Q-SYS

Designing for VisionSuite

Best practices for creating Intelligent High-Impact Spaces

About VisionSuite

Q-SYS VisionSuite delivers best-in-class experiences through two complementary types of intelligence:

Speaker Spotlight:

Intelligent speaker switching that dynamically points cameras at active speakers with shots created in real-time based on audio coordinate data

Presenter Spotlight:

Autonomous presenter tracking that creates responsive presentation environments thanks to computer vision identification and vision-based triggers

Leverage the full pan-tilt-zoom capabilities of NC Series cameras for optimal shot composition for any high-impact space, enabling flexibility even when room layouts change and dynamic use cases emerge.

This is all delivered through VisionSuite Designer’s unified graphical interface and powered by the VSA-100 VisionSuite AI Accelerator, so that you can deploy intelligent high-impact spaces with unparalleled simplicity.

Speaker Spotlight

Speaker Spotlight delivers front-row experiences for remote participants by automatically framing active speakers in the room.

Unlike preset-based systems that force fixed seating layouts and pre-defined shots, Speaker Spotlight creates shots in real-time based on the active speaker's location, ensuring the far-end can follow the conversation with immersive and engaging close-up shots.

Combined with Voice Activity Detection, VisionSuite mitigates false camera switches by differentiating ambient noise from speech, eliminating unnecessary distractions.

Active Speak Framing

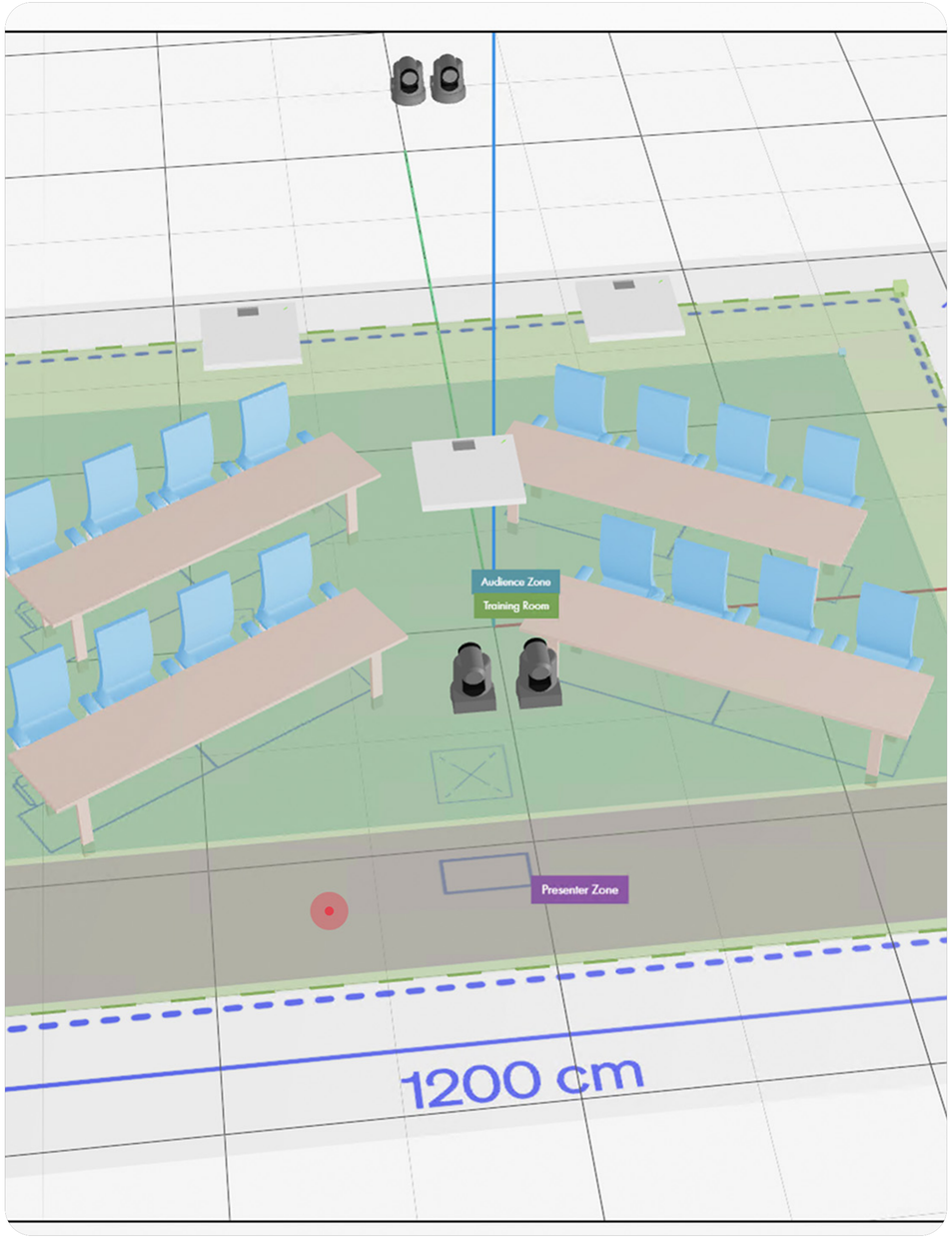
Speaker Spotlight transforms audio data into intelligent camera control through real-time coordinate processing.

Using the XYZ coordinates sent from flagship beamforming ceiling microphones, VisionSuite pinpoints speaker locations in 3D space, automatically moving cameras to frame the active speaker.

When multiple microphones are in use, VisionSuite consolidates audio data into a single unified understanding for the entire room, improving accuracy by triangulating coordinates from multiple sources.

VisionSuite then translates these coordinates into optimal camera shots, to frame the active speaker without the need for pre-configured presets.

With the addition of Voice Activity Detection, VisionSuite ensures that only speech triggers camera switching, while configurable logic allows for smooth transitions that follow the conversation naturally.



Presenter Spotlight

Presenter Spotlight enables truly hands-free presentations, following presenters autonomously.

Thanks to computer vision algorithms, VisionSuite delivers predictive and robust presenter tracking that keeps up with the presenter's movements, even in complex use cases where presenters may face away, cross each other, walk behind furniture, or move quickly.

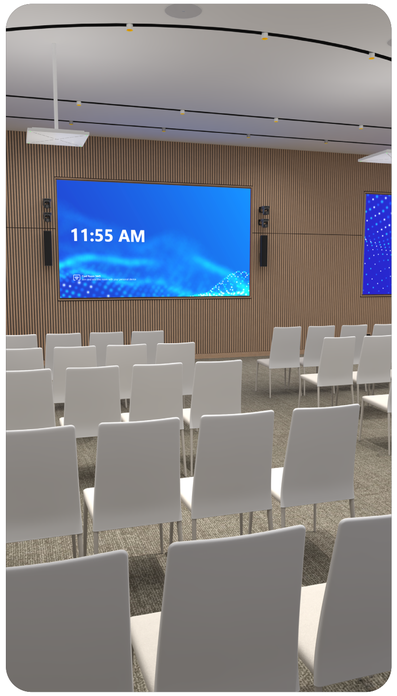
With full-body recognition and adaptive shots, VisionSuite creates unique IDs for every person based on facial features, hair, clothing, height, and more; making sure the 'VIP' is always kept in frame. Visual detections can also be fused with audio detections for advanced room awareness.

Trigger Zones and Tracking Zones enable responsive spaces that can automatically switch cameras, adjust lights, or power on displays as presenters enter and leave zones.

