

# Riproduzione del suono, 3d-sound e Spazializzazione Binaurale



UNIVERSITÀ DEGLI STUDI  
DI MILANO

16/12/2008 Corso di Informatica Applicata alla  
Musica

Davide A. Mauro



# Qualche domanda...

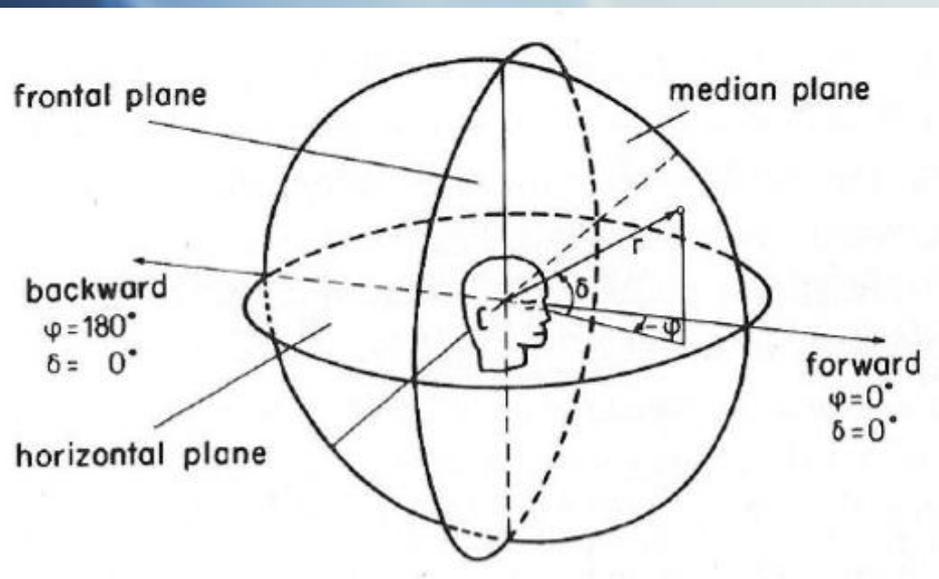
- Che cos'è uno spazio sonoro tridimensionale?
- In che modo il nostro apparato uditivo è in grado di percepire la direzione di provenienza di un suono?

# Ancora...

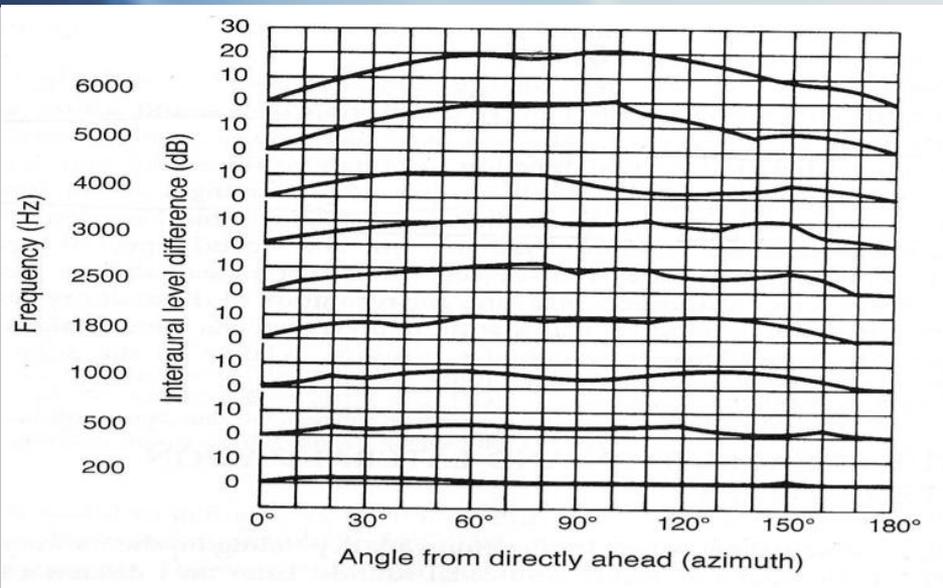
- Come possiamo capire se un suono proviene da:
  - Destra-Sinistra
  - Alto-Basso
  - Davanti-Dietro
  - (Distanza)

# Come funziona...

- Tre parametri chiamati “Localization Cues”:
  - ILD: Interaural Level Difference
  - ITD: Interaural Time Difference
  - DDF: Direction Dependent Filtering

