

# Appunti sul suono spazializzato

---

A partire dagli appunti della *Masterclass su suono immersivo*, View Conference, Torino, 23 ottobre 2018

## Relatori

- Gianni Ricciardi - WANTmusik
- Matteo Milani - Unidentified Sound Object

## Introduzione

---

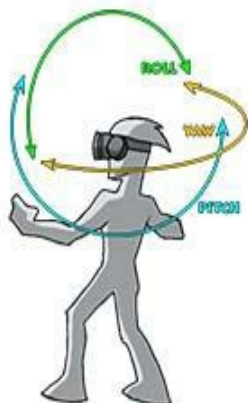
Principale differenza fra le esperienze VR -> **gradi di libertà (DOF)**

**3 DOF** -> solo roll, pitch and yaw. Rappresenta l'esperienza ottenuta con normali video 360°. I visori di questo tipo sono equipaggiati con piattaforma inerziale e monitor (Oculus Go) o sfruttano lo smartphone (Samsung Gear VR).

**6 DOF** -> consente la localizzazione dell'utente nello spazio. Rappresenta l'esperienza di realtà virtuale completa. E' quindi necessario un sistema in grado di tracciare i movimenti traslatori dell'utente. Esistono due tecniche:

- **outside-in tracking**: tracciamento attraverso sensoristica esterna, come Oculus Rift e HTC Vive;
- **inside-out tracking**: tracciamento attraverso sensori interni al dispositivo (camere infrarossi, ecc), come Oculus Quest e Razer OSVR.

3 degrees of freedom (3-DoF)



6 degrees of freedom (6-DoF)



# Suono spazializzato

---

La realtà virtuale comporta l'immersività anche dal punto di vista del suono: l'orecchio umano ha una risoluzione di pochi *cm* in termini di [localizzazione di un suono](#) nello spazio.

I primi esperimenti sul suono spazializzato risalgono agli anni 70. Ad esempio:

- Sviluppo di [microfoni binaurali](#) come il celebre [Neumann KU100](#);
- Registrazioni [olofoniche](#) pubblicate da [Hugo Zuccarelli](#) negli anni 80;
- Esperimenti di sintesi binaurale e studi sull'[HRTF](#).

Per una precisa percezione del suono spazializzato è necessario l'ascolto in cuffia, per evitare cross-talking fra orecchio sx e dx e le interferenze fra i due segnali acustici.

Un problema comune in questo tipo di contenuto audio è la difficoltà nel distinguere fra un suono frontale ed uno posteriore. Nel caso di suoni reali l'ambiguità è risolta quando l'ascoltatore effettua leggeri movimenti della testa. Nel caso di una registrazione olofonica la posizione del suono è solidale a quella della testa: il suono registrato sarà sempre localizzato nello stesso punto da chi ascolta, indipendentemente dalla posizione della sua testa.

Nella realtà virtuale è necessario invece che il suono rimanga sempre fisso nella stessa posizione dello spazio, e di conseguenza venga localizzato in punti diversi a seconda della posizione della testa. Oltre un sistema di **head tracking** è quindi necessario un insieme di tecniche di registrazione e riproduzione che consentono di ottenere questo tipo di contenuto audio.

## Ambisonics

---

Il primo step è quello di registrare un suono ricreando completamente la sfera acustica 360°. Il formato [Ambisonics](#) nasce esattamente con questo scopo intorno alla metà degli anni 70.

**Ambisonics** è un contenitore, *open-source*, che comprende un insieme di metodi per *registrare, mixare e riprodurre* una sfera di suono a 360°, proveniente da diverse direzioni attorno ad un punto centrale (*sweet spot*) in cui si trova l'ascoltatore.

Ambisonics rappresenta una *bolla di suono*, liscia, stabile e continua, completa di elevazione. È *speaker-agnostic*: può essere decodificato per qualsiasi sistema di speaker.

PROS	CONS
flessibile, evocativo, future-proof	Ascolto solo in un punto preciso, richiede decodifica, file pesanti

### Campi di applicazione

Film/video - Mix/post-produzione - VR - Installazioni sonore - Ricerca scientifica - Composizione musicale

### Hardware ambisonic

Un microfono in grado di registrare audio Ambisonics è chiamato anche [Soundfield microphone](#), ed è stato brevettato da Michael Gerzon e Peter Craven nel '75. Attualmente la [SoundField](#) è un'azienda che produce HW e SW per l'audio ambisonics.

