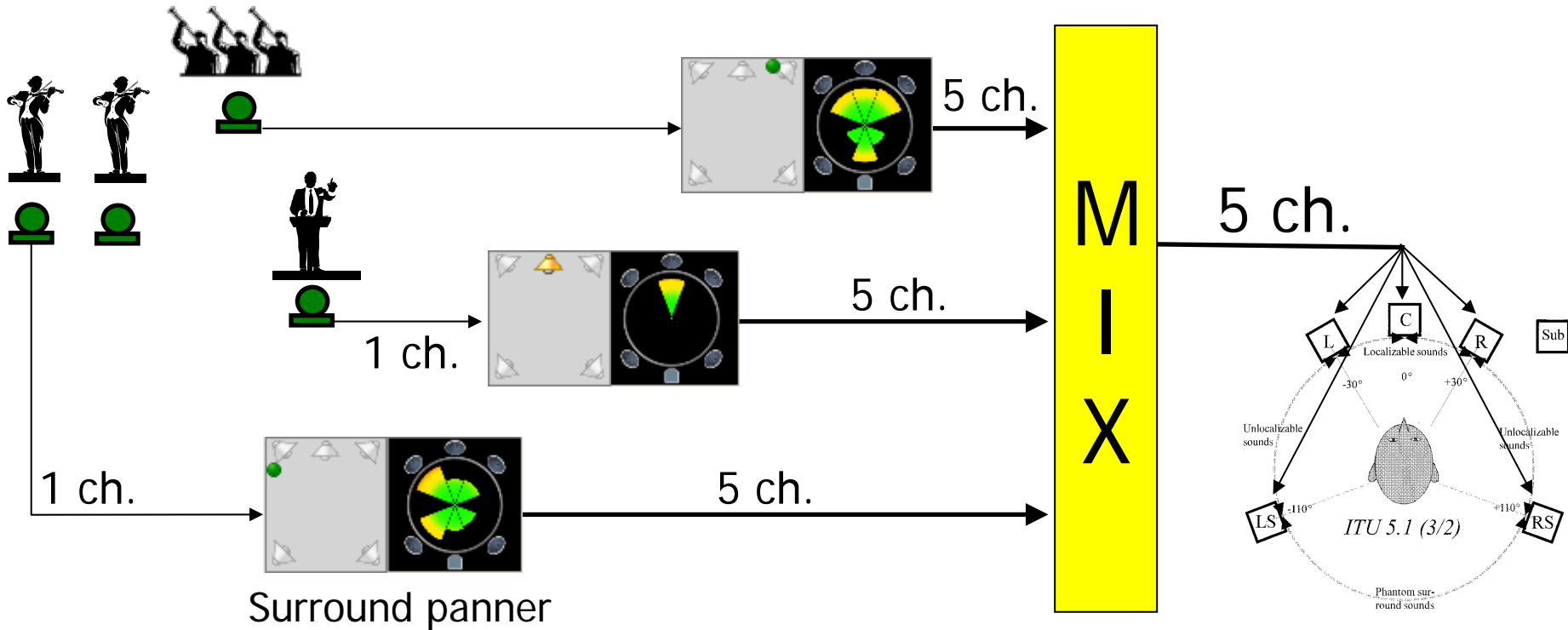


Array di microfoni per registrazioni musicali in formato 5.1 Surround ed Ambisonics

Angelo Farina, Paolo Martignon,
Enrico Armelloni, Andrea Capra
Dip. Ing. Industriale, Università di Parma



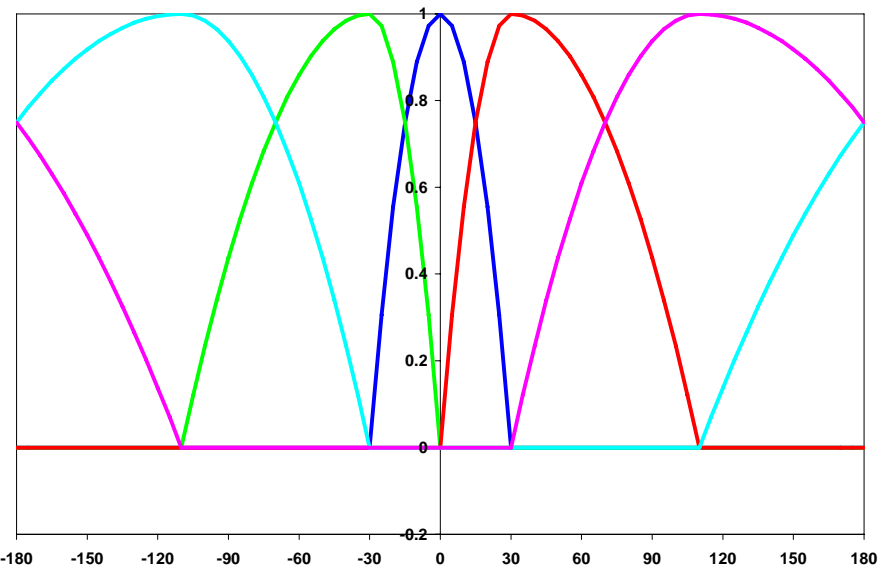
Metodo tradizionale (registrazioni mono “panned”)



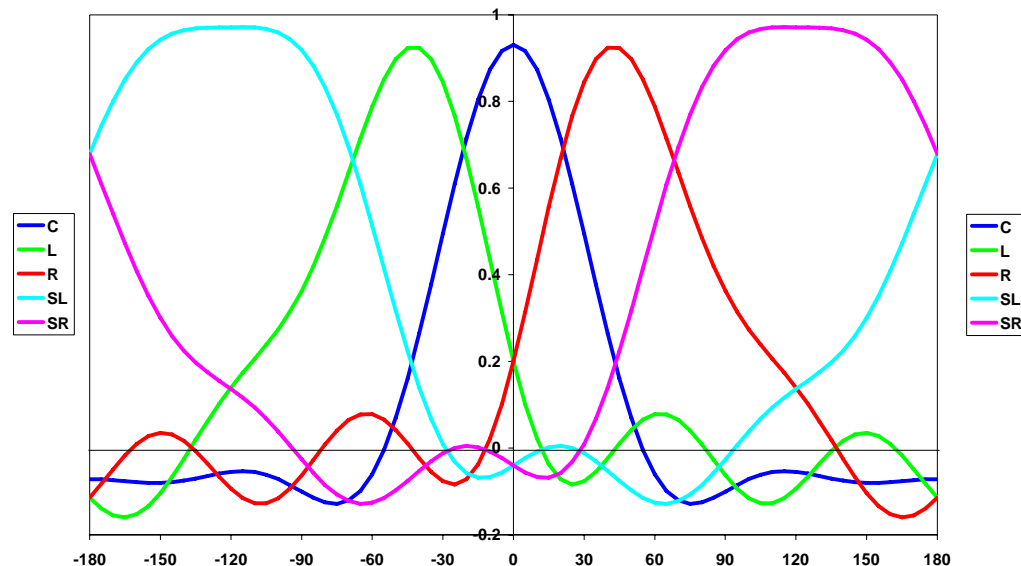
- Ciascuna traccia mono registrata rappresenta una sorgente in una diversa posizione, che viene posizionata mediante una appropriata legge di “panning” multicanale

Leggi di “panning”

- Si è visto sin dagli albori che non conviene posizionare le tracce mono sui singoli canali del surround in modo discreto, ma che conviene utilizzare appropriate leggi di “panning” in modo da alimentare sempre piu’ di un altoparlante per volta



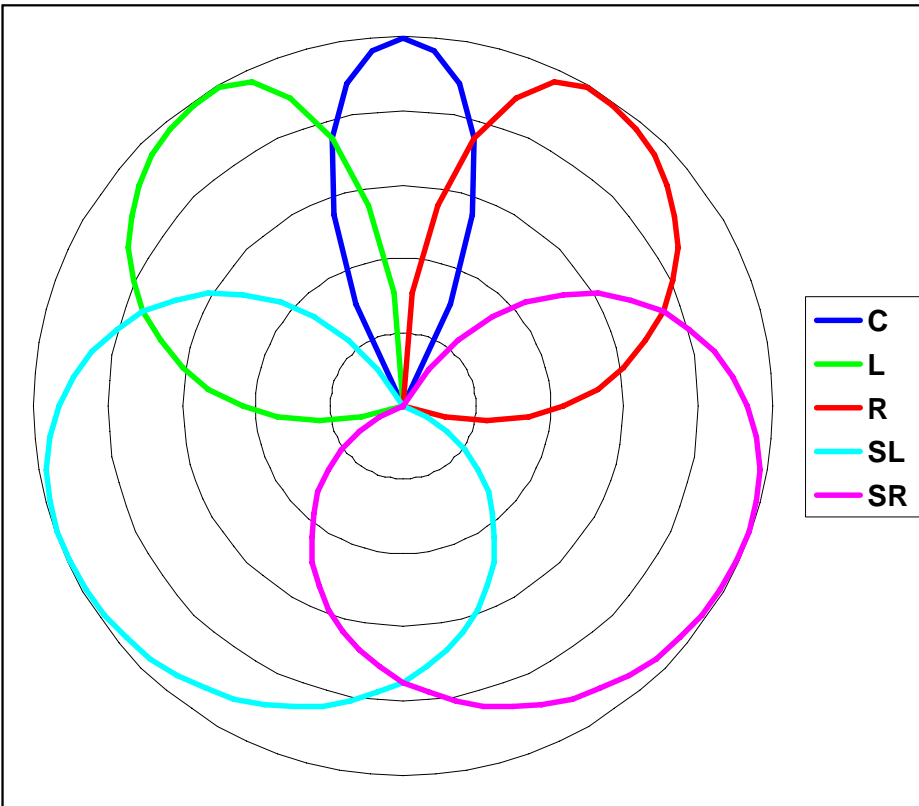
“Pairwise Panning” a potenza costante - il segnale viene inviato a due altoparlanti per volta



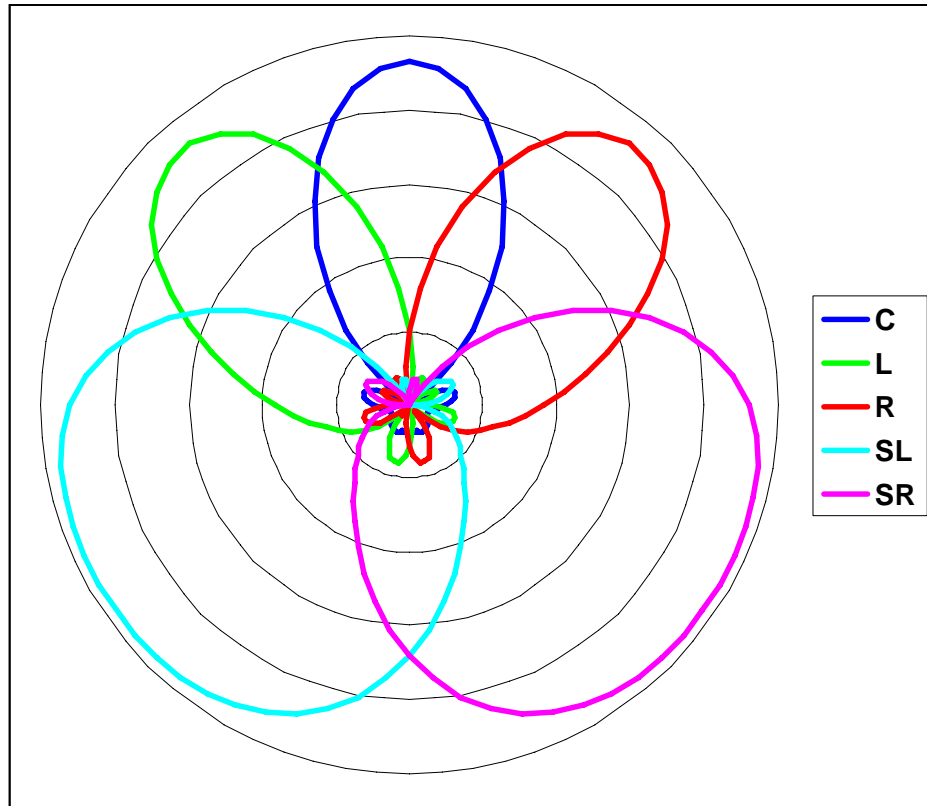
Panning basato sulla teoria di Peter Craven - il segnale viene inviato sempre a tutti 5 gli altoparlanti, con opportuni guadagni

Leggi di “panning”

- Un modo alternativo per visualizzare le leggi di panning consiste nell’ipotizzare l’esistenza di 5 microfoni virtuali con opportuni “pattern” di direttività



“Pairwise Panning”



