A fresh look at Spatial and Next-Gen Audio

Is audio ready for the metaverse?

KINICHO
HEARING IS BELIEVING

A quick intro

- Who are we?
- Why are we here?
- What is this talk about?

Stefan Kazassoglou

@stefan_Kinicho

stefan.k@kinicho.com

Garry Haywood

@Kinicho

garry@kinicho.com

KINICHO
HEARING IS BELIEVING

Towards a practical audio for the metaverse and beyond

- Share some insights from our journey
- Start a conversation in the developer community
- Spatial Audio has this potential to be big, but today it's not really practical

KINICHO
HEARING IS BELIEVING

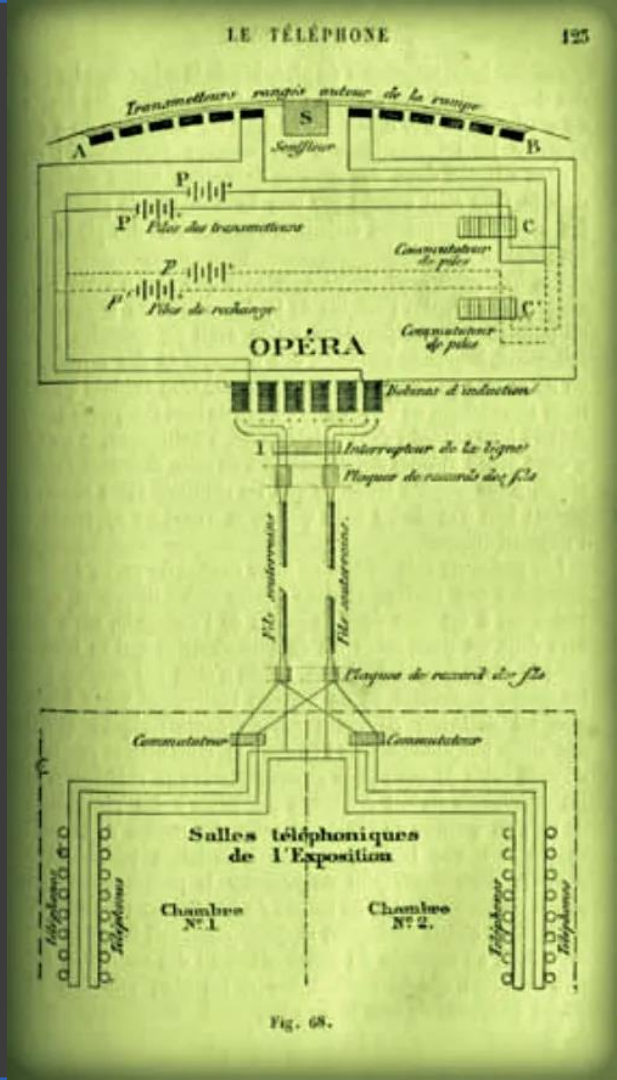
Audio technology & the trajectory of realism

- Quality of recording and reproduction
  - Back and forth - e.g CD/Redbook (1981) -> Lossless streaming (2021)
- Development towards the sound image
- Share exceptional experiences with the sense of presence

KINICHO
HEARING IS BELIEVING

Théâtrophone and binaural audio relays

- Several primitive microphone capture the Paris Opera stage
- Primitively mixed into 2 channels
- Relayed 2km to Exhibit Hall in the World Expo
- Split into a primitive cross-talk matrix for multiple pairs for binaural consumption

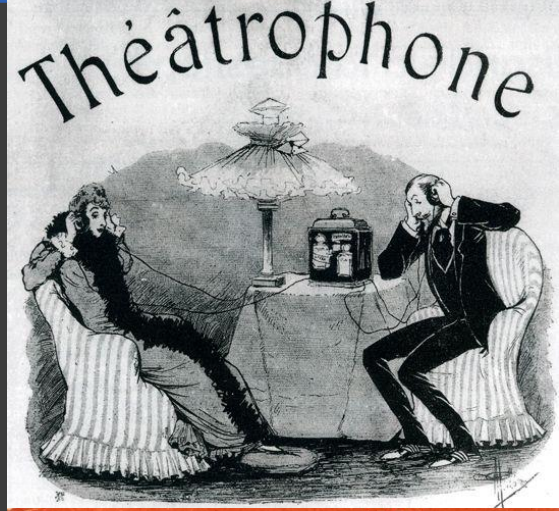


KINICHO
HEARING IS BELIEVING

Théâtrophone and binaural audio relays

- Several primitive microphone capture the Paris Opera stage
- Primitively mixed into 2 channels
- Relayed 2km to Exhibit Hall in the World Expo
- Split into a primitive cross-talk matrix for multiple pairs for binaural consumption

