

VisionSuite Environmental Recommendations

Overview

Q-SYS VisionSuite allows us to gather real-time context of what's happening in a room such as who is actively speaking or where people move to in a space and provide the optimal camera shot based on these events. These looks to serve not only the far-end experience with a more engaging camera feed but also the near-end by providing automation based on where participants are located within the space.

VisionSuite consists of two main features: Speaker Spotlight and Presenter Spotlight. Speaker Spotlight leverages intelligent microphones capable of detecting the direction of arrival (DoA) of speech, allowing the system to identify and frame the active speaker dynamically. Presenter Spotlight uses advanced computer vision algorithms to recognize individuals in motion, enabling seamless camera tracking and room automation based solely on participant positioning.

Because this system relies on microphones as the ears and cameras as the eyes of the room, it is critical that environmental factors such as acoustics and lighting meet certain standards. Suboptimal room conditions can negatively affect input data accuracy, leading to degraded system performance.

All of the information below is applicable to both the new VisionSuite Designer and also the legacy VisionSuite solutions of Seervision and Automatic Camera Preset Recall (ACPR).

Environmental Recommendations

Acoustic Conditions

Acoustics are vital for both intelligibility and the accurate detection of speaker location. While audio algorithms such as Noise Reduction (NR), Acoustic Echo Cancellation (AEC), and Voice Activity Detection (VAD) can mask many acoustic imperfections, The interpretation of the DoA, provided by intelligent microphones, is a lot more sensitive to the acoustic performance of the space.

Reflections

To ensure accurate DoA data, reverberation time (RT60) should not exceed a maximum target level of 0.6-0.8s [1][5]. However, an ideal range between 0.3 to 0.5 seconds provides the best performance and is an expectation shared by leading unified communication providers [3][4][5].

Excessive reverberation can impair localization and beamforming. Ideal RT60 values are achieved through treatments such as wall-mounted acoustic panels ($NRC \geq 0.9$) [3], acoustic ceiling tiles, and carpeted flooring [1]. While larger venues like auditoriums may inherently exceed this range [1], the microphones employed remain conferencing-grade and should be treated with these acoustic expectations in mind.

Microsoft Teams suggests at least two walls should be treated and floors to be carpeted or otherwise acoustically treated [2][5].

Noise Floor

Background noise should not exceed NC-35 [1][5] or 35–45 dBA SPL [3][4] to ensure optimal microphone performance and accuracy in voice tracking, assuming a normal speaking voice is about 70dB SPL A-weighted [4]. An optimal signal-to-noise ratio to aim for is 30dB, a minimum target is 20dB [4]. Sources such as HVAC systems and AV equipment can mask speech or cause false triggers. We also should recognize that this is noise floor measured from the microphones perspective. The room itself may sound great, but as these microphones are installed in the ceiling, we need to consider what impact that may have on the DOA data.

Ceiling Microphone Placement

Speaker Spotlight performance depends on proper installation of ceiling microphones. Correct placement ensures full coverage and minimizes false-positive voice detection. The below is the minimum requirements from the manufacturer specification sheets and from our own testing.

- Maximum distance to active talker: 4.5m (15ft) [6][7]
- Minimum distance to active talker: 1m (3ft) [6]
- Maximum mounting height: 3m (10ft) [7]
- Minimum distance from loudspeakers: 1.2m (4ft) OR 3m (10ft) from high SPL loudspeakers. [7]

Lighting Conditions

Lighting plays a critical role in supporting computer vision accuracy for Presenter Spotlight. Best practices from leading unified communication providers and our own testing suggest vertical illumination between 400–500 lux [1][3][5] on participants' faces and 700–900 lux on horizontal work surfaces [3]. Use flicker-free, high CRI (≥ 85) LED lighting [3][5] with a 4000–5000K colour temperature [1][5] to ensure natural skin tones and prevent visual anomalies. It is recommended to target a uniformity number 0.6U0 as a minimum [5][8] but a value of >0.8 is strongly recommended [3].

Overhead-only lighting should be avoided as it creates facial shadows and affects camera exposure. Instead, use soft, frontal lighting at 30°–45° angles from the camera line of sight. Windows must be fitted with blinds or curtains to eliminate backlighting, and lighting should be evenly distributed across the space. In stage applications, ensure that the stage wash cancels any shadows caused by spotlights to prevent false vision triggers.

References

- [1] AVIXA Performance Verification Guide: <https://cdn.avixa.org/production/docs/default-source/default-document-library/avsystemsperformanceverificationguider.pdf>
- [2] Microsoft Teams Room Design Guidance: [Step 3 - Design and build a Signature Teams Rooms meeting space or update an existing space - Microsoft Teams](#)
- [3] Cisco Webex Room Design Best Practices: <https://www.cisco.com/c/dam/en/us/td/docs/telepresence/endpoint/technical-papers/workspace-best-practices.pdf>
- [4] Zoom Rooms Design Guidelines: [Zoom Rooms Design Guide](#)
- [5] Google Meet Room Design Guide: [Room Design Guide](#)
- [6] Sennheiser TCC2 Documentation: Sennheiser — [Headphones, Microphones, Wireless Systems](#)
- [7] Shure MXA920 Documentation: [MXA920 - Ceiling Array Microphone - Shure USA](#)
- [8] Glamox Meeting Room Application Guide: [Conference and meeting rooms](#)