

# REGISTRAZIONE DELLA RISPOSTA ACUSTICA DELLE SALE DA CONCERTO PER I POSTERI

ANGELO FARINA<sup>1</sup>, ANDREA AVANZINI<sup>2</sup>

<sup>1</sup> *Dipartimento di Ingegneria Industriale, Università di Parma, ITALIA*  
[farina@unipr.it](mailto:farina@unipr.it)

<sup>2</sup> *Dipartimento di Ingegneria Industriale, Università di Parma, ITALIA*  
[a.avanzini@inwind.it](mailto:a.avanzini@inwind.it)

Il titolo di questo articolo richiama quello di una famosa pubblicazione di Michael Gerzon sulla rivista JAES, Vol. 23, Numero 7, pp. 569 (1975): "Recording Concert Hall Acoustics for Posterity" [1]. Dopo più di 25 anni il problema posto in quell'articolo è ancora aperto, particolarmente circa la tecnica migliore per catturare le caratteristiche "spaziali" del suono dentro ad un teatro esistente. Viene qui presentata una nuova tecnica, compatibile con tutti i formati di surround conosciuti.

## INTRODUZIONE

Quando, durante la notte del 29 Gennaio 1996, il famoso e rinomato "Gran Teatro La Fenice" di Venezia bruciò, uno dei migliori teatri d'opera del mondo improvvisamente sparì. Il suo suono meraviglioso era comunque almeno parzialmente salvo perché, proprio due mesi prima, alcune misurazioni acustiche erano state effettuate utilizzando la tecnica della risposta all'impulso binaurale [2].

La disponibilità di queste risposte binaurali fu molto importante durante la progettazione della ricostruzione del teatro e dimostrò l'importanza di acquisire e salvare per i posteri la risposta acustica delle sale da concerto.

M. Gerzon [1] per primo propose di iniziare una sistematica raccolta di risposte all'impulso tridimensionali misurate in antichi teatri ed in sale da concerto, onde valutare il comportamento acustico di queste ultime e preservarlo per le future generazioni. La sua proposta ha trovato una favorevole risposta solo in tempi molto recenti con la pubblicazione della "Carta di Ferrara" [3] e con la nascita di un gruppo internazionale di ricercatori che si sono accordati sulla metodologia sperimentale da seguire per raccogliere questi dati [4].

Finora solo di un modesto numero di teatri si ha una caratterizzazione completa mediante risposte tridimensionali all'impulso.

Tuttavia le tecniche finora proposte per la registrazione di risposte all'impulso "3D", contenenti sia informazioni spaziali che temporali, sono decisamente critiche se si utilizzano tali dati acquisiti per riproduzioni surround, attraverso la tecnica di auralizzazione (basata sulla convoluzione).

In effetti i due metodi attualmente utilizzati (misure binaurali con una testa artificiale orientata verso la sorgente sonora e misure B-format ottenute utilizzando un microfono Soundfield) sono entrambi inadeguati per ottenere realmente una riproduzione di alta qualità attraverso sistemi di riproduzione multicanale "standard" (ITU 5.1). Altri alternativi arrays di altoparlanti sono stati sviluppati (basati sulla cancellazione del cross-talk quando si riproduce materiale binaurale e sulla decodifica Ambisonics-like quando si riproduce materiale B-format); in alcuni casi queste due tecniche possono essere riunite assieme per una miglior riproduzione 3D (Ambiophonics, [5]).

Recentemente è stata proposta una tecnica completamente alternativa, detta "2.5-D": essa è basata sulla teoria della Wave Field Synthesis (WFS) e sull'uso di un microfono Soundfield che si muove circolarmente mediante un braccio rotante [6]. Anche questa tecnica, comunque, è inadeguata per un diretto impiego delle risposte all'impulso misurate su di una configurazione surround standard (ITU 5.1).

In questo articolo è proposto un nuovo metodo di misura, che ingloba tutte le precedenti tecniche conosciute in un approccio unico e coerente: tre microfoni differenti sono collocati su di un braccio rotante (una testa artificiale binaurale, una coppia di microfoni con diagramma di direttività cardioide in configurazione ORTF ed un microfono Soundfield) e viene misurato un set di risposte all'impulso in ogni posizione angolare.

La Figura 1 mostra una schematizzazione di questo schema microfonico.

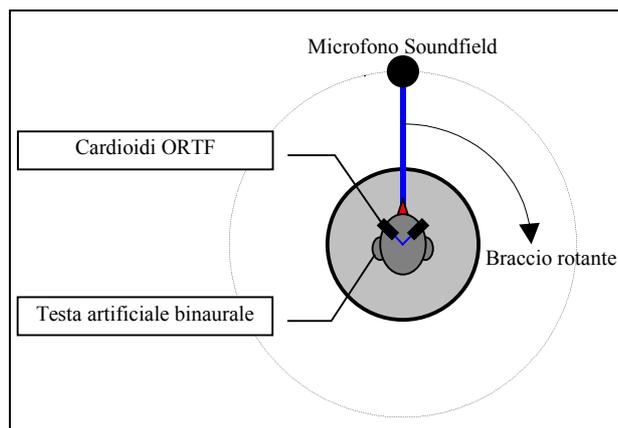


Figura 1: Schema dei microfoni.

I risultati di questi set di misure sono compatibili con i metodi di misurazione nelle sale da concerto già proposti (binaurale, B-format e WFS), ma aggiunge la possibilità di ottenere formati di ripresa microfonica surround, quali OCT ed INA, ed introduce la prospettiva di utilizzare anche il metodo del Binaural Room Scanning [7] o i microfoni circolari d'ordine elevato di Poletti [8].

Il presente articolo descrive i dettagli della implementazione di questa nuova tecnica di misurazione e fornisce i primi risultati sperimentali, ottenuti da misure effettuate in alcune sale da concerto.

## 1 IL METODO DI MISURA

Questo capitolo descrive le particolarità del metodo di misurazione, l'attrezzatura hardware e software utilizzata e la procedura seguita.

Sebbene la maggior parte di questi aspetti non sia intrinsecamente nuova, la combinazione di essi in un approccio coerente fornisce un metodo generale, dal quale possono essere ottenuti tutti i formati multicanale conosciuti.

### 1.1 Il segnale di test e la deconvoluzione

La tecnica della eccitazione-deconvoluzione utilizzata per l'acquisizione della risposta all'impulso è il metodo dello sweep sinusoidale logaritmico, come già proposto da uno degli autori del presente articolo [9]. Studi indipendenti hanno dimostrato che questo metodo è migliore rispetto ai precedenti [10,11].

Un buon compromesso tra intervallo di frequenze misurate, durata dello sweep e rapporto segnale-rumore è stato raggiunto con la scelta dei seguenti parametri:

Frequenza iniziale	22 Hz
Frequenza finale	22 kHz
Durata della sweep	15 s
Silenzio tra due sweep	10 s
Tipo di sweep	LOG

L'inusuale lunghezza del silenzio tra due sweep consecutivi è dovuto al tempo necessario al piatto rotante per arrivare nella successiva posizione. La rotazione è comandata da un segnale impulsivo, automaticamente generato a metà dell'intervallo di silenzio dal secondo canale della scheda audio.

La scelta dei parametri sopra riportati permette la misurazione di risposte all'impulso aventi ampia larghezza in frequenza, un buon range dinamico (approssimativamente di 90dB) e che sono sostanzialmente immuni dal rumore di fondo eventualmente presente durante le misurazioni.

La deconvoluzione è ottenuta mediante una convoluzione lineare (e non circolare) con un opportuno filtro inverso, che è automaticamente generato insieme al segnale di test. Come spiegato da uno degli autori [9], il filtro inverso è semplicemente il segnale di test rovesciato sull'asse del tempo, opportunamente equalizzato in ampiezza per compensare la diminuzione di 6 dB per ottava causata dallo sweep logaritmico.

La deconvoluzione lineare è efficace nell'evitare che il comportamento non-lineare dei trasduttori possa causare artefatte distorsioni armoniche della risposta all'impulso misurata.

Dal momento che la scheda audio utilizzata è in grado di lavorare a 96 kHz-24 bits, c'è sufficiente distanza tra la massima frequenza generata e la frequenza di Nyquist, sicché il filtro anti-aliasing non deve intervenire e la risposta all'impulso ottenuta non soffre di distorsione di fase ad alte frequenze.

Anche l'ampiezza del segnale di test generato è stato opportunamente equalizzato in ampiezza, al fine di compensare l'irregolare risposta in frequenza dell'altoparlante; in questo modo la potenza sonora emessa ha uno spettro sufficientemente piatto su tutto l'intervallo di frequenze.

Le Figure 2 e 3 mostrano rispettivamente il segnale di test equalizzato (è stato utilizzato il programma CoolEditPro per la registrazione e la riproduzione) e l'interfaccia utente del software utilizzato per la deconvoluzione. La deconvoluzione è incredibilmente veloce (circa il 20% della durata del segnale registrato), grazie all'utilizzo delle nuove ed altamente efficienti routines FFT Intel Integrated Performance Primitives v. 3.0.

