

Spatial Conferencing Decision Sheet

When to build it, who can hear it, and the Web Audio recipe — on one page.

What it buys you (and what it doesn't)

Win	Each voice sits at its on-screen tile. Angular separation lets the brain unmask overlapping talkers — less listening effort on long, crosstalk-heavy calls.
Not	Pure 'immersion' is not the pitch. No benefit in 1:1 calls, town-halls with one speaker, or for users on a single mono speaker.
Rule	Strongest when competing voices are at similar levels — normalize mics first.

Step 1 — can your users even hear it? (device gate)

- Needs two channels reaching two ears: wired stereo headphones, wired stereo speakers, or built-in stereo speakers.
- Classic Bluetooth drops to mono when the mic is live — no spatial effect. Teams disables it on Bluetooth.
- LE Audio Bluetooth can carry stereo on a call — supported as devices ship it.
- Detect the output device. If it can't do stereo, fall back to the mono mix silently.

Step 2 — how the platforms do it (2026)

- Teams: tile position → stereo pan. Wired stereo only; no 1:1, no meetings > 100.
- Zoom: gallery/immersive seat → stereo field. Zoom Rooms / Workplace desktop; no wireless.
- Google Meet: stereo for presentations; Beam anchors each voice to the life-size 3D person.
- FaceTime: HRTF + head tracking via AirPods; optional personalized HRTF from a head scan.

Step 3 — build it on WebRTC (Web Audio API)

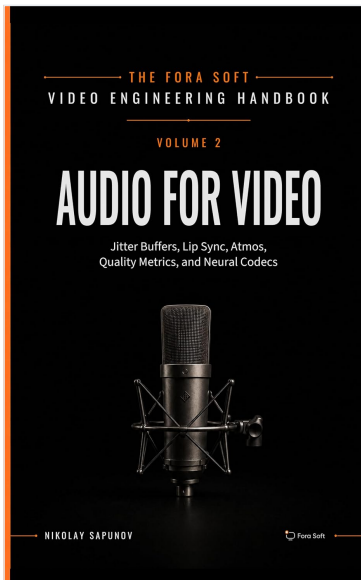
- One PannerNode per remote voice. panningModel = HRTF (binaural) or equalpower (cheap).
- Position = remote tile minus my position. Glide with setTargetAtTime (~20 ms), never snap.
- Map gallery columns to azimuth: $x = \sin(\text{angle})$, $z = -\cos(\text{angle})$.
- Use an SFU (each voice = its own track); spatialize per client, not on the server (MCU).
- Watch 3GPP IVAS (Release 18): carries 3-4 placed voices at 24.4 / 32 kbit/s in the codec itself.

Pre-ship checklist

- ☐ Output device detected; clean mono fallback when stereo isn't available.
- ☐ Spatial audio off for 1:1 calls and single-presenter town-halls.
- ☐ Positions glide, not snap, when tiles re-lay-out.
- ☐ Microphones normalized so competing voices sit at similar levels.
- ☐ Per-client spatialization on an SFU; not a per-listener server mix.

Companion to: 'Spatial Audio in Conferencing and Group Calls'

www.forasoft.com/learn/audio-for-video/articles-audio/spatial-audio-in-conferencing-group-calls



GO DEEPER · VOLUME 5 OF 7

Audio for Video: A Complete Guide to Sound in Video Products

From Jitter Buffers and Lip Sync to Atmos, Quality Metrics, and Neural Codecs

By Nikolay Sapunov, CEO at Fora Soft · 1,288 pages · Kindle & paperback on Amazon

[Read it on Amazon →](#)
amazon.com/dp/B0HH6XPY46

This sheet tells you when spatial audio helps a group call, which devices can hear it and how to build it with Web Audio. Volume 5 covers what happens to sound in the network and after it: jitter buffers and NetEq, topology and bitrate decisions, spatial audio and quality metrics.

WHAT YOU WILL FIND INSIDE

- Spatial audio rendering
- Jitter buffers and NetEq
- Bitrate and topology decisions for calls
- Audio quality metrics for conferencing

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI

Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF

Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery

Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing

Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs

Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video

Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Want group calls where every voice sits at its own tile?

We will add per-client spatial rendering on your SFU, detect stereo output devices, and fall back cleanly to mono where needed. 250+ video products shipped since 2005.

[Book a 30-min call →](#)
[WhatsApp →](#)
[Email us →](#)