

## **NUOVI METODI E STRUMENTI DI MISURA PER LO STUDIO DELL’ACUSTICA DEI TEATRI STORICI ITALIANI**

Angelo Farina

Universita’ degli Studi di Parma, Dip. Ing. Industriale, PARMA

### **1. Riassunto**

In questo articolo viene presentata una nuova metodica di misurazione della risposta all’impulso “spaziale” dentro teatri storici, al fine di valutarne le proprietà acustiche, e di preservarne la sonorità per i posteri.

Dai segnali registrati é anche possibile visualizzare la direzione di provenienza del suono, mettendo in evidenza i pattern di riflessione che si sviluppano entro la sala, che vengono sovrapposti ad una fotografia panoramica della stessa.

La stessa metodica matematica di filtraggio numerico utilizzata per registrare i segnali a partire da quelli catturati da un array di microfoni, consente poi di riprocessare i segnali stessi al fine di realizzare la riproduzione sonora su un array di altoparlanti: questo consente di virtualizzare l’acustica del teatro storico, anche in caso esso vada distrutto, come accaduto purtroppo di sovente.

Il metodo proposto é l’evoluzione di una ricerca durata quasi 20 anni, che ha fatto impiego inizialmente della tecnica binaurale, poi di quella Ambisonics di primo ordine, per passare successivamente ad Ambisonics di ordine elevato, ed approdare infine all’attuale tecnica denominata “Spatial PCM Sampling”, o SPS, che viene descritta per la prima volta nel presente articolo.

### **2. Introduzione**

L’importanza di registrare le caratteristiche acustiche di un teatro storico a fini della sua preservazione per i posteri sono risultate improvvisamente evidenti a seguito del tragico incendio che distrusse il teatro La Fenice di Venezia nella notte del 29 gennaio 1996.

Fortuitamente poche settimane prima all’interno del teatro erano state effettuate misure acustiche da parte dell’ing. Tronchin, impiegando strumentazione e tecnica di rilievo innovative (per l’epoca) [1]. La possibilità di utilizzare le risposte all’impulso binaurali registrate nel teatro per realizzare l’ascolto virtuale della sala ormai distrutta consente oggi di confrontare la accuratezza della ricostruzione, e la sua fedeltà all’originale.

A metà degli anni '90 la tecnica di registrazione e riproduzione binaurale era l'unica in grado di catturare l'informazione spaziale, oltre alla risposta temporale e timbrica ottenibili già da tempo mediante registrazioni monofoniche della risposta all'impulso di un teatro.

Tuttavia la tecnica binaurale ha anche dei limiti abbastanza severi: non consente di analizzare in modo oggettivo i segnali registrati, al fine di individuare l'effettiva direzione di provenienza delle singole onde riflesse, né di riprodurre il suono virtuale del teatro mediante sistemi di altoparlanti tridimensionali.

Per superare questi limiti, occorre registrare risposte all'impulso con un maggior numero di canali. Inizialmente è stato reso operativo il metodo suggerito da Gerzon [2], mettendo a punto una tecnica di misurazione di risposte all'impulso a 4 canali, in formato Ambisonics (B-format), che utilizza opportune sonde microfoniche in grado di campionare il segnale di pressione e le tre componenti cartesiane del segnale di "particle velocity" [3]. Grazie a questa tecnica, la soc. israeliana Waves ha registrato una vastissima collezione di "fotografie acustiche tridimensionali" dei più famosi teatri del mondo [4].

Si è comunque visto che estrarre la completa informazione spaziale da una risposta all'impulso B-format è in generale problematico. Sicuramente la classica metodica Ambisonics vi riesce solo in minima parte, e sono pertanto state sviluppate metodiche di analisi dei segnali B-format basate su metodiche analitiche avanzate, quali il metodo SIRR di Ville Pulki [5], oppure Harpex-B di Svein Berge [6].

Anziché ricorrere a tali "sporchi trucchi", comunque, l'approccio più robusto consiste sicuramente nell'utilizzare una tecnica di campionamento dotata di maggior "sampling rate" spaziale, facendo impiego di una sonda microfonica dotata di un maggior numero di sensori.

L'autore ha proposto in particolare l'utilizzo di una array sferico di microfoni a pressione [7], unito all'uso di una sorgente sonora in grado di sviluppare una direttività controllabile, in modo da campionare una completa matrice tridimensionale rappresentativa della funzione di trasferimento spaziale del teatro analizzato.

In pratica, il metodo Ambisonics viene esteso includendo un maggior numero di armoniche sferiche (High Order Ambisonics).

Nell'impiego pratico, tuttavia, si è scoperto che le armoniche sferiche di ordine elevato possono essere misurate accuratamente solo in un ristretto intervallo di frequenze, limitando pertanto la possibilità di effettuare una ricostruzione virtuale a banda larga e ad alta risoluzione spaziale [8].

Recentemente pertanto l'autore ha proposto di abbandonare la rappresentazione dell'informazione spaziale mediante armoniche sferiche, ed utilizzare invece una rappresentazione "diretta" basata sul campionamento dello spazio tramite una suddivisione quasi regolare dello spazio in fasci. In pratica, è l'equivalente della discretizzazione di una forma d'onda come successione di impulsi (PCM, pulse code modulation), anziché come sovrapposizione di segnali sinusoidali (analisi di Fourier).

In pratica la misura richiede l'impiego di una array microfonico sferico [9]. Anziché derivare da esso i 16 segnali corrispondenti alle armoniche sferiche sino al terzo ordine, i 32 segnali provenienti dalle capsule sono processati mediante una matrice di filtri FIR (32x32), in modo da sintetizzare 32 microfoni ultradirettivi, che puntano nelle direzioni delle 32 capsule originali. In pratica, si tratta di una tecnica di beamforming superdirettivo.