La maggior parte dei microfoni Ambisonics sono equipaggiati con 4 capsule a condensatore (solitamente cardioidi) che puntano nelle direzioni Left-Front, Left-Back, Right-Front e Right-Back e registrano in A-Format, ovvero l'output raw delle 4 capsule.



Alcuni microfoni codificano anche un metadata contentene la direzione di puntamento (upright, endfire, upside-down).

Alcuni mic Ambisonics:

- [Soundfield SPS200](#)
- [Sennheiser Ambeo](#)

- [Rode/Soundfield NT-SF1](#)
- [Zoom H3-VR](#)

Il segnale raw trasdotto da un microfono ambisonics è chiamato A-Format.

I microfoni ambisonici possono essere differenti per diversi parametri (numero e diagramma polare di capsule, configurazione, ecc), di conseguenza l'A-Format non può essere considerato uno standard.

### First-order vs Higher-order Ambisonics

A partire da una registrazione raw in A-Format, il segnale viene normalmente trasformato in un altro formato, detto B-Format.

Il B-Format è il formato standard per la rappresentazione dell'audio Ambisonics. Il segnale sferico viene codificato in una serie di segnali che rappresentano ognuno una porzione del volume sferico.

Il formato di base è chiamato First Order Ambisonics (FOA), è composto da quattro canali e la detiene la risoluzione spaziale minima per un segnale audio 360.

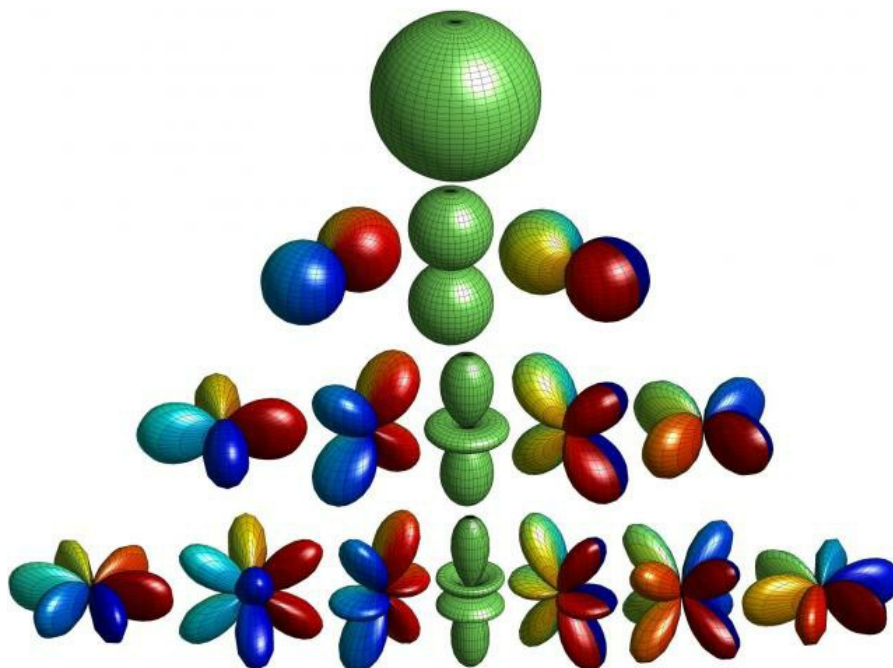
I quattro canali si dividono in:

- W - diagramma polare omnidirezionale - *sound pressure*
- X - diagramma polare a 8 (bidirezionale) - *front-minus-back*
- Y - diagramma polare a 8 (bidirezionale) - *left-minus-right*
- Z - diagramma polare a 8 (bidirezionale) - *up-minus-down*

La scarsa risoluzione del FOA ha portato allo sviluppo di formati di ordine maggiore, che mediante l'utilizzo di un numero maggiore di canali è in grado di rappresentare l'informazione di un numero maggiore di porzioni del volume (*spherical harmonics*).

Ordine Ambisonic	Numero di canali
FOA	4
2OA	9
3OA	16
4OA	25
5OA	36
6OA	49

Nello schema sottostante una rappresentazione visuale degli *spherical harmonics* del segnale Ambisonics dal primo al terzo ordine.