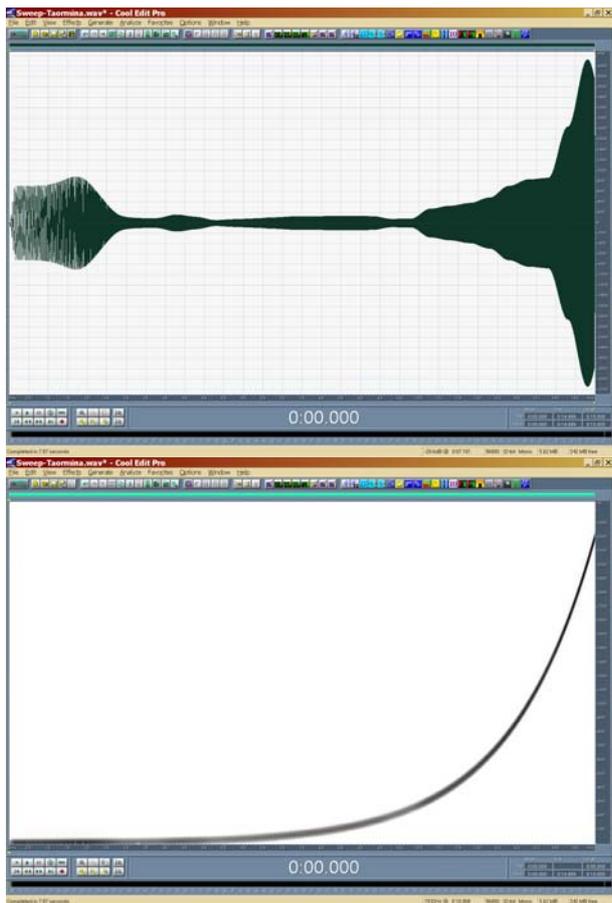


Figura 2: Il segnale di test equalizzato.

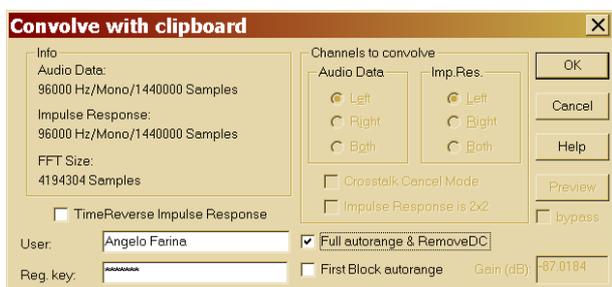


Figura 3: Il convolutore veloce utilizzato per la deconvoluzione.

## 1.2 La sorgente sonora

Generalmente una sorgente sonora omnidirezionale è preferibile per rilevare le risposte all'impulso in una sala. Sebbene ciò non corrisponda all'effettivo diagramma di direttività delle sorgenti sonore reali (quali gli strumenti musicali o i cantanti e gli oratori umani), l'utilizzo di una sorgente sonora omnidirezionale è raccomandata dagli standard correnti (per esempio la norma ISO 3382) ed evita di eccitare strani effetti dell'ambiente, come può capitare utilizzando altoparlanti altamente direttivi (anormale

esaltazioni di alcuni eco e focalizzazioni per alcune particolari orientazioni della sorgente).

Uno speciale ed ultracompatto altoparlante dodecaedrico è stato appositamente costruito per gli scopi di questa ricerca, utilizzando 12 altoparlanti accorpati in una struttura di piccole dimensioni (il diametro complessivo è approssimativamente di 200 mm). Ovviamente questa unità non è in grado di produrre una significativa potenza acustica al di sotto dei 120 Hz; perciò, al fine di estendere il range operativo alle basse frequenze, è stato aggiunto un subwoofer, incorporandolo dentro al contenitore cilindrico di trasporto, il quale inoltre contiene l'amplificatore di potenza (300 W RMS) e serve da base di supporto per il dodecaedro.

La Figura 4 mostra una immagine di questa speciale sorgente sonora omnidirezionale.



Figura 4: L'altoparlante dodecaedrico e il subwoofer.

La risposta acustica dell'altoparlante è stata misurata dentro ad una camera anecoica, mediando il suono irradiato su di una circonferenza completa. Dal momento che lo spettro in terzi d'ottava rilevato quando all'altoparlante veniva inviato un rumore rosa perfettamente piatto era decisamente irregolare, è stata necessaria una appropriata equalizzazione del segnale di test.

La Figura 5 mostra il confronto tra la potenza sonora irradiata dall'altoparlante prima e dopo l'equalizzazione, la quale è stata ottenuta applicando direttamente al segnale di test un opportuno filtraggio in terzi d'ottava, quello richiesto per rendere piatta in frequenza la risposta.

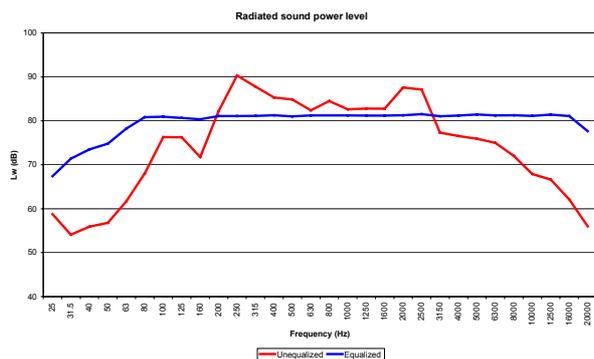


Figura 5: Gli spettri della potenza sonora irradiata.

Si può notare dal grafico come l'equalizzazione in frequenza sia perfettamente in grado di rendere piatta la risposta dell'altoparlante tra gli 80 ed i 16000 Hz, con un dolce decremento al di fuori di questo intervallo. Dopo l'equalizzazione il livello di potenza sonora complessivamente irradiata (sotto forma di rumore rosa) è di circa 97 dB.

### 1.3 I microfoni

Tre diverse sonde microfoniche sono state utilizzate:

- una coppia di cardioidi di alta qualità in configurazione ORTF (microfoni Neumann K-140), distanziati di 180mm e formanti tra loro un angolo di 110°;
- una testa artificiale binaurale (Neumann KU-100);
- una sonda microfonica pressione-velocità B-format 4-canali (Soundfield ST-250).

Tutti questi microfoni sono stati installati su di una tavola rotante, in modo tale che il centro di rotazione coincida con il centro della testa artificiale ed il punto di intersezione degli assi dei due microfoni a cardioide (che sono collocati sopra la testa artificiale); mentre il microfono Soundfield è stato posto esattamente ad 1 metro dall'asse di rotazione, giusto di fronte alla testa artificiale.

Il piatto rotante (Outline ET-1) è stato programmato per fermarsi ogni 10°; di conseguenza, mediante una completa rotazione, vengono acquisiti 36 set discreti di risposte all'impulso, uno per ciascuna delle posizioni del piatto rotante.

Le Figure 6, 7 e 8 mostrano il set-up microfonico.



Figura 6: I microfoni sul piatto rotante.



Figura 7: La configurazione dei microfoni.



Figura 8: Testa e microfoni Neumann.

#### 1.4 Il Computer e la scheda audio

Il metodo di misura richiede l'utilizzo di una scheda audio di alto livello, dotata di 8 ingressi analogici a 24 bits-96 kHz, comprendente preamplificatori microfonici controllati in modo digitale (per assicurare un accurato controllo del guadagno in ingresso ed una calibrazione sia relativa che assoluta delle registrazioni). Attualmente queste richieste possono essere soddisfatte solo da unità esterne, collegate al computer per mezzo di una interfaccia PCI.

Ciò impedisce l'utilizzo di qualsiasi computer portatile ed obbliga alla scelta dell'unico computer senza ventola attualmente disponibile, che spicca per la sua completa silenziosità: il Signum Data Futureclient.

Il modello utilizzato per questa ricerca è equipaggiato di un processore P-IV a 1.8 GHz, è dotato di 512 Mbytes di RAM e di un hard disk ad alta velocità (7200 RPM). Ciò consente una perfetta operatività del sistema quando si registrano 8 canali e si riproducono simultaneamente 2 canali, tutti a 96 kHz-24 bits.

La scheda audio scelta per questo lavoro è una Aadvark Pro-Q10.

Le Figure 9 e 10 mostrano l'equipaggiamento, fronte e retro, che è fissato dentro ad una coppia di flight-cases, onde agevolarne il trasporto.



Figura 9: Il PC raffreddato a liquido (FutureClient) ed il comando del piatto rotante.



Figura 10: La scheda audio Aadvark Pro-Q10.

#### 1.5 La misura

Il programma CoolEditPro è stato utilizzato per la riproduzione dei segnali di test e contemporaneamente per la registrazione degli 8 canali microfonici. Il segnale di test è stato ripetuto 36 volte, corrispondenti alle 36 posizioni che il braccio rotante assume durante una completa rotazione.

La seguente immagine (Figura 11) mostra una sessione multitraccia, durante una misurazione effettuata con il metodo appena descritto.



Figura 11: Sessione di misura in multitraccia.

Ogni misurazione richiede circa 15 minuti (25 secondi per 36 ripetizioni); dopo che la misurazione è stata completata, altri 10 minuti sono richiesti per salvare tutte le tracce nell'hard disk (a 32 bits, per utilizzare tutto il range dinamico a disposizione). Durante questo tempo la sorgente e/o i microfoni sono spostati in una nuova posizione.

## 1.6 I dati acquisiti

Sinora sono stati misurati 9 famosi teatri con il metodo descritto in precedenza, come riportato dalla seguente tabella:

N	Teatri	N. sorgenti/ ricevitori
1	Uhara Hall, Kobe, Japan	2/2
2	Noh Drama Theater, Kobe, Japan	2/2
3	Kirishima Concert Hall, Kirishima, Japan	3/3
4	Teatro Greco di Siracusa, Italia	2/1
5	Teatro Greco-Romano di Taormina, Italia	3/2
6	Auditorium di Parma, Italia	3/3
7	Auditorium di Roma (Sala 700), Italia	3/2
8	Auditorium di Roma (Sala 1200), Italia	3/3
9	Auditorium di Roma (Sala 2800), Italia	3/5

Il numero delle sale da concerto e dei teatri misurati aumenterà velocemente: si è programmato di raggiungere il numero di almeno 30 diverse sale in meno di 6 mesi.