Si misura dunque una risposta all'impulso a 32 canali, a banda larga (20 Hz – 20 kHz), che preserva pertanto l'intera informazione spaziale e tempo-frequenziale.

É comunque agevole ottenere anche, parallelamente, una classica risposta all’impulso B-format, utilizzabile per il calcolo dei parametri acustici secondo la norma ISO 3382, che richiede l’uso di un microfono omnidirezionale e tre microfoni “a figura di 8”.

I 32 segnali ultradirettivi costituiscono il campionamento PCM spaziale: si può infatti pensare che essi derivino da 32 microfoni ultradirettivi, la cui risposta polare é in sostanza una “delta di Dirac spaziale”, cioè presenta un forte picco nella direzione di puntamento, e va rapidamente a zero verso tutte le altre direzioni.

La fig. 1 pone a confronto la tecnica PCM applicata alla discretizzazione di una forma d’onda nel tempo ed alla discretizzazione della distribuzione dell’energia sonora nello spazio.

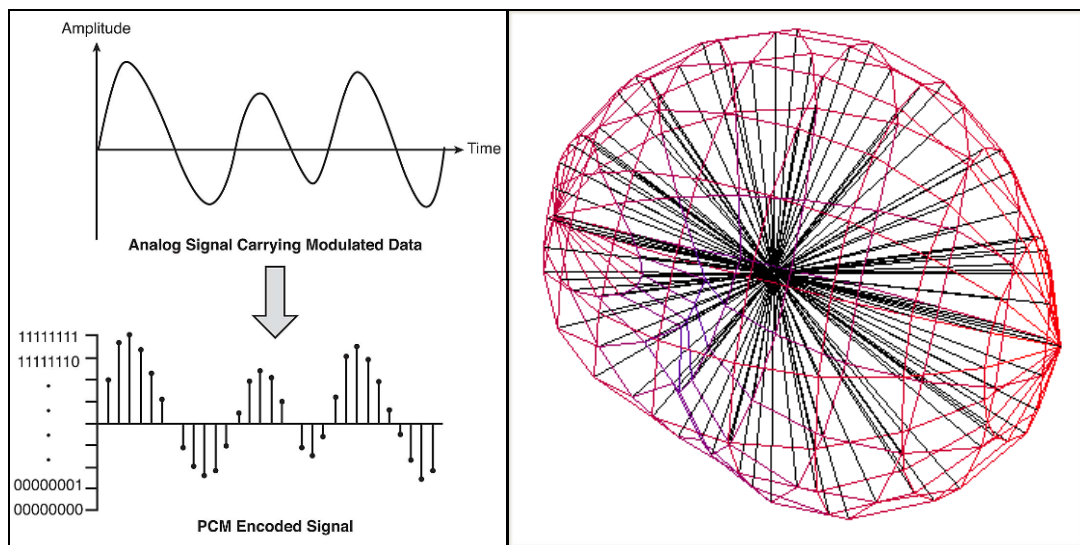


Figura 1: Campionamento PCM nel tempo (a sinistra) e nello spazio (a destra)

La fig. 2 mostra invece la trasformazione nel dominio della frequenza dei segnali campionati con metodica PCM: per una forma d’onda, troviamo la classica analisi di Fourier, per cui il segnale può venire rappresentato come somma di un numero discreto di sinusoidi, con opportuna ampiezza e fase. Analogamente, l’informazione PCM spaziale può venire convertita nella rappresentazione del campo sonoro come sovrapposizione di un numero discreto di armoniche sferiche, anche in questo caso assegnando di ciascuna di esse un opportuno valore di ampiezza.

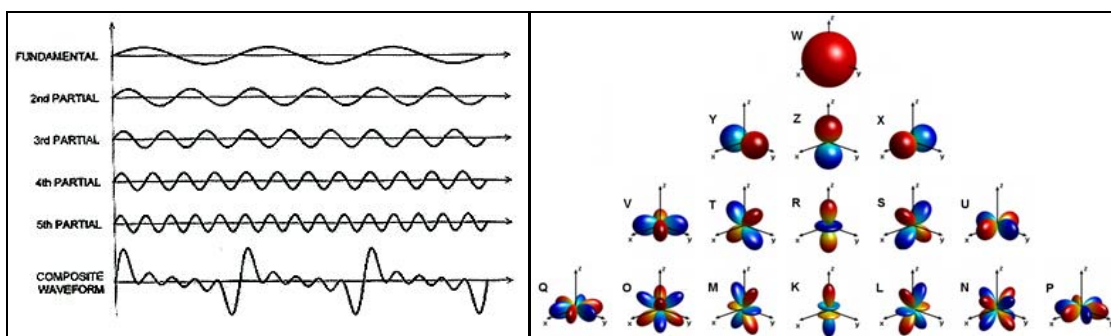


Figura 2: Analisi di Fourier (a sinistra) e in armoniche sferiche (a destra)

Anche nel dominio spaziale si possono applicare le normali operazioni di filtraggio numerico utilizzate per il filtraggio digitale dei segnali campionati. Si possono pertanto utilizzare filtri FIR spaziali, al fine di modificare l'informazione spaziale originaria, ad esempio facendo ruotare il campo sonoro, o deformandolo a piacimento.

Ma, ai fini dell'utilizzo nei teatri storici, l'aspetto più interessante consiste nella possibilità di analizzare l'informazione spaziale in modo visivo, realizzando dei brevi filmati che mostrano l'evolversi nel tempo della distribuzione spaziale dell'energia sonora che arriva nel punto di ascolto [10].

Questo approccio è stato impiegato per la prima volta quest'anno per visualizzare le riflessioni sonore dentro il teatro della Scala di Milano.