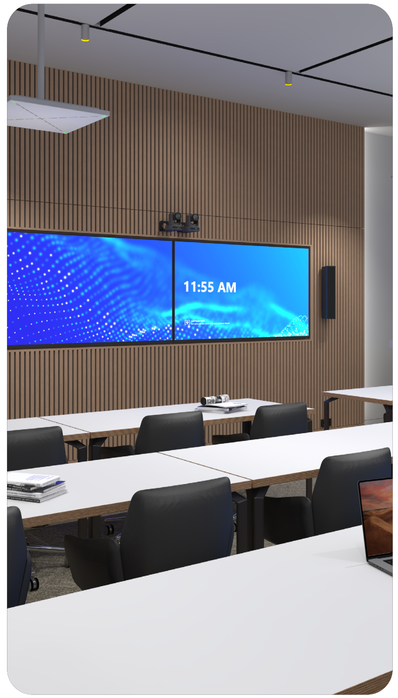
Intelligence for High-Impact Spaces

VisionSuite is perfect for medium and large high-impact spaces with requirements for multiple cameras, where closer camera shots and an immersive hybrid collaboration experience are a must-have.

Our flexible solutions can meet the needs of any complex use case and room type with requirements for speaker tracking, presenter tracking, and room automation.

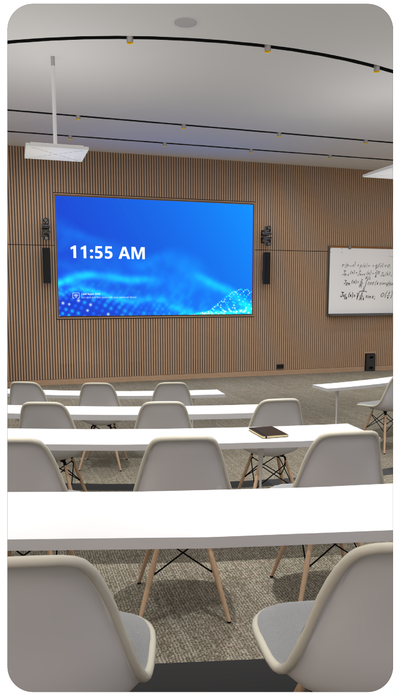


ALL-HANDS SPACES



TRAINING ROOMS

BOARDROOMS



LECTURE HALLS

DIVISIBLE ROOMS
(coming soon)

VisionSuite Quick Facts

These are the current VisionSuite specifications as of launch at v10.0.2

Number of cameras supported: 14 Speaker Spotlight and 2 Presenter Spotlight cameras

Number of ceiling mics supported: Up to 8 (MXA920 / TCC2) per room

Number of VisionSuite rooms supported: 1 VisionSuite room per Core / Q-SYS Design

Maximum presenter tracking distance: 20m / 65ft (NC20x60 at Half Body shot or wider)

Minimum camera distance: 2m / 6.5ft

Maximum recommended active speaker distance from mic: 4m / 13ft

TIP:

As the VSA-100 tends to produce a high-pitch noise and requires adequate ventilation, ensure to place the VSA-100 in a purpose-built server room instead of inside a credenza rack or lectern in the room.

Layouts and Furniture

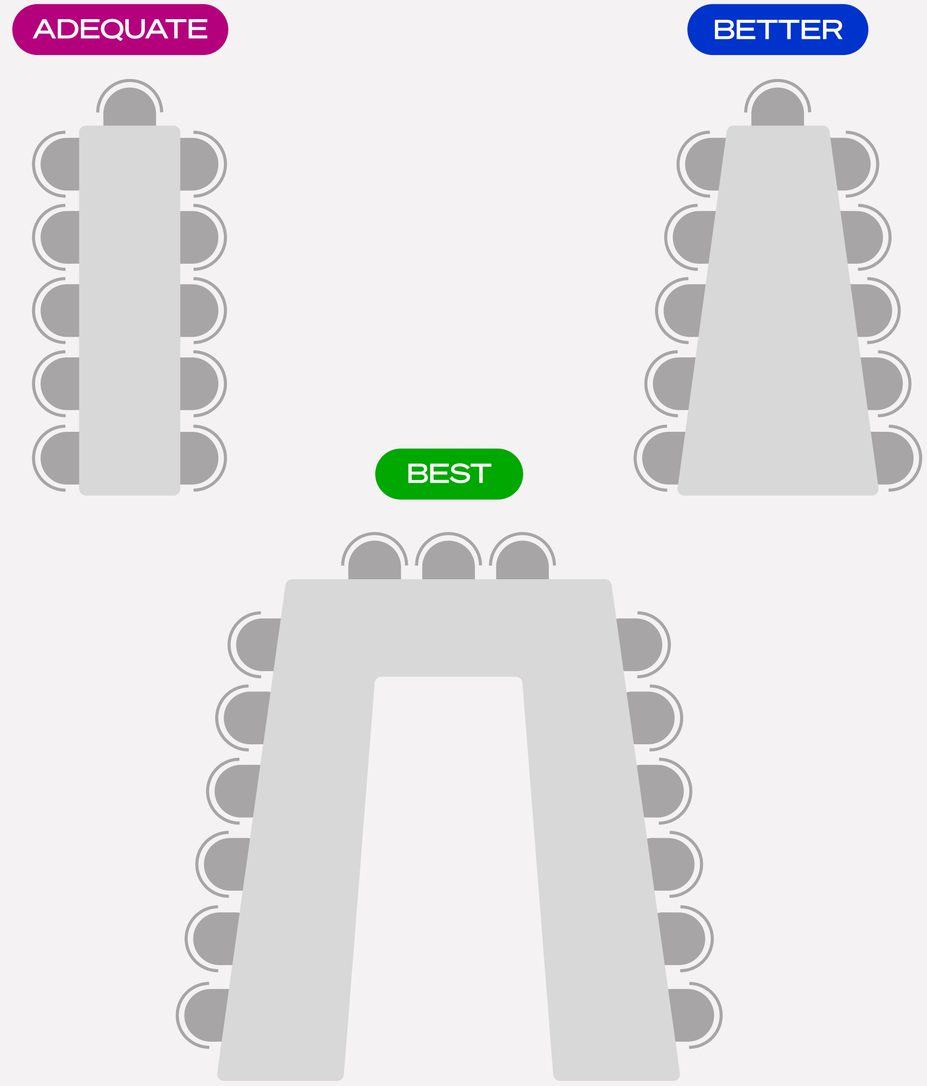
Always start by understanding the user’s needs and the use case within a given room. This will help to dictate the room dimensions and the furniture layout to optimize it for the desired experience.

In conference rooms, trapezoidal, V-shaped, or U-shaped meeting tables that allow for optimal viewing angles and better interaction with far-end participants can go a long way in aiding camera placement decisions.

These layouts help to orient participants around the screen, especially when tables are tapered to facilitate camera angles.

When choosing meeting tables, we recommend **80 cm or 30”** of seatway between each user around a table. This way, the audio coordinates of different participants will be easier to distinguish.

In training spaces, prioritize forward-facing table layouts like tiered rows or chevron arrangements that naturally direct participants toward cameras, avoiding grouped tables or circular tables that produce side-profile or back-of-head shots.