Panning di Peter Craven

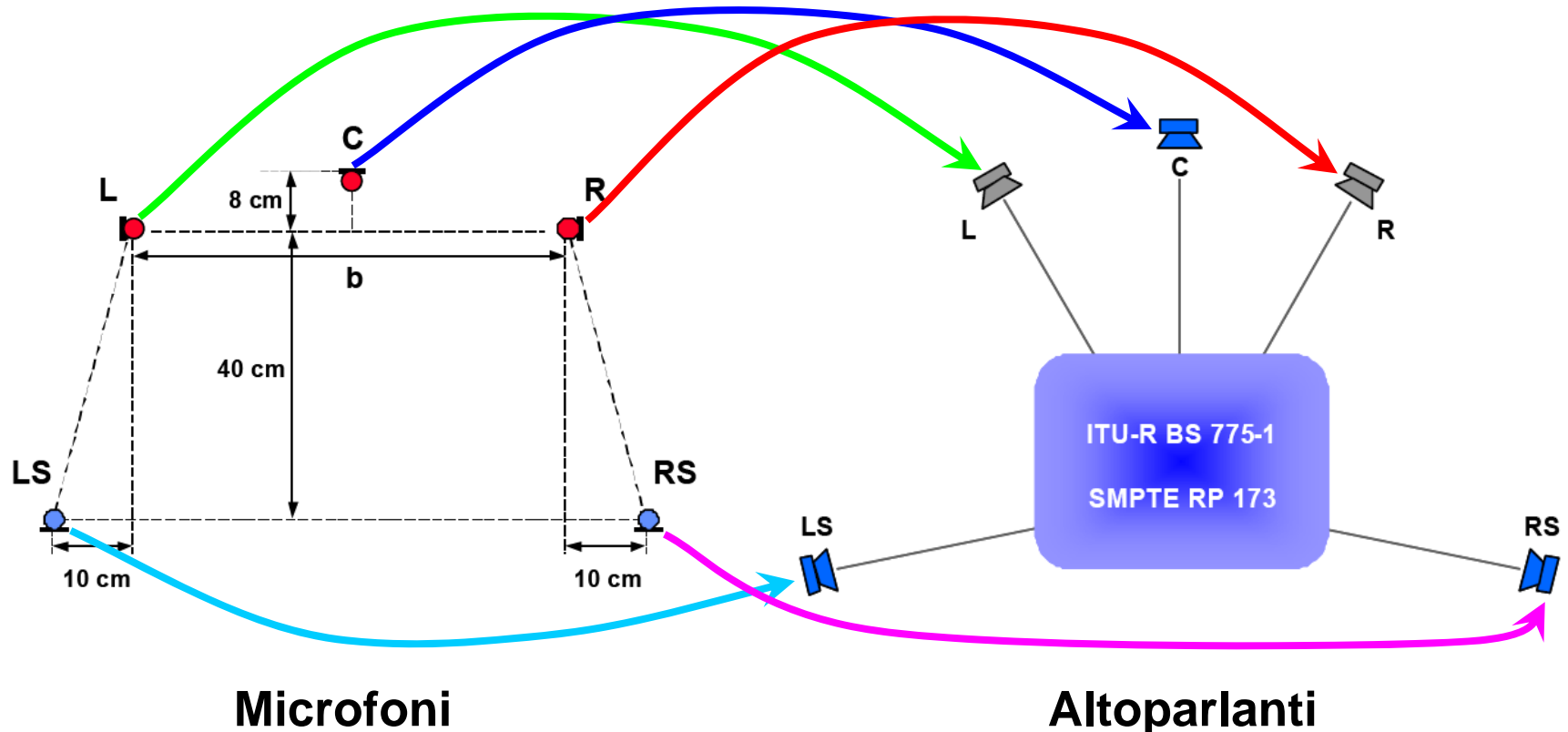
Array “discreti” di microfoni



- Sono stati proposti svariati standard di posizionamento di 5 microfoni discreti con opportuna caratteristica direttiva

Array discreti di microfoni

- In un sistema discreto, il segnale di ogni singolo microfono va ad alimentare un singolo altoparlante



Array “discreti” di microfoni

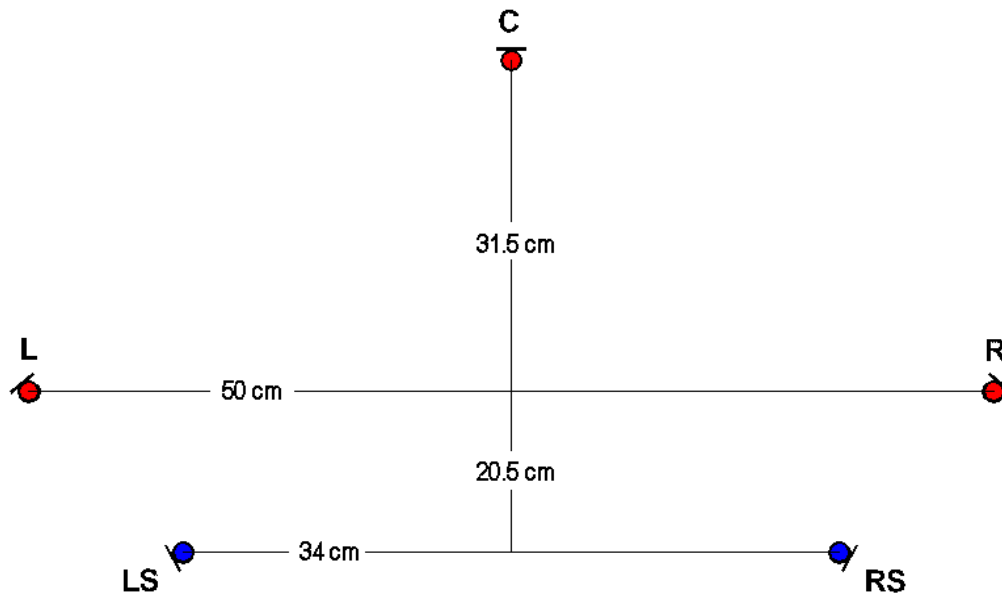
- Williams MMA

**Schema del sistema
microfonico Williams MMA**

C : Cardioide, 0°

L, R : Cardioide, $\pm 40^\circ$

LS, RS : Cardioide, $\pm 120^\circ$



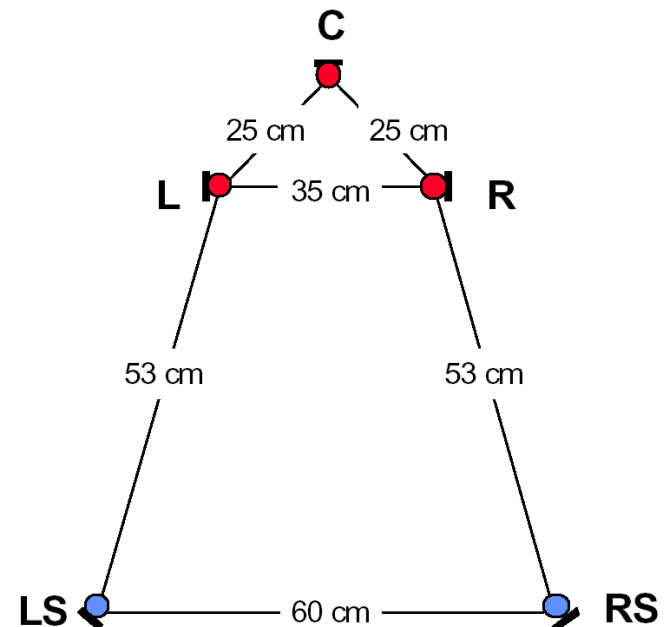
- INA-5

**Schema del sistema
microfonico INA-5**

C : Cardioide, 0°

L, R : Cardioide, $\pm 90^\circ$

LS, RS : Cardioide, $\pm 150^\circ$

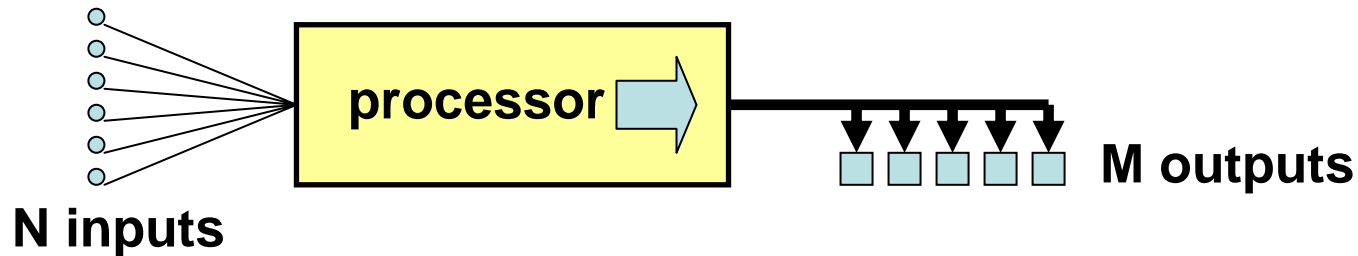


Array discreti di microfoni

Perchè hanno avuto ben poco successo?

- Perchè consentono poco controllo in fase di registrazione
- Forniscono un angolo di ripresa fissato, a meno che siano realizzati con una struttura snodata, peraltro difficile da regolare sul campo
- Non consentono facilmente di mixare i segnali provenienti da tracce mono registrate a breve distanza dalla sorgente (ad esempio voci dei cantanti, strumenti singoli, etc.)
- I pattern direttivi degli usuali microfoni per riprese stereo (cardioidi, ipercardioidi) non corrispondono all'esigenza di avere una legge di "panning" omogenea rispetto alla geometria non regolare del layout degli altoparlanti di un sistema surround 5.1

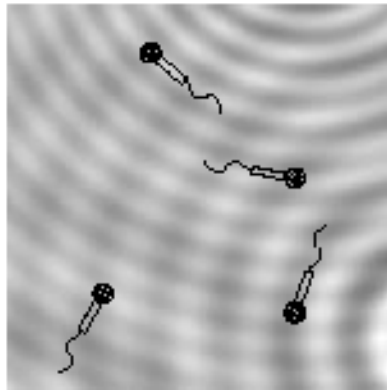
Array microfonici avanzati



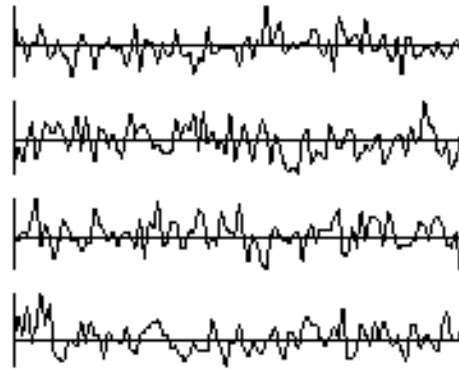
- I segnali di N microfoni x_i , vengono mixati e processati, producendo M segnali y_j
- Qualunque sia la tecnica di processamento impiegata, i segnali in uscita sono esprimibili come:

$$y_j = \sum_{i=1}^N h_{ij} \otimes x_i$$

Possibili tipi di segnali in uscita



Acoustic field



Sampled acoustic field

$x_1(t)$

$x_2(t)$

$x_3(t)$

$x_4(t)$

Filtri precalcolati

$$y_j = \sum_{i=1}^N h_{ij} \otimes x_i$$

- $y_j(t)$ È una delle M uscite del sistema, e può essere:
- il segnale riferito ad una specifica sorgente posta nel punto P_{source}
 - un'onda piana che viaggia in una certa direzione
 - una armonica sferica del campo sonoro incidente nel centro dell'array
 - il segnale corrispondente ad un microfono virtuale direttivo

Tecniche di processamento tradizionali

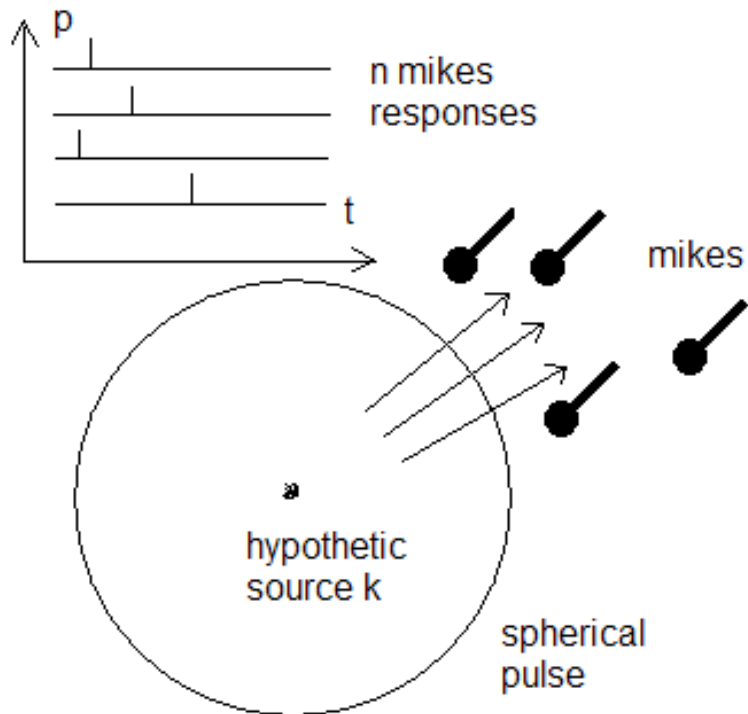
- I filtri numerici h_{ij} sono calcolati per via teorica, risolvendo le complesse equazioni del campo acustico, ed assumendo che i microfoni abbiano un comportamento ideale e siano tutti identici
- Volendo rimuovere l'ultima restrizione, si può aggiungere un filtro compensatore della sensibilità e della risposta in frequenza di ciascun trasduttore

Nuova tecnica di processamento

- Nessuna teoria: i filtri h_{ij} sono ricavati direttamente da una serie di misurazioni di risposta all'impulso, mediante una tecnica ai minimi quadrati
- In pratica, si crea un sistema di equazioni lineari, impiegando come coefficienti i valori misurati delle risposte dei singoli microfoni.
- I risultati ottenuti sono pertanto “massimamente simili” a quelli voluti
- Questo metodo porta automaticamente a compensare anche per le differenze fra i trasduttori e per gli artefatti di origine acustica (schermature, diffrazioni, riflessioni, etc.)

Esempio: focalizzazione su una sorgente puntiforme

Si misurano le risposte all'impulso $c_{k,i}$, cioè la risposta del ***i-esimo*** microfono in presenza della ***k-esima*** onda sferica (proveniente dalla sorgente ***k***, fra le ***P*** possibili sorgenti)



$$C = \begin{bmatrix} c_{1,1} & c_{2,1} & c_{k,1} & c_{P,1} \\ c_{1,2} & c_{2,2} & c_{k,2} & c_{P,2} \\ c_{1,i} & c_{2,i} & c_{k,i} & c_{P,i} \\ c_{1,N} & c_{2,N} & c_{k,N} & c_{P,N} \end{bmatrix}$$

E ricaviamo I filtri inversi che focalizzano su tale sorgente mediante l'inversione della matrice C suddetta:

$$h_{i,k} = H = \frac{1}{C}$$

Esempio: focalizzazione su una sorgente puntiforme

Ci sono due possibili strategie per la ricerca dei filtri inversi: strategia locale e globale.

Nella strategia **locale**, si prendono in esame solo le risposte all'impulso relative alla sorgente k , senza preoccuparsi di come il sistema risponderà ai segnali provenienti dalle altre P possibili sorgenti.