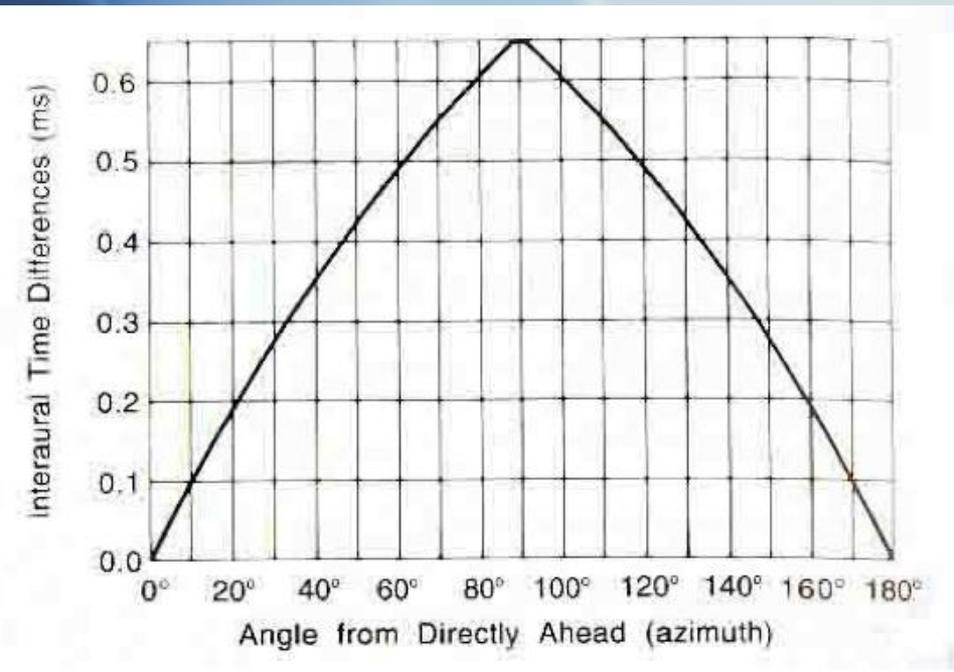


# ILD



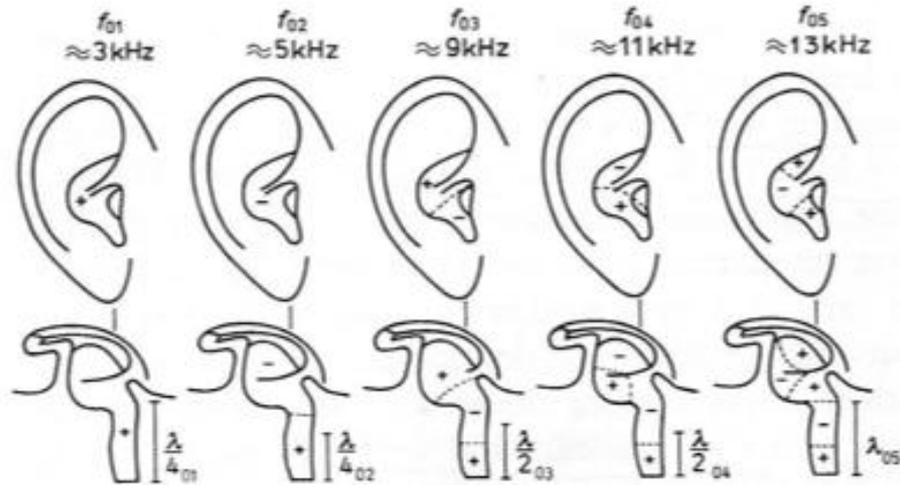
- *Interaural Level Difference* (espresso in dB), differenze di intensità fra gli stimoli alle due orecchie
- ILD varia in funzione dell'azimuth e della frequenza
- (frequency dependent)
- Utile per le alte frequenze

# ITD



- *Interaural Time Differences*, differenze di tempo di arrivo degli stimoli alle due orecchie.
- Varia in funzione dell'azimuth
- (frequency independent)
- Utile per le basse frequenze