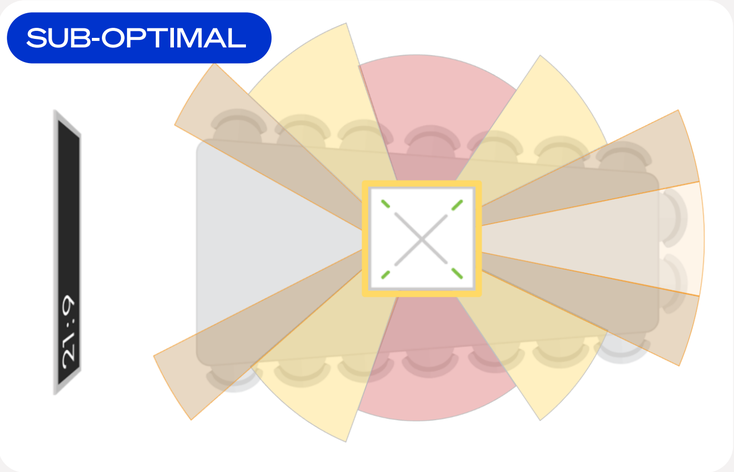
Microphone Placement

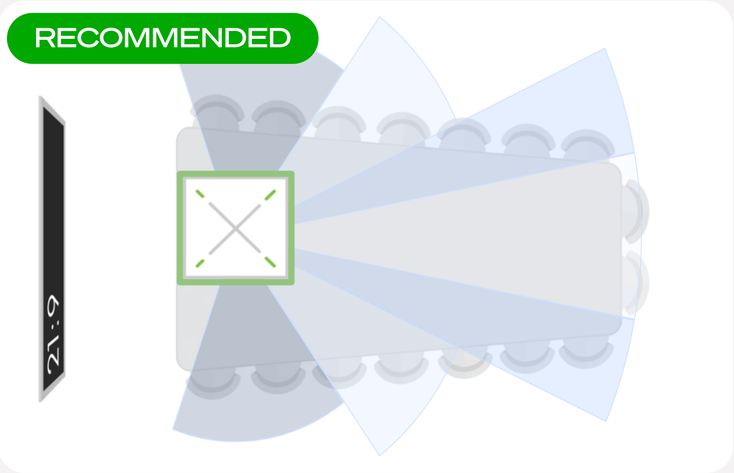
We advise seeking assistance from your chosen microphone manufacturer to help with the number of ceiling array microphones needed and their optimal placement.

The Shure MXA920 and Sennheiser TCC2 that VisionSuite supports have the same coverage area of **9 x 9 meters (30 x 30 ft).** We recommend not pushing their limits and instead being more conservative, having at least one microphone covering at most every **8 x 8 meters (26 x 26 ft)** of space.

Coverage areas that push the limits of a microphones' radius will lead to inaccuracies in the XYZ coordinates that VisionSuite receives, and much lower confidence for speaker positions.

Microphones should ideally be placed in the front third or two-thirds of the room; where in-room participants speak toward, which is usually where the displays showing the far-end are located. This also avoids situations where the nearest microphones to a talker could be behind them.





Microphone Placement

Correct microphone placement ensures full coverage and minimizes false-positive audio triggers. Note that maximum distance refers to the direct distance from the microphone to the active speaker. Higher ceiling mounting will therefore reduce the effective floor coverage area.

See below the minimum requirements from manufacturer's specification sheets and from our own testing:

Maximum ceiling mic distance to active speaker: 4.5m / 15ft

Minimum ceiling mic distance to active speaker: 1m / 3ft

Maximum ceiling microphone mounting height: 3.7m / 12ft

Minimum distance from loudspeakers: 1.2m / 4ft or 3m / 10ft from high SPL loudspeakers

Minimum distance from HVAC: 1.2m / 4ft

TIP:

As VisionSuite consolidates audio coordinates into a unified understanding when multiple microphones are used, speaker location data will be even more accurate in areas where the coverage from multiple microphones overlaps.

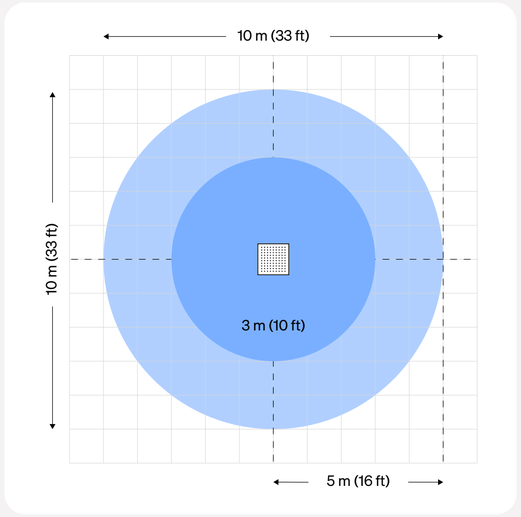
Microphone Placement

Make sure to follow the manufacturer's latest speaker tracking guidelines for improved accuracy.

SHURE [SOURCE]

If you need to frame 1 talker at a time with the room's camera, set things up so that the microphone is about 3 meters (10ft) from the different talker positions. The MXA920 performs well with talkers that are up to 4.9 meters (16ft) from the microphone, but moving closer helps room cameras with single-talker framing.

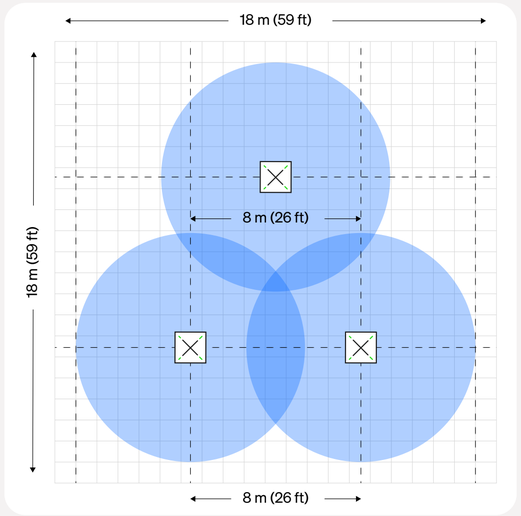
Position talkers so that they are no more than 90 degrees from facing the microphone.



SHURE

SENNHEISER [SOURCE]

For larger spaces requiring multiple TCC2 units, the distance between units should not exceed 7 to 8 meters (23 to 26 ft) to ensure proper coverage overlap. This creates approximately 1-2 meters of overlap for continuous coverage.



SENNHEISER

NOTE:

Do not mix and match microphone models in VisionSuite rooms. Mixing different models (e.g. TCC2 and MXA920 in the same room) will cause unpredictable performance due to different coordinate reporting methods.

Microphone Differences

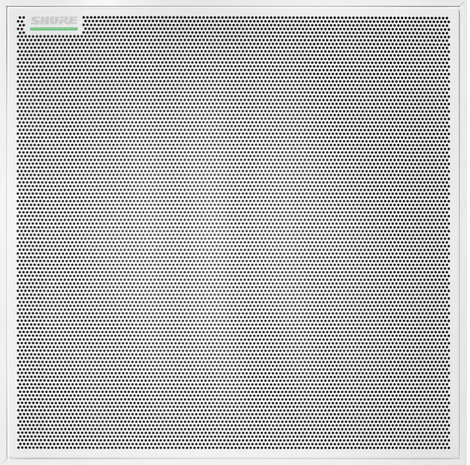
It’s important to note the differences in functionality between the MXA920 and TCC2.

The Shure MXA920 provides eight dynamic beams that can independently report speaker positions. This allows it to send multiple speaker locations simultaneously, allowing VisionSuite to keep up even in lively group discussions.

The Shure MXA920 also reports talker height data in the Z axis, helping VisionSuite to aim the cameras for standing participants speaking. To ensure greater reliability and shot consistency, we nevertheless recommend fixing the speaker height within VisionSuite Designer (even when using the MXA920) to achieve the best results.

In contrast, the Sennheiser TCC2 uses a single dynamic beam which automatically aligns to the position of the person speaking. This means that when multiple participants speak simultaneously, the TCC2 will report the position of the loudest speaker, instead of reporting multiple speaker locations.

The TCC2 also does not provide height data in the Z axis. This means that in VisionSuite Designer, users must configure a fixed height (e.g. seating height) if a vision-driven overview camera won’t be used to provide persistent speaker locations.