- 1st streaming audio subscription service
- Electrophone Co in London, 30 years relayed performances from 14 Theatres to approx 2k subscribers at it's peak



KINICHO
HEARING IS BELIEVING

The arrival of the loud-speaker

- 1-to-1 relationship of early recording
- Engraved analogue of soundwaves
- Playback by a single speaker

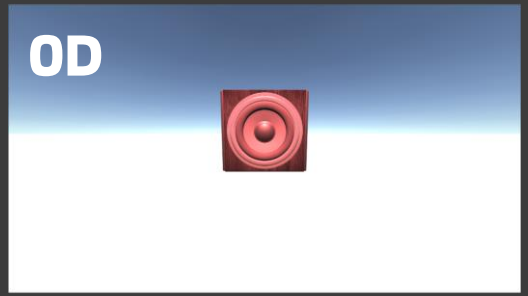
- Radio broadcasting - “wireless”

KINICHO
HEARING IS BELIEVING

The arrival of the loud-speaker

- 1-to-1 relationship of early recording
- Engraved analogue of soundwaves

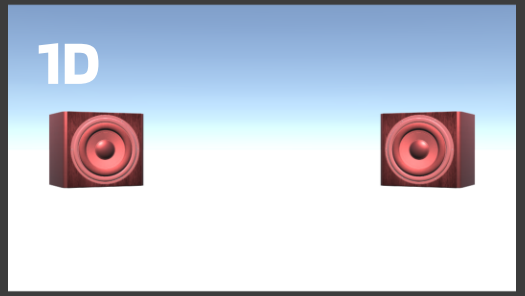
- Radio broadcasting - “wireless”
- Electrostatic improves the quality
- Listening to the ‘sum’ of a performance



KINICHO
HEARING IS BELIEVING

The spectre of Stereo's phantom image

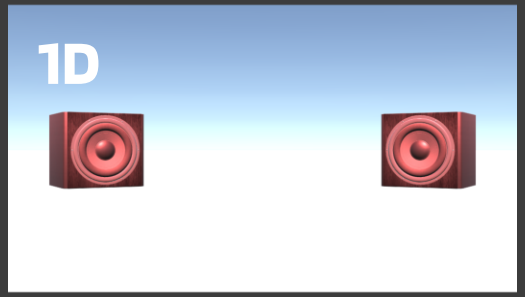
- Blumlein's Pair
- Haas or Precedence Effect
- The illusion of a sound stage
- Cross-talk required (absent from headphones)



KINICHO
HEARING IS BELIEVING

The Space Hack

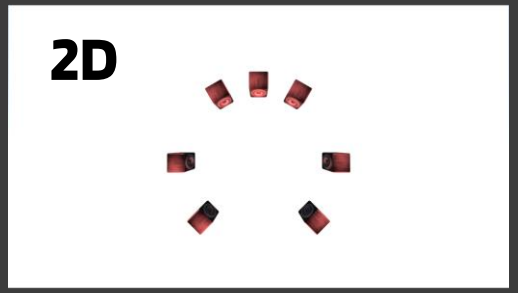
- Effects chains - reverb
- Multi-track recording
- overdubbing - virtual audio
- Smart use of the phantom image