Nell'ultima parte di questo articolo viene infine spiegato come sia possibile utilizzare le registrazioni di risposte all'impulso in formato PCM spaziale al fine di operare l'auralizzazione, ovvero l'ascolto di suoni di vario tipo (musica, parlato, etc.) preventivamente registrati in ambiente anecoico, simulandone la propagazione nel teatro campionato e presentando il risultati agli ascoltatori mediante un idoneo sistema di altoparlanti, installati entro una sala d'ascolto insonorizzata.

La tecnica di Spatial PCM Sampling costituisce dunque una valida alternativa alle tecniche di registrazione di risposte all'impulso spaziali precedentemente impiegate (binaurale, Ambisonics di primo ordine, High Order Ambisonics): a parità di numero di microfoni impiegati, essa fornisce una più estesa riposta in frequenza e massimizza l'informazione spaziale catturata. Inoltre, la sua semplice definizione matematica la rende accessibile a tutti, senza la complessità concettuale derivante dall'uso delle armoniche sferiche: i segnali registrati possono essere processati in modo assolutamente standard, con un normale programma di editazione multitraccia, o addirittura semplicemente con un mixer, senza dover ricorrere a decoder dedicati o a speciali "plugins" per l'elaborazione dei segnali armonici sferici.

### **3. Sintesi di microfoni virtuali da un array microfonico**

In questo capitolo viene spiegato come ottenere un segnale corrispondente a quello che verrebbe catturato da un microfono dotato di un pattern di direttività arbitrario, a partire dal filtraggio numerico dei segnali catturati da un array microfonico costituito da  $M$  capsule.

La base di questo metodo, che non si basa su alcuna formulazione teorica per rappresentare la propagazione sonora nello spazio, è quella di aver misurato la risposta all'impulso multicanale dell'array microfonico a partire da un elevato numero  $D$  di direzioni diverse, più o meno uniformemente distribuite sulla superficie sferica. Chiamiamo  $[c]$  la matrice  $M \times D$  di queste risposte all'impulso misurate.

La fig. 3 mostra tale situazione, in cui i segnali degli  $M$  microfoni sono campionati quando il suono viene generato dalla  $D$ -esima sorgente sonora.

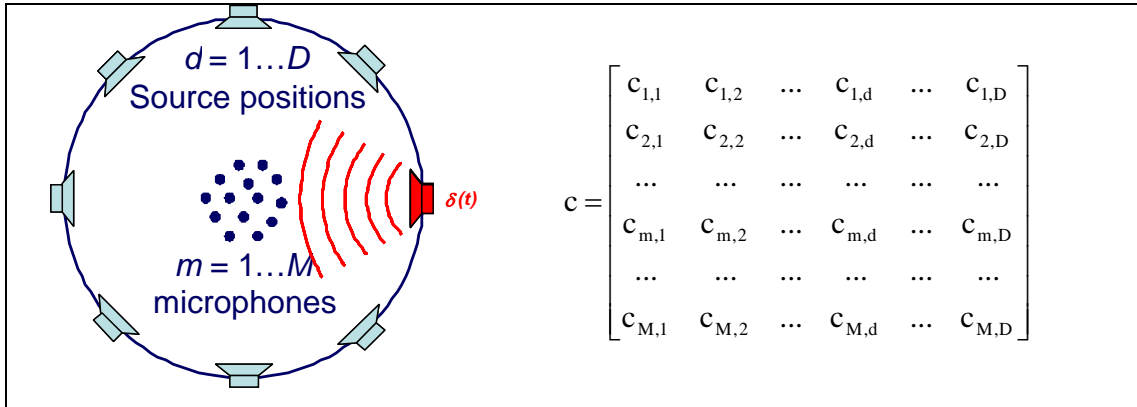


Figura 3: misura delle risposte all’impulso di M microfoni a partire da un numero D di posizioni della sorgente

In questa sede é irrilevante quale forma abbia il polar pattern del microfono virtuale da sintetizzare (potrebbe esser un semplice cardioide, o addirittura un omni). Quel che conta é definire la risposta che tale microfono avrà in corrispondenza di ciascuna delle D direzioni da cui é stato caratterizzato l’array microfonico. Chiamiamo  $\{p_d\}$  il vettore delle risposte all’impulso prescritte per il microfono virtuale nelle D direzioni di misura.

A noi serve un vettore di M filtri  $\{h_m\}$ , tali che, applicandoli agli M segnali delle capsule, e sommando i risultati, si ottenga, per ciascuna delle D direzioni, il segnale del microfono virtuale voluto  $p_d$ , come mostrato in fig. 4.

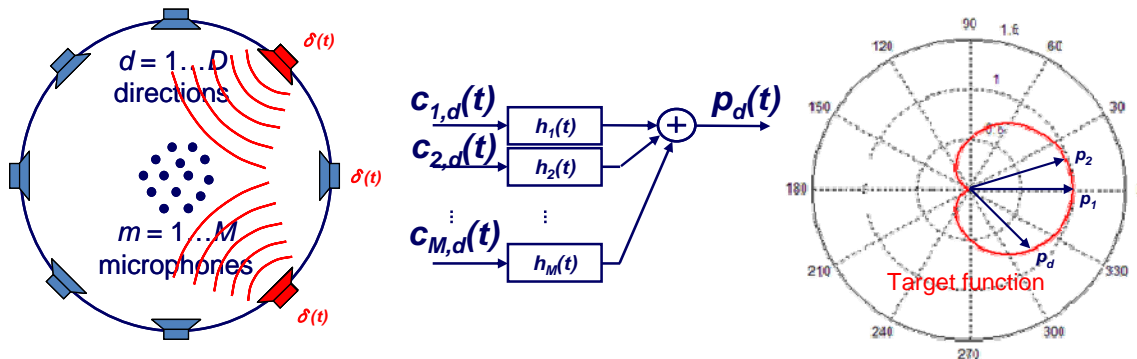


Figura 4: definizione della matrice di filtri h tali da produrre un microfono virtuale con il pattern di direttività voluta

$$(1) \quad \sum_{m=1}^M c_{m,d} * h_m \Rightarrow p_d \quad d = 1..D$$

Poiché sono note sia le risposte all’impulso misurate  $\{c\}$ , sia le risposte prescritte per il microfono virtuale  $\{p\}$ , é sufficiente passare nel dominio della frequenza, in modo che la convoluzione diventi una semplice moltiplicazione, e ci troviamo davanti ad un sistema di D equazioni in M incognite, fortemente sovracondizionato ( $D \gg M$ ).

