker, and in a large room where the use of multiple ATND1061 units is required, communication among multiple people solely using their natural voices can be challenging. To solve this problem, we have added a new feature to provide voice lift functionality.

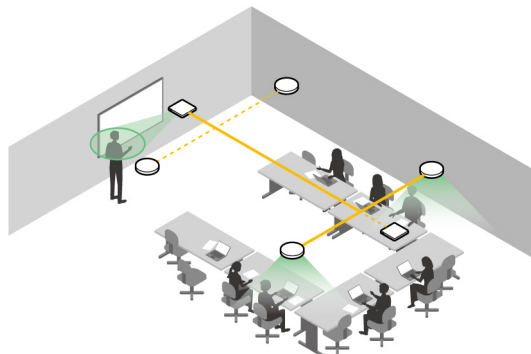## Differences with general sound reinforcement functions

General sound reinforcement

The main purpose of general sound reinforcement functions is to transmit the sound of one person's voice to multiple people. This is why loud speakers are used to deliver clear audio at a sufficient volume even to people located far away from the speakers.

Voice lift

The purpose of the voice lift function is to create an environment where multiple people can communicate clearly using their natural voices. By enabling the volume in the room to remain uniform, people can feel as if they're speaking to each other from a close distance, even though they're actually far away.
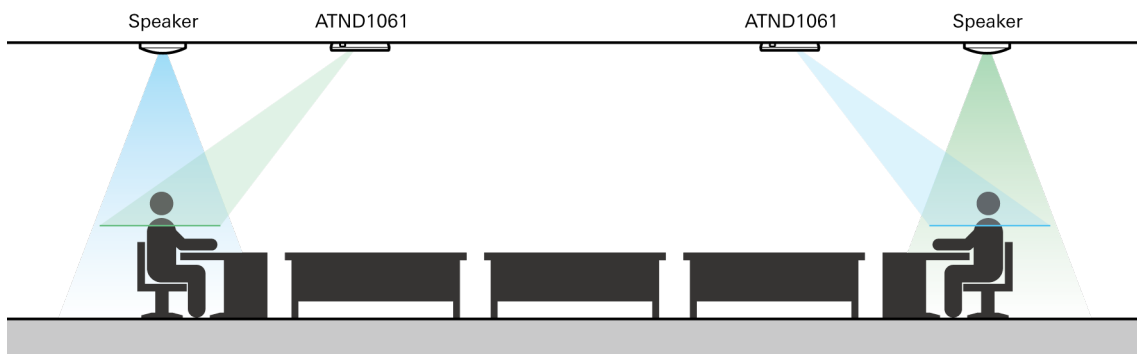
Room conditions

Using the voice lift function enables hybrid conferences to be held even in large rooms that require the use of multiple microphones. It is said that in a typical meeting, it becomes difficult to hear natural voices from a distance of 7 m (23') away from the talker, which is why the voice lift solution was designed to be effective in large rooms of more than 10 m (33') on a side.

- The voice lift function may be unnecessary in small rooms where it is unnecessary to use speakers for sound reinforcement. For this reason, we do not recommend using voice lift for small rooms. It can also be difficult to use voice lift in rooms with high noise levels or excessively long reverberation times.

- Because low-range frequencies in particular spread easily (as they are difficult to attenuate) and have low directivity, a high-pass filter (HPF) may be used to effectively suppress feedback. For this reason, the ATND1061 features a 300 Hz HPF by default in voice lift mode.

- The appropriate frequency will depend on the room environment.

Zoning

Zoning is a method used to increase the clarity of the entire room by emitting sound from speakers located far away from the talker. This method involves dividing a space into two or more zones and outputting the talker's voice only from speakers located far away from the talker.

Calculating PAG/NAG

Calculating PAG/NAG is helpful for voice lifting without causing system feedback.
NAG (Needed Acoustic Gain): Gain required by the system
PAG (Potential Acoustic Gain): Gain that can be increased without causing feedback

Obtaining PAG and NAG

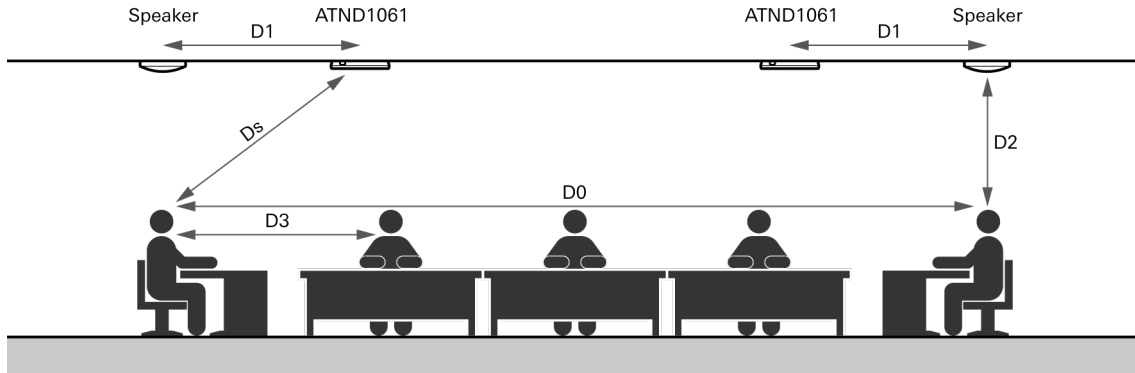PAG = 20*log((D0*D1)/(D2*Ds))-10*log(NOM)–FSM
NAG = 20*log(D0/D3)
FSM = Feedback Stability Margin
NOM = Number of Open Microphones

- When "PAG − NAG > 0", the system is stable. However, when the PAG is less than the NAG, there may be lack of clarity, or feedback may occur.