Lo scopo del presente articolo non è quello di presentare un esauriente studio comparativo dei dati acquisiti, che verrà effettuato quando la raccolta di risposte all'impulso sarà completata e tutti i risultati verranno compiutamente analizzati.

Ciononostante, la seguente figura (Figura 12) mostra un set di 36 risposte all'impulso ottenute nell'Auditorium di Parma, ciò per dare una idea circa la modalità con cui i dati vengono memorizzati: per ogni coppia di microfoni (i Neumann ORTF in questo caso). Le 36 risposte all'impulso ottenute durante la rotazione sono poste l'una dopo l'altra e l'intera sequenza è salvata come file WAV a 32 bits (float).

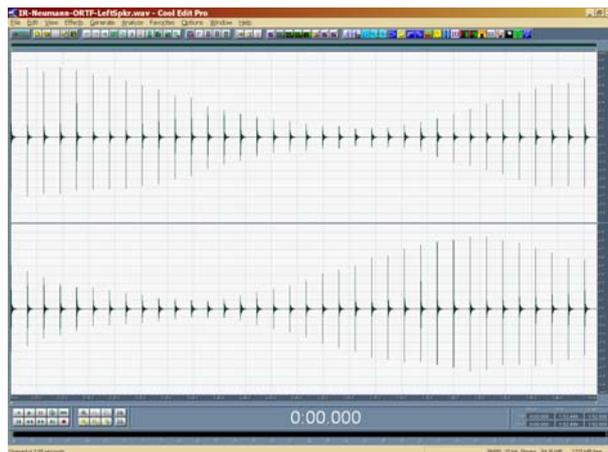


Figura 12: Un set di risposte all'impulso ottenute (con 36 posizioni microfoniche).

## 2 L'ESTRAZIONE DI PARAMETRI ACUSTICI OGGETTIVI

Fondamentalmente il calcolo di parametri acustici oggettivi è basato sulla norma ISO 3382-1997. La maggior parte dei parametri è ottenuta dalla risposta all'impulso prelevata con un microfono omnidirezionale, che sostanzialmente è il canale W del microfono Soundfield, nella sua posizione iniziale (0 gradi).

Tuttavia i parametri spaziali richiedono l'elaborazione di risposte all'impulso stereo; di conseguenza anche le coppie di tracce binaurali e WY devono essere elaborate.

Questa ricerca è principalmente rivolta all'acquisizione ed all'analisi delle proprietà spaziali del campo sonoro, con l'obiettivo di creare delle ricostruzioni surround multicanale realistiche; di conseguenza il maggior sforzo è dedicato all'analisi dei parametri spaziali.

Il risultato estremamente innovativo, reso disponibile da questa nuova tecnica di misura, è la possibilità di ottenere e visualizzare diagrammi polari dei parametri acustici spaziali, mostrando la loro variazione col ruotare del ricevitore.

### 2.1 Il tempo di riverbero

Si utilizza il canale W (omnidirezionale) della risposta all'impulso B-format. La risposta all'impulso è inizialmente integrata all'indietro, seguendo il metodo di Schroeder ed applicando la rimozione del rumore secondo la norma ISO 3382.

Successivamente il Tempo di Riverbero  $T_{30}$  è calcolato per mezzo di una regressione lineare sulla curva di decadimento nell'intervallo tra  $-5$  e  $-35$  dB al di sotto del livello stazionario che si ha prima che inizi il decadimento. Occorre notare che di solito queste risposte all'impulso sono così "pulite" e prive di rumore che sarebbe possibile misurare direttamente il  $T_{60}$

(nell'intervallo da -5 a -65 dB), ma la norma ISO 3382-1997 non lo consente (fu scritta quando la misurazione di risposte all'impulso con un così grande range dinamico era molto difficile da ottenere).

La Figura 13 mostra un tipico andamento della risposta all'impulso e della curva di decadimento (ricavata con una integrazione all'indietro) ottenute in uno dei teatri oggetto della presente ricerca.

L'immagine mostra che il livello della pressione sonora totale integrata è circa 90 dB al di sopra del rumore stazionario di fondo presente dopo che la risposta all'impulso si è estinta.

### 2.2 I parametri temporali monofonici

Sebbene il Tempo di Riverbero sia il più importante criterio per valutare il comportamento acustico di una sala, è spesso opportuno avere una migliore e più dettagliata idea sulla distribuzione dell'energia acustica.

A questo scopo la norma ISO 3382 suggerisce l'utilizzo di 4 criteri temporali monoaurali: C<sub>50</sub>, C<sub>80</sub>, D, T<sub>s</sub>.

C<sub>50</sub> è l'Indice di Chiarezza su 50ms, calcolato applicando la seguente formula alla risposta all'impulso ottenuta da un microfono in pressione omnidirezionale, partendo dall'istante di arrivo del suono diretto:

$$C_{50} = 10 \cdot \lg \left[ \frac{\int_0^{50ms} p^2(\tau) \cdot d\tau}{\int_{50ms}^{\infty} p^2(\tau) \cdot d\tau} \right] \quad (1)$$

C<sub>80</sub> è simile, ma l'estremo temporale di integrazione è portato da 50 ms a 80 ms. Solitamente l'indice C<sub>50</sub> è considerato più rappresentativo della chiarezza del parlato, mentre il C<sub>80</sub> è più indicato per valutare la chiarezza di uno strumento musicale.

L'Indice di Definizione D è sostanzialmente simile a C<sub>50</sub>, ma è espresso in % invece che in dB, ed ha la seguente formulazione:

$$D = \frac{\int_0^{50ms} p^2(\tau) \cdot d\tau}{\int_0^{\infty} p^2(\tau) \cdot d\tau} \cdot 100 \quad (2)$$

Infine il Tempo Baricentrico T<sub>s</sub> è definito come:

$$T_s = \frac{\int_0^{\infty} \tau \cdot p^2(\tau) \cdot d\tau}{\int_0^{\infty} p^2(\tau) \cdot d\tau} \quad (3)$$

Esso ha il vantaggio di evitare una brusca separazione tra l'energia precoce (considerata utile) e quella ritardata (considerata dannosa), che è insita nelle definizioni di C e D.

Il calcolo di tutti questi succitati parametri e del tempo di riverbero è ottenuto grazie ad un opportuno plug-in, sviluppato con lo scopo di rendere automatico il calcolo dei parametri acustici della norma ISO 3382. La Figura 13 mostra l'interfaccia utente di questo plug-in.

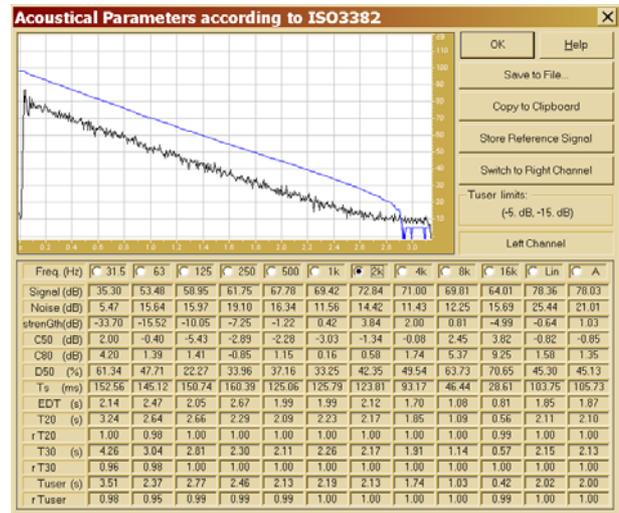


Figura 13: I parametri acustici secondo norma ISO 3382

### 2.3 Il livello di pressione sonora assoluto e relativo

Poiché la potenza acustica della sorgente sonora è stata accuratamente calibrata grazie alle misure in camera anecoica ed avendo memoria del guadagno applicato ai preamplificatori microfonici, è possibile conoscere con una accettabile precisione (+/- 1 dB) il livello assoluto di pressione sonora ottenuto durante la misurazione.

Inoltre, poiché la deconvoluzione di tutte le risposte all'impulso di un certo teatro viene effettuata utilizzando sempre lo stesso fattore di scala, le ampiezze visualizzate delle risposte all'impulso preservano fra loro una scala relativa.

La differenza fra il livello di pressione sonora SPL e il livello di potenza sonora irradiata L<sub>w</sub> permette di calcolare un parametro acustico decisamente rilevante, il Sound Strength G (in italiano a volte tradotto con "sonorità"):

$$G = SPL - L_w + 31 \quad \text{dB} \quad (4)$$

Il fattore correttivo di +31 dB deriva dalla definizione di G, che è definito come la differenza tra l'SPL misurato dentro la sala e l'SPL teorico misurato in campo libero, ad una distanza di 10 metri dalla sorgente.

### 2.4 I parametri spaziali binaurali (IACC)

Seguendo la teoria di Y. Ando [12], il primo parametro binaurale è lo IACC (Inter Aural Cross Correlation), definito come il massimo valore della funzione normalizzata di cross-correlazione fra i segnali derivanti dai microfoni di una testa binaurale:

$$\rho(\tau) = \frac{\int_{-\infty}^{\infty} h_d(\tau) \cdot h_s(\tau + t) \cdot d\tau}{\sqrt{\int_{-\infty}^{\infty} h_d^2(\tau) \cdot d\tau \cdot \int_{-\infty}^{\infty} h_s^2(\tau + t) \cdot d\tau}} \quad (5)$$

Altri parametri sono  $\tau_{IACC}$  e  $w_{IACC}$ , definiti rispettivamente come il ritardo in ms del massimo valore della funzione normalizzata di cross-correlazione, e come la larghezza del picco (al 10% del massimo), sempre in ms.

Un apposito plug-in è stato creato per misurare i parametri basati sullo IACC. Questo plug-in inoltre calcola la durata del ritardo temporale tra il suono diretto e la prima riflessione ed il  $T_{sub}$  (subsequent reverberation time), in conformità alla teoria di Ando.