La tecnica numerica di soluzione di tale sistema di equazioni esula dagli scopi del presente articolo, e comunque é chiaramente descritta in [9]. Analogamente in [9] viene dettagliatamente descritta l’effettuazione delle misure di taratura della sonda microfonica eseguite in camera anecoica, che hanno consentito di determinare la matrice  $\{c\}$ .

I filtri di processamento che sintetizzano i microfoni virtuali possono solitamente essere abbastanza corti, tipicamente 2048 o 4096 coefficienti, lavorando con una frequenza di campionamento di 48 kHz.

Un moderno PC elabora il filtraggio FIR in maniera molto efficiente, grazie all'algoritmo di convoluzione partizionata [11]. Questo rende possibile applicare il vettore di filtraggio  $\{h\}$  in tempo reale, anche quando il numero  $M$  di capsule microfoniche é abbastanza elevato: nel caso qui presentato,  $M=32$ , ed un normale PC portatile dell'ultima generazione riesce facilmente a generare simultaneamente i segnali filtrati di decine di microfoni virtuali operando in tempo reale ed a bassa latenza. Di fatto, dunque abbiamo vettori di filtraggio  $\{h\}$  contenenti 32 filtri FIR di 2048-4096 campioni cadauno. Ed il PC convolve i 32 segnali di ingresso dalle capsule, con 32 diversi vettori di filtraggio, corrispondenti a 32 distinti microfoni virtuali, che puntano nelle stesse 32 direzioni in cui sono posizionate le capsule. In pratica, abbiamo una matrice di filtraggio  $[h]$  di  $32 \times 32$  filtri FIR.

La fig. 5 illustra la disposizione delle capsule dell'array sferico impiegato, ed i vettori di puntamento dei microfoni virtuali.

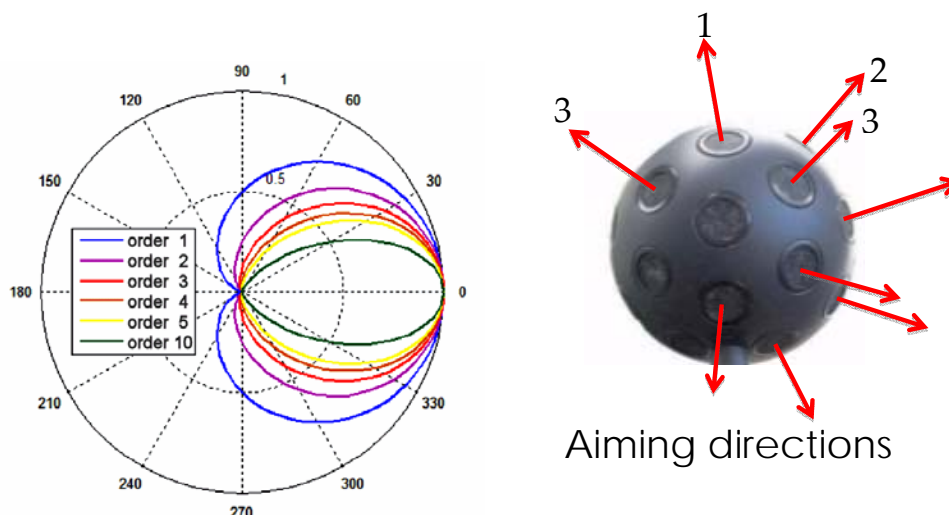


Figura 5: pattern dei microfoni virtuali e loro direzioni di puntamento

I pattern mostrati in figura 5 corrispondono a microfoni di tipo cardioide di vario ordine, definiti dalla relazione:

$$(2) \quad Q(\varphi, n) = [0.5 + 0.5 \cdot \cos(\varphi)]^n$$

Dopo varie prove, si é verificato che l'ordine ottimale onde ottenere una sufficiente risoluzione angolare, mantenendo una completa "copertura" dell'intera superficie sferica con 32 microfoni virtuali é  $n=4$ .

Questo metodo, unito alle ben note metodiche di eccitazione della sala mediante un altoparlante alimentato con segnale Exponential Sine Sweep [12,13], consente di misurare, in ogni punto della stessa, una risposta all'impulso a 32 canali in formato SPS (Spatial PCM Sampling).

#### 4. Strumentazione impiegata

La teoria precedentemente esposta é divenuta utilizzabile nella pratica solo dopo che si é resa disponibile sul mercato strumentazione di registrazione multicanale di alta qualità e facilmente portatile.

In particolare solo nel 2010 é comparso un sistema microfonico sferico con prestazioni sufficientemente buone da consentirne l’utilizzo a fini di misurazioni scientifiche, superando le limitazioni intrinseche della precedente generazione di sonde microfoniche tridimensionali, concepite unicamente per registrazioni musicali, e quindi non calibrabili né impiegabili ai fini di misura.

In particolare, la MH Acoustics (USA) ha iniziato la commercializzazione del primo prototipo d Eigenmike, visibile sulla sinistra in fig. 6. Il Centro Ricerche RAI, in collaborazione con l’Università’ di Parma, ha iniziato la sperimentazione utilizzando uno dei primissimi esemplari prodotti (S/N 005). Su specifica richiesta dei tecnici RAI, la MH Acoustics ha successivamente sviluppato la prima serie di produzione di Eigenmike, visibile nella parte destra di fig.6. Attualmente il Centro Ricerche RAI ha acquistato 3 di questi microfoni, e l’Università’ di Parma ne ha appena ordinato un quarto.