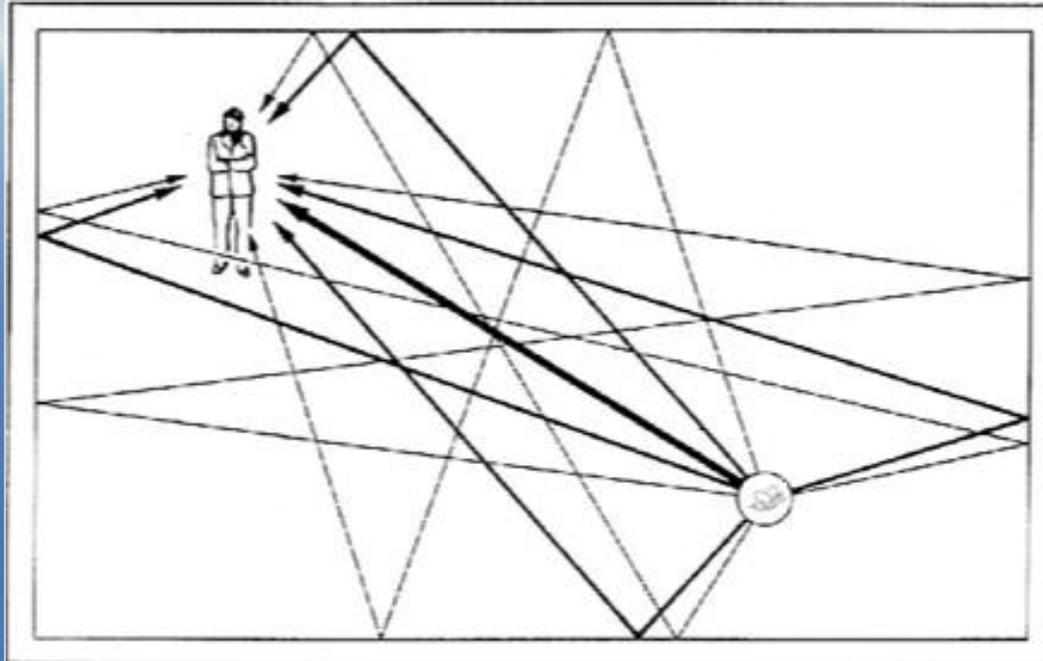
# DDF



- *Direction Dependent Filtering*, filtraggio effettuato dall'orecchio esterno, dalla testa e dal torso, dipendente dalla direzione di provenienza dello stimolo.
- Diversi modi di risonanza della pinna

# L' Ambiente

- Una precisazione: Stiamo parlando di ascolto in *free-field*, ovvero senza l'interazione dell'ambiente esterno
- Riverbero: Centrale per la percezione di “spazialità e ambiente” (es canali nei sistemi dolby) può dare più profondità anche a un semplice stereo



# Tecniche: Una prima idea

- Più altoparlanti da disporre attorno all'ascoltatore più possibilità di ricreare posizionamento spaziale?
- Posso registrare anche in mono e poi “assegnare”(panpot) agli altoparlanti
- Percezione delle sorgenti virtuali
- Sì ma... In questo modo NON ho informazioni su di uno spazio acustico reale