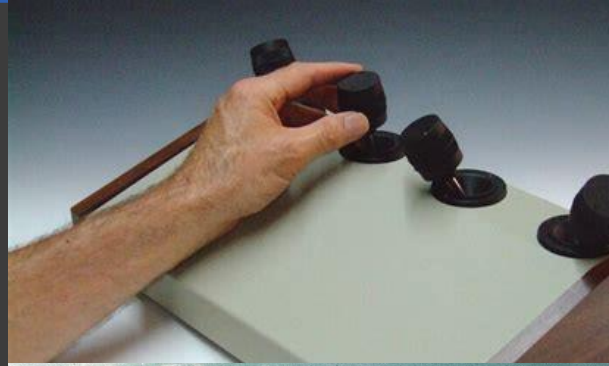


KINICHO
HEARING IS BELIEVING

Surrounded - a step to immersion

- Multi-track leads to Quadrophonic
- Perimetal manipulations
- Dolby solve 'surround sound' with rear-ambience
- How to listen with headphones?



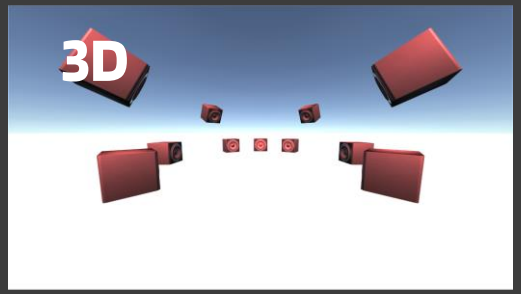


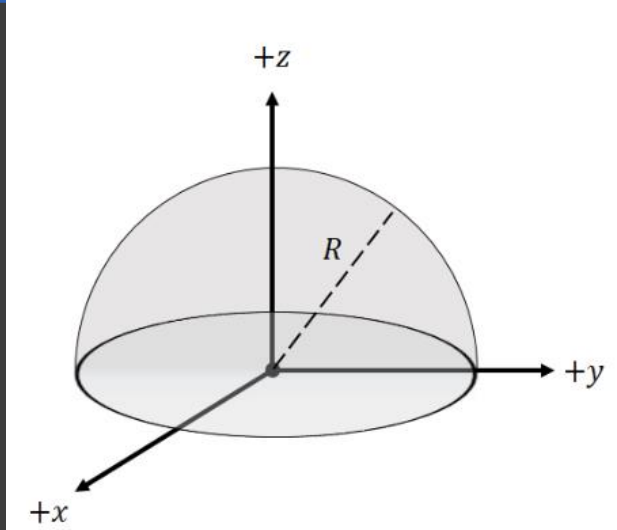
KINICHO
HEARING IS BELIEVING

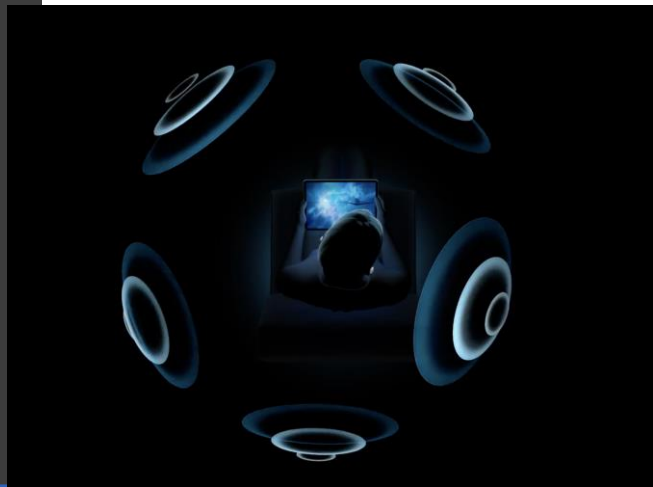
Elevation and partial periphony

- Adding height
- 3D - partial periphony
- Still on the perimeter of the space

- Apple – Atmos 5.1.4 > 1o Ambisonics > HRTF







KINICHO
HEARING IS BELIEVING

MPEG.H

- Hierarchical speaker system
  - 2.0
  - …
  - 7.1.4
  - …
  - 22.2

```
kAudioChannelLayoutTag_AudioUnit_6_0 = kAudioChannelLayoutTag_Hexagonal
ruct SpeakerLayoutTable : AudioChannelSet // save us some typing

  static LayoutTagSpeakerList* get() noexcept
  {
      static LayoutTagSpeakerList tbl[] = {
          // list layouts for which there is a corresponding named AudioChannelSet first
          { kAudioChannelLayoutTag_Mono, { centre } },
          { kAudioChannelLayoutTag_Stereo, { left, right } },
          { kAudioChannelLayoutTag_MPEG_3_0_A, { left, right, centre } },
          { kAudioChannelLayoutTag_ITU_2_1, { left, right, centreSurround } },
          { kAudioChannelLayoutTag_MPEG_4_0_A, { left, right, centre, centreSurround } },
          { kAudioChannelLayoutTag_MPEG_5_0_A, { left, right, centre, leftSurround, rightSurrou
          { kAudioChannelLayoutTag_MPEG_5_1_A, { left, right, centre, LFE, leftSurround, rightS
          { kAudioChannelLayoutTag_AudioUnit_6_0, { left, right, leftSurround, rightSurround, c
          { kAudioChannelLayoutTag_MPEG_6_1_A, { left, right, centre, LFE, leftSurround, rightS
          { kAudioChannelLayoutTag_DTS_6_0_A, { leftSurroundSide, rightSurroundSide, left, righ
          { kAudioChannelLayoutTag_DTS_6_1_A, { leftSurroundSide, rightSurroundSide, left, righ
          { kAudioChannelLayoutTag_AudioUnit_7_0, { left, right, leftSurroundSide, rightSurroun
          { kAudioChannelLayoutTag_AudioUnit_7_0_Front, { left, right, leftSurround, rightSurro