Figura 6: prima e seconda versione della sonda Eigenmike<sup>(TM)</sup>



Figura 7: composizione del sistema Eigenmike<sup>(TM)</sup>

Come mostrato nella fig. 7, il sistema si compone di una sonda microfonica con elettronica integrata all’interno della sfera (preamplificatori, convertitori e sistema di instradamento dei segnali campionati su cavo Ethernet), ed una “scheda audio” Firewire

da collegare al computer portatile, che riceve il flusso dati dal cavo Ethernet proveniente dal microfono (provvedendo altresì ad alimentarlo, sempre attraverso il cavo Ethernet, ed a inviare allo stesso i comandi di regolazione del guadagno dei preamplificatori).

La scheda audio, denominata EMIB, è alimentata a 12V, quindi anche a batteria, ed è dotata di normali drivers, che la fanno “vedere” a computer Windows, Mac e Linux come periferica ASIO a 32 canali di ingresso e 32 canali di uscita.

La EMIB dispone anche di due uscite analogiche (solitamente usate per monitorare in cuffia il segnale di due dei microfoni virtuali che stanno venendo sintetizzati) ed 8 canali digitali su cavo ottico ADAT, molto utili per il collegamento ad un mixer digitale, per applicazioni di broadcasting in tempo reale (tipica applicazione utilizzata dalla RAI).

Ai fini della misura della risposta all'impulso, invece, l'uscita analogica della EMIB viene impiegata per generare il segnale Exponential Sine Sweep, che viene inviato all'altoparlante.

Nel corso delle prime prove di utilizzo della nuova tecnica, si è fatto impiego di un altoparlante dodecaedrico Look Line D300, accoppiato ad un subwoofer Audio Pro 100, mediante un opportuno cross-over regolato a circa 100 Hz.

Il segnale ESS è stato pre-equalizzato, in modo da rendere piatto il livello di potenza generato dall'altoparlante effettuando una misura intensimetrica secondo ISO 9614/2. In questo modo, la intera catena di misura non produce alcuna apprezzabile colorazione timbrica sulle risposte all'impulso misurate: questo è estremamente importante allorché esse vengono impiegate come filtri di auralizzazione, mediante appositi programmi di convoluzione in tempo reale.

La fig. 8 mostra il sistema altoparlante-microfono durante le misure effettuate al Teatro alla Scala.



Figura 8: Altoparlante e microfono nel teatro alla Scala



## 5. Programmi di elaborazione impiegati

Occorre distinguere tre fasi:

1. Effettuazione delle misure di risposta all’impulso
2. Conversione dei segnali generando i 32 microfoni virtuali (SPS)
3. Elaborazione delle risposte all’impulso SPS

La fase di misurazione é stata effettuata senza far ricorso alla possibilità di convertire direttamente in tempo reale i segnali provenienti dalle capsule microfoniche nei 32 segnali dei microfoni virtuali SPS. Questa scelta deriva dal fatto che la tecnica era ancora in fase di prima sperimentazione, e si é preferito registrare direttamente i segnali originali delle capsule (chiamati in gergo “A-format”), in modo da poter provare poi differenti metodiche di filtraggio per la conversione dei segnali nei 32 microfoni virtuali SPS, o anche per generare invece i normali segnali B-format del primo ordine (WXYZ, ovvero il segnale di pressione W ed i tre segnali di velocità particellare X,Y,Z, corrispondenti a tre microfoni a figura di 8 orientati secondo i tre assi isometrici definiti dalla norma ISO 2731).

Per la fase di misurazione, é stato impiegato il tradizionale programma Adobe Audition [14], integrato dalla suite di plugins Aurora [15].

Nella seconda fase si é operata la conversione fra i segnali originali A-format ed i segnali dei 32 microfoni virtuali ultradirettivi. Tale conversione richiede di effettuare una “convoluzione matriciale” fra i 32 canali di ingresso ed i 32 canali di uscita, utilizzando una matrice di filtri “completa”, come definita al cap. 3.

A tal fine, si é fatto impiego del “plugin VST” Xvolver [16], sviluppato presso l’Universita’ di Parma, e mostrato in fig. 9. Esso é stato “ospitato” dentro il programma AudioMulch [17], che si é fatto carico di leggere e scrivere i files WAV a 32 canali.