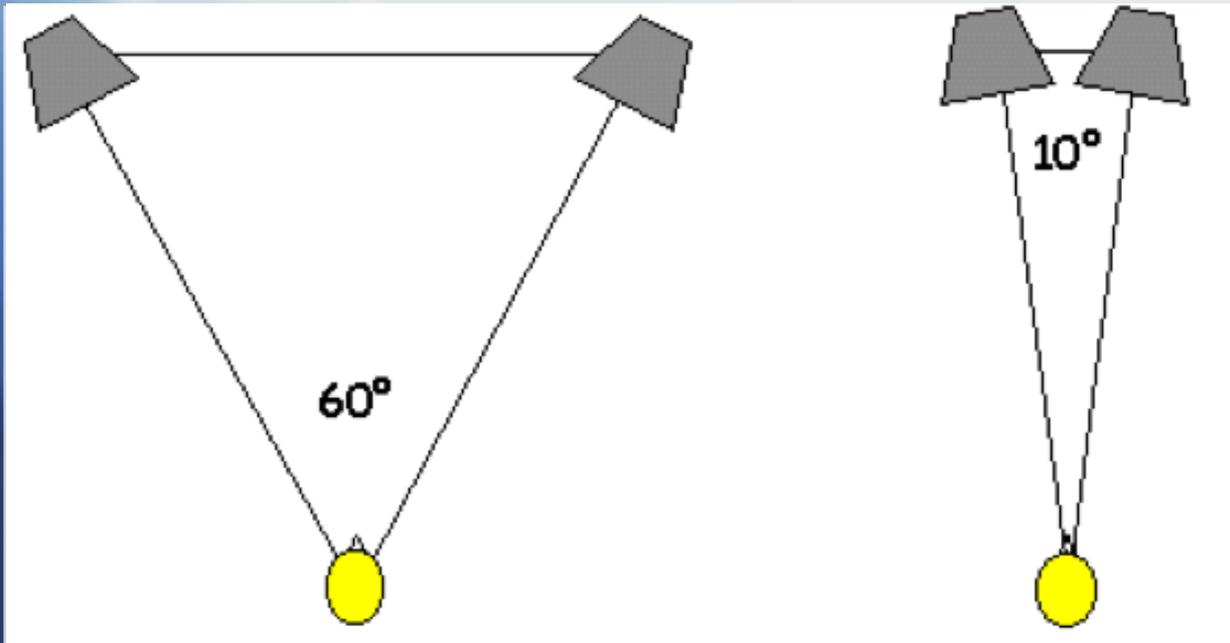
# Panning e Sorgenti Virtuali

# Precisazione

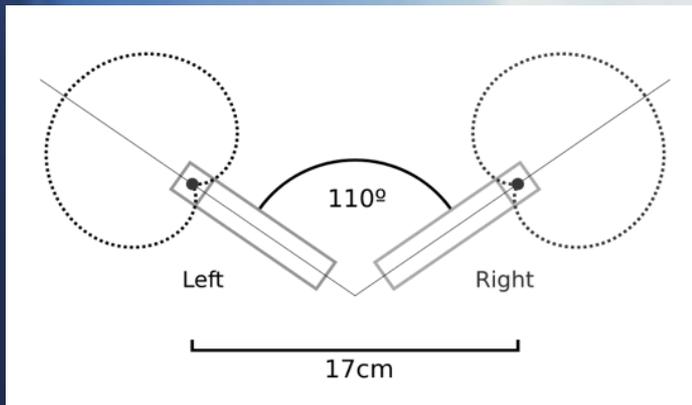
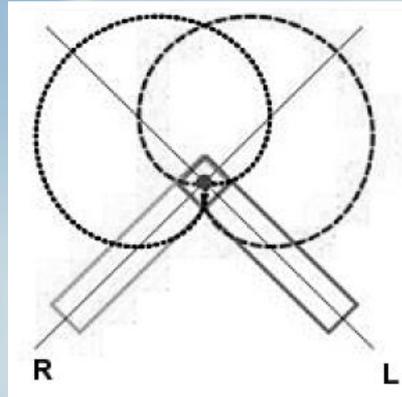
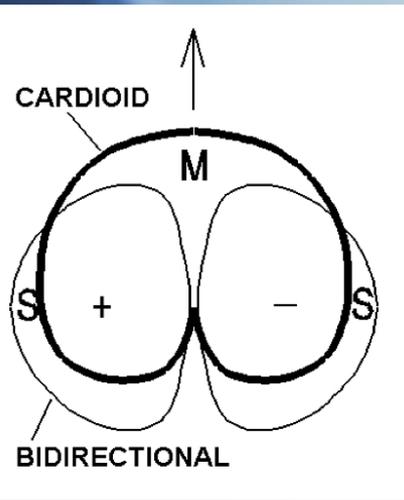
- Le prime due tecniche che citeremo funzionano bene se l'ascoltatore si trova esattamente nel punto previsto per l'ascolto
- Altrimenti... La sorgente si sposta verso l'altoparlante più prossimo all'ascoltatore

# Tecniche: Stereo

- Riproduzione “1-D”
- Riproduzione tramite 2 altoparlanti equidistanti dal punto di ascolto, varie configurazioni es Stereo Dipole

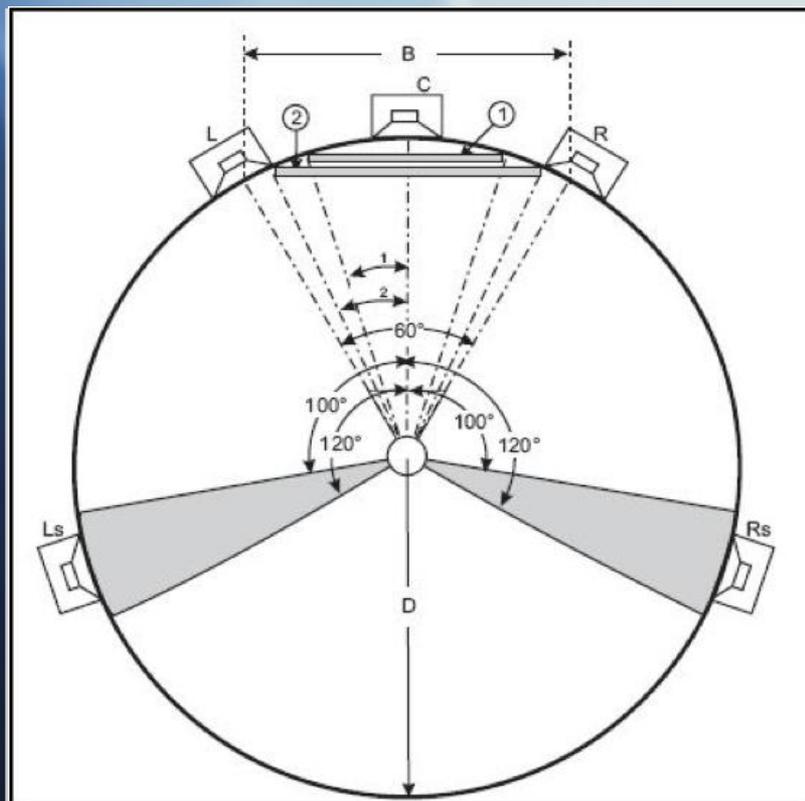


# Stereo: Microfonazione



- X-Y Coincidente: 2 Microfoni direzionali (cardioidi)
- M-S: Middle Side: 1 Cardioide + 1 Bidirezionale (pattern a 8)
- ORTF: 2 Cardioidi