          { kAudioChannelLayoutTag_MPEG_7_1_C, { left, right, centre, LFE, leftSurroundSide, ri
          { kAudioChannelLayoutTag_MPEG_7_1_A, { left, right, centre, LFE, leftSurround, rightS
          { kAudioChannelLayoutTag_Atmos_7_1_4, { left, right, leftSurroundSide, rightSurroundS
          { kAudioChannelLayoutTag_Ambisonic_B_Format, { ambisonicW, ambisonicX, ambisonicY, am
          { kAudioChannelLayoutTag_Quadraphonic, { left, right, leftSurround, rightSurround } }
          { kAudioChannelLayoutTag_Pentagonal, { left, right, leftSurroundRear, rightSurroundRe
          { kAudioChannelLayoutTag_Hexagonal, { left, right, leftSurroundRear, rightSurroundRea
          { kAudioChannelLayoutTag_Octagonal, { left, right, leftSurround, rightSurround, centr

          // more uncommon layouts
          { kAudioChannelLayoutTag_StereoHeadphones, { left, right } },
          { kAudioChannelLayoutTag_MatrixStereo, { left, right } },
          { kAudioChannelLayoutTag_MidSide, { centre, discreteChannel0 } },
          { kAudioChannelLayoutTag_XY, { ambisonicX, ambisonicY } },
          { kAudioChannelLayoutTag_Binaural, { left, right } },
          { kAudioChannelLayoutTag_Cube, { left, right, leftSurround, rightSurround, topFrontLe
          { kAudioChannelLayoutTag_MPEG_3_0_B, { centre, left, right } },
          { kAudioChannelLayoutTag_MPEG_4_0_B, { centre, left, right, centreSurround } },
          { kAudioChannelLayoutTag_MPEG_5_0_B, { left, right, leftSurround, rightSurround, cent
          { kAudioChannelLayoutTag_MPEG_5_0_C, { left, centre, right, leftSurround, rightSurrou
          { kAudioChannelLayoutTag_MPEG_5_0_D, { centre, left, right, leftSurround, rightSurrou
          { kAudioChannelLayoutTag_MPEG_5_1_B, { left, right, leftSurround, rightSurround, cent
          { kAudioChannelLayoutTag_MPEG_5_1_C, { left, centre, right, leftSurround, rightSurrou
          { kAudioChannelLayoutTag_MPEG_5_1_D, { centre, left, right, leftSurround, rightSurrou
          { kAudioChannelLayoutTag_MPEG_7_1_B, { centre, leftCentre, rightCentre, left, right,
          { kAudioChannelLayoutTag_Emagic_Default_7_1, { left, right, leftSurround, rightSurrou
```

KINICHO
HEARING IS BELIEVING

The speaker model as technical debt

- Everything is speaker first
- Workflow borrows from stereo's advances
- The phantom image haunts spatial audio

KINICHO
HEARING IS BELIEVING

Headphone overtake

- Mobile-led
- Ipod & MP3 ecosystem - < quality
- 2010 headphones generate > $ than domestic hifi
- “Wireless” convenience - < quality
- Streaming media – dominance of Atmos

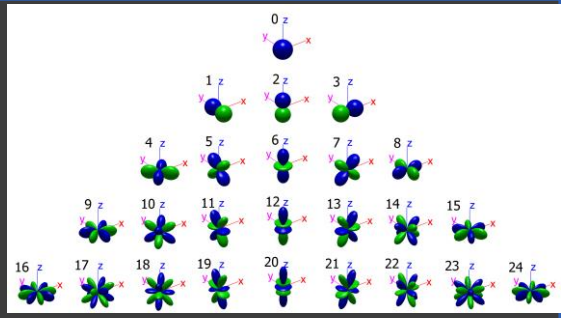
KINICHO
HEARING IS BELIEVING

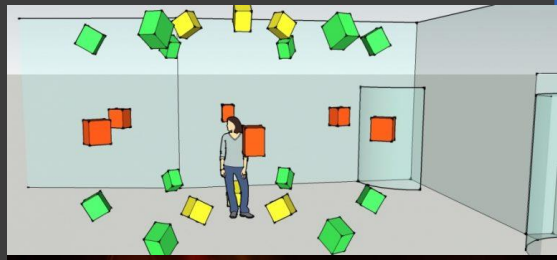
Ambisonic is speaker agnostic

- Revival from VR
- Spherical Harmonics to derive angles
- Distribute energy about the surface of a sphere
- Map the energy onto speakers
- Directional only
- Limited Angular resolution

  resolution = 180°/(order+1)

  1st Order 90°, 3rd Order 45°, 7th Order 22.5°

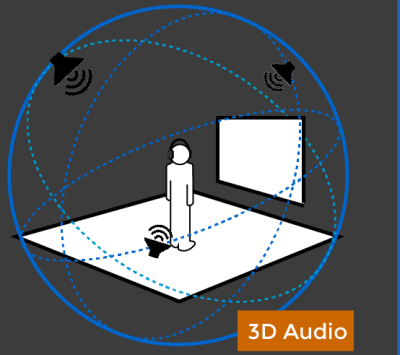


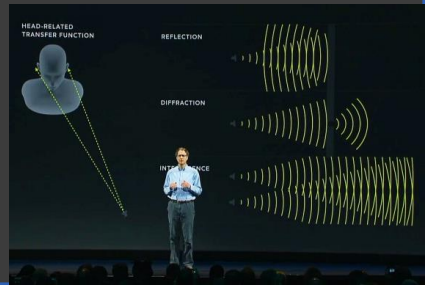


KINICHO
HEARING IS BELIEVING

Head-Related Transfer Function (HRTF)

- Collection of Head-Related Impulse Responses
- Use convolution filtering
- Emulation
- All other spatial acoustic physics baked into sound design
- Timbral shift + artefacts