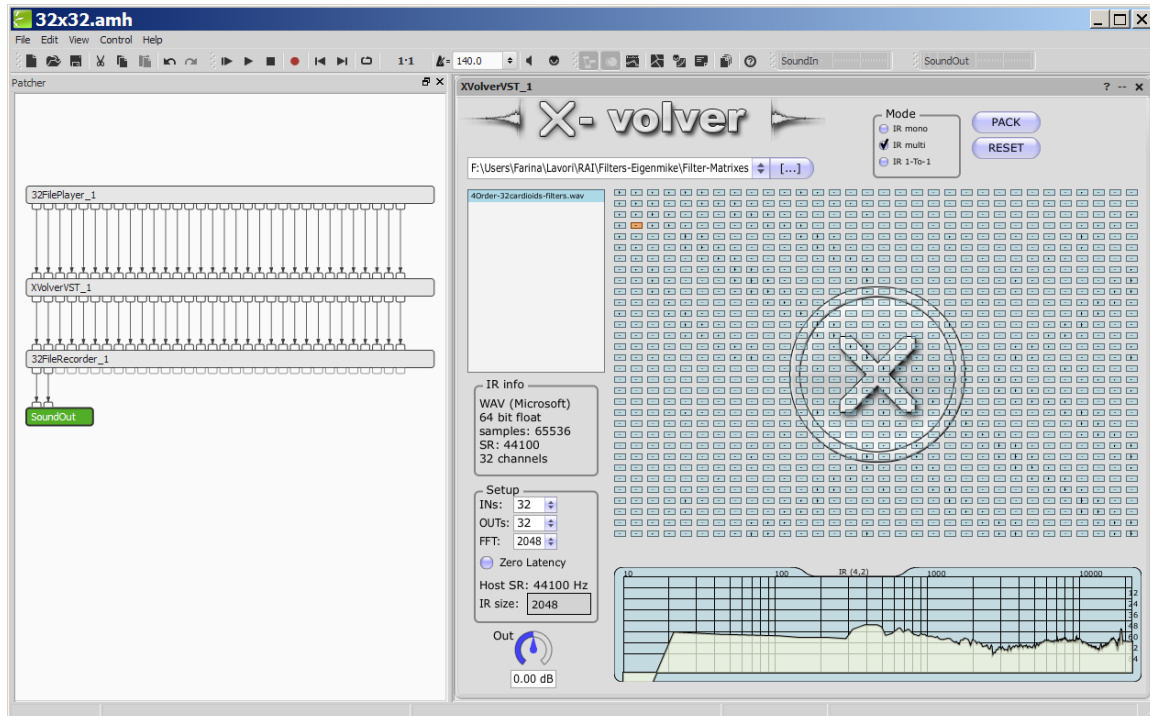


Figura 9: Interfaccia grafica di AudioMulch e del plugin Xvolver

Sempre utilizzando Xvolver ed una seconda matrice di filtri FIR di 32x4 filtri, si sono generati i 4 segnali “B-format” WXYZ, che sono poi stati utilizzati per il calcolo dei

classici parametri acustici secondo ISO3382, utilizzando il modulo Aurora Acoustical Parameters.

I segnali B-format ottenuti con il sistema Eigenmike<sup>(TM)</sup> sono risultati significativamente migliori di quelli ottenibili con i classici microfoni Soundfield o similari, grazie all'uso di una sonda con un numero di capsule molto più elevato (32 anziché 4) e di qualità superiore (sono capsule di precisione in classe I realizzate appositamente dalla Sennheiser per la MH Acoustics).

In particolare questo è vero per i segnali di “particle velocity”: la precedente esperienza ha infatti mostrato come sia difficile registrare i segnali X,Y,Z facendo impiego di strumentazione quali i microfoni Soundfield, le tradizionali sonde intensimetriche di tipo p-p, o le moderne sonde intensimetriche p-v (Microflown). Nessuno di tali sistemi consente infatti di avere simultaneamente un range di frequenza esteso, un buon rapporto segnale-rumore, un perfetto accoppiamento di guadagno e fase con il segnale di pressione, ed elevata accuratezza geometrica nella forma dei polar-pattern.

Per l'elaborazione delle risposta all'impulso a 32 canali in formato SPS, infine, è stato sviluppato un codice Matlab che analizza i segnali dei microfoni virtuali, realizzando una mappatura a colori “dinamica” che mostra, allo scorrere del tempo, l'evolversi della distribuzione spaziale del suono nella sala. Il successivo capitolo spiega nei dettagli questa tecnica di elaborazione.

## 6. Visualizzazione della distribuzione spaziale dell'energia sonora

Occorre anzitutto disporre di uno “sfondo grafico” che rappresenti la proiezione piana dell'intero orizzonte sferico “visto” dal punto in cui è collocato il microfono.

Si tratta dunque di eseguire una fotografia panoramica a 360 gradi, cosa attualmente agevolmente realizzabile impiegando una comune macchina fotografica digitale compatta, dotata di obiettivo sufficientemente grandangolare, con la quale si scatta una cinquantina di foto puntandola tutt'attorno, in modo da inquadrare gran parte del panorama sferico osservabile.

Per trasformare l'insieme delle foto in unica immagine piana, si deve fare uso della Proiezione di Mercatore, molto impiegata per la cartografia nautica, come illustrato in fig. 10.

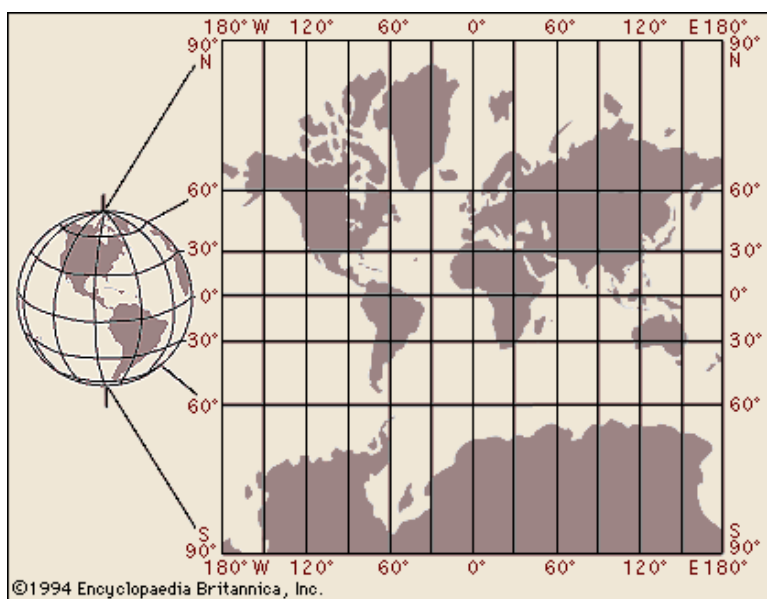


Figura 10: Proiezione di Mercatore

Se la proiezione é eseguita correttamente, si ottiene una immagine con rapporto di forma 2:1, in cui i pixels orizzontali corrispondono all’angolo di Azimuth (da -180 a +180 gradi) ed i pixel verticali corrispondono all’angolo di elevazione (da -90 a +90 gradi), in perfetta corrispondenza con longitudine e latitudine terrestri.