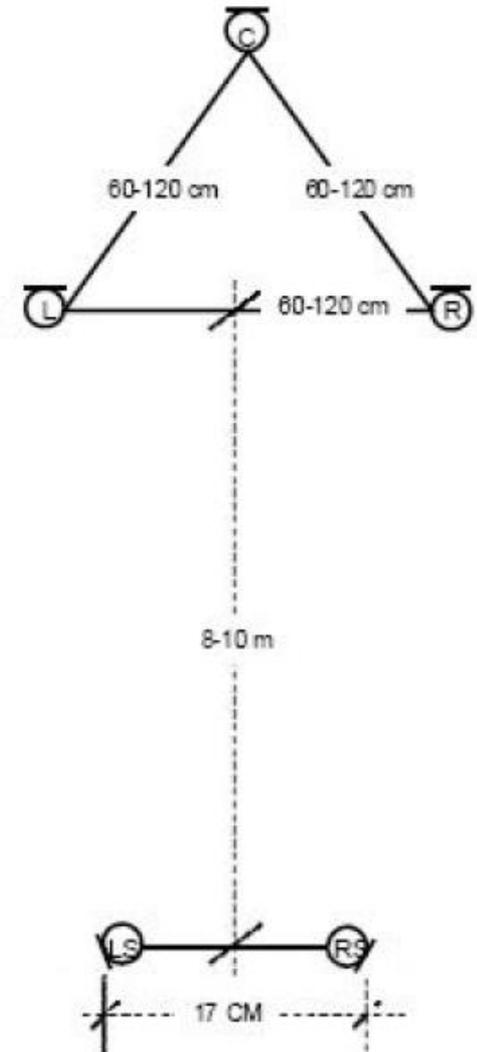
# Tecniche: Multicanale discreto



- Ad esempio home theater 5.1
- 1 Canale centrale (parlato)
- 2 Canali frontali
- 2 Canali posteriori (riverbero)
- 1 Canale LFE (Subwoofer)
- Riproduzione “2-D”

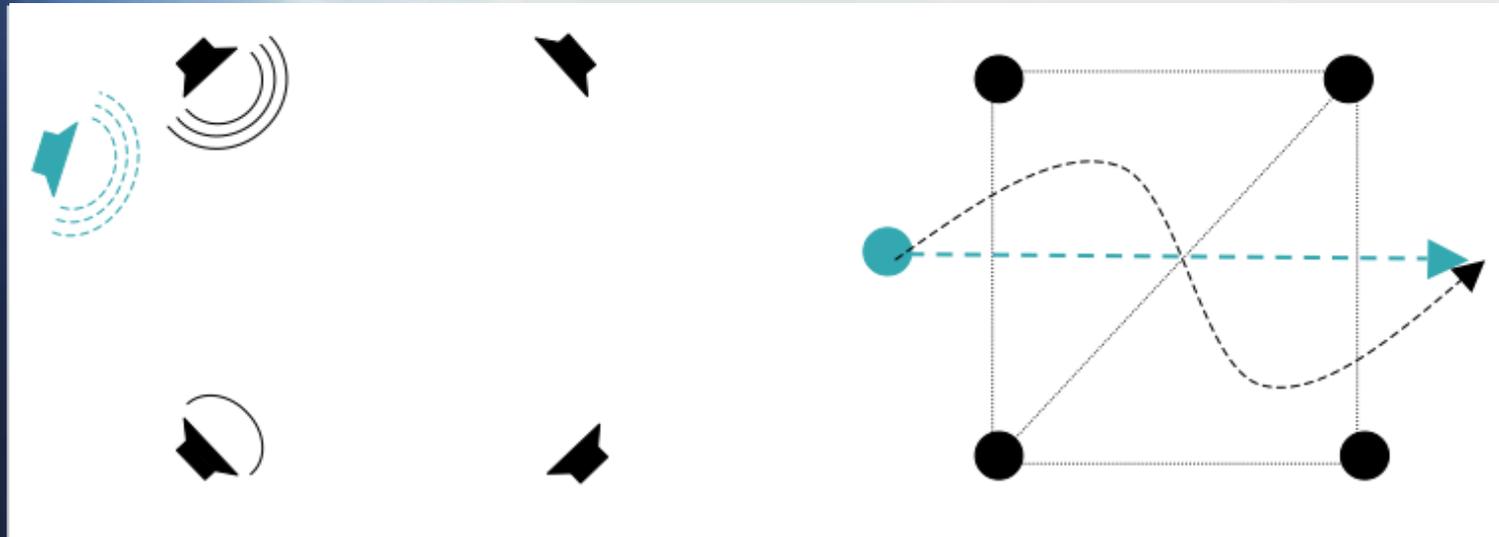
# Multicanale discreto: Microfonazione

- Anche qui molte tecniche:  
Es Decca-Tree+ORTF
- Esistono microfoni specifici
- Tool specifici per il mixaggio