La Figura 14 mostra l'interfaccia utente di questo plug-in.

Convenzionalmente questa misura viene effettuata con la testa artificiale orientata in direzione della sorgente sonora. Nel nostro caso invece la testa è orientata verso 36 differenti direzioni, con passi di 10°. Di conseguenza si ottengono 36 valori di IACC ed è quindi possibile creare un diagramma polare di IACC.

La disponibilità di questi diagrammi polari è una novità ed ancora si deve comprendere quali informazioni possano essere estratte da essi. Ciò che immediatamente appare, comunque, è che sale con circa lo stesso valore di IACC "standard" possono avere diagrammi polari abbastanza differenti, ciò dimostra che le proprietà "surround" di una sala non possono essere completamente descritte solo dal vecchio e singolo valore di IACC "standard".

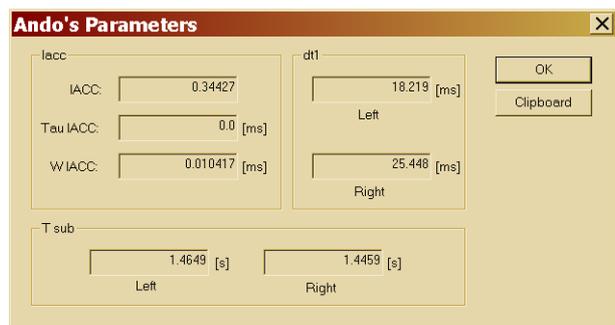


Figura 14: Il plug-in che calcola i parametri di Ando.

Ciò è dimostrato dal confronto dei diagrammi polari riportati in Figura 15, che mostrano l'Auditorium di Parma e quello di Roma (Sala 1200). Nel secondo il suono appare essere maggiormente polarizzato, mentre nel primo esso è più diffuso.

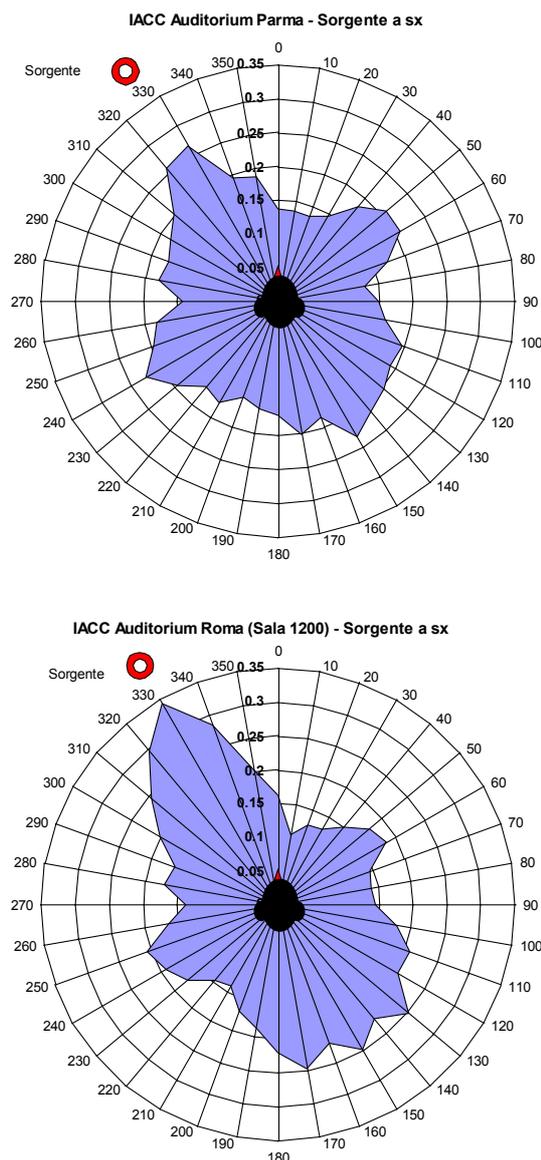


Figura 15: Diagrammi polari di IACC negli Auditorium di Parma e Roma (Sala 1200).

## 2.5 I parametri spaziali B-format (Lateral Fraction)

La norma ISO 3382 definisce due descrittori spaziali ottenuti da una risposta all'impulso B-format (più precisamente dai canali W ed Y di una risposta all'impulso B-format), indicati rispettivamente con LF e con LFC.

La Lateral Fraction LF (in italiano Frazione della prima energia Laterale) è il rapporto tra il primo suono laterale ed il suono omnidirezionale:

$$LF = \frac{\int_{0ms}^{80ms} h_Y^2(\tau) \cdot d\tau}{\int_{0ms}^{80ms} h_W^2(\tau) \cdot d\tau} \quad (6)$$

Per applicare correttamente la formula appena scritta ad una misurazione effettuata con il microfono Soundfield, occorre precisare che l'asse X di quest'ultimo deve essere orizzontale ed orientato verso la sorgente sonora, l'asse Y orizzontale, ortogonale all'asse X ed orientato in direzione dell'orecchio sinistro e l'asse Z verticale, rivolto verso il soffitto. Inoltre, è necessario compensare il fatto che il canale W (omnidirezionale) ha un guadagno di 3 dB in meno rispetto ai canali X, Y e Z.

LFC è il secondo parametro, esso è definito come:

$$LFC = \frac{\int_{0ms}^{80ms} h_Y(\tau) \cdot h_W(\tau) \cdot d\tau}{\int_{0ms}^{80ms} h_W^2(\tau) \cdot d\tau} \quad (7)$$

In questo caso il numeratore ci da l'intensità sonora, mentre il denominatore il valore efficace della pressione sonora. In sostanza, LFC è un parametro abbastanza prossimo alla definizione dell'indice pressione-intensità, normalmente utilizzato nelle applicazioni di sistemi di misura di intensità sonora (ISO9614).

Anche per questi parametric basati sul B-format è stato sviluppato uno speciale plug-in, la cui interfaccia utente è mostrata nella seguente figura (Figura 16).

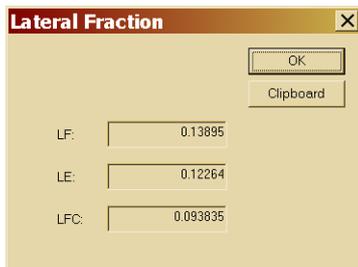


Figura 16: Il plug-in che calcola la Lateral Fraction.

Si noti che questo plug-in calcola anche il parametro Lateral Efficiency LE (Efficienza Laterale) di Jordan [13], la cui definizione assomiglia a quella di LF, ma con un estremo iniziale dell'integrale a numeratore pari a 25ms anziché a 5ms.

Poichè il microfono Soundfield può essere "virtualmente ruotato" attorno al suo asse, è facile, attraverso una singola risposta all'impulso B-format, ottenere un diagramma polare completo di LF. In realtà il microfono non sta semplicemente ruotando su se

stesso, poiché è posto su di una circonferenza di 1 metro di raggio. Di conseguenza, prendendo per ogni posizione del microfono la sua orientazione radiale, è possibile costruire un diagramma polare modificato, che mostra la variazione di LF (o di 1-LF) lungo il percorso di circonferenza descritto dal microfono nel suo movimento.

La seguente figura (Figura 17) mostra questi diagrammi polari per le stesse due sale di cui in precedenza si è analizzato lo IACC.

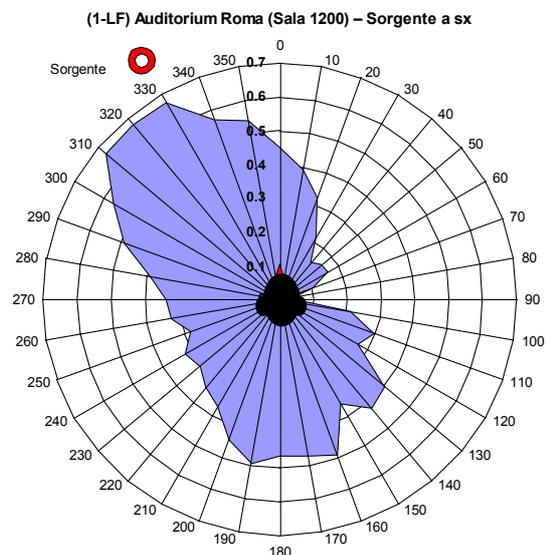
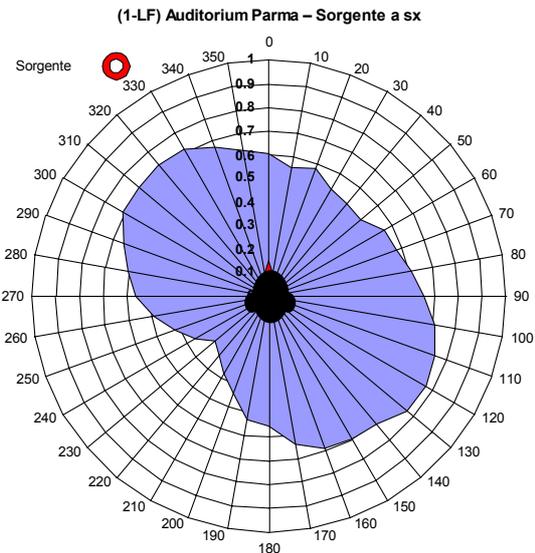


Figura 17: Diagrammi polari di (1-LF) negli Auditorium di Parma e Roma.

Si noti che il parametro (1-LF) ha la stessa polarità di IACC, cosicchè i diagrammi polari di Figura 17 sono direttamente confrontabili con quelli di Figura 15. Anche in questo caso è abbastanza evidente come il

campo sonoro sia molto più diffuso nell'Auditorium di Parma, mentre in quello di Roma il suono sia più polarizzato. In quest'ultimo, inoltre, vi è un piccolo settore angolare ove LF è circa unitario (e di conseguenza 1-LF è zero).