Possiamo quindi invertire una ad una le N risposte all'impulso, imponendo che si verifichi che:

$$\mathbf{h}_{i,k} \otimes \mathbf{c}_{k,i} = \delta \quad (\text{Dirac's delta function})$$

Dopo aver ottenuto i filtri inversi $\mathbf{h}_{i,k}$, usando ad esempio la tecnica di inversione con regolarizzazione di Kirkeby, calcoliamo il segnale proveniente dalla sorgente k -esima come:

$$\mathbf{y}_k = \mathbf{x}_1 \otimes \mathbf{h}_{1,k} + \mathbf{x}_2 \otimes \mathbf{h}_{2,k} + \dots + \mathbf{x}_N \otimes \mathbf{h}_{N,k}$$

Esempio: focalizzazione su una sorgente puntiforme

Con la strategia globale, I filtri vengono calcolati in modo da controllare la risposta del sistema non solo quando la sorgente si trova nella posizione k , ma anche quando si trova in tutte le altre P posizioni. Per ciascuna posizione, si impone che il sistema produca una una risposta v_k predefinita. Nel caso di focalizzazione, si impone che:

$$\sum_{i=1}^N h_{i,1} \otimes c_{1,i} = 0$$

$$\sum_{i=1}^N h_{i,2} \otimes c_{2,i} = 0$$

.....

$$\sum_{i=1}^N h_{i,k} \otimes c_{k,i} = \delta$$

.....

$$\sum_{i=1}^N h_{i,P} \otimes c_{P,i} = 0$$

Il secondo membro costituisce la risposta predefinita v_k , con $k=1..P$

Una volta determinati i filtri inversi che focalizzano ciascuna sorgente escludendo tutte le altre, possiamo estrarre il segnale di ciascuna di esse con la relazione:

$$y_K = \sum_{i=1}^N x_i \otimes h_{i,k}$$

Inversione ai minimi quadrati

- Il metodo dei minimi quadrati consente di determinare i valori dei filtri h_{ik} .
- Si definisce un “errore quadrato totale” ε_{tot} come:

$$\varepsilon_{\text{tot}} = \sum_{k=1}^P \left[\sum_{i=1}^N (h_{ik} \otimes c_{ki}) - v_k \right]^2$$

- Per ciascuna posizione della sorgente k , si crea un sistema lineare di N equazioni minimizzando ε_{tot} , imponendo che:

$$\frac{\partial \varepsilon_{\text{tot}}}{\partial h_{ik}} = 0 \quad (i = 1 \dots N)$$

Limiti operativi

- Basse frequenze: lunghezza d'onda molto più grande dell'array - l'approccio locale fornisce una risoluzione molto modesta - l'approccio globale ha problemi di convergenza numerica dovuti al fatto che tutti i segnali sono molto simili.
- Medie frequenze: lunghezza d'onda confrontabile con le dimensioni dell'array - l'approccio locale dà luogo a deboli lobi parassiti - l'approccio globale funziona perfettamente, eliminando i lobi parassiti e fornendo un lobo principale stretto e coerente.
- Alte frequenze: la lunghezza d'onda diviene confrontabile con la spaziatura fra i microfoni - errore di aliasing spaziale – la disposizione causale dei microfoni consente all'approccio locale di fornire ancora una risoluzione accettabile, ma rende di nuovo numericamente instabile la soluzione con l'approccio globale

Array lineare a spaziatura variabile



- Registrazione con Adobe Audition
- Calcolo dei filtri inversi e convoluzione degli stessi mediante i plugins Aurora

- 16 microfoni omnidirezionali montati con spaziatura esponenziale variabile da 20 mm a 120 mm
- sistema a 16 canali:
2 Behringer A/D converters + RME Hammerfall digital sound card

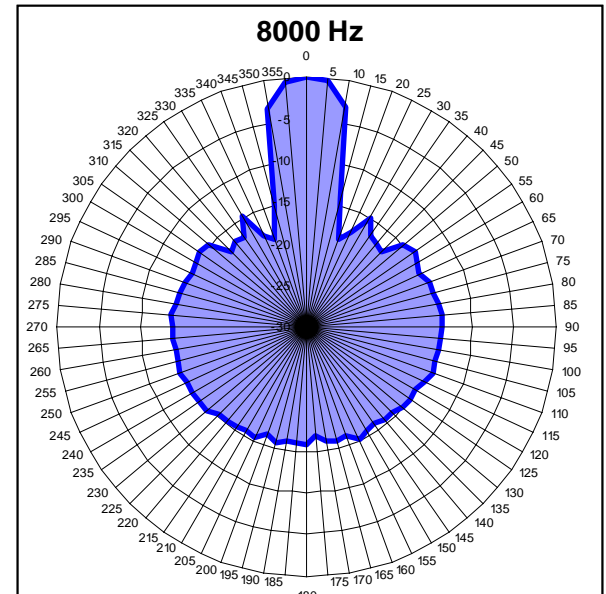
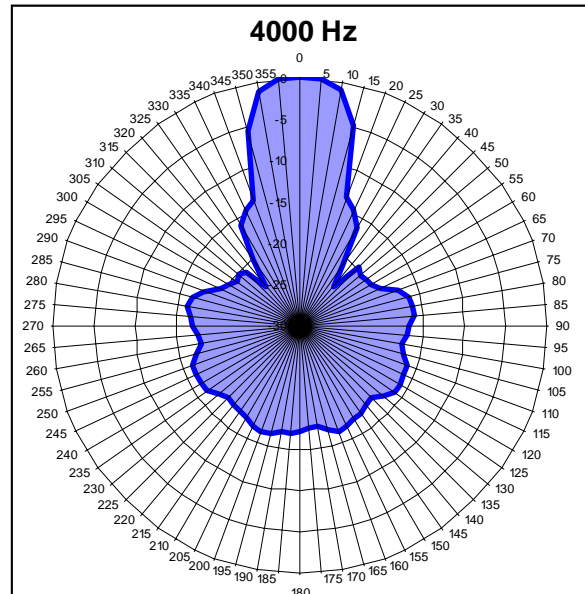
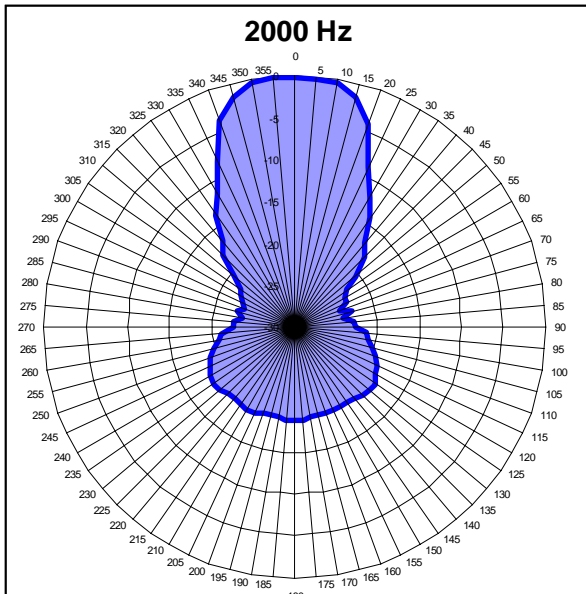
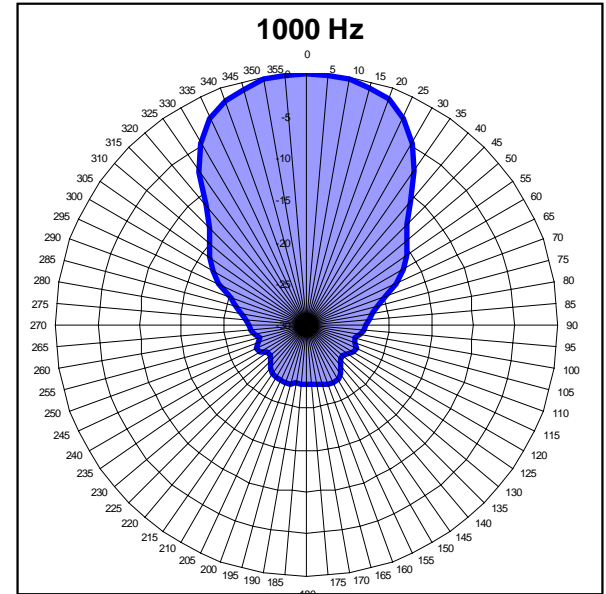
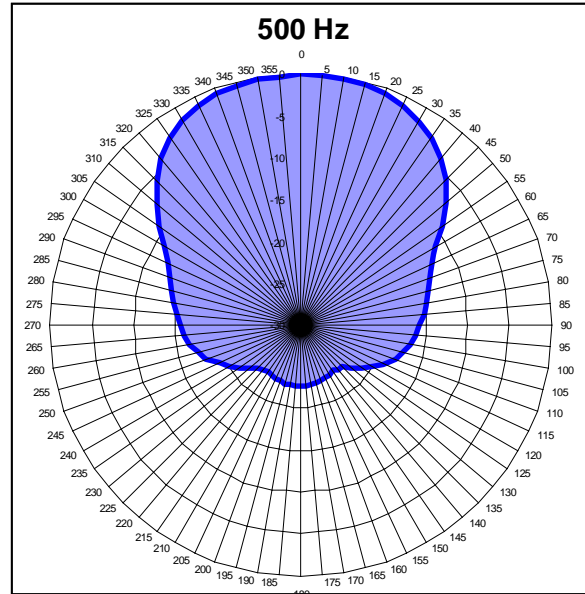
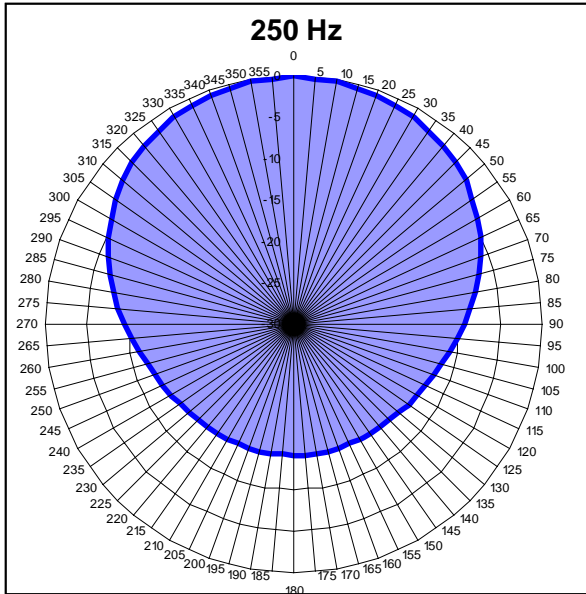


Calibrazione dell'array lineare



- Tavola Rotante Outline
- Altoparlante Mackie HR24
- 72 risposte all'impulso misurate con sine sweep logaritmico (Audition+Aurora).
- La tavola rotante viene fatta avanzare da un impulso generato sul canale 2
- I filtri inversi sono stati calcolati con l'approccio locale (utilizzando separatamente 16 volte il plugin "Kirkeby4")

Curve di direttività dell'array lineare

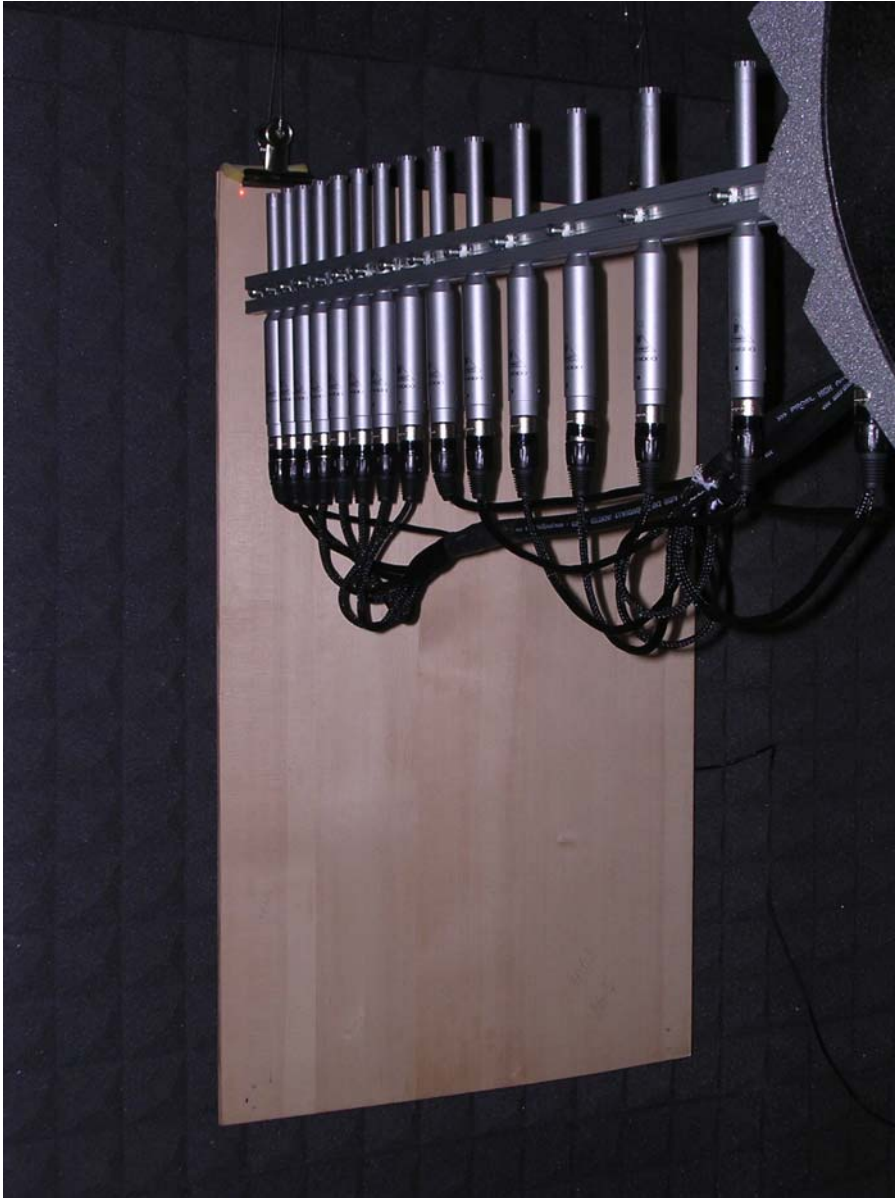


Utilizzo dell'array lineare



- L'array viene movimentato tramite un sistema di scansione X-Y
- un vibrometro laser Polytec è montato tangente all'array
- Il sistema viene impiegato per mappare i modi di vibrare di una tavola di legno (Abete della Val di Fiemme)
- Un computer con Labview e scheda national controlla i motori passo del sistema X-Y
- L'apparato, denominato "modografo", è attualmente in funzione presso il laboratorio del CNR-IVALSA di San Michele all'Adige

Utilizzo dell'array lineare



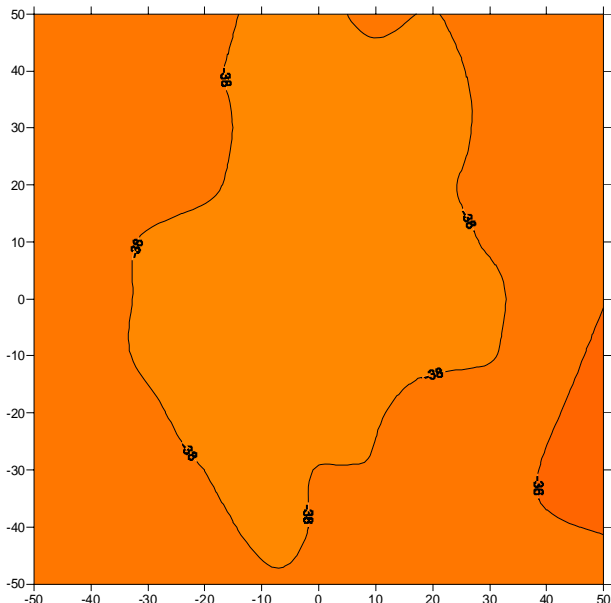
Il pannello di legno è eccitato da un piezo

Si ottengono due tipi di risultati:

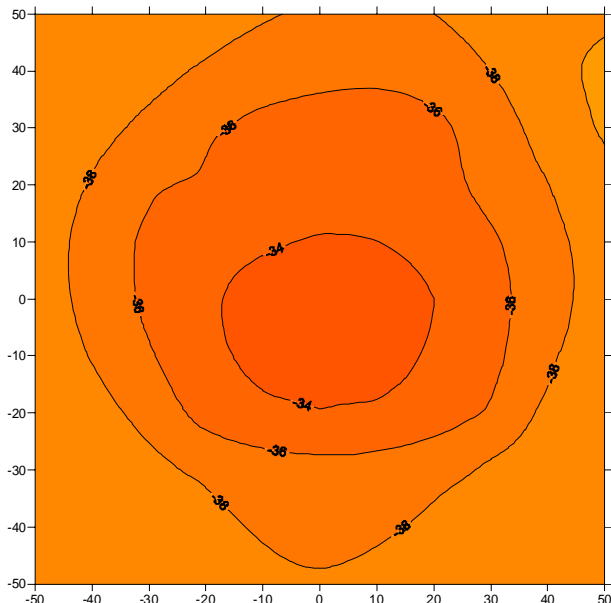
- Lo spettro in frequenza mediato sulla superficie dei parametri acustici più interessanti: pressione sonora, velocità delle particelle, intensità sonora, potenza irradiata
- Una mappa a colori, a ciascuna frequenza, che mostra il modo di vibrare della tavola

Mappature di prova (pannello con altoparlante)

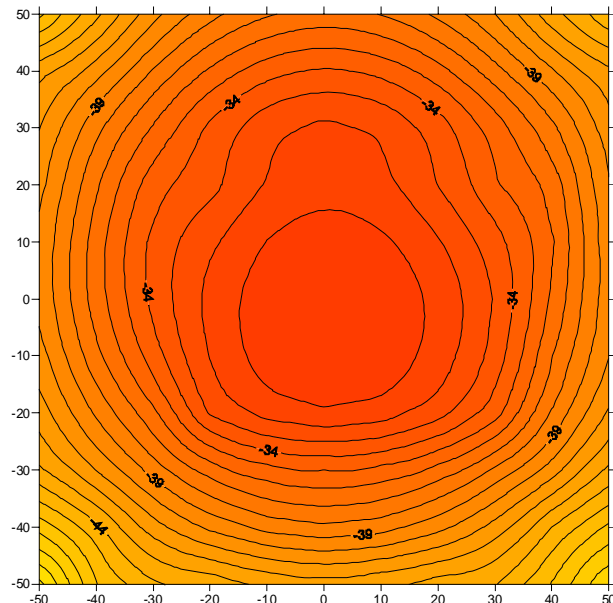
250 Hz



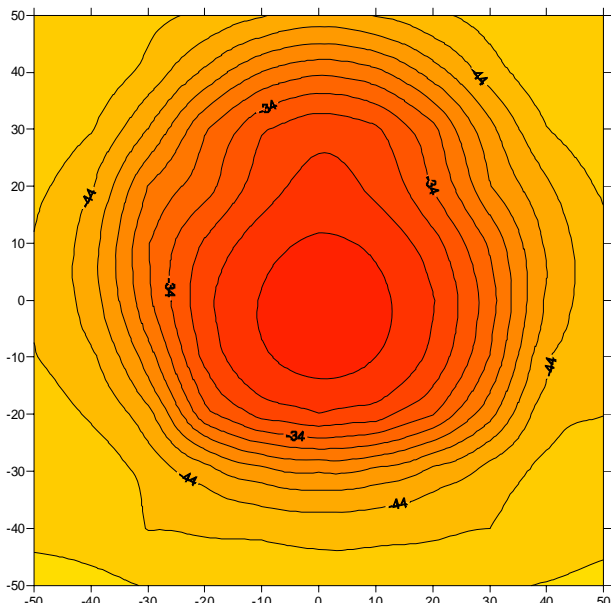
500 Hz



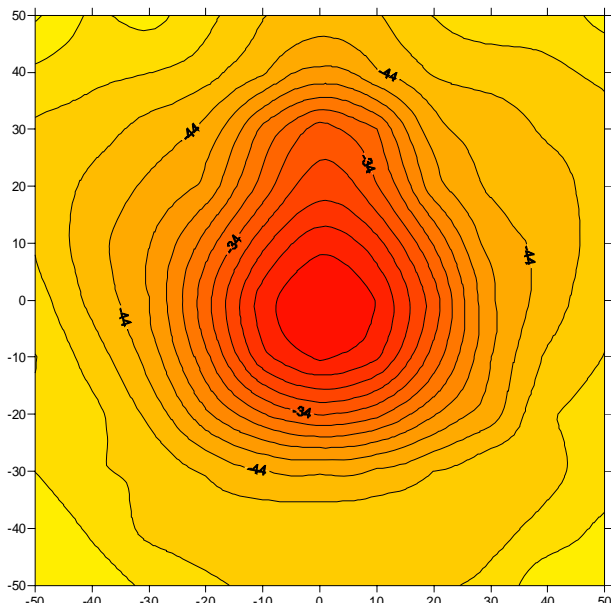
1000 Hz



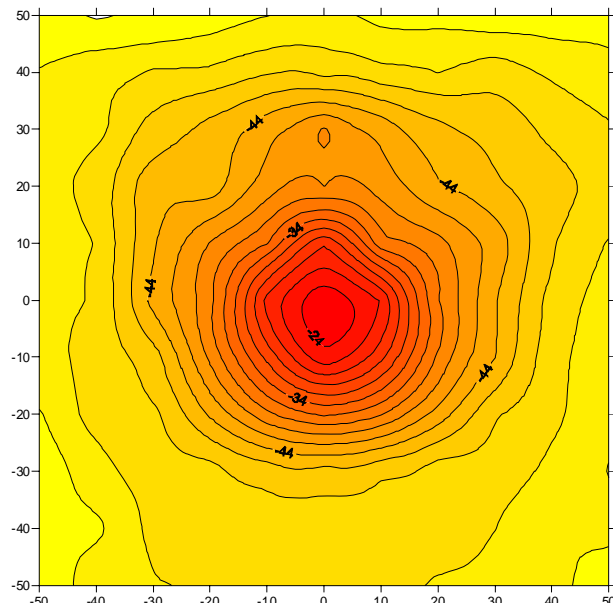
2000 Hz



4000 Hz



8000 Hz



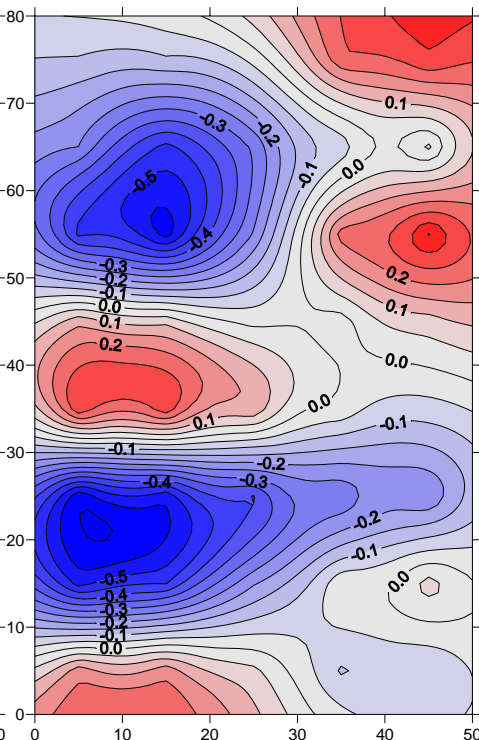
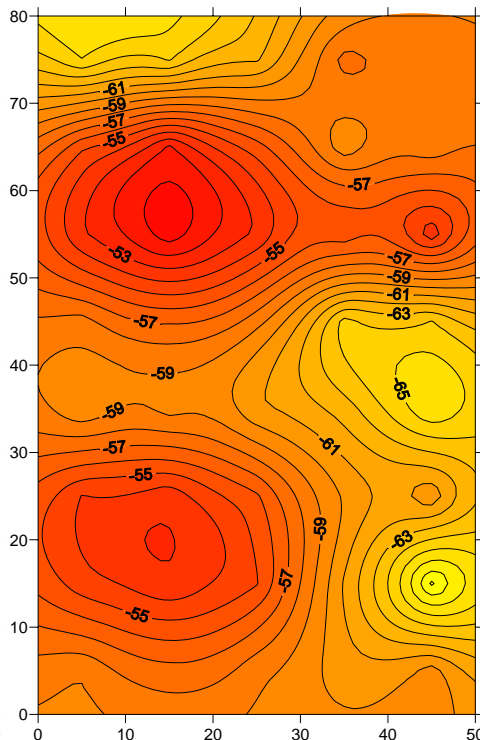
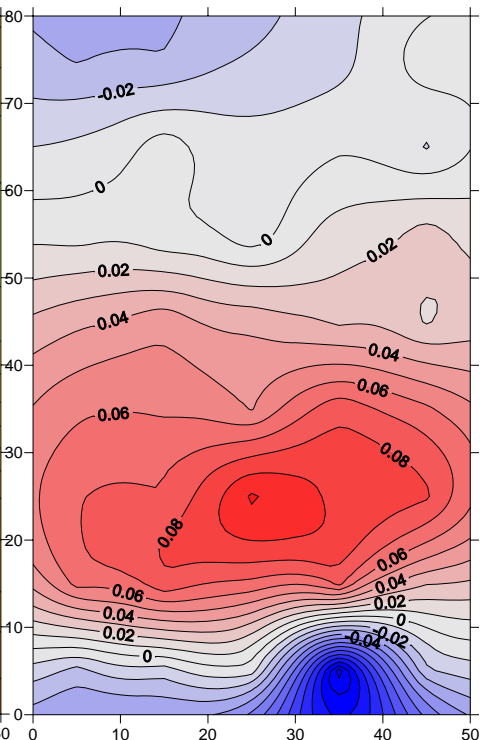
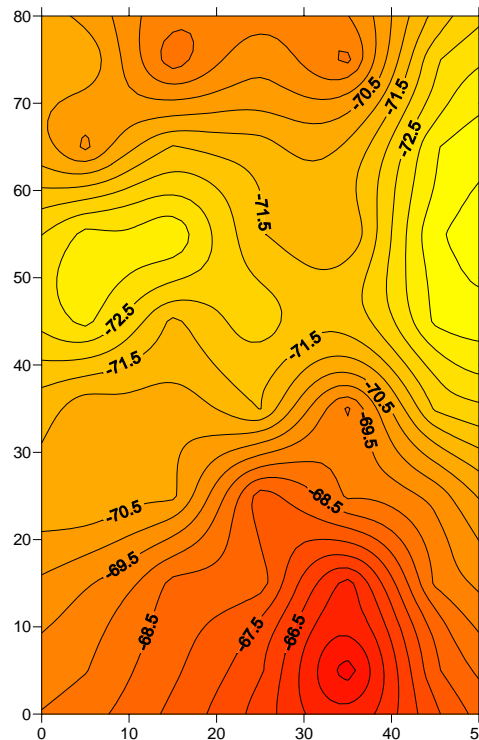
Mappature effettive su tavola di legno armonico

468 Hz

469 Hz

1359 Hz

1312 Hz



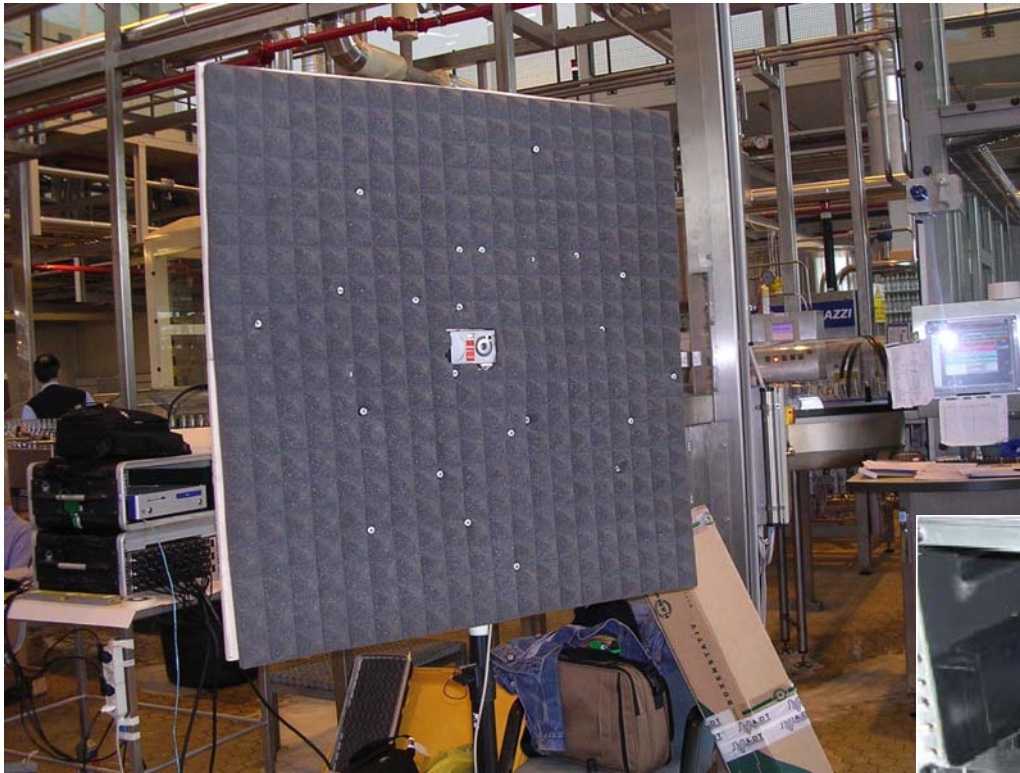
SPL (dB)

velocity (m/s)

SPL (dB)

velocity (m/s)

Array planare (“acoustic camera”)

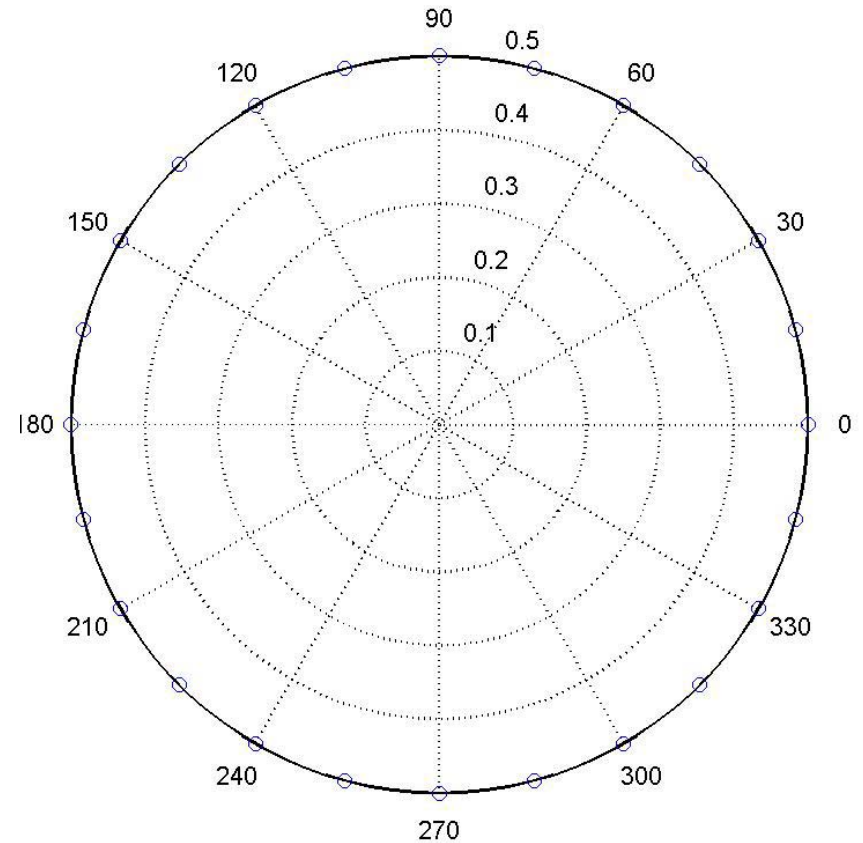
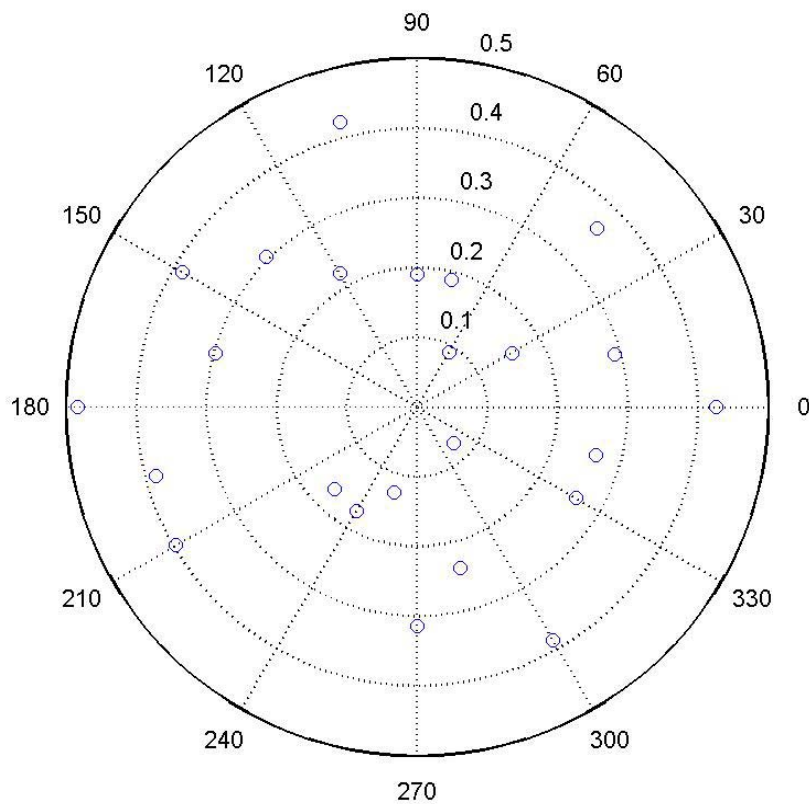


- 24 microfoni omnidirezionali montati su un pannello ripiegabile di 1m x 1m, con disposizione pseudo-random ottimizzata
- sistema audio a 24 canali: Behringer A/D converters + RME Hammerfall digital audio card

- Calcolo dei filtri inversi ed elaborazione grafica off-line in Matlab

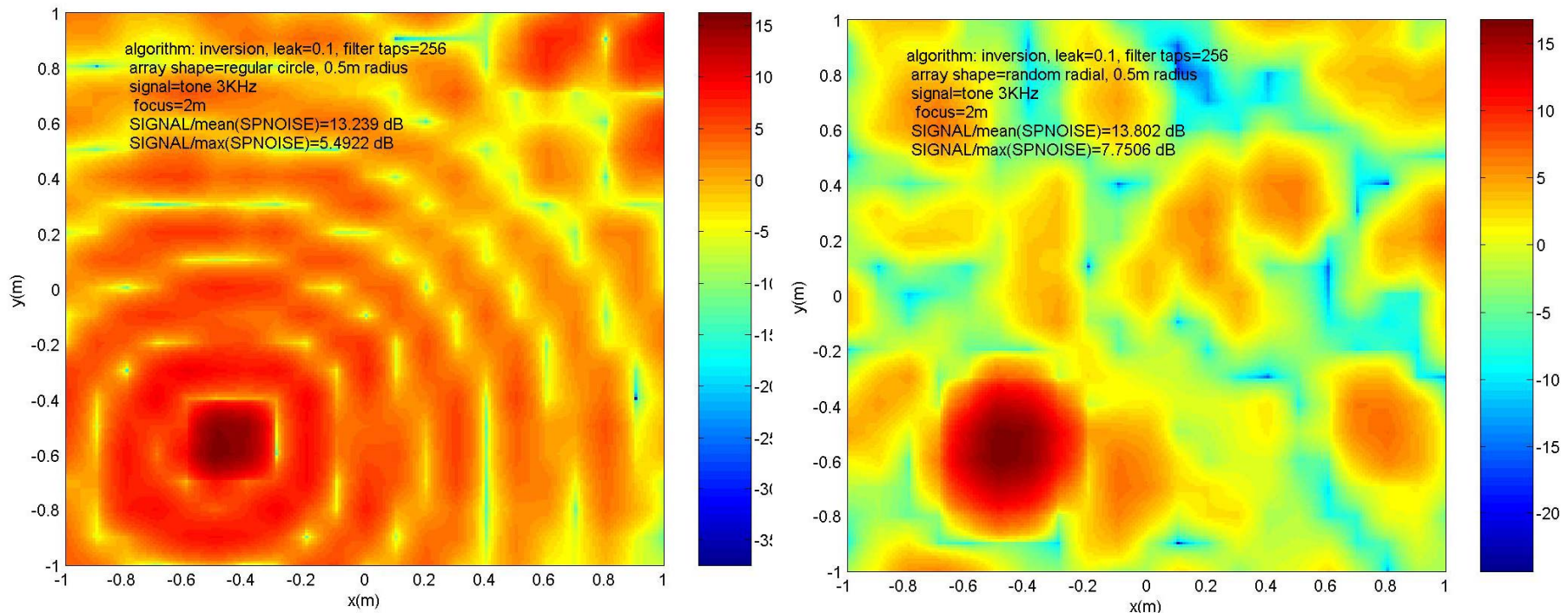


Confronto fra array randomizzato ed array circolare regolare



Confronto fra array randomizzato ed array circolare regolare

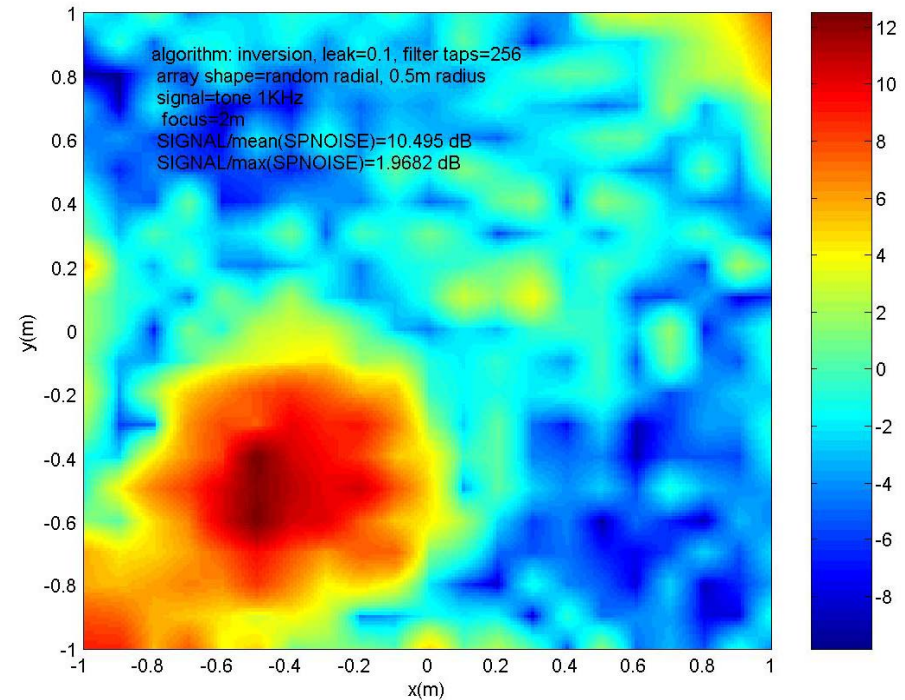
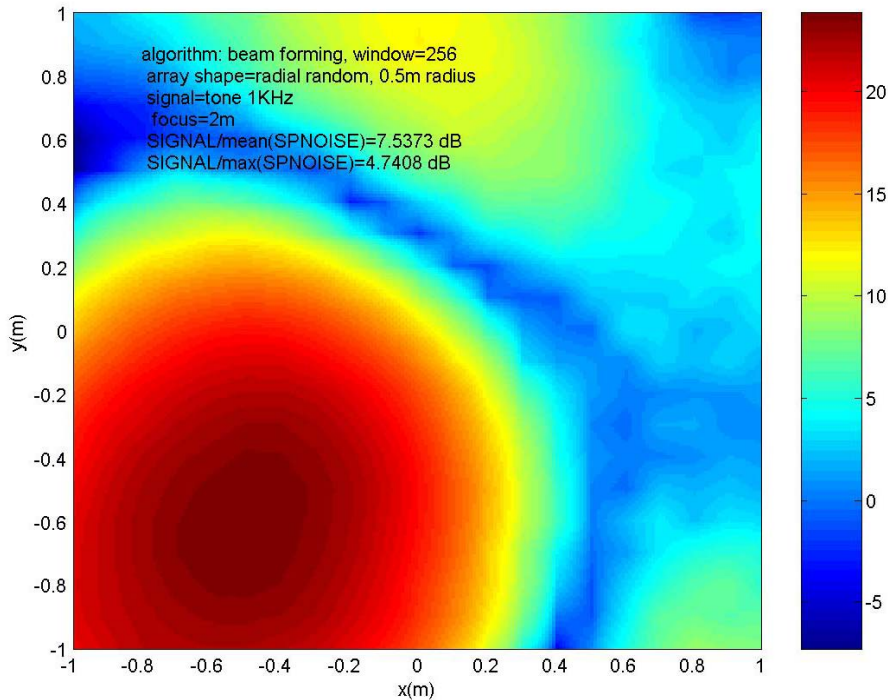
Simulazione Matlab - 3 kHz



La posizione ottimale dei microfoni è stata scelta generando 10000 posizionamenti casuali, e scegliendo quello che ha fornito il miglior rapporto picco/rumore

Approccio locale - Approccio globale

Random Array - Simulazione Matlab - 1 kHz



Localizzazione sorgenti di rumore



- L'array è equipaggiato di una WebCam
- Si crea una mappa a colori focalizzando l'array su ciascun "pixel" di una matrice 20x30 di possibili sorgenti disposte su un piano

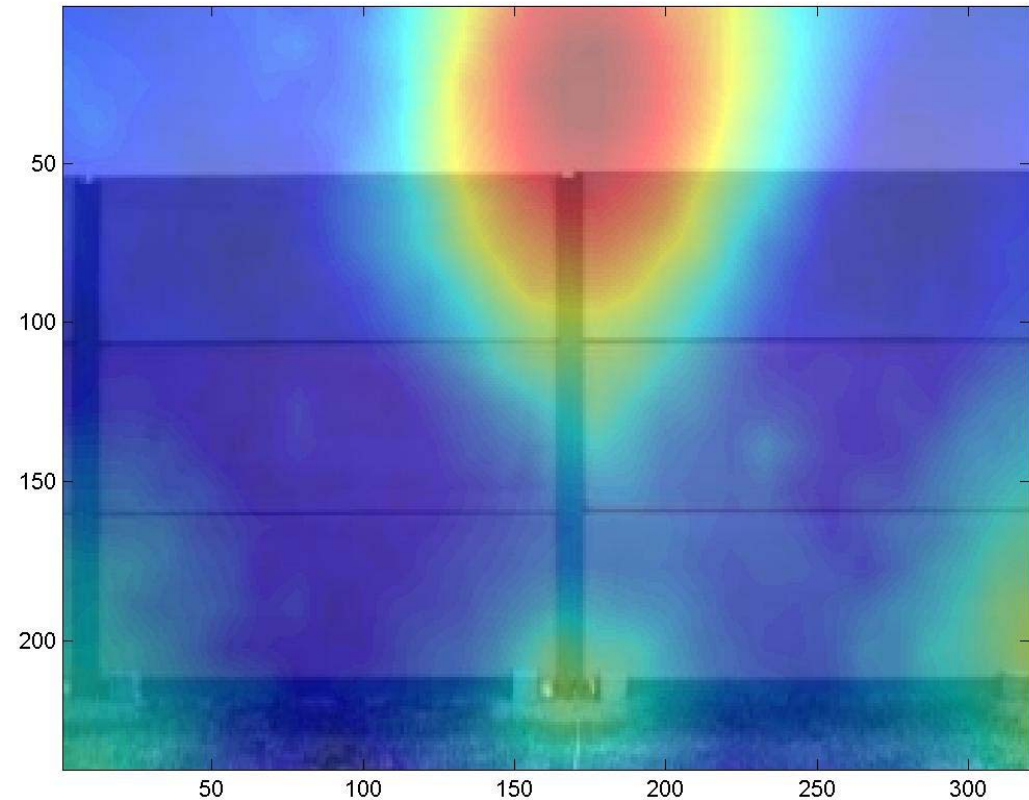
- Si calcola il livello sonoro in dB(A) del segnale proveniente da ciascuna sorgente
- La mappa dei livelli sonori così calcolati è infine sovrapposta all'immagine della WebCam



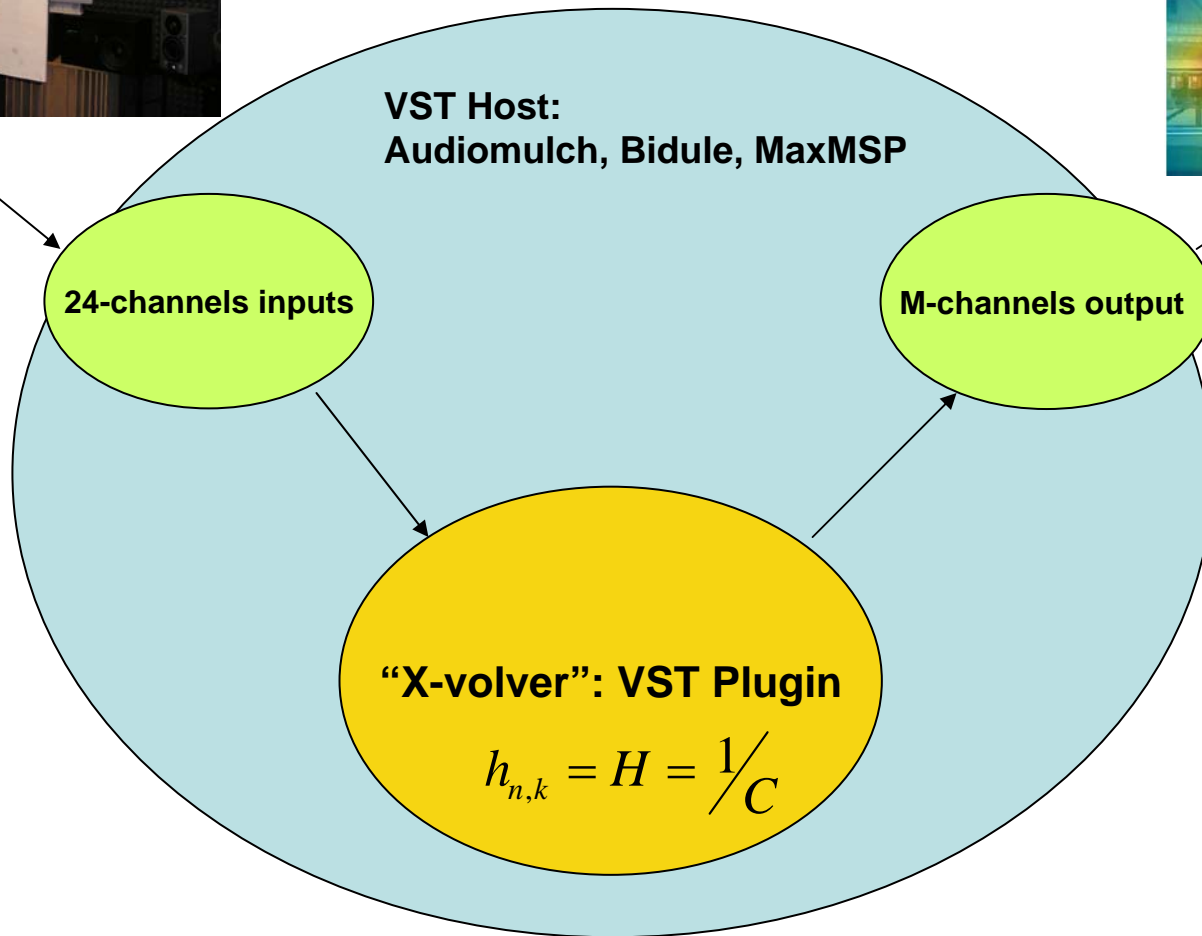
Applicazione all'aperto



- L'array è posizionato dietro una barriera antirumore
- La mappa a colori mostra il suono che trafila dai buchi alla base dei montanti



Il futuro: funzionalità in realtime

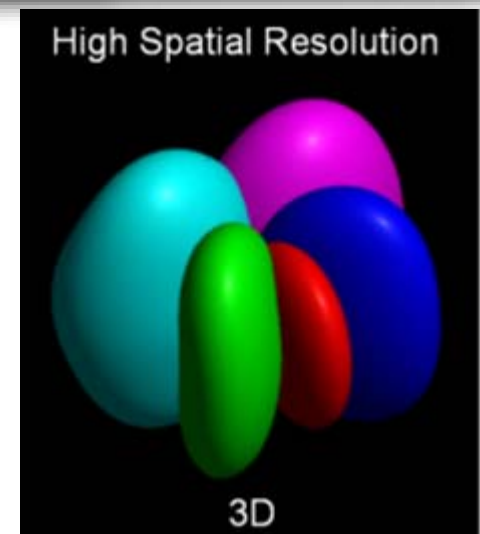
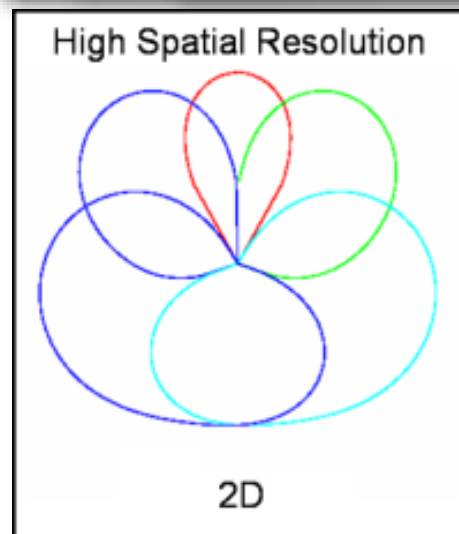


“X-volver” VST plugin

The screenshot displays the X-Volver 16x1 VST plugin interface within the AudioMulch environment. The window title is "X-Volver 16x1 - AudioMulch". The interface includes a menu bar (File, Edit, View, Control, Help), a toolbar with various controls, and a main workspace. On the left, a rack of 16 auxiliary inputs (AuxIn1 to AuxIn8) is visible, with a "multivolver_1" instance connected to the "SoundOut" output. The main workspace shows the "multivolver_1" instance with a large "X" logo and a waveform display. A red box highlights the "16 inputs" label, and another red box highlights the "1 output" label. The interface also features a "Main Mixer" section with a "Dry/Wet" knob and a "Reset" button. At the bottom, there are sections for "File info" (WAVEX, 32 bit float, 16384 samples, SR: 48000, mono) and "Options" (IR mono, IR multi, IR 1-To-1, FFT size: 512, IR size: 1024). The CPU load is shown as 5.76% at the bottom left.

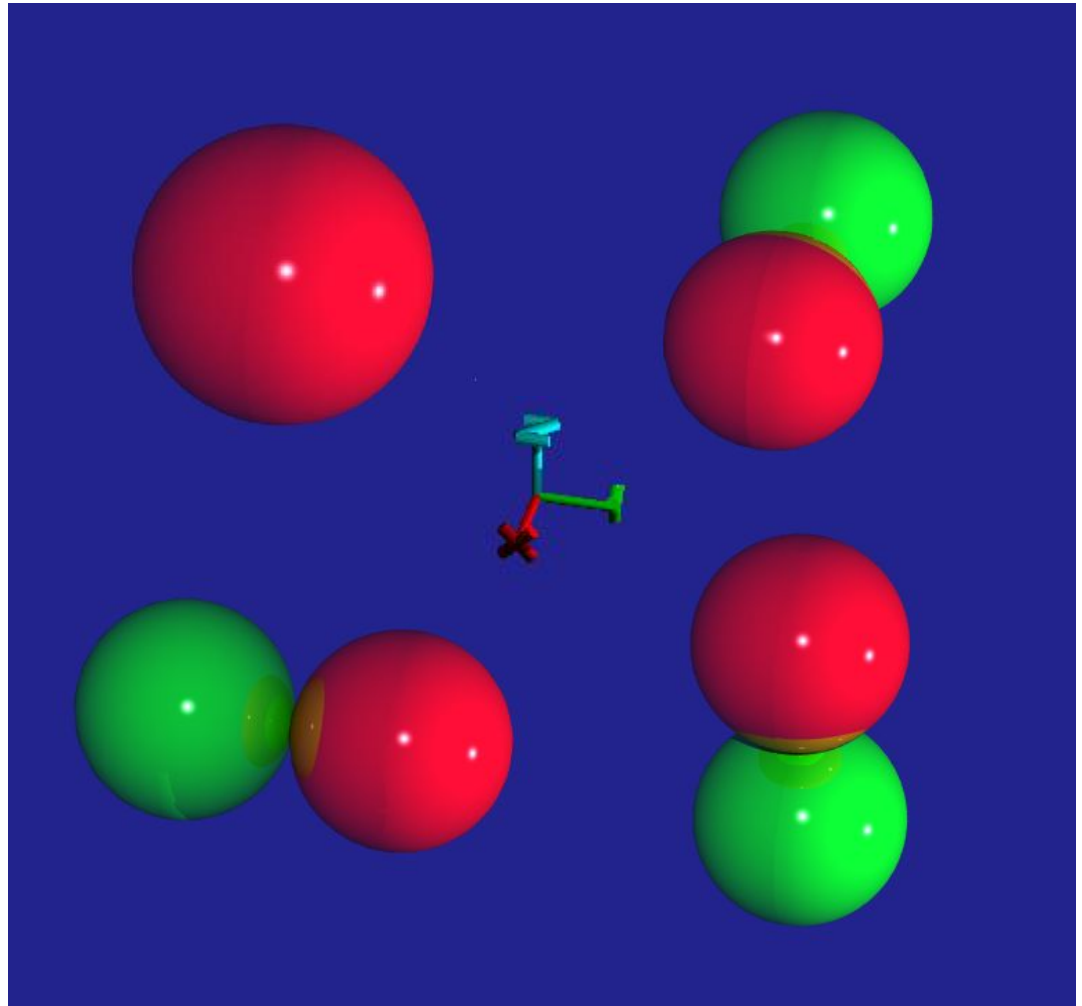
Array 2D per surround

- La Trinnov produce un sistema ad array orizzontale per registrazioni “surround”



Array tridimensionali

- Il primo sistema di misura ad array tetraedrico tridimensionale è stato il microfono Soundfield



Array tridimensionali

- Ora sono disponibili anche array tetraedrici alternativi (Trinnov, DPA) con processamento digitale dei segnali



Misura delle risposte all'impulso del DPA-4



84 IRs were measured, uniformly scattered around a sphere



Calcolo dei filtri inversi

- Sono richiesti in totale 16 filtri inversi (4 inputs, 4 outputs)
- Per ciascuna delle 84 misure effettuate è possibile definire una risposta ideale dell'array, corrispondente alle direttività di ordine 0 ed 1 (W,X,Y,Z)
- Si possono dunque porre $84 \times 4 = 336$ condizioni:

$$\left. \begin{aligned} c_1 \otimes h_{1,W} + c_2 \otimes h_{2,W} + c_3 \otimes h_{3,W} + c_4 \otimes h_{4,W} &= \text{out}_{k,W} \\ c_1 \otimes h_{1,X} + c_2 \otimes h_{2,X} + c_3 \otimes h_{3,X} + c_4 \otimes h_{4,X} &= \text{out}_{k,X} \\ c_1 \otimes h_{1,Y} + c_2 \otimes h_{2,Y} + c_3 \otimes h_{3,Y} + c_4 \otimes h_{4,Y} &= \text{out}_{k,Y} \\ c_1 \otimes h_{1,Z} + c_2 \otimes h_{2,Z} + c_3 \otimes h_{3,Z} + c_4 \otimes h_{4,Z} &= \text{out}_{k,Z} \end{aligned} \right\} \mathbf{k = 1 \dots 84}$$

Implementazione realtime

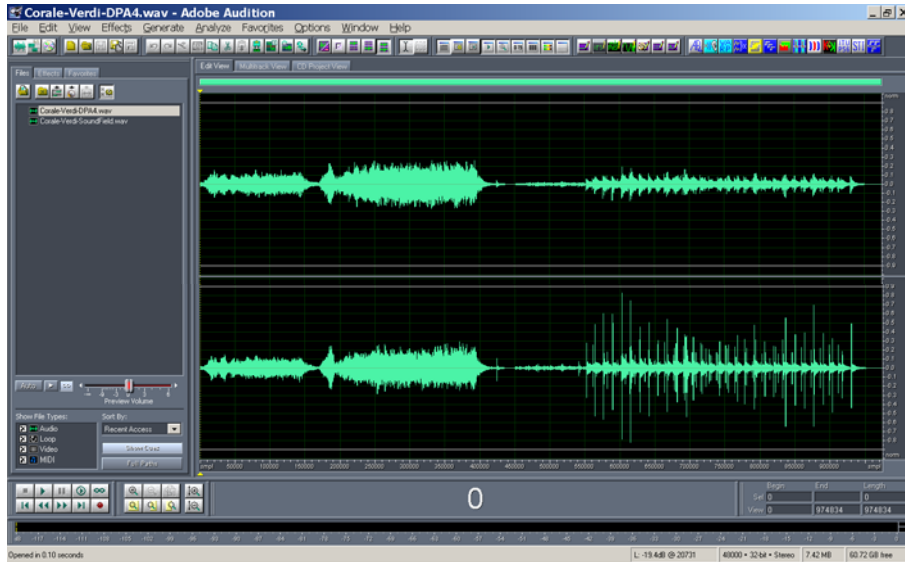
The screenshot displays the X-Volver 4x4 - AudioMulch software interface. The main window, titled "multivolver_1", features a grid of 16 impulse response (IR) waveforms labeled INV01.wav through INV16.wav. A red box highlights the top-left corner of the grid with the text "4 inputs". A large red bracket on the right side of the grid is labeled "4 outputs". The interface includes a "Main Mixer" section with volume controls (0.00 dB Dry, -oo dB Wet) and a "Reset" button. At the bottom, there are sections for "File info" (WAV (Microsoft), 32 bit float, samples: 1024, SR: 48000, mono) and "Options" (Mode: IR mono selected, FFT size: 512, IR size: 1024). The software's menu bar (File, Edit, View, Control, Help) and transport controls are visible at the top.

Confronto fra 3 diversi sistemi

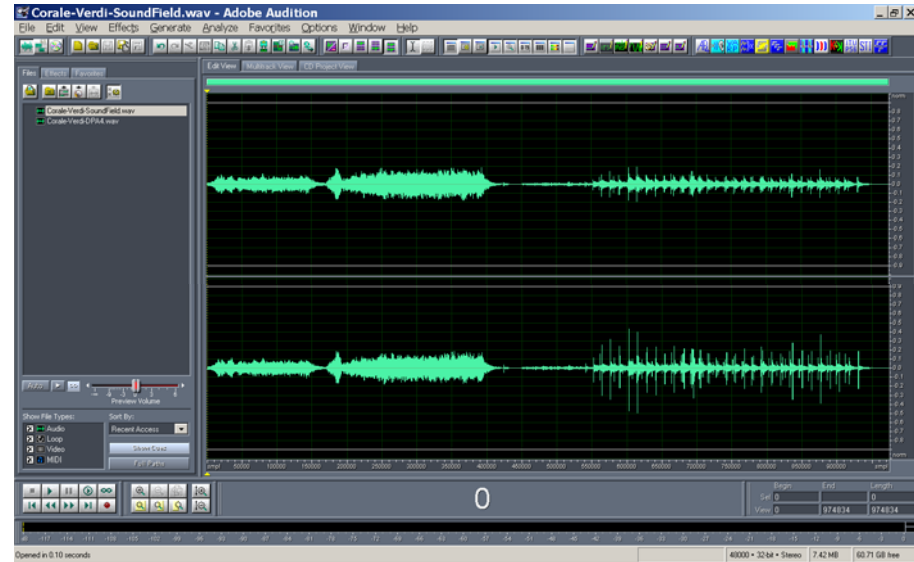
- 2 crossed Neumann K-140 were compared with a pair of virtual cardioids derived from B-format signals, recorded either with a Soundfield ST-250 and with the new DPA-4



Sound samples



DPA-4

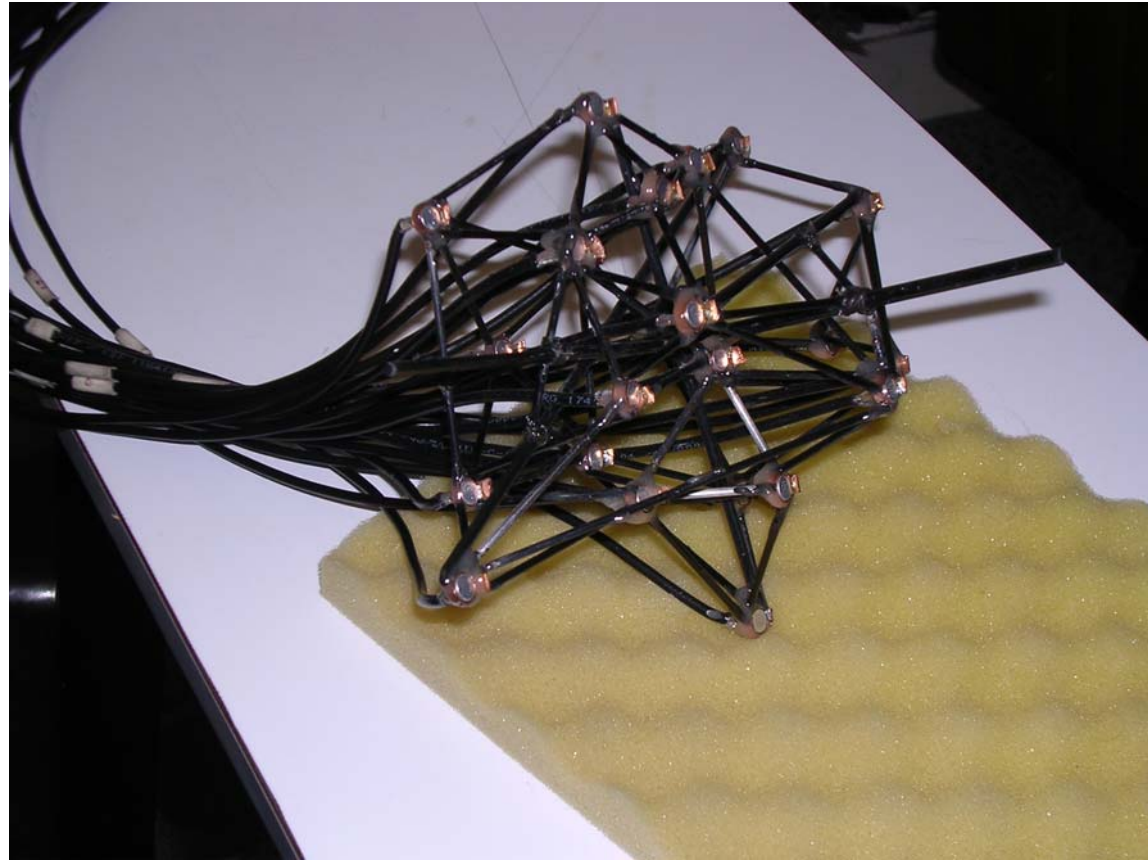


Soundfield

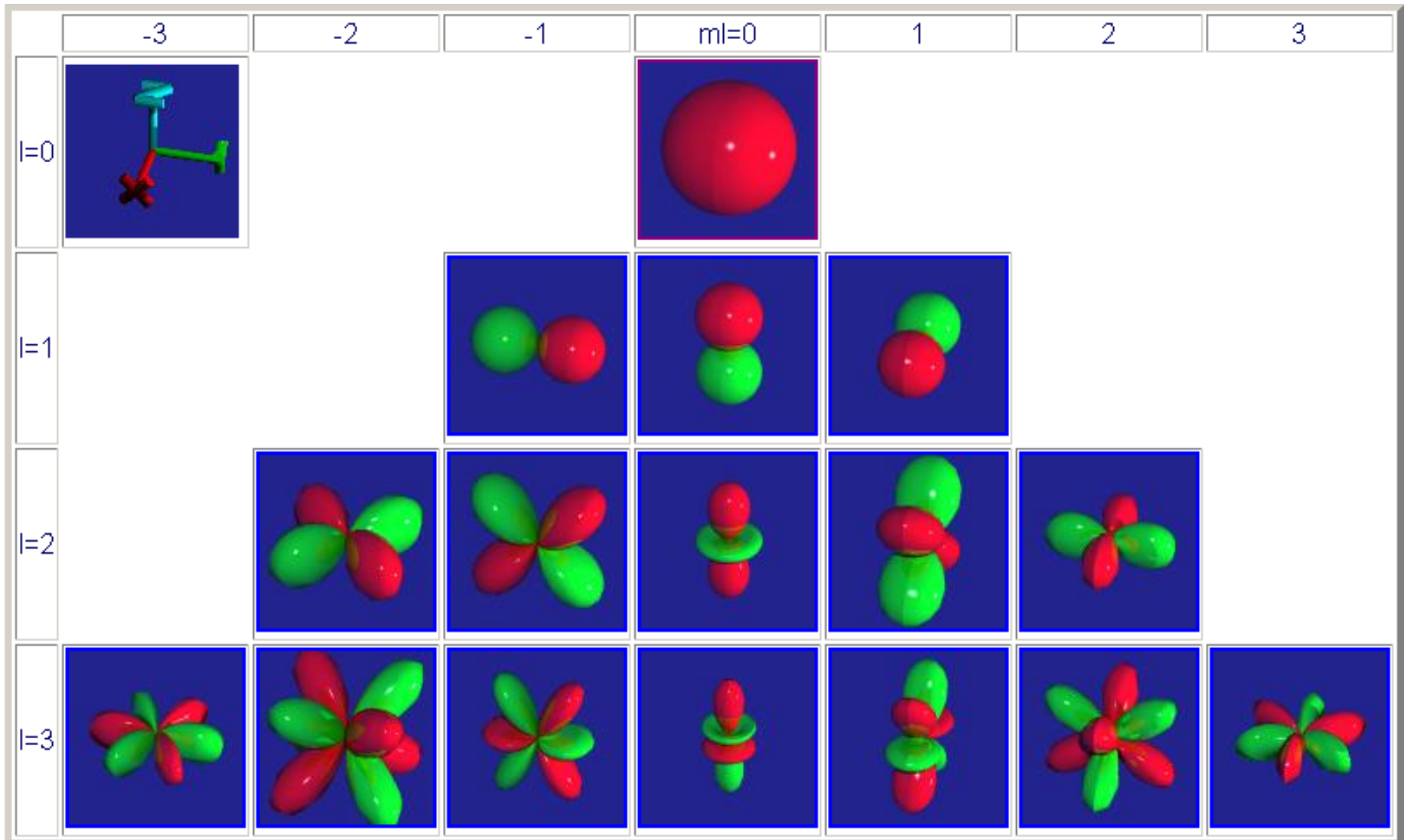
- The new DPA-4 outperforms the Soundfield in terms of stereo separation and frequency response, and is indistinguishable from the “reference” Neumann cardioids

Array tridimensionali di ordine elevato

- La Trinnov ha sviluppato un array 3D costituito da 24 capsule a spaziatura pseudocasuale

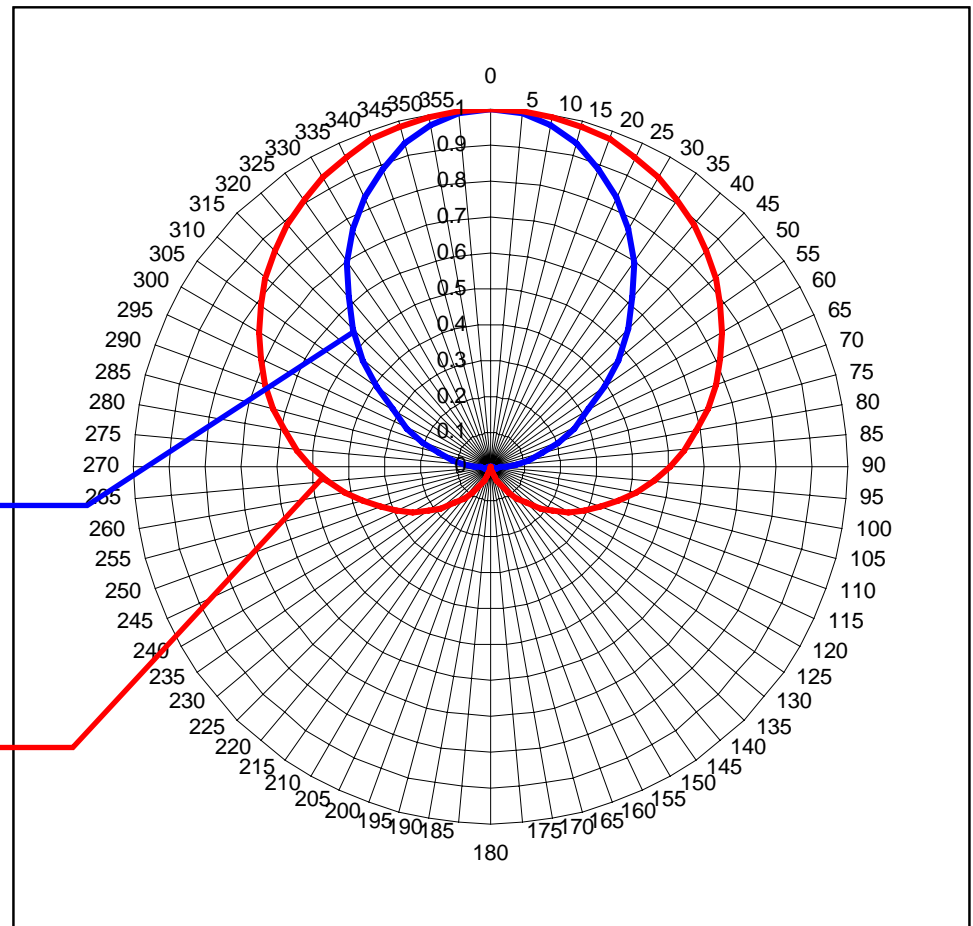


Risposte direttive del 3° ordine



Utilizzo dei segnali armonici sferici

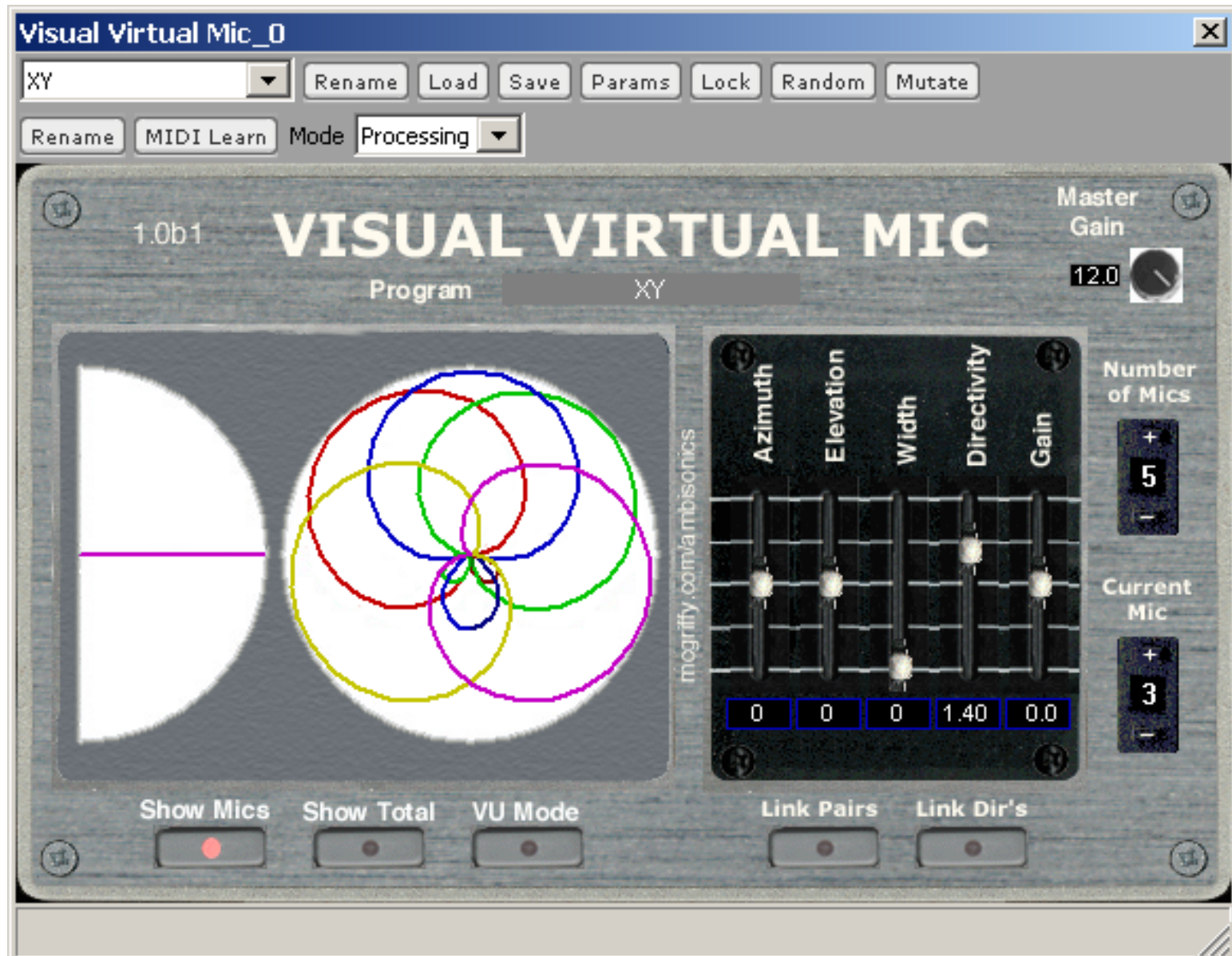
- E' possibile sintetizzare un numero arbitrario di microfoni, con direttività ed orientazione scelti a piacimento



**Cardioide avanzato
(ordini 0, 1, 2 e 3)**

**Cardioide classico
(ordini 0 e 1)**

Visual Virtual Microphone



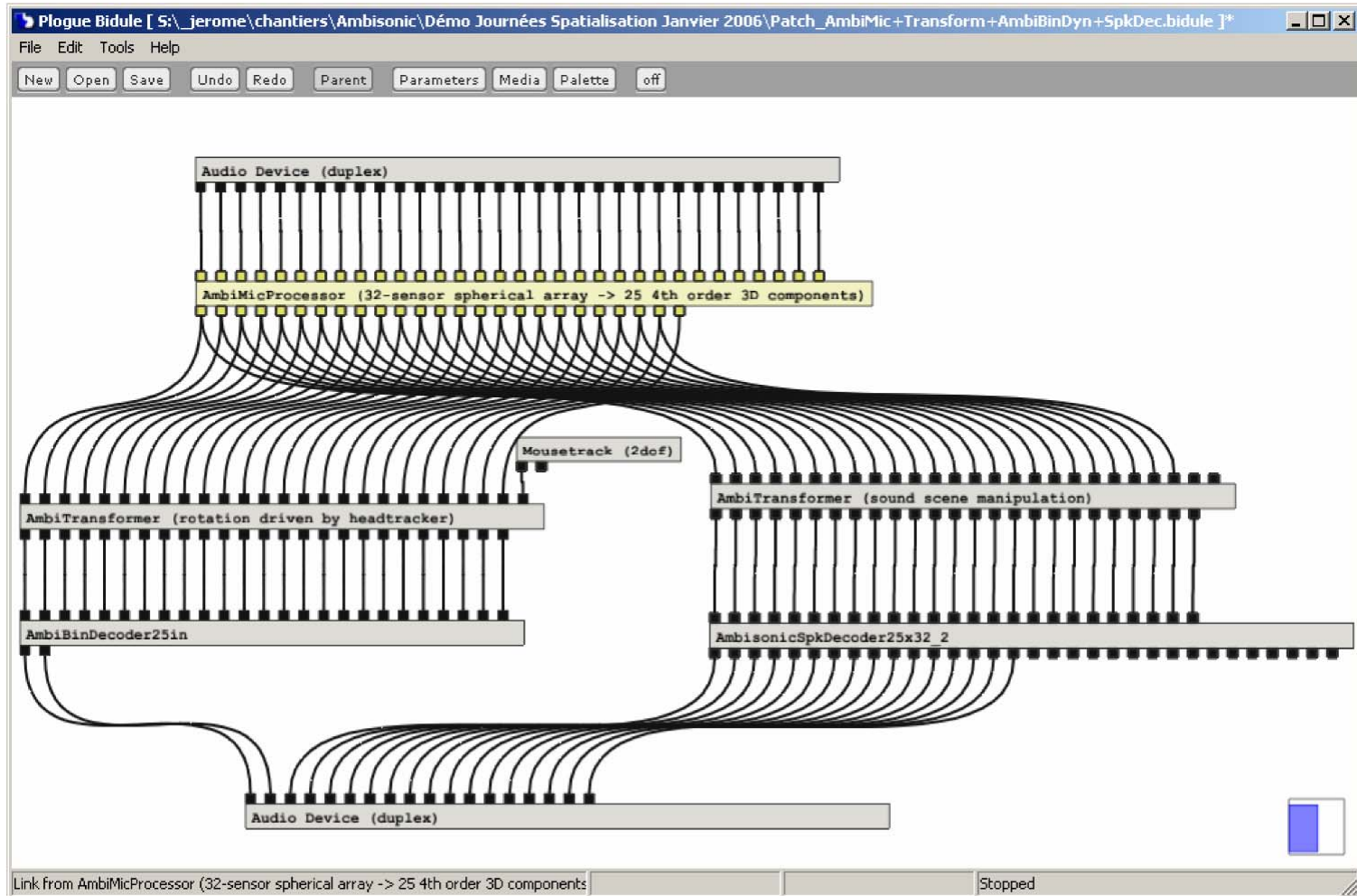
Ambisonics Decoders

- Un decoder Ambisonics è semplicemente un array di microfoni virtuali, ciascuno dei quali alimenta un altoparlante



Programmi Host VST multicanale

- I programmi utilizzabili come host multicanale (sino a 32 canali) sono plogue Bidule, AudioMulch e Max/MSP



Audio Mulch

Surround-Panner - AudioMulch

File Edit View Control Help

1 · 1 120.0

VSTPanner5_1_ver1_2_1

VST Panner Surround 5.1 ver. 1.2 by Elvio Megliecca

Gain 0.00 LFE 0.00

45°

120°

Monitor Setup

Left/Right 30 45°

Left/Right 45

Surround 110 120°

Surround 120

SoundIn

AuxIn3

SoundOut

AuxOut1

AuxOut2

AuxOut3

Use Input File

File: C:\...\midasstudios-stringquartet-w.wav

Auto Rewind duration: 177.74 (sec)

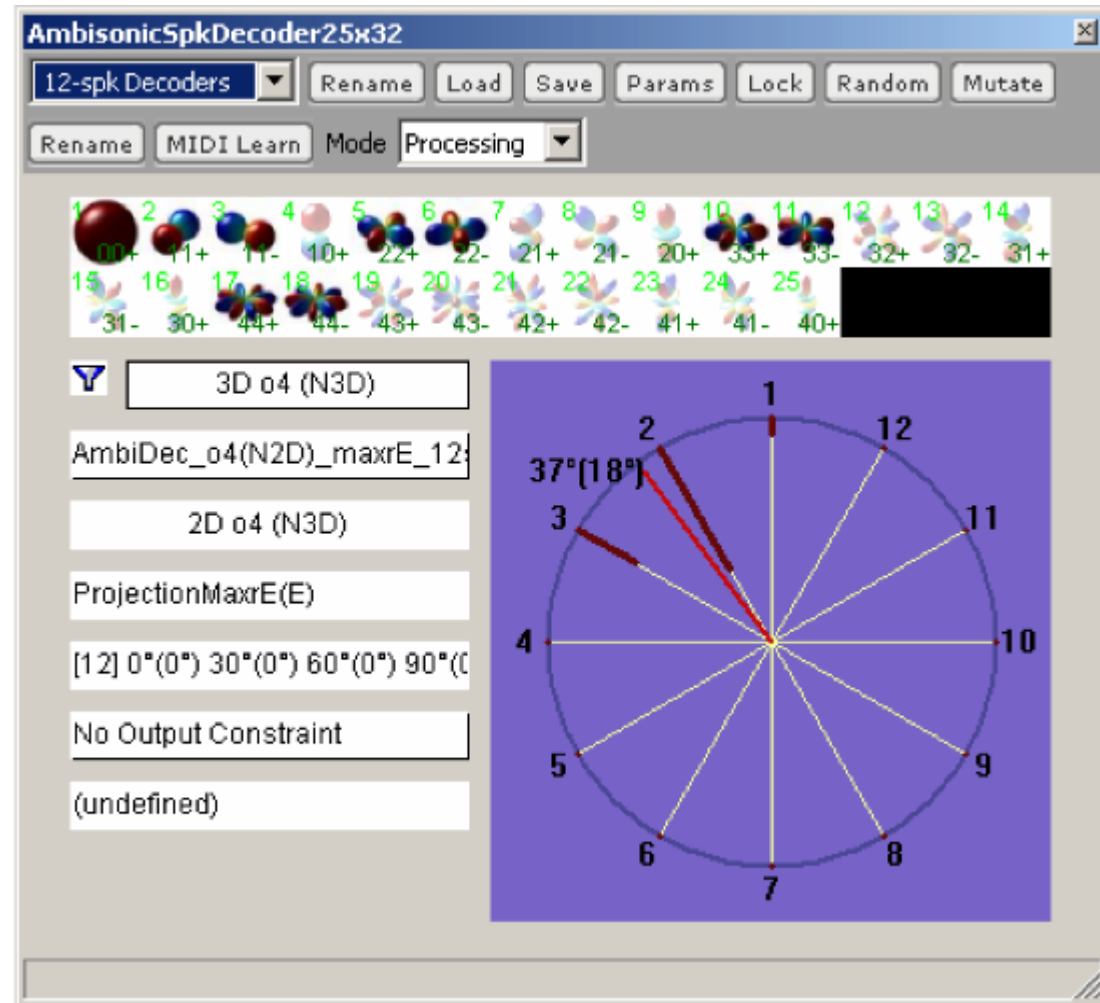
Loop Delay: 0 (seconds)

1-3 2-1 2-3 3-1 3-3 4-1 4-3 5-1 5-3 6-1 6-3 7-1 7-3 8-1

CPU load: 26.60 Disk read: 89 kB/sec

High Order Ambisonics (HOA)

- Jerome Daniel (France Telecom) ha sviluppato un decoder Ambisonics del 4° ordine!



High Order Ambisonics (HOA)

- France Telecom dispone anche di array sferici a 32 capsule (4° ordine)



Conclusioni

- Si possono facilmente realizzare array microfonicici di ordine elevato per ottenere registrazioni surround ad alta definizione angolare
- La tecnica di filtraggio numerico sviluppata non richiede la soluzione di complesse teorie, ed essendo basata su misure sperimentali compensa automaticamente le imperfezioni dei microfoni
- L'implementazione della tecnica di filtraggio richiede semplicemente di effettuare un adeguato numero di convoluzioni
- A tal fine è stato sviluppato il plugin VST multiconvolutore X-volver

Sviluppi futuri

- E' in corso di assemblaggio un array tridimensionale randomizzato costituito da 24 microfoni di alta qualità (B&K 4188)



Esso consentirà di effettuare registrazioni di armoniche sferiche sino al 3° ordine, e richiederà il processamento di una matrice di 16x24 filtri