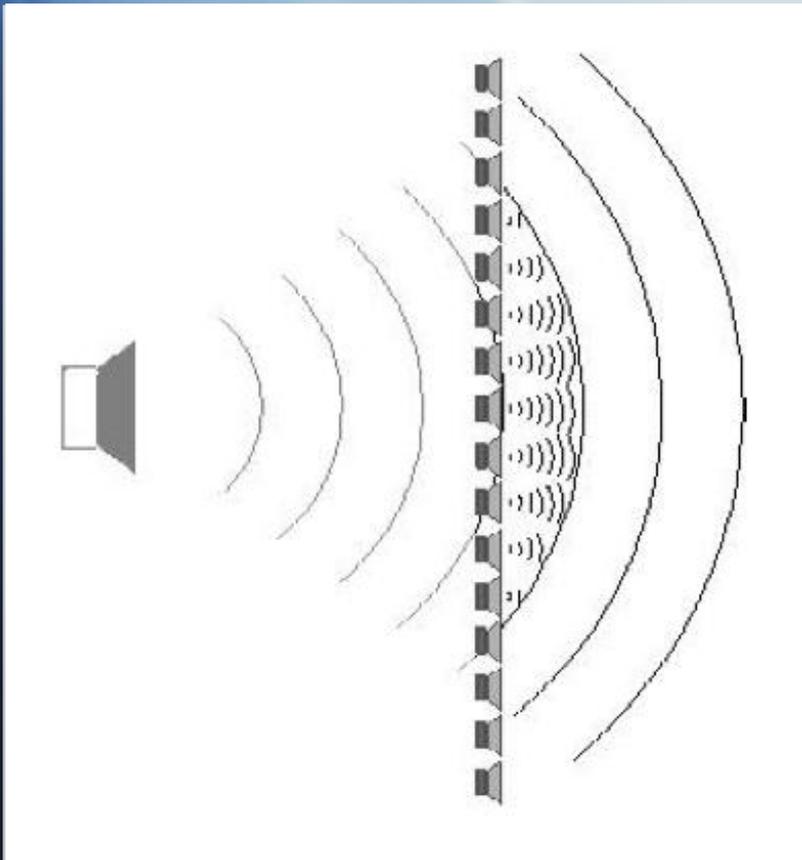


# Multicanale discreto

- Non perfetta localizzazione delle sorgenti.
- Risente molto della configurazione e del punto di ascolto

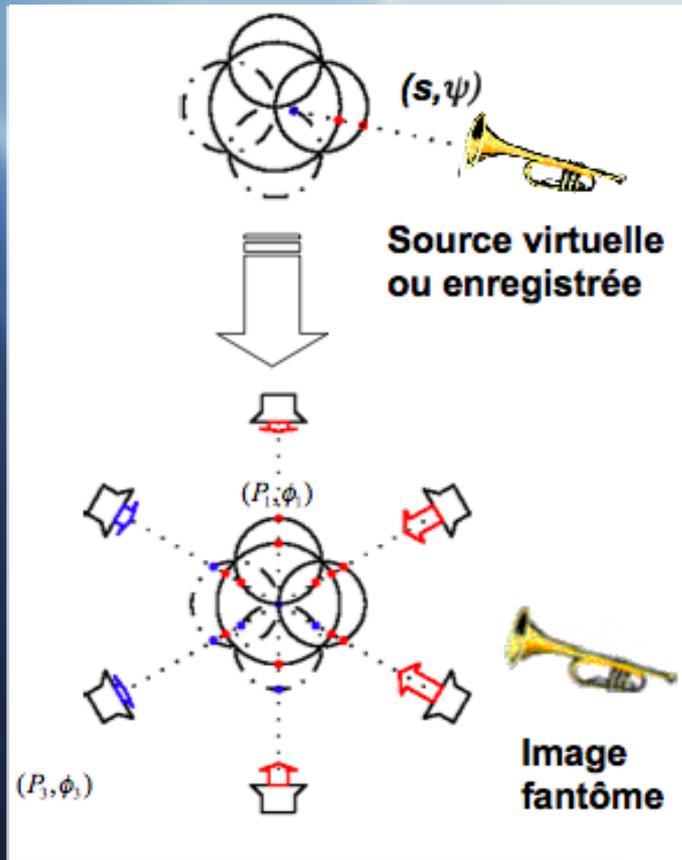


# Tecniche: WFS



- Wavefield Synthesis: Ricreazione del fronte d'onda
- Principio di Huygens: Scomposizione del fronte d'onda. Sovrapposizione di onde elementari
- Necessita di ambiente anecoico e molti altoparlanti disposti in array (distanza 10-15 cm max)
- La localizzazione non dipende dalla posizione dell'ascoltatore