La gran parte del software di elaborazione fotografica, anche quello professionale, pur fornendo le funzioni di “Photo-Stitching” che consentono di assemblare fra loro tutte le immagini scattate, non opera la corretta conversione trigonometrica che porta ad una immagine conforme alle specifiche della Proiezione di Mercatore. Fortunatamente il centro di ricerca Microsoft ha rilasciato una applicazione gratuita, chiamata Microsoft Image Composite Editor [18] che invece opera la proiezione con estrema accuratezza.

La fig. 11 mostra tale software in azione per la ricomposizione del panorama sferico dentro un palco del teatro alla Scala di Milano.

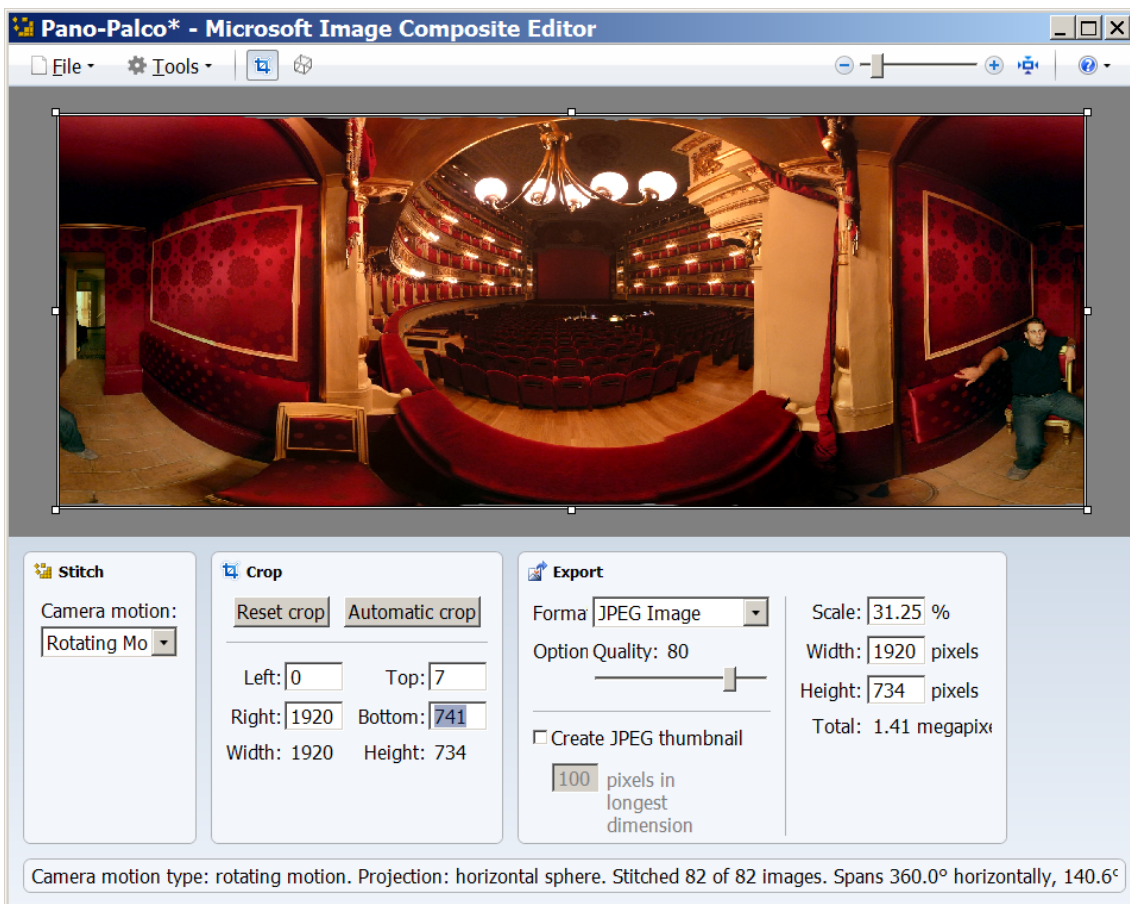


Figura 11: Programma Microsoft ICE in azione

Una volta ottenuta l’immagine di sfondo in formato 2:1, con una risoluzione tipicamente pari a 960x480 pixel, si provvede a proiettare sulla stessa gli assi di puntamento dei microfoni virtuali, come mostrato in fig. 12.

Ovviamente alcuni microfoni virtuali sono plottati due volte, in quanto la periodicità del sistema di coordinate angolari li fa apparire sia sull’estremo sinistro, che sull’estremo destro dell’immagine.