L'analisi dei risultati mostra una poca significatività del parametro LFC (che è sempre decisamente piccolo, indipendente dalla sala e dalla orientazione della sonda microfonica) e una debole dipendenza dalla orientazione del microfono per LE.

Dunque LF si conferma essere il parametro più sensibile fra quelli che si basano sulle risposte all'impulso B-format, sebbene sia chiaro quanto una classifica della impressione spaziale di una sala basata su LF non coincida necessariamente con una classifica basata sullo IACC. La seguente tabella confronta i valori di IACC e di (1-LF) per i due casi già riportati nelle Figure 15 e 17:

Auditorium	IACC	1-LF
Parma	0.266	0.725
Roma	0.344	0.676

Da questa tabella, guardando IACC, Parma sembrerebbe avere una maggiore impressione spaziale rispetto a Roma, mentre guardando LF si otterrebbe un giudizio opposto.

Ciò significa che l'informazione circa la diffusione del suono ottenuta mediante questi due descrittori possa essere ingannevole e che una corretta valutazione su quale delle due sale abbia un campo sonoro più avvolgente non possa essere effettuata solamente in base ai parametri ottenuti indirizzando i microfoni verso la direzione della sorgente sonora, ma invece richieda che si analizzi la variazione di tali parametri spaziali quando i microfoni ruotano verso tutte le direzioni.

L'esperienza soggettiva di ascolto degli autori indica chiaramente che il suono nell'Auditorium di Parma è decisamente più diffuso di quello della "Sala 1200" dell'Auditorium di Roma; le stesse conclusioni appaiono evidenti se si confrontano fra loro i diagrammi polari sia di Figura 15 che di Figura 17.

## 2.6 Critica ai parametri della norma ISO 3382

L'applicazione dei parametri della norma ISO 3382 a queste risposte all'impulso ha evidenziato come questa norma, sebbene sia stata pubblicata nel 1997, già richieda una sostanziale revisione. In pratica tre questioni richiedono un approfondimento:

- La norma non dà precise indicazioni per le misurazioni basate sullo sweep, né discute i motivi che rendono il metodo dello sweep preferibile a quello basato su MLS (invarianza temporale, non linearità, tolleranza ad errori di clock, etc.);
- Quasi tutti i parametri sono detti essere relativi all'energia acustica, ma essi sono attualmente

calcolati utilizzando i valori della pressione efficace. Invece i veri valori della intensità attiva e della densità di energia sonora sono disponibili da una misurazione B-format; ed è ben noto come, in un campo sonoro parzialmente reattivo, i veri parametri energetici possano differire in modo significativo da quelli stimati, se basati sulla pressione efficace.

- La definizione dei parametri spaziali (siano essi ricavati da risposte all'impulso binaurali o B-format) presuppone una ben specifica orientazione del microfono, che deve essere rivolto verso la sorgente sonora. Ciò non ha senso in presenza di più sorgenti, o in sale dotate di sistemi di rinforzo del suono. Anche nel caso di una sorgente puntiforme questi parametri danno risultati contraddittori.

## 3 L'AURALIZZAZIONE DEI DATI ACQUISITI

Questo capitolo analizza la possibilità di utilizzare i risultati della presente ricerca per creare delle riproduzioni udibili del comportamento acustico delle sale originarie, presentate agli ascoltatori per mezzo di cuffie o di altoparlanti.

Il metodo alla base della auralizzazione è la convoluzione: le risposte all'impulso sono utilizzate come lunghissimi filtri FIR ed applicate a registrazioni "asciutte" (cioè anecoiche) di musica o di parlato. La convoluzione è una tecnica di filtraggio molto efficiente, in modo particolare se implementata con opportuni algoritmi (vecchi) su processori veloci (nuovi): come chiaramente dimostrato da A. Torger e A. Farina [14], un Pc dotato di un processore dell'ultima generazione può effettuare la convoluzione in tempo reale e con bassa latenza di dozzine di canali con più risposte all'impulso di centinaia di migliaia di coefficienti ognuna. Ed i risultati ottenuti con i più semplici algoritmi inizialmente sviluppati negli anni '60 [15] sono migliori di quelli ottenuti utilizzando versioni più recenti [16], che sembrano essere preferibili sotto il punto di vista del numero totale di moltiplicazioni richieste, ma che sono molto meno ottimizzati per l'architettura di gestione della memoria dei processori moderni.

Lo scopo del presente lavoro è quello di creare dei set di risposte all'impulso adatte ad essere utilizzate da questi convolutori software, generando dei risultati in ognuno dei formati attualmente disponibili atti alla riproduzione multicanale e tentando così di ricreare nel modo più fedele possibile le caratteristiche spaziali del campo sonoro originario.

### 3.1 Le risposte all'impulso ORTF-stereo

Questo è il procedimento più banale, finalizzato ad una presentazione stereo "standard" dei risultati dell'auralizzazione. Esso si basa sulla disponibilità di un

certo numero di registrazioni monofoniche “asciutte” (anecoiche), una per ogni sezione dell’orchestra e per ogni cantante.

Ogni registrazione mono deve essere convoluta con una ben specifica risposta all’impulso stereo, ottenuta da una coppia di microfoni a cardioide in configurazione ORTF. In linea di principio ognuna di queste risposte all’impulso dovrebbe essere misurata mediante l’esatto posizionamento della sorgente sonora.

In realtà le misurazioni sono generalmente effettuate con sole tre posizioni della sorgente sul palcoscenico (Sinistra, Centro e Destra); ciò limita il numero di sorgenti “virtuali” indipendenti che possono essere collocate sulla scena acustica.

In pratica, comunque, è possibile trarre vantaggio dal fatto che, per ogni posizione della sorgente, la misurazione ORTF viene realizzata con 36 differenti orientazioni dei microfoni (con variazioni di 10°). Ciò significa che alcuni piccoli aggiustamenti della posizione della sorgente virtuale (di 10 o 20 gradi) possono essere ottenuti selezionando la risposta all’impulso ORTF proveniente da una direzione diversa da 0°. Ciò ovviamente non è perfettamente rigoroso, ma è effettivamente e soggettivamente indistinguibile dalla convoluzione con risposte all’impulso ORTF ottenute con una orientazione del microfono a 0° ed il corretto posizionamento della sorgente.

Ovviamente i risultati della convoluzione di tutte le registrazioni anecoiche sono sommate in un unico singolo file di uscita stereo, il quale è perfettamente adatto per essere riprodotto da un normale sistema stereo con due altoparlanti.

### 3.2 Le risposte all’impulso binaurali (binaural room scanning)

L’approccio binaurale di base è sostanzialmente lo stesso del precedente metodo ORTF, salvo che ora si utilizzano le risposte all’impulso binaurali. In questo modo il risultato della convoluzione è un file avente due canali, adatto ad una riproduzione in cuffia.

Possono però essere utilizzati due metodi per migliorare decisamente l’effetto surround ottenuto: per la riproduzione con altoparlanti deve essere aggiunto un opportuno filtro per la cancellazione del cross-talk e per la riproduzione con cuffie un sensore head-tracking può comandare un convolutore in tempo reale, cambiando le risposte all’impulso con le quali si sta facendo la convoluzione quando l’ascoltatore ruota la propria testa. In merito alla creazione di filtri per una ottima cancellazione del cross-talk, sono stati pubblicati diversi articoli in anni recenti [17,18].

Riguardo invece il processo di head-tracking in tempo reale, alcune soluzioni sono state proposte da LakeDsp [19] e Studer [7], ma richiedono apposite e costose workstations DSP. Uno degli autori dell’articolo sta lavorando ad un nuovo ed economico sistema per

l’auralizzazione in tempo reale, che utilizza un sistema di head tracking di qualità non professionale ed un nuovo convolutore software, altamente efficiente ed a bassa latenza, implementato su scheda DSP.

### 3.3 Le risposte all’impulso B-format (Ambisonics)

In questo caso ogni sorgente monofonica anecoica è convoluta con l’appropriata risposta all’impulso B-format. Di conseguenza, dopo il mixaggio di tutte queste convoluzioni, si ottiene una uscita 4-canali B-format.

La riproduzione di un segnale B-format mediante un opportuno array di altoparlanti richiede un “decoder” Ambisonics, onde avere il giusto segnale per ogni altoparlante.

La creazione di un “decoder” software è stata ideata da uno degli autori [20] ed è stata in seguito perfezionata da colleghi nell’Università di York, i quali recentemente hanno reso pubblica una serie di plug-ins VST [21], permettendo così l’elaborazione e la decodifica di segnali B-format per mezzo di diversi impianti di altoparlanti.

In conclusione, l’auralizzazione Ambisonics semplicemente richiede la disponibilità di un convolutore multicanale (con 1 ingresso e 4 uscite), un mixer B-format ed un decodificatore Ambisonics B-format. Il primo di questi viene attualmente sviluppato dalla Waves Ltd, gli altri due sono già disponibili (vedi [21]).

### 3.4 ITU 5.1 surround (da alcune risposte all’impulso B-format)

Alla base della riproduzione ITU 5.1 sta innanzitutto la scelta della configurazione dei microfoni da utilizzare per pilotare i 5 altoparlanti principali [22]. Sono state proposte molte disposizioni, ed in un recente progetto di confronto, chiamato “Verdi Projekt”, sono state comparate la maggior parte di esse [23].

Noi ora qui ne consideriamo solo tre, che hanno ottenuto buoni risultati nel succitato studio comparativo: Williams MMA [24], OCT [22] and INA [25].