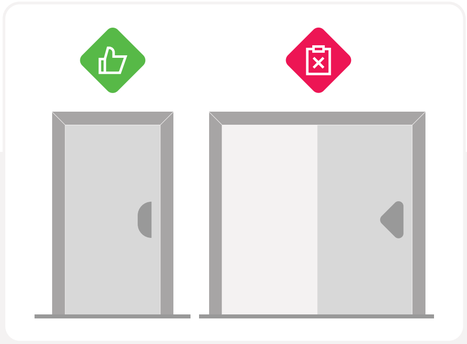


Optimizing Audio Coordinates

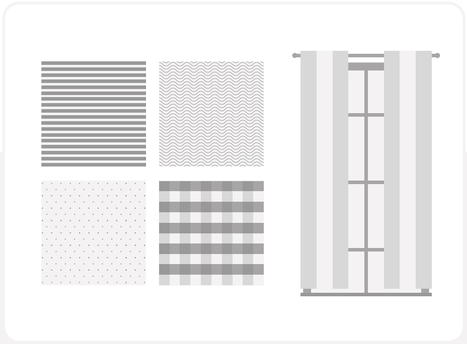
Room acoustics play a critical role in ensuring the audio location information is as accurate as possible.

The better the quality of the positional data received from the ceiling microphones, the fewer errors will occur when triangulating coordinate positions of active speakers.

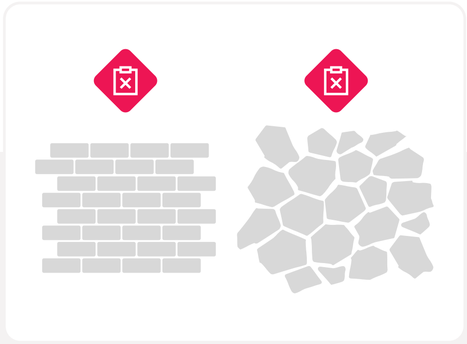
Below some recommendations to optimize reverberant rooms with echo, reflections, and other noise sources:



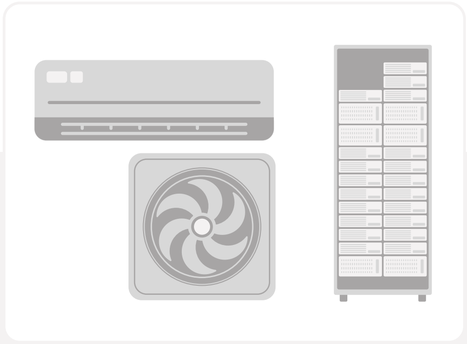
avoid sliding doors and close all windows



use acoustic panels, carpeting, and curtains



minimize acoustic reflections from glass, metal, and hard wall surfaces



minimize static noise from HVAC or server racks

NOTE:

While the VAD built into VisionSuite will differentiate between speech and noise to mitigate false switches, it will not perform noise suppression, nor will it enable VisionSuite to function reliably in environments with continuous high ambient noise levels or persistent acoustic interference.

Optimizing Audio Coordinates

The interpretation of direction of arrival (DoA) data provided by beamforming ceiling microphones is more sensitive to the acoustic performance of the space. 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.

Minimizing acoustic reflections is critical for accurate coordinate detection. Reflective surfaces interfere with precise speaker localization and should be treated appropriately. We strongly recommend engaging with an acoustical consultant to properly assess and address acoustic reflections in VisionSuite rooms.

To ensure accurate coordinate audio data, reverberation time (RT60) should not exceed a maximum target level of 0.6-0.8s. Aiming for an ideal range between 0.3 to 0.5 seconds will however provide the best performance.

Background noise should also not exceed NC-35 or 35–45 dBA SPL to ensure optimal microphone performance and accuracy in speaker tracking, assuming a normal speaking voice is about 70dB SPL A-weighted.

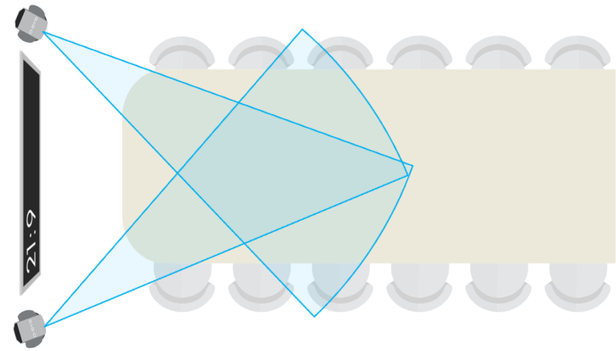
An optimal signal-to-noise ratio to aim for is 30dB, with a minimum target being 20dB. This is measured from the microphone’s perspective.

Camera Placement

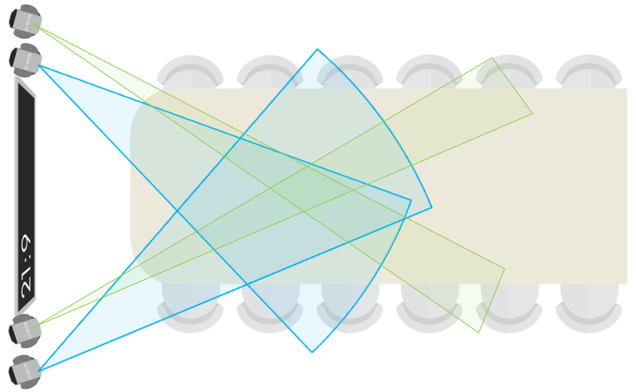
For VisionSuite Speaker Spotlight, a rule of thumb is to have at least one PTZ camera dedicated to each section of the room or meeting table. In the case of a conference room table, covering the space diagonally is recommended for achieving more frontal shots.

To eliminate distracting camera movements between adjacent speakers on the same side of the table, we recommend using at least two NC Series PTZ cameras per ‘Speaker Zone’. This enables the secondary camera to move before switching the feed, ensuring smooth transitions.

ADEQUATE



RECOMMENDED



NOTE:

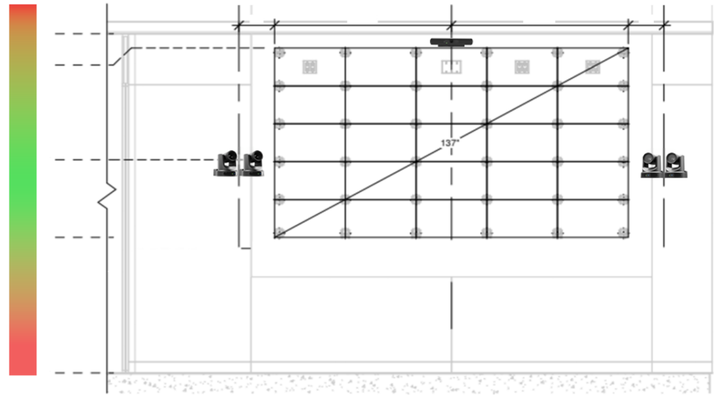
Cameras must be placed in static locations. Mounting them on furniture with wheels, or moving them in any way will undo the calibration.

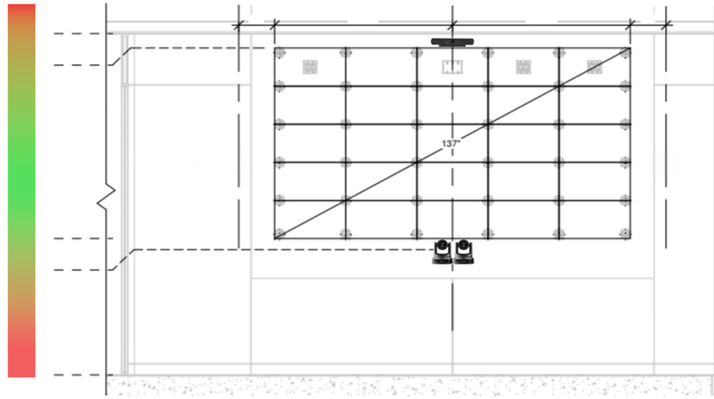
A notable exception are motorized lifts, which are fine as long as the cameras always return to the exact same position in which they were originally set up and calibrated.

Camera Placement

Placing cameras in the direction where participants look is fundamental. We recommend placing cameras below or next to displays so that participants naturally look in the direction of the cameras when they speak to the far-end participants.

Wherever practical, cameras should be placed near eye level. If placed below a display, around **120cm (47”)** is a good height. If placed at either side of a display, a height ranging between **180cm to 230cm (70”– 90”)** is recommended.





TIP:

Add an audience-facing static overview camera (e.g. NC-110) to not only provide a ‘silence’ overview shot, but to also aid VisionSuite with sensor fusion (see next page).

Ensure the camera’s FoV doesn’t overlap with the presenter area as to not occlude audience detections with the presenter’s head (if the overview camera is not high enough). We recommend to ceiling mount the overview camera directly above the presenter area or even slightly in front for optimal performance.

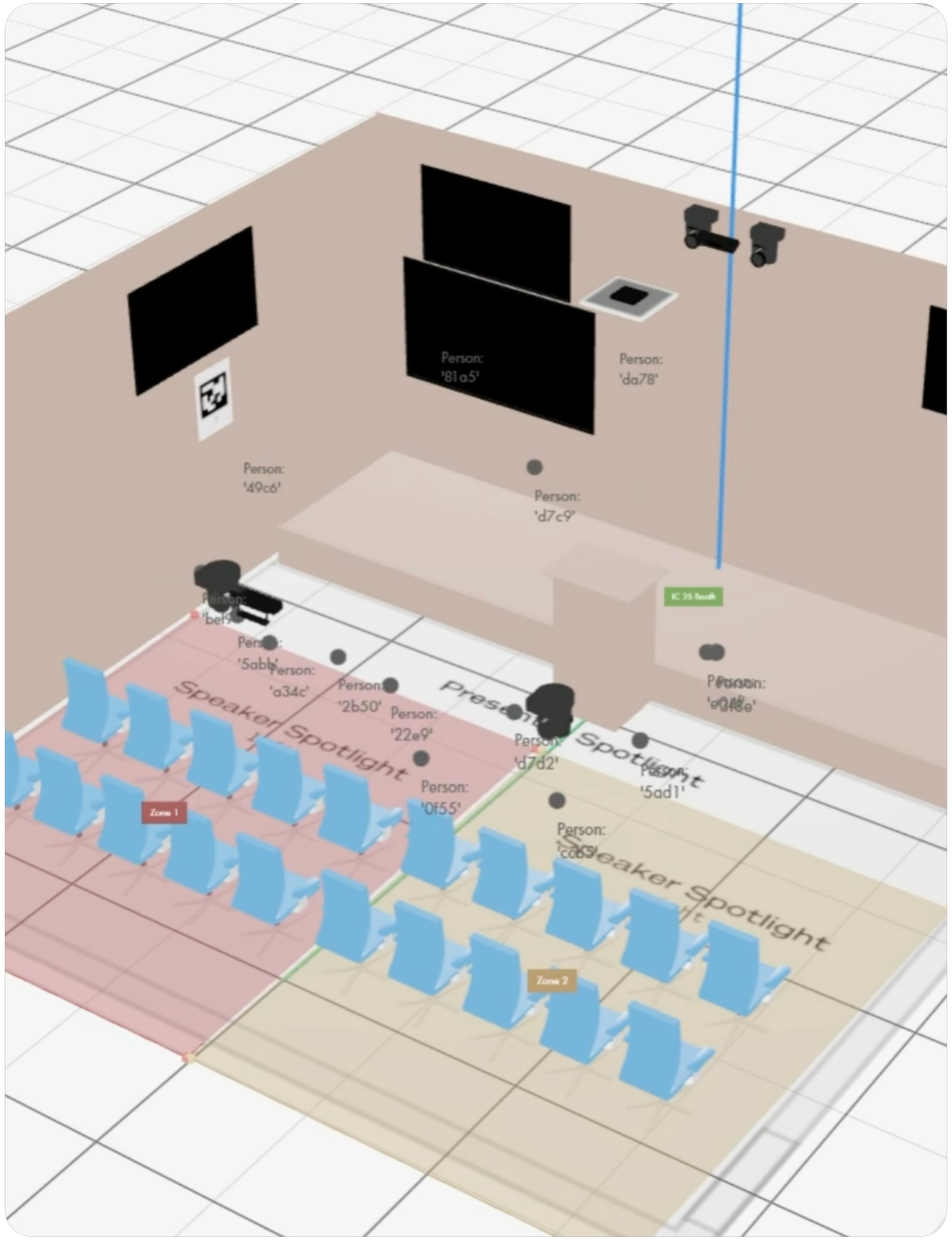
Enhancing Speaker Spotlight

We strongly recommend using an audience-facing vision-driven overview camera in Speaker Spotlight applications.

This dedicated camera will aid in improving room awareness, providing more contextual understanding of participant's locations within the room, and enhancing speaker location confidence when fused with the audio coordinate data. There will be a noticeable difference in Speaker Spotlight performance of systems with and without a vision-driven overview camera for the audience.

Using Speaker Spotlight without a vision-driven overview camera will mean that when a person stops talking, their detection will eventually disappear in VisionSuite, forgetting their precise location and producing less centered shots and more frequent camera switches.

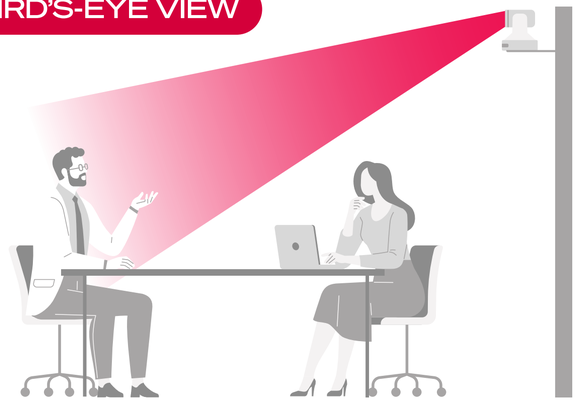
However, such an overview camera is only practical in rooms where participants are no more than **7m / 23ft** away (e.g. if using an NC-110). For larger spaces not already using Presenter Spotlight cameras, consider deploying two ceiling-mounted audience-facing overview cameras; one at the front, and one towards the middle of the room to extend the effective coverage range.



Camera Placement

It is important to strike the right balance in terms of camera height. The higher your cameras are mounted, the more unnatural their perspective will look. However, cameras need to be high enough to provide clear lines of sight for unobstructed views of participants.

BIRD'S-EYE VIEW



TOO LOW

OPTIMAL PLACEMENT

NOTE:

If there will be a presentation area at front of a room, it is best to avoid:

- Positioning audience-facing cameras at hip-height to avoid the presenter blocking the camera's view when presenting
- Installing floor-to-ceiling displays, mirrors, or LED walls behind presenters to avoid false detections
- Having glass walls or open areas behind the presenter where background detections could interfere with tracking

Camera Roles

VisionSuite supports different camera roles within VisionSuite Designer depending on the automation functionality desired. Understanding each camera role is therefore essential to delivering a room's requirements.