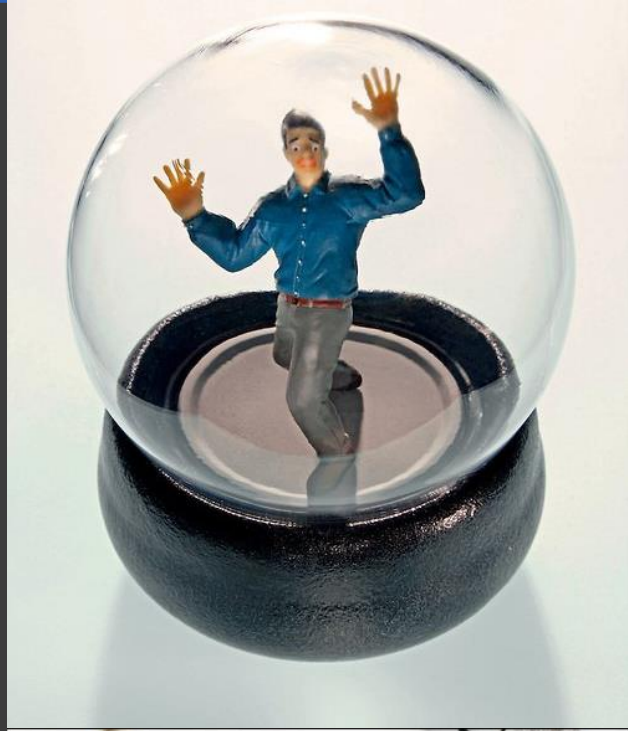




KINICHO
HEARING IS BELIEVING

Fixing HRTF

- **Better filtering**
- **Multiple effects – near/far**
- **Personalisation**
  - Elusive ‘High St’ model
  - Ear scanning – with machine learning
- **Reduce the confusions?**



KINICHO
HEARING IS BELIEVING

Volumetric Metaverse

- More about experience than interface
- 3D world metaphor
- Spatial audio will be the ‘presence’ maker
- Stream Media – through shared experiences

KINICHO
HEARING IS BELIEVING

Sonic Reality of the Metaverse

- Interactable 6 DoF Space – users and objects
- Continuous Movement
- Agency, Attention and focus
- UGC
- Generative media - AI

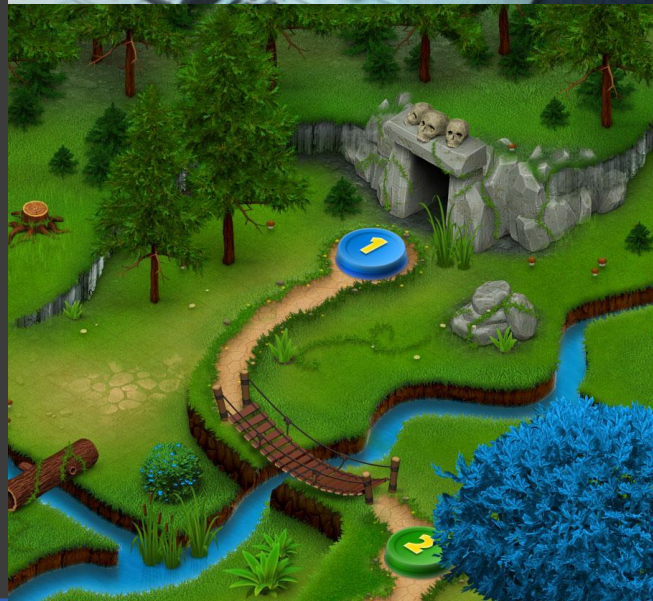
- Persistence
- Duration

KINICHO
HEARING IS BELIEVING

End of the Mix

- Inherited from the stereo workflow
- Where everything is mapped out
- Audio is tied to path

- The Metaverse will be unmixed



KINICHO
HEARING IS BELIEVING

Spatial Audio 2.0

- Simulation-based - Sonic reality
- Real-time
- High definition volumetric spaces
- Organic sounding / analogue modelled
- Ear-centric
- binaural that's more natural

KINICHO
HEARING IS BELIEVING

NextGen Audio

- continues with Sonic Reality paradigm
- But with more play-back options
- E.g. Speaker

KINICHO
HEARING IS BELIEVING

So Speakers are OK then?

- Next-gen will be playback agnostic
- No problem with speakers, it's the speaker-based workflow

KINICHO
HEARING IS BELIEVING

Legacy formats

- We'll still have these formats
- Back-catalogue
- New creations
- Component to the metaverse experience

KINICHO
HEARING IS BELIEVING

Refactor Audio

- Thinking about the audio future
- Transition from the speaker model
- New Work-flows
- New Tools

- What are we building for the future?

KINICHO
HEARING IS BELIEVING

Contacting us

Stefan Kazassoglou

@stefan_Kinicho

stefan.k@kinicho.com

Garry Haywood

@Kinicho

garry@kinicho.com

KINICHO
HEARING IS BELIEVING