Le seguenti figure (Figure 18, 19 e 20) mostrano la configurazione microfonica di questi tre set-up.

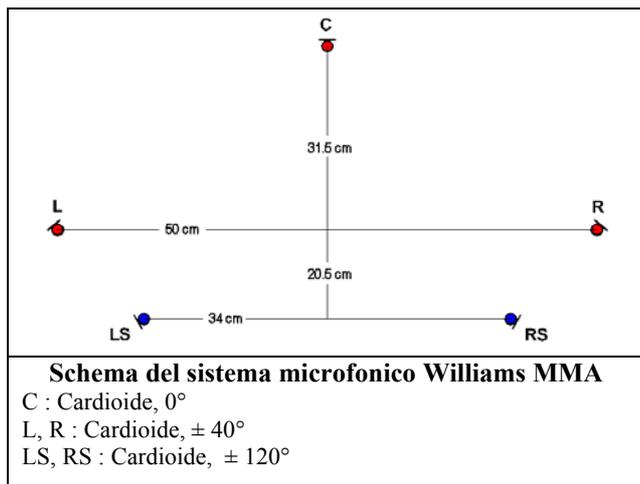


Figura 18: Schema microfonico Williams MMA.

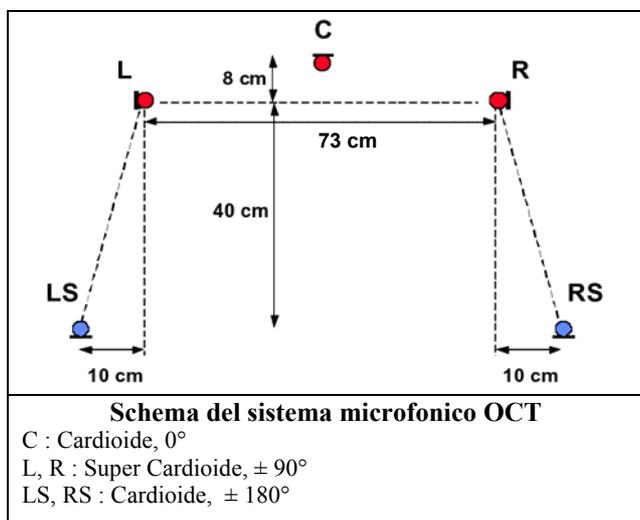


Figura 19: Schema microfonico OCT.

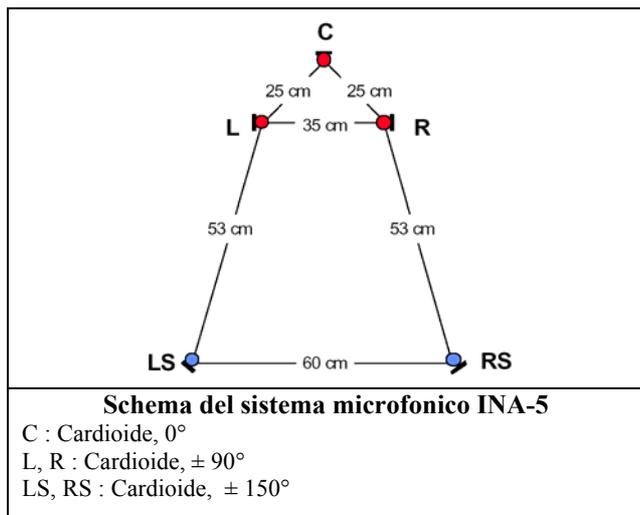


Figura 20: Schema microfonico INA.

Per ognuno degli schemi raffigurati qui a lato è possibile selezionare un insieme di 5 delle 36 posizioni ove il microfono Soundfield è collocato, prendendo quelle più vicine alle posizioni imposte dalla configurazione scelta.

In seguito, dalla risposta all'impulso B-format ottenuta in ognuna di queste 5 posizioni selezionate, una risposta all'impulso singola (monofonica) viene sintetizzata grazie al programma Visual Virtual Microphone, sviluppato da David McGriffy e disponibile gratuitamente su Internet [26]. La Figura 21 mostra l'interfaccia utente di questo programma, quando viene utilizzato per estrarre la risposta a supercardioide per il canale R di un set-up OCT a partire dalla risposta all'impulso proveniente dalla posizione a 20° e con la sorgente sonora sul lato sinistro del palcoscenico.

Si tenga presente che le misurazioni effettuate con il microfono Soundfield che ruota assumono intrinsecamente un angolo in verso orario (dovuto al fatto che il piatto rotante gira solo in questo verso), mentre di solito nelle applicazioni surround si utilizzano angoli antiorari.

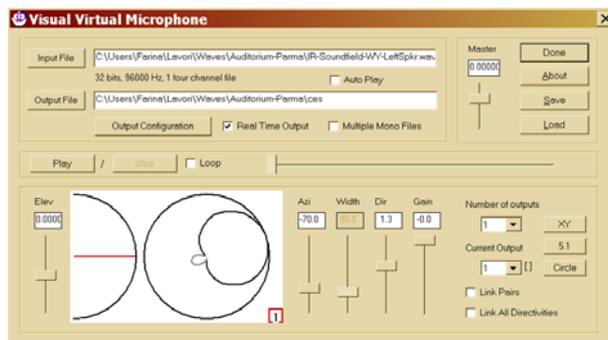


Figura 21: Visual Virtual Microphone.

Poichè il microfono nella posizione in cui si trova è già inclinato di 20° verso destra ed il canale R del set-up OCT richiede una orientazione del supercardioide destro di 90°, una ulteriore rotazione di 70° deve essere richiesta al programma Visual Virtual Microphone.

La seguente figura (Figura 22) mostra una schematizzazione geometrica di quanto appena descritto.

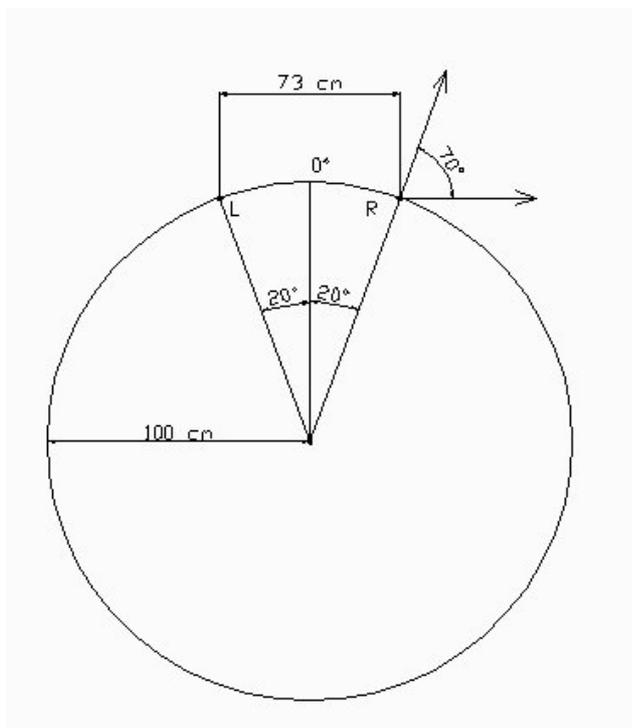


Figura 22: Schema del procedimento adottato per avere gli orientamenti desiderati dei microfoni virtuali.

Nel caso in cui il set-up microfonico prescelto richieda una posizione del microfono che non si trova sulla circonferenza di 1 metro di raggio descritta dal microfono Soundfield, è possibile utilizzare il metodo WFS (vedi paragrafo 3.6) per estrapolare la risposta all'impulso nella posizione desiderata.

Infine ogni sorgente monofonica anecoica è convoluta con la risposta all'impulso a 5-canali ottenuta dalla corrispondente posizione della sorgente sonora sul palcoscenico; i risultati di tutte queste convoluzioni sono mixati in una singola traccia finale 5-canali, che è adatta per una riproduzione su un impianto di altoparlanti secondo lo standard ITU 5.1.

### 3.5 I microfoni virtuali ad alta direttività di Mark Poletti

Durante la rotazione del sistema microfonico, anche i due cardioidi utilizzati per le registrazioni ORTF descrivono una piccola circonferenza avente un raggio di circa 110 mm, come mostrato dalla seguente figura (Figura 23).

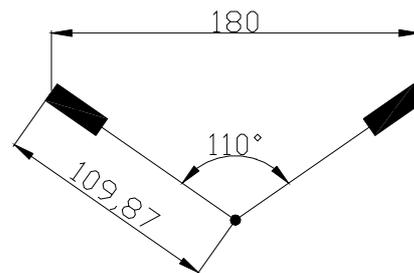


Figura 23: Geometria dei microfoni ORTF.

Consideriamo per semplicità uno solo dei due microfoni: egli cattura, durante la sua completa rotazione, 36 risposte all'impulso. Da questo insieme di dati è possibile ottenere le risposte di un altro insieme di microfoni coincidenti di vario ordine, idealmente collocati nel centro di rotazione, utilizzando una versione modificata della teoria di M. Poletti [8].

Alla base di questo metodo vi è la definizione di una classe di diagrammi di direttività orizzontale aventi una forma a “multifoglia”. L'ordine 0 (zero) è un microfono omnidirezionale, l'ordine 1 è dato da due microfoni con diagramma “a forma di 8” (come nell'Ambisonics solo orizzontale), gli ordini 2 e 3 hanno diagrammi di direttività “multifoglia” corrispondenti rispettivamente al coseno di due e tre volte l'angolo:

$$\begin{aligned}
 D_0 &= 1 \\
 D_{1,n} &= \cos\left(\vartheta + n \cdot \frac{\pi}{2}\right) \quad n = 0,1 \\
 D_{2,n} &= \cos\left(2 \cdot \vartheta + n \cdot \frac{\pi}{4}\right) \quad n = 0,1 \\
 D_{3,n} &= \cos\left(3 \cdot \vartheta + n \cdot \frac{\pi}{6}\right) \quad n = 0,1
 \end{aligned} \tag{8}$$

Le risposte di questi microfoni virtuali possono essere pensate come la decomposizione in armoniche cilindriche del campo sonoro nella posizione centrale o come una analisi spaziale di Fourier del campo sonoro fatta lungo la coordinata angolare  $\vartheta$ .