Vision-Driven Cameras:
Their logic is controlled based on computer vision identification and/or visual event triggers. These cameras can host Trigger Zones, Exclusion Zones, and Tracking Zones.

Tracking Camera:
Follows a presenter autonomously using full-body recognition and predictive control

Overview (Conductor) Camera:
Static camera that monitors a dedicated area, aiding the tracking camera by 'telling it' what's outside of its periphery (e.g. to help deliver multi-presenter automations) or aiding Speaker Spotlight with room awareness by providing visual context of participant's locations.

Audio-Driven Camera:
Their movement and switching is triggered by audio activity in Speaker Zones.

Speaker Camera:
Points at active speakers dynamically using XYZ audio coordinates from ceiling mics

Overview Camera:
Provides a fixed shot of participants during quiet periods or when the far-end speaks

---

Static Cameras:
Can be triggered by vision events or audio events, and are not powered by any intelligence.

Content Camera:
Dedicated static camera used for close-up shots of whiteboards or other content

Camera Placement

When placing cameras for presenter tracking with Presenter Spotlight, ideally they should be centered on the presentation area, and not skewed to one side. This will help to provide frontal shots and simplify Trigger Zone, Tracking Zone and Exclusion Zone shapes. Also make sure that no columns, pillars, or large furniture are in the way, blocking the view of wherever the presenter might move.

Our recommended camera distances for accurate computer vision identification are below:

	Recommended Distance for Presenter Tracking (HALF BODY SHOT OR WIDER)	Recommended Distance for Overview/Conductor (NO OPTICAL ZOOM)	Maximum Distance for Computer Vision
NC-12x80	11-15m / 36-49ft	14-16m / 46–52ft	25m / 82ft at 100% zoom
NC-20x60	15-20m / 49-65ft	18-20m / 59–65ft	38m / 125ft at 100% zoom
NC-110 (CONDUCTOR CAMERA ONLY)	N/A	3-7m / 10–23ft	7m / 23ft at 100% zoom

TIP:
VisionSuite supports software-definable dynamic camera roles across configurations. This means a camera's role can change from audio-driven to a vision-driven tracking camera when the room config changes.

Lighting Essentials

VisionSuite's computer vision technology works reliably in standard office lighting conditions. However, following these recommendations will maximize camera image quality and optimize the far-end experience.

Illumination

400-500 lux vertical illumination on participants' faces

Color Temperature

4000-5000K for natural appearance

Color Rendering

CRI ≥ 85 for accurate skin tones

Uniformity

Target 0.6 Uo for consistent coverage across the space

Choose flicker-free LED lighting to prevent visual artifacts. Maintaining uniformity ensures all participants appear evenly lit, eliminating distracting bright/dark spots that can affect auto-framing and tracking performance.

Also try to avoid mixing different light sources (i.e. fluorescent lights and LEDs) as their color profiles may vary, and it could diminish the camera's ability to accurately represent colors.

Optimizing Image Quality

To optimize the camera's performance, we also recommend that shades are used on all external-facing windows to stabilize the room's light levels and prevent backlighting or glare, which can affect the image quality.

Bear in mind that current generation of NC Series cameras do not support HDR. As a result, cameras may experience worse tracking performance in front of bright displays or LED walls. Users must decide whether to expose for the presenter's face or expose for the display content, but not both simultaneously.

Consider also using curtains or frosting the glass on any internal windows that may face corridors in order to mitigate Auto Framing inconsistencies, avoid distractions for the far-end, and improve the overall acoustics.

In stage applications, ensure that the stage wash cancels any shadows caused by spotlights to prevent false vision detections or triggers that may not be addressed with Exclusion Zones.

Limitations of VisionSuite

At v10.0.2 VisionSuite does not support:

More than one VSA-100 per Q-SYS Design or more than one VisionSuite Designer component per Q-SYS Design

More than one VisionSuite room per Q-SYS Design (e.g. divisible rooms)

Divisible Rooms or multiple VisionSuite rooms per design can be delivered via the legacy VisionSuite plugins (ACPR and Seervision)

Other microphones beyond the MXA920 or TCC2

e.g. Sennheiser TCCM, AudioTechnica 1061, Discrete Mics, etc are not supported

These are still compatible with the Automatic Camera Preset Recall (ACPR) plugin

SVs1 and SVs4 are not compatible at 10.0 with VisionSuite Native

In the meantime, the legacy VisionSuite plug-ins must be used

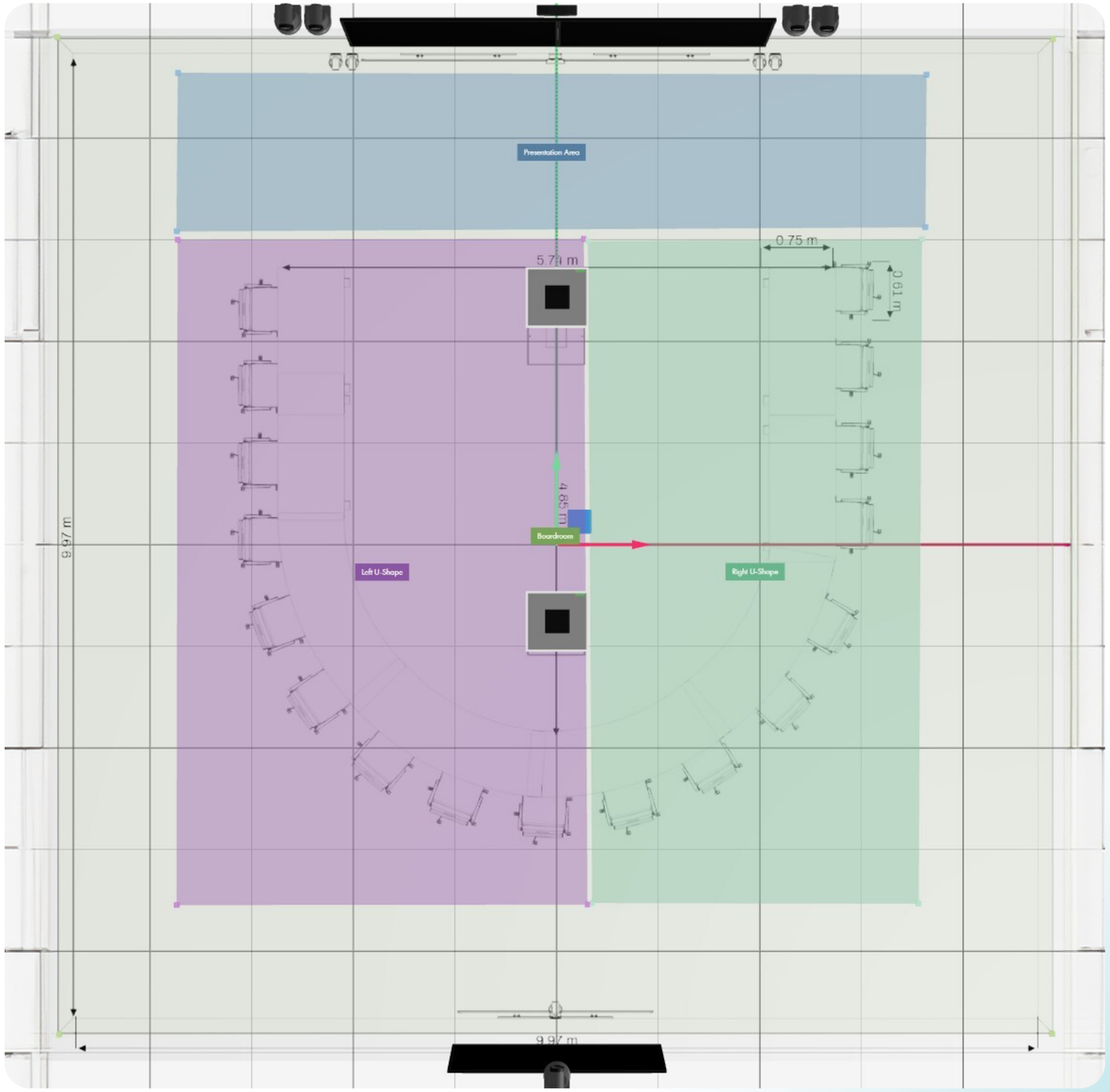
VisionSuite does not support any privacy or auto privacy use cases yet

VisionSuite does not directly support third-party NDI, HDMI, or SDI switchers or cameras.