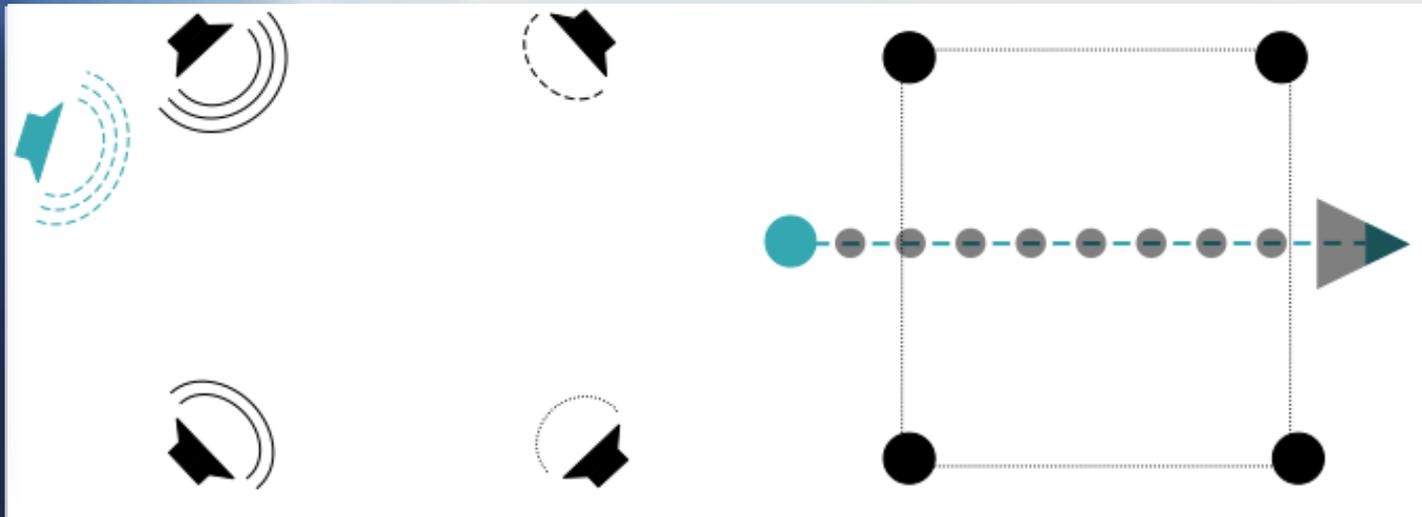
# Tecniche: Multicanale con matrici di conversione



- Ambisonic:
- Utilizzo di microfoni specifici
- Decodifica su un qualsiasi sistema di altoparlanti a meno di alcuni vincoli:
  - Non è importante la posizione quanto la direzione
  - Numero minimo per restituire le caratteristiche N-dimensionali

# Ambisonic

- 4 canali per 1 ordine (3 senza elevazione)
- 9 canali per 2 ordine (7 senza elevazione)
- Miglior impressione di spazialità
- Migliore “stabilità” e precisione per la localizzazione



# Tecniche: Spazializzazione Binaurale

- Riproduzione “3-D”
- Idea: ”I principali veicoli dell'informazione uditiva al soggetto umano sono le sue due orecchie”
- Non si cerca di ricreare un paesaggio sonoro tridimensionale attraverso il posizionamento di sorgenti sonore nelle tre dimensioni ma si crea un paesaggio sonoro virtuale attraverso la simulazione del comportamento dell'apparato uditivo fornendo all'ascoltatore un segnale a due canali che contenga al suo interno i parametri necessari affinché il nostro orecchio possa determinare la direzione di provenienza del suono in uno spazio tridimensionale”

# Spazializzazione Binaurale

- Quindi --> Ricreazione dei parametri
- Idonea alla riproduzione in cuffia
- Realizzabile tramite diverse tecniche:
  - Filtraggi e linee di ritardo
  - Convoluzione

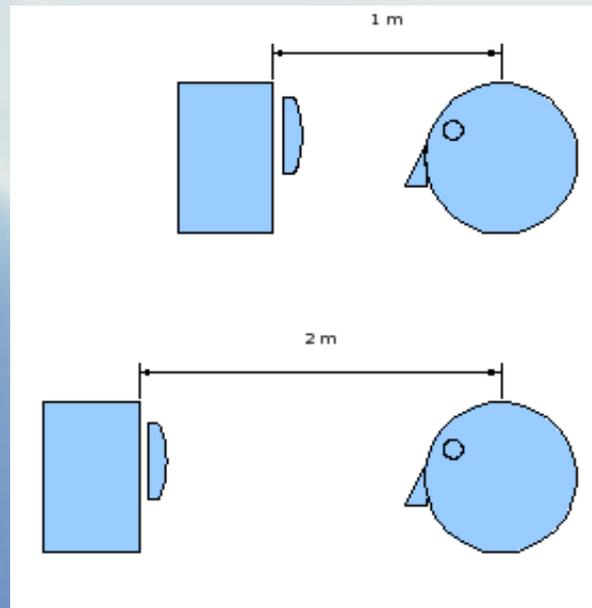
# Spazializzazione Binaurale

- Convoluzione con risposta all'impulso
- Dummy Head per le misurazioni e campionamenti di HRTF (Head Related Transfer Function)