La seconda spiegazione suggerisce un modo semplice per calcolare le risposte desiderate: i segnali provenienti dalle 36 posizioni microfoniche sono semplicemente moltiplicati per un insieme di 36 fattori di peso, ottenuti dall'equazione (8) di sopra, e poi sommati.

Ciò ovviamente non fornisce il risultato ricercato, indipendente dalla frequenza e lineare in fase: come chiaramente dimostrato da Poletti, questi microfoni virtuali “grezzi” presentano una risposta in frequenza decisamente irregolare in modulo e in fase, che può comunque essere compensata in seguito.

Poletti ha inoltre ricavato le espressioni teoriche delle funzioni di trasferimento, che possono essere usate per creare degli opportuni filtri di equalizzazione.

Comunque, una soluzione più intelligente e pratica è semplicemente quella di misurare queste funzioni di trasferimento “grezze” in una camera anecoica ed in seguito ricavare, per ogni microfono virtuale, il proprio filtro inverso per mezzo del metodo di inversione di Kirkeby [18]. Ciò ha inoltre il vantaggio di compensare la risposta specifica del microfono utilizzato ed il suo diagramma di direttività dipendente dalla frequenza (che solo in modo approssimativo corrisponderà al diagramma a cardioidi teorico).

Una volta che le risposte dei microfoni sono ottenute, esse possono essere utilizzate come filtri di convoluzione da applicare ai segnali anecoici monofonici corrispondenti alle diverse posizioni della sorgente. Dopo aver mixato i risultati, un decodificatore Ambisonics di ordine elevato (3° ordine in questo esempio) è necessario per ricavare i segnali da inviare ad un normale array multicanale di altoparlanti (generalmente posizionati regolarmente attorno ad un cerchio circostante lo “sweet spot”); tale array comporta una localizzazione ed una separazione dei canali decisamente migliore di quanto faccia il sistema Ambisonics “standard” (di 1° ordine).

Un secondo modo possibile di utilizzare questi segnali di ordine elevato è quello di pilotare un array 5.1 ITU standard mediante la sintetizzazione di 5 opportuni diagrammi di direttività asimmetrici, come suggerito da D. de Vries ed altri [27].

### 3.6 L'approccio WFS circolare

Le 36 misurazioni B-format effettuate lungo la circonferenza con 1 metro di raggio sono esattamente l'insieme dei dati richiesti per utilizzare il metodo della Wave Field Synthesis (WFS), descritto da D. de Vries ed altri in un loro articolo [6].

Alla base del metodo vi è il Principio di Huygens: conoscendo la pressione sonora e la velocità delle particelle su di una superficie chiusa, è possibile ricreare dentro di essa lo stesso campo sonoro che era presente nello spazio originario, utilizzando un adeguato array di altoparlanti, posti esattamente in corrispondenza delle posizioni del microfono. La teoria comunque permette anche di “espandere” o “restringere” la geometria dell'array trasduttore in virtù del fatto che il campo sonoro è decomposto in fronti d'onda che avanzano.

La WFS è una riduzione bidimensionale di questa teoria generale, in cui i microfoni sono collocati lungo una linea chiusa attorno all'area di ascolto e di conseguenza l'espansione-compressione può avvenire solo sul piano orizzontale. Ciò limita pure la quantità di “movimento” che può essere effettuata. Comunque, partendo da un array avente 1 metro di raggio, è relativamente semplice ricavare i segnali per una batteria di altoparlanti adatta per una sala di ascolto di medie dimensioni e “stiracchiare” tale batteria in modo tale che gli altoparlanti siano collocati su 4 arrays lineari

anziché su un array circolare. La seguente figura (Figura 24, tratta da [27]) mostra una schematizzazione di questo processo.

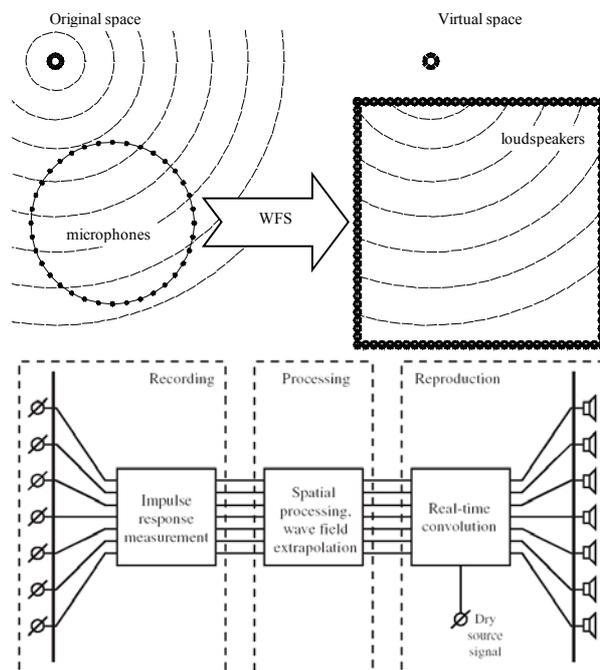


Figura 24: Schema della elaborazione WFS.

L'“elaborazione spaziale” richiesta per ottenere le risposte all'impulso adatte a questo tipo di riproduzione a partire dalle risposte all'impulso misurate non è banale, e può essere compresa solo dopo uno studio approfondito del materiale pubblicato (e non pubblicato) dalla Technical University di Delft. Finora gli autori non sono stati in grado di creare un semplice plug-in che effettui facilmente questa trasformazione spaziale, ma ciò è in programma per il futuro.

Ovviamente questa teoria richiede una piccola distanza spaziale tra due posizioni microfoniche consecutive onde ridurre l'aliasing spaziale che si verifica quando si campionano i fronti d'onda. Dal momento che in questo caso il numero dei microfoni è abbastanza limitato (36), ciò si traduce in una rigida limitazione dell'intervallo di frequenze che non provoca aliasing spaziale. Al di sopra di questa soglia (che è di circa 1kHz per la geometria qui utilizzata) non è più possibile in alcun modo ricostruire fedelmente i fronti d'onda. Per evitare allora alterazioni e colorazioni è consigliabile randomizzare la fase, in modo tale che la somma delle uscite dei vari altoparlanti formanti l'array non causi interferenza troppo a lungo e riduca al minimo la somma energetica (come nell'Ambisonics).

La randomizzazione della fase può essere ottenuta dalla convoluzione del segnale che alimenta ciascun altoparlante con un differente “burst” di rumore bianco,

oppure utilizzando altoparlanti a fase incoerente (DML, altoparlanti a modo distribuito).

### 3.7 Metodi ibridi (Ambiophonics)

Il metodo Ambiophonics è una soluzione ibrida, tesa a compensare reciprocamente i difetti di due sistemi di base: la riproduzione con cancellazione del cross-talk di tracce binaurali mediante altoparlanti ravvicinati (Stereo Dipolo) e surround 3D ottenuto dalla convoluzione di microfoni virtuali orientati in corrispondenza agli altoparlanti.

La seguente figura (Figura 25) mostra un tipico array Ambiophonics, (stereo dipolo frontale più impianto surround a 8 altoparlanti).

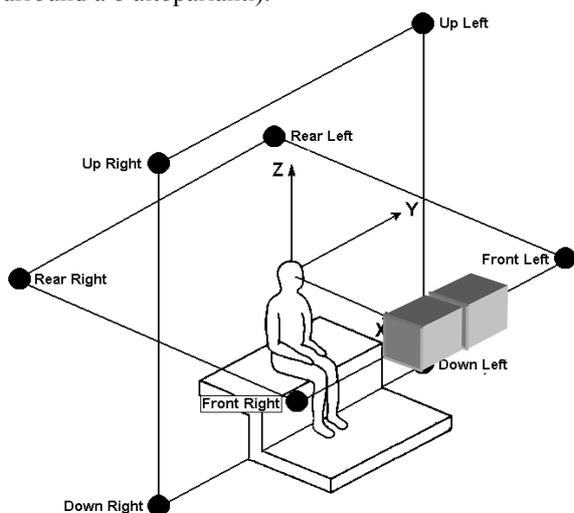


Figura 25: Ambiophonics array.

La teoria per ottenere i segnali per questi altoparlanti è già stata presentata nei capitoli precedenti e il procedimento di assemblaggio dell'intero sistema è stato esaurientemente descritto da uno degli autori [5].

L'unico punto che merita qui un approfondimento è il fatto che, in un sistema Ambiophonics, gli altoparlanti Stereo Dipolo devono fornire solo il suono diretto e le prime riflessioni provenienti dal palcoscenico, mentre gli altri altoparlanti "surround" devono fornire le ulteriori riflessioni ed il riverbero.

Ciò significa che le risposte all'impulso necessitano di essere opportunamente modificate: da un lato le ORTF, che sono utilizzate dallo Stereo Dipolo, devono essere dolcemente tagliate dopo il suono diretto; dall'altro lato le risposte all'impulso B-format, da cui si ricavano i canali surround, devono essere private del suono diretto. Queste operazioni sono abbastanza delicate perché, se fatte in modo improprio, possono causare un infelice mescolamento dei due sistemi di base o possono introdurre ritardi artificiali che alterano la distanza temporale tra il suono diretto e la susseguente riverberazione.