EXAMPLE

Boardroom

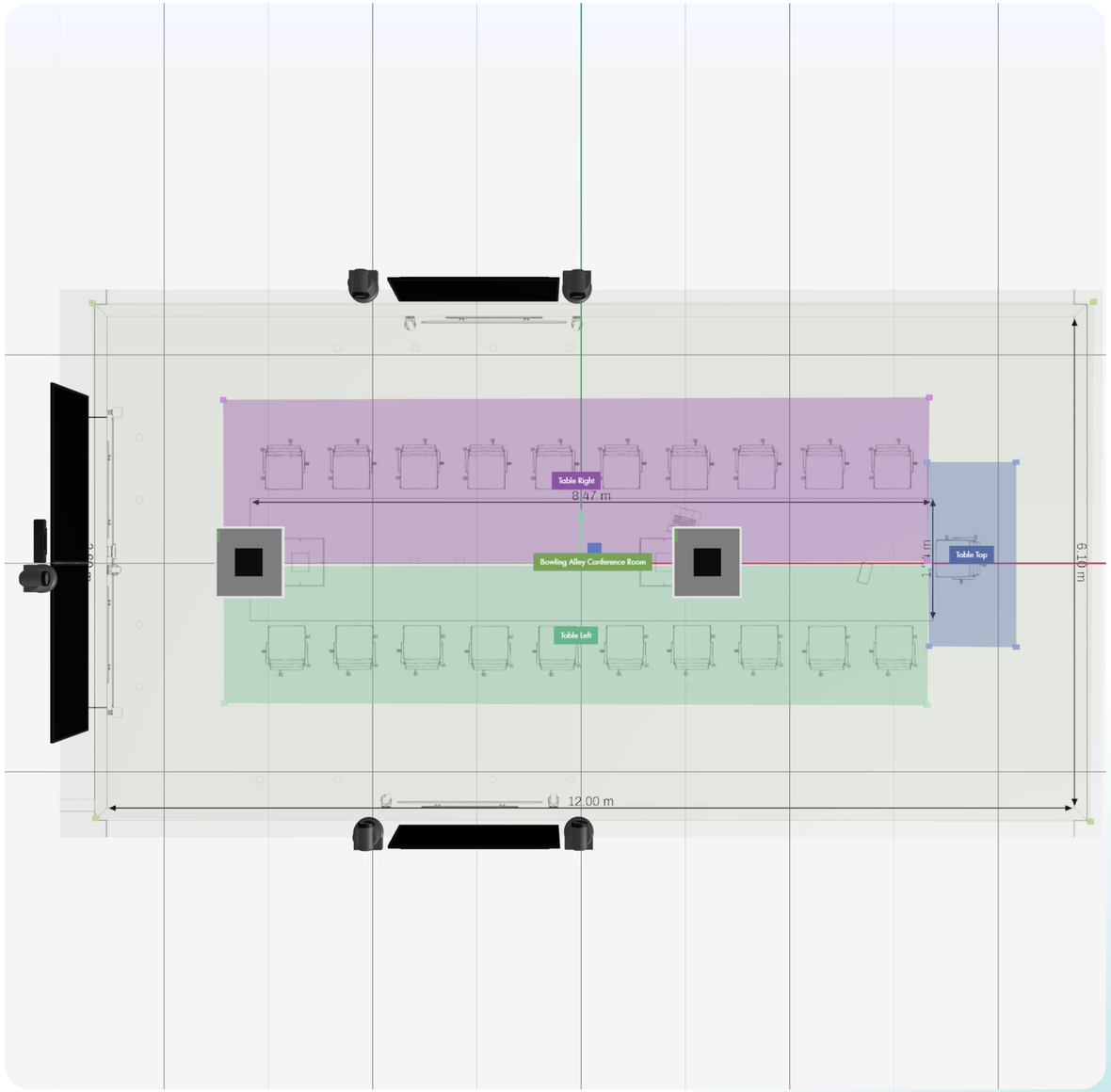
- **Choose the MXA920** if mixed sitting/standing meetings will be frequent (TCC2 lacks Z axis data)
- **Ensure 80cm/30”** minimum spacing between seats for accurate speaker coordinate detection
- **Place microphones in the front 2/3rds** of the room where participants naturally speak towards
- **Install cameras covering the table diagonally** for optimal frontal participant shots
- **Place cameras in pairs** to eliminate distracting movements when switching between speakers
- **Add a vision-driven audience facing overview camera** for persistent speaker location detections



EXAMPLE

‘Bowling Alley’ Meeting Room

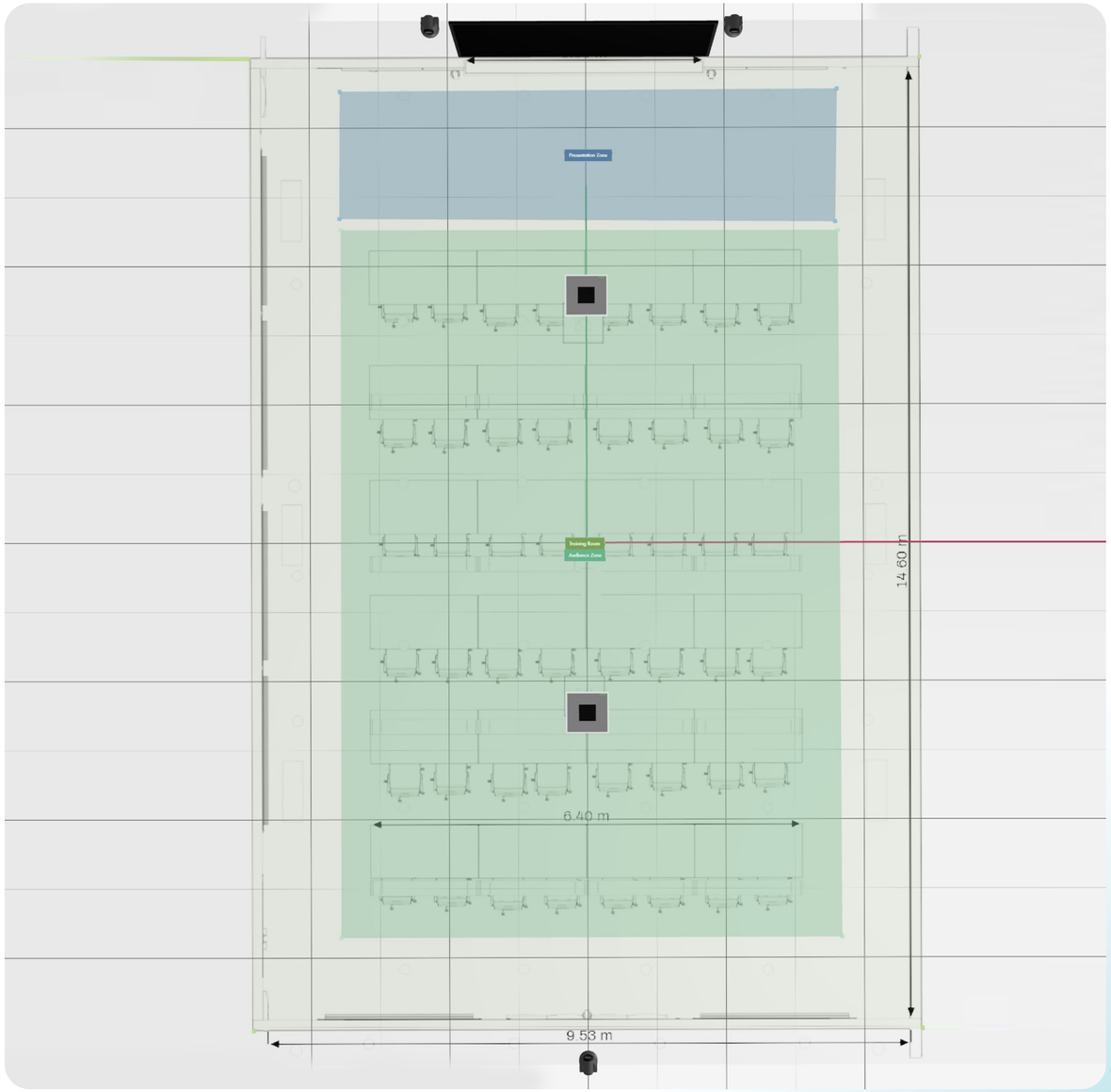
- **Avoid glass walls and control reverberation** to prevent acoustic challenges with audio coordinates
- **Place a dedicated NC20x60** for framing executive leadership at the head of the table
- **Mount side cameras along long walls** for when participants speak across the table
- **Place side displays next to cameras** to ensure frontal shots when participants speak to the far-end
- **Mount cameras in pairs** to eliminate distracting movements when switching between speakers
- **Choose the MXA920** if mixed sitting/standing meetings will be frequent (TCC2 lacks Z axis data)