- These calculations assume omnidirectionality. An additional 4 to 6 dB of margin can be obtained by using the voice lift function of the ATND1061.

Example of application in an actual room

If calculated using the numerical example as a reference, PAG − NAG = −2.6 dB. Because the voice lift function of the ATND1061 provides a margin of approximately 5 dB, this results in the calculation "−2.6 dB + 5 dB = 2.4 dB", which satisfies "PAG − NAG > 0".

Item	Description	Numerical example
D0	Distance between the talker and the person farthest from the talker	8
D1	Distance between the ATND1061 and the speaker	2.5
D2	Distance between the person farthest from the talker and the speaker	1.8
D3	Distance between the talker and the person closest to the talker	3
Ds	Distance between the talker and the ATND1061	2
FSM	Feedback stability margin	6
NOM	Number of open microphones	2

When the talker(s) and speaker are near each other

In the example above, if the talker seated in the middle is also expected to speak, not only does D0 become smaller, but the microphone will also be directed toward the speaker, resulting in a reduced margin and increased risk of feedback. This situation would then require the following measures to be taken (specific measures will differ depending on the actual environment).

- Increase the distance from the microphone to the speaker.
- Exclude the talker seated in the middle from the microphone pickup area because that talker is sufficiently audible even to the person seated farthest away and does not require voice lifting.
- Apply Processing Levels 3 to 5.

Unique technology for suppressing feedback

A new 7th beam is used internally for voice lifting. Unlike normal sound pickup beams, this beam is different in that it moves within the entire sound pickup range (excluding Exclusion Zones, i.e., Coverage Zones and Priority Zones). Once focused on a talker, it stays there for 70 ms before moving on to the next talker. The movement of the beam can be changed to suppress feedback.

Additionally, because voice lifting requires that any sound collected by a microphone is output from a speaker located in the same room, feedback is more likely to occur. In particular, in the case of speaker-tracking beamforming microphones such as the ATND1061, the distance between the beam and speaker may become closer depending on the location of the talker, further increasing the risk of feedback. To address this problem, Audio-Technica has developed its proprietary AFBC (Adaptive Feedback Canceller) algorithm to create a voice lift solution that is resistant to feedback.
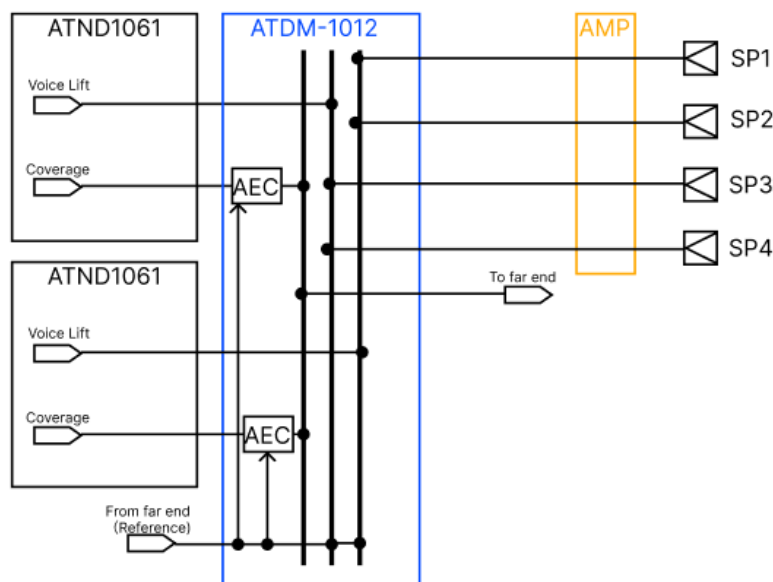
- This feedback-suppression algorithm uses an adaptive filter to prevent feedback even when a speaker is located nearby the talker. This algorithm enables the ATND1061 to be used in situations where talkers speak in front of speakers, such as in seminars.

- Furthermore, the ATND1061 is equipped with notch filter and frequency shift algorithms that are commonly used to suppress feedback, and which can be configured as needed using five patterns to suit the room environment.

- If the distance between the microphone and speaker is far and the PAG/NAG calculation is sufficient, settings 1 and 2 are suitable, while in other cases where the speaker and talker are close, settings 3 to 5 are suitable.

Setting	Description	Latency
1	No DSP processing	18 ms
2	Notch filter	28 ms
3	AFBC + notch filter	33 ms
4	AFBC + notch filter + NLP	33 ms
5	AFBC + notch filter + NLP + frequency shift	33 ms

Using the voice lift function and AEC (Acoustic Echo Cancellation) in combination

When delivering audio remotely, such as via online lectures or online meetings held in a large room, the voice lift and AEC must be used in combination. Although AEC is a processing technique that prevents microphones from picking up amplified far-end audio, which would otherwise return to the far end, when used in conjunction with the voice lift, it causes further mixing with local sounds. As a result, problems such as poor AEC processing or increased reverberation may occur.

Because we do not recommend using voice lift and AEC together, AEC cannot be used in voice lift mode with the ATND1061. If their use in combination is required, AEC function of a meeting app or external device must be used. In such cases, observe the following:

- Separate the signal systems for audio to be output from the speaker and the audio to be sent to the far end.
  Example: When using the ATDM-1012 or ATDM-0604a, this can be performed as shown below.

- If AEC must be used after the voice lift functions, we recommend using Processing Levels 1 to 3. Using a level of 4 or higher would cause the echo component to also deteriorate, possibly preventing echoing from disappearing.