# Tecnica

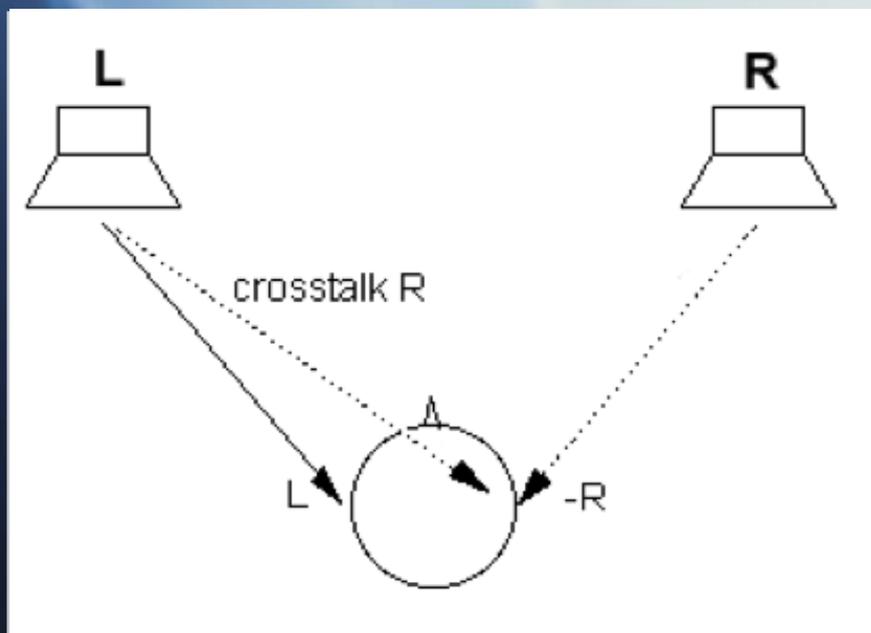
- Praticamente: Uso della dummy e di altoparlanti per inviare il segnale di misura, registrazione dai microfoni, processing.



# Tecnica

- Misurazione con Sweep seno-logaritmica (derivata da TDS Time Delay Spectrometry)
- Convoluzione con segnale TRM (Time Reversal Mirror)
- Tecnica robusta, permette di trattare le non-linearità del sistema

# Problemi... Crosstalk



- Con altoparlanti ogni canale raggiunge entrambe le orecchie
- Causa problemi nella creazione dell'immagine spaziale
- Esistono tecniche di cancellazione : Es. segnale ritardato e in controfase--> Efficaci per una zona di ascolto molto ridotta (tende ad 1 punto centrale)
- Combinato con Spazializzazione Binaurale --> Stereo Transaurale

# Problemi...

- Modalità di ascolto: Altoparlanti, Cuffie
- Computazionali: Es. convoluzione è molto costosa se fatta nel dominio del tempo --> Portata nel dominio delle frequenze (FFT) diventa trattabile
- Efficacia: Individualizzare le risposte all'impulso (es dummy con calchi orecchie) permette migliori risultati ma... come fare?

# Per cosa possiamo usare queste tecniche?

- Idea: RV (Realtà Virtuali)
- Di cosa abbiamo bisogno?
  - Posizione dell'ascoltatore e orientamento (head tracking)
  - Posizione della sorgente sonora
  - Geometria e caratteristiche dello spazio

# BREVE BIBLIOGRAFIA

- J. Blauert “*Spatial Hearing, the Psychophysic of Human Sound Localization*”, the MIT Press, Cambridge, Massachussetts, 1996
- D. R. Begault “3-D Sound for Virtual Reality and Multimedia”, Academic Press, Cambridge, MA, 1994
- W. A. Yost, G. Gourevitch “*Directional Hearing*”, Springer, New York, 1987
- Sunil Bharitkar, Chris Kyriakakis “*Immersive audio signal processing*”, New York, Springer, 2006
- R. E. Runstein, D. M. Ruber “*Modern Recording Techniques*”, Howard W. Sams, Indianapolis, IN , 1989
- Udo Zolzer “*DAFX : digital audio effects*”, Chichester, John Wiley, 2002
- A. Farina, P. Galaverna, M. Giabbani “*Il processo di auralizzazione: metodologia ed esemplificazione*”, 2000.
- A. S. Bregman “*Auditory Scene Analysis*”, MIT press, Cambridge, MA, 1990

# PER INFORMAZIONI:

- LIM
- [dm733602@silab.dsi.unimi.it](mailto:dm733602@silab.dsi.unimi.it)