EXAMPLE

Training Room

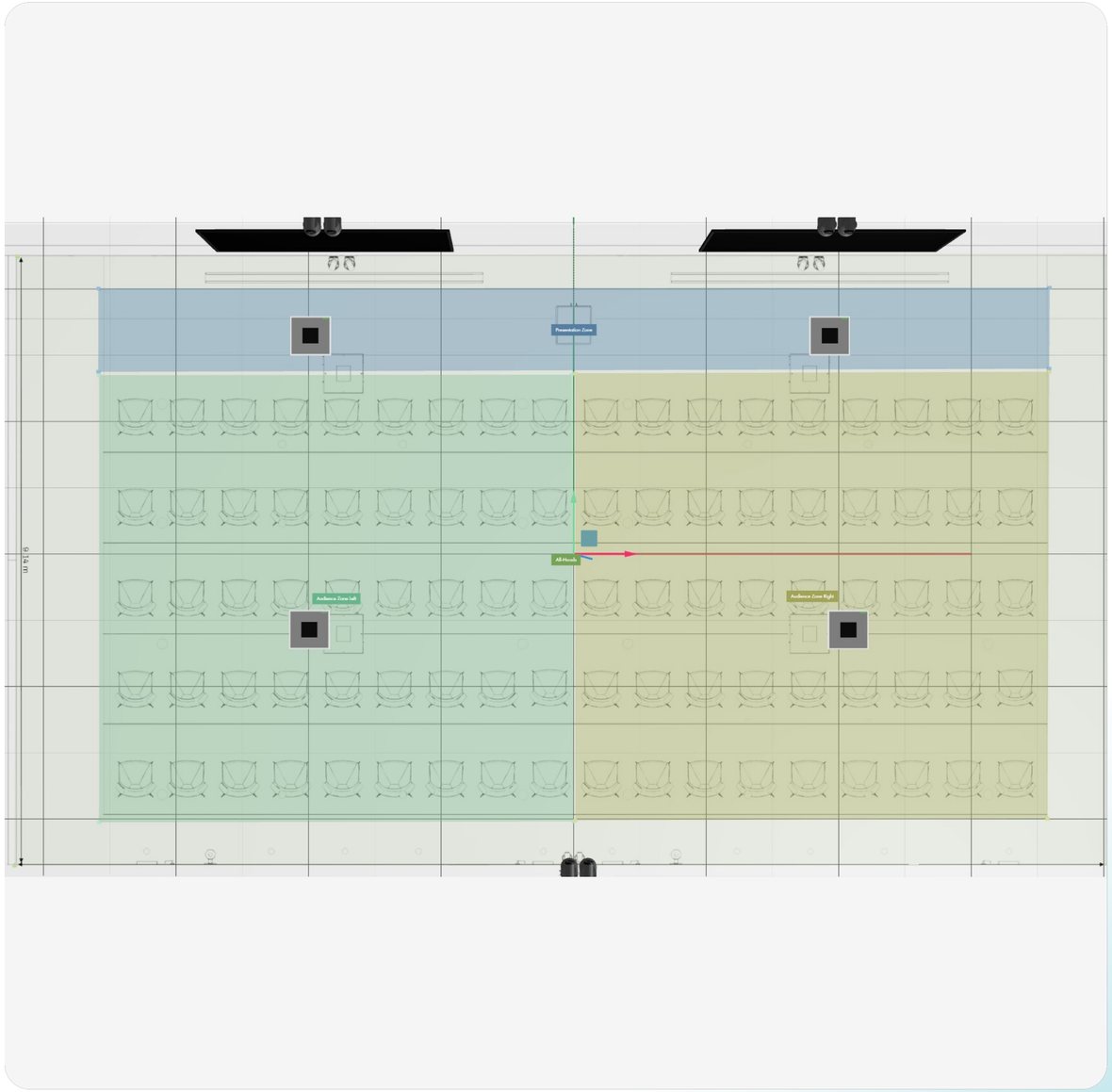
- **Create one large audience zone** to accommodate flexible seating layouts
- **Verify presenter tracking distance** stays within 20m / 65ft maximum for NC20x60 cameras (Half Body)
- **Leverage software-definable camera roles** if rotating the room's orientation between layouts
- **Skip the audience-facing conductor camera** due to room depth limitations (>7m effective range)



EXAMPLE

All-Hands Space

- **Avoid floor-to-ceiling displays** behind the presentation area to prevent false vision detections
- **Minimize having open areas behind presenters** where background activity could interfere
- **Do not mount ceiling mics above** 3.7m / 12ft from participants to maintain coordinate accuracy
- **Handheld / lavalier mics can be used for audio** but ceiling mics will be needed to provide location data
- **Ensure to treat room acoustics** as large spaces often exceed 0.8s RT60 without treatment
- **Configure Voice Lift carefully** for subtle sound reinforcement only, ensuring adequate loudspeaker distance from ceiling mics to avoid feedback
- **Add stage wash lighting** to cancel any harsh shadows from spotlights that interfere with presenter IDs



EXAMPLE

Lecture Hall

- **Create dedicated audio zones** for each seating tier in the case of tiered seating to handle height variations
- **Add separate mic zones** when audience level height differences exceed 1.5m / 5ft
- **Do not mount ceiling mics above 3.7m / 12ft from** participants to maintain coordinate accuracy
- **Handheld / lavalier mics can be used for audio** but ceiling mics will be needed to provide location data
- **Verify presenter tracking distance** stays within 38m / 135ft maximum for NC20x60 cameras (Max Zoom)
- **Use a podium/lectern as the defacto starting point for tracking where presenters naturally gravitate towards**
- **Add dedicated static cameras** for seamless content switching to whiteboards during lectures