Una ultima considerazione riguarda la scelta delle risposte all'impulso che comandano l'array "surround". In un precedente lavoro [5] esse erano tutte ottenute da una singola risposta all'impulso B-format, semplicemente utilizzando Visual Virtual Microphone ed orientando il microfono virtuale nella direzione del corrispondente altoparlante. Ora, grazie alla disponibilità di diverse risposte all'impulso B-format lungo una circonferenza, è possibile selezionare, per ogni altoparlante, non solo la direzione del microfono virtuale, ma anche la sua corrispondente posizione lungo la circonferenza.

Ciò migliora in modo significativo il risultato perché in questo modo le risposte all'impulso sono prelevate in posizioni diverse e sono mutualmente incoerenti. Ciò evita interferenza ed alterazioni dovute alla interazione fra segnali provenienti da diversi altoparlanti, tutti alimentati con segnali strettamente correlati.

## 4 CONCLUSIONI

Questo articolo ha presentato una nuova ed avanzata tecnica di misura, che consente di acquisire il maggior numero possibile di informazioni acustiche dentro ad un teatro esistente. Questo metodo si basa sulla misurazione di un grande numero di risposte all'impulso per mezzo di un set-up microfonico ruotante.

A partire dal set di dati ottenuti è possibile ricavare un insieme di risposte all'impulso adatte per la riproposizione di uno spazio acustico virtuale, mediante le tecnologie di riproduzione attualmente disponibili. Con riferimento in modo particolare alla riproduzione delle proprietà spaziali del campo sonoro, si può notare che i dati ottenuti permettono l'auralizzazione dei risultati utilizzando:

- Una riproduzione stereo standard mediante una coppia di altoparlanti;
- Una riproduzione binaurale mediante cuffie acustiche, con head tracking;
- Una riproduzione mediante altoparlanti posti a breve distanza l'uno dall'altro (per mezzo di filtri che cancellano il cross-talk) – Stereo Dipolo;
- Una riproduzione Ambisonics mediante un array di altoparlanti 2D o 3D;
- Una riproduzione "surround" ITU 5.1, in conformità a set-ups microfonici standard (OCT, INA, Williams MMA, etc.);
- Una riproduzione multicanale ad alta direttività mediante il metodo dell'array circolare di Mark Poletti;
- L'auralizzazione ad ampia area mediante l'approccio della Wave Field Synthesis (WFS);
- Una qualsiasi combinazione dei summenzionati metodi, dando origine a tipologie di surround ibride e ad alto livello (Ambiophonics, Panorambiophonics e derivazioni).

Questo metodo fornisce di conseguenza il miglior modo possibile per conservare le proprietà acustiche di sale famose e preziose, quali sale da concerto e teatri, e di preservarle per le future generazioni. I dati ottenuti possono essere utilizzati per ricostruzioni acustiche tramite gli attuali sistemi surround (auralizzazione), senza però limitarne un uso futuro, restando ancorati ai limiti della tecnologia attualmente disponibile.

D'altra parte, i dati in possesso possono essere già utilizzati per elaborazioni ad alta qualità di registrazioni anecoiche, surclassando le attuali unità "artificiali" di riverberazione e spazializzazione, se impiegate assieme ad un valido convolutore software.

### RINGRAZIAMENTI

Questa ricerca è stata sovvenzionata e logisticamente sostenuta dalla Waves Ltd ([www.waves.com](http://www.waves.com)), come parte dello sviluppo di un nuovo plugin software per la riverberazione artificiale, basato sul campionamento di risposte acustiche all'impulso e capace di elaborare in tempo reale un "mix" surround multicanale.

La calibrazione dell'altoparlante e le misure effettuate nei teatri in Giappone sono state possibili solo grazie all'aiuto del Prof. Yoichi Ando e dei colleghi della Università di Kobe in Giappone (Kosuke Kato, Takuya Hotehama, Yosuke Okamoto), i quali hanno acconsentito all'utilizzo dei loro laboratori ed aiutato nel corso delle misurazioni. Inoltre utili discussioni e scambi di informazioni tecniche con questi colleghi hanno consentito agli autori di migliorare la tecnica di misura.

Gli studi su varie tecniche di riproduzione e su avanzate soluzioni multicanale ibride sono stati sovvenzionati dall'Ambiophonics Institute, dove si effettuano esperimenti di ascolto con diversi formati.

Gli autori desiderano ringraziare sentitamente gli Enti proprietari e/o gestori dei nove teatri oggetto delle misurazioni, che gentilmente hanno inoltre concesso l'autorizzazione alla pubblicazione dei dati acquisiti, e il Prof. L. Tronchin per l'aiuto prestato durante le misure.

### BIBLIOGRAFIA

- [1] Michael Gerzon - "Recording Concert Hall Acoustics for Posterity", *JAES* Vol. 23, Numero 7, p. 569 (1975).
- [2] L. Tronchin, A. Farina - "The acoustics of the former Teatro "La Fenice", Venice", *JAES* Vol. 45, Numero 12, p. 1051 (1997).
- [3] "Carta di Ferrara", *CIARM*, <http://acustica.ing.unife.it/ciarm/Carta.htm> .
- [4] "Guidelines for acoustical measurements inside historical opera houses: procedures and validation", *CIARM*, <http://acustica.ing.unife.it/ciarm/download.htm>

- [5] A. Farina, R. Glasgal, E. Armelloni, A. Torger - "Ambiophonic Principles for the Recording and Reproduction of Surround Sound for Music" - *19th AES Conference on Surround Sound, Techniques, Technology and Perception* - Schloss Elmau, Germany, 21-24 Giugno 2001.
- [6] E.Hulsebos, D.de Vries, and E. Bourdillat - "Improved Microphone Array Configurations for Auralization of Sound Fields by Wave-Field Synthesis", *JAES* Vol. 50, Numero 10, p. 779 (2002).
- [7] A. Karamustafaoglu, U. Horbach, R. Pellegrini P. Mackensen, G. Theile - "Design and Applications of a Data-based Auralisation System for Surround Sound", *Pre-prints of the 106th AES Convention* (1999).
- [8] M. A. Poletti - "A Unified Theory of Horizontal Holographic Sound Systems", *JAES* Vol. 48, Numero 12, p. 1049 (2000).
- [9] A. Farina - "Simultaneous measurement of impulse response and distortion with a swept-sine technique", *110th AES Convention*, Paris, 18-22 Febbraio 2000.
- [10] S. Müller, P. Massarani - "Transfer-Function Measurement with Sweeps", *JAES* Vol. 49, Numero 6, p. 443 (2001).
- [11] G. Stan, J.J. Embrechts, D. Archambeau - "Comparison of Different Impulse Response Measurement Techniques", *JAES* Vol. 50, Numero 4, p. 249 (2002).
- [12] Y. Ando - "Concert hall acoustics", Springer Series in electrophysics, Berlin, 1985.
- [13] V.L. Jordan - "A group of objective acoustical criteria for concert halls", *Applied Acoustics*, Vol. 14 (1981).
- [14] A. Torger, A. Farina - "Real-time partitioned convolution for Ambiophonics surround sound", *2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* - Mohonk Mountain House New Paltz, New York, 21-24 Ottobre 2001.
- [15] T. G. Stockham Jr. - "High-speed convolution and correlation", *AFIPS Proc. 1966 Spring Joint Computer Conf.*, Vol 28, Spartan Books, pp. 229 - 233 (1966).
- [16] W.G. Gardner - "Efficient convolution without input-output delay", *JAES* Vol. 43, Numero 3, pp. 127-136 (1995).
- [17] O. Kirkeby, P. A. Nelson, H. Hamada - "The "Stereo Dipole" - A Virtual Source Imaging

- System Using Two Closely Spaced Loudspeakers” – *JAES* Vol. 46, Numero 5, pp. 387-395 (1998).
- [18] O.Kirkeby, P.A. Nelson, P. Rubak, A. Farina – “Design of Cross-talk Cancellation Networks by using Fast Deconvolution” - *106th AES Convention*, Munich, 8-11 Maggio 1999.
- [19] Lake DSP Huron Workstation, [HTTP://www.lakedsp.com](http://www.lakedsp.com).
- [20] A. Farina, E. Ugolotti - “Software Implementation Of B-Format Encoding And Decoding”, *Pre-prints of the 104rd AES Convention*, Amsterdam, 15 - 20 Maggio 1998.
- [21] A.Field - “B-dec High resolution First Order Ambisonic B-format decoder”, University of York, [http://www.york.ac.uk/inst/mustech/3d\\_audio](http://www.york.ac.uk/inst/mustech/3d_audio).
- [22] G. Theile – “Multichannel Natural Music Recording Based on Psychoacoustic Principles” - *AES 19th International Conference*, Maggio 2001.
- [23] Roland Jacques, MultiMedia Projekt VERDI, TU Ilmenau Laboratory, Germany (2002) - <http://www.stud.tu-ilmeneau.de/~proverdi/daten/umlenn.html>
- [24] Williams, M.; Le Du, G. – “Multichannel Microphone Array Design”, *Pre-prints of the 108th AES Convention* (2000).
- [25] Herrmann, U., Henkels, V., Braun, D. – “Comparison of 5 surround microphone methods”, *Proceedings 20th Tonmeistertagung*, (ISBN 3-598-20361-6), pp. 508-517 (1998).
- [26] D. McGriffy - “Visual Virtual Microphone”, [HTTP://mcgriffy.com/audio/ambisonic/vvmic](http://mcgriffy.com/audio/ambisonic/vvmic)
- [27] E. Hulsebos, T. Schuurmans, D. de Vries and R. Boone – “Circular microphone array for discrete multichannel audio recording”, *Pre-prints of the 114th AES Convention*, Amsterdam, 22-25 Marzo 2003.