Lo standard B-Format ha due varianti, che differiscono per l'ordine dei canali:

- ACN Order (AmbiX)
- Furse-Malham Ordering (FuMa) (lo standard *de facto* per gli ordini  $\leq 3$ )

NOTA: lettura consigliata - [Towards usability of Higher Order Ambisonics in Digital Audio Workstations](#), Matthias Kronlachner, 2012

## Plugins

---

Per la codifica/decodifica e gestione del suono spazializzato sono necessari plugin, disponibili per le più usate DAW in commercio:

- [Sennheiser Ambeo Suite](#) - suite
- [ambiX by Mathias Kronlachner](#) - suite di produzione audio spazializzato
- [Soundfield by Rode](#) - ambisonics recordings post-processor
- [SoundField SurroundZone2](#) - ambisonics recordings post-processor
- [Plugins by Noisemakers](#) - diversi plugins
- [HARPEX](#) - spatial audio formats cross-coding algorithm
- [Blue Ripple Sound O3A](#) - suite (con effetti spazializzati)
- [Waves 360° Ambisonics Tools](#)

# FB360 Spatial Audio Workstation

---

[Facebook360 Spatial Audio Workstation](#) (*FB260 SAW*) è una suite di strumenti completa e gratuita per il design e la codifica di contenuti audio spazializzati per il VR.

**Spatial Audio Workstation 8-channel:** formato inventato da FB360, 8 canali perché sta a metà fra FOA e 2OA, ha più risoluzione sul piano orizzontale - *NON STANDARD*

Il formato di uscita è chiamato Two Big Ears (.tbe).

NOTA: lettura consigliata - [Conversion from Ambisonics to TBE format](#) dal blog di Angelo Farina, 2017

FB360 mette anche a disposizione un SDK che supporta l'integrazione del formato .tbe in diversi ambienti (Android, Unity...).

A questo [link](#) si trova il gruppo di discussione su FB riguardo l'utilizzo della *FB360 SAW*.

## Recap: formati e standard

---

- **A-Format:** raw ambisonics. Contiene i dati audio *raw* provenienti dalle capsule di un mic - *NON STANDARD*
- **B-Format:** formato standard di rappresentazione dell'audio ambisonics. Il numero di canali dipende dall'ordine dell'ambisonics (vedi paragrafo precedente). Si divide in FuMa e AmbiX a seconda dell'ordine in cui sono codificati i canali (FuMa: WXYZ, AmbiX: WYXZ) - *STANDARD*
- **Two Big Ears (.tbe):** formato di uscita dell'audio processato con FB360 SAW. Può contenere 8 canali (solo spazializzato) o 10 canali (spazializzato + una traccia stereo fissa) - *NON STANDARD*

## Altri progetti e riferimenti

---

### Carne y Arena

[Installazione di realtà virtuale](#) che comprende anche l'utilizzo di "mixed audio reality": suono in cuffia spazializzato + infrasuoni riprodotti da subwoofer [Meyer Sound](#).

Regia di Alejandro Inarritu, Sound Design di Randy Thom e Martin Hernandez.

# FMOD Audio Engine

**FMOD** è un sound engine proprietario usato soprattutto per lo sviluppo di audio per video games e applicazioni interattive.

## Link vari

### Teoria

- [Ambisonics Explained: A Guide for Sound Engineer](#) - Waves Audio
- [Differences between A Format and B Format](#) - postperspective.com
- [An Introduction to Ambisonics](#) - creativefieldrecording.com
- [Binaural and ambisonic sound](#) - Gareth Fry

### Pratica

- [FB360 Spatial Workstation Workflow](#) - Facebook
- [Resonance Audio Project](#) - Google
- [A beginner's guide to spatial audio in 360-degree video](#) - NPR
- [Your Detailed Guide to Spatial Audio for 360 Video](#) - nofilmschool.com
- [First-order ambisonics in REAPER](#) - Google
- [Samsung VR Best Practices](#) - Samsung
- [Ambisonic Audio with 360 Video in Premiere Pro CC 2017](#) - Jason Levine
- [Add Spatial Audio To Oculus Video \(Gear VR\)](#) - Facebook
- [How to Create 360° Mixes from Stereo Mix Stems](#) - Waves Audio