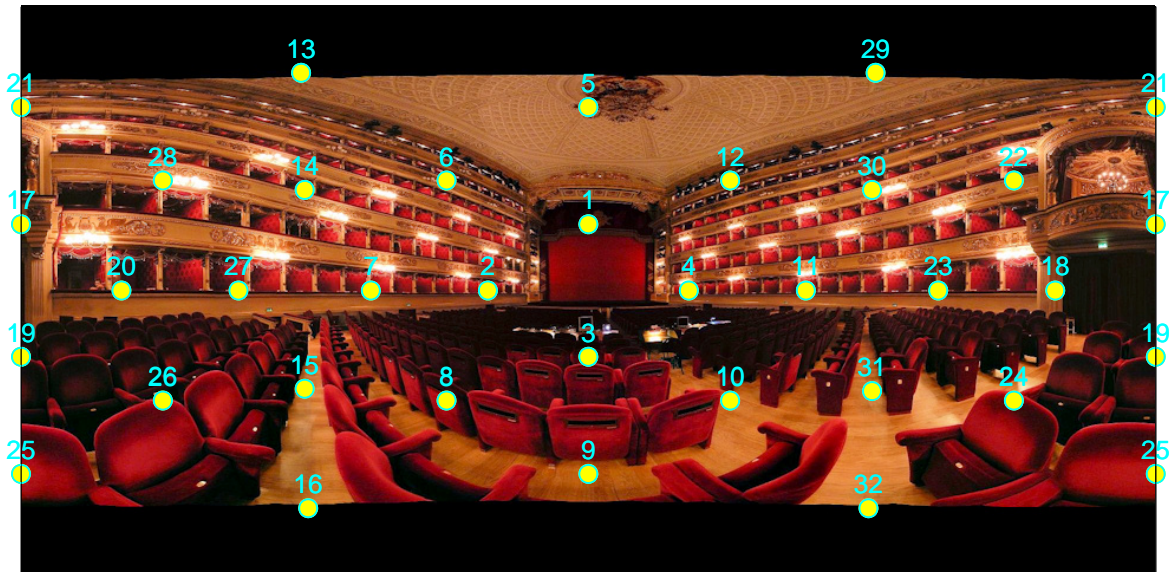


Figura 12: “copertura” del panorama sferico con i 32 microfoni virtuali

Si nota come la distribuzione dei microfoni virtuali sull’immagine sia ragionevolmente uniforme, considerando la peculiarità della proiezione di Mercatore, che tende ad allargare in modo cospicuo gli oggetti posti in vicinanza dei poli.

A questo punto, é possibile realizzare delle mappe isolivello, andando a graficare il valore di un qualunque parametro acustico computato a partire dalla risposta all’impulso “mono” di ciascun microfono virtuale.

Questa tecnica apre prospettive sinora sconosciute, ad esempio si ha la possibilità di ottenere una mappa spaziale della Chiarezza C80, o del Tempo Baricentrico  $t_s$ . Allo stato attuale della ricerca, tuttavia, é stato ritenuto prioritario effettuare una mappatura dinamica della ampiezza della risposta all’impulso, allo scorrere del tempo.

Si ottiene in questo modo non una semplice mappa statica, ma una sorta di animazione video, che mostra, al trascorrere del tempo in ms, i successivi arrivi di energia sonora sulla sonda microfonica. Si riesce in tal modo a visualizzare la provenienza del suono diretto, e delle prime riflessioni discrete. Dopo oltre 100 ms, il campo diventa diffuso, ma la sua mappatura fornisce comunque una indicazione sulla sua effettiva isotropicità.

La fig. 13 mostra l’interfaccia grafica della applicazione Matlab a tale scopo sviluppata. Si nota l’arrivo di una forte riflessione proveniente dalla parete laterale del proscenio. Una piu’ estesa discussione di questa tecnica di visualizzazione é stata recentemente presentata in [10] ed in [19],

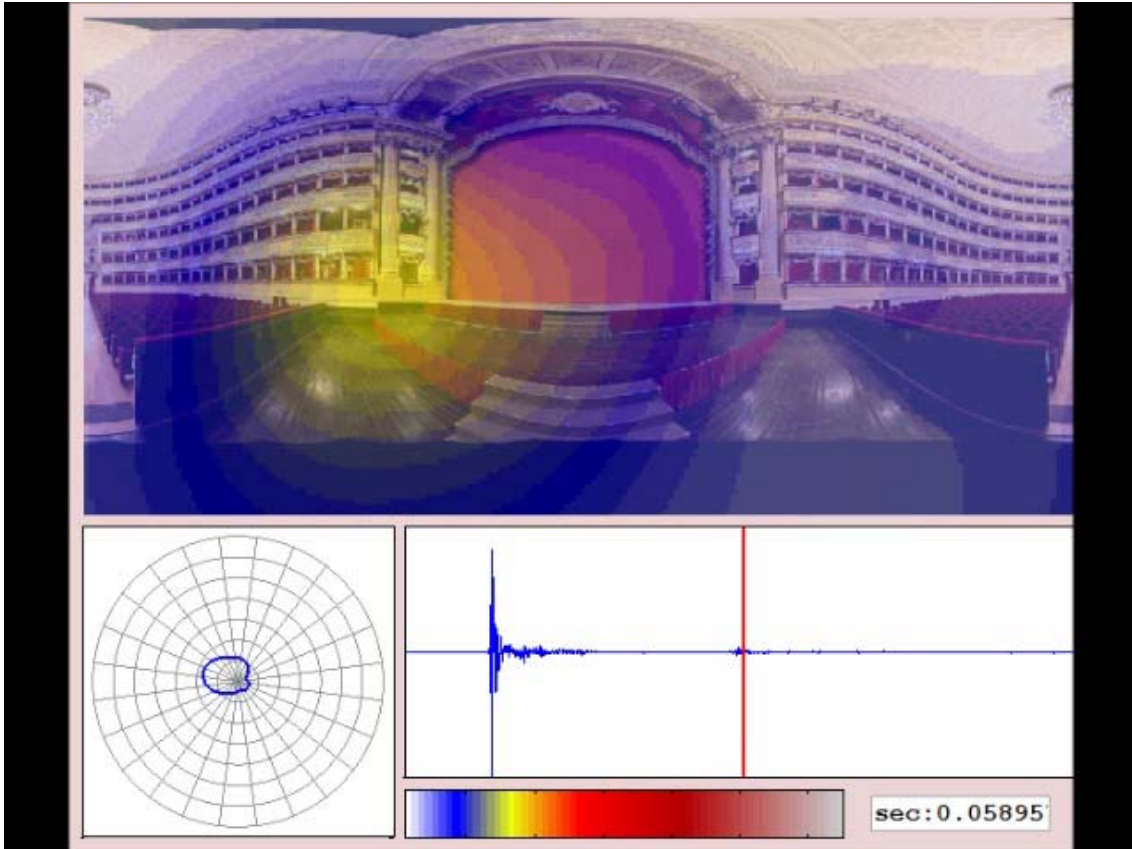


Fig. 13 – programma Matlab per la mappatura dinamica dell’energia sonora

### 7. Prima sperimentazione nel teatro alla Scala di Milano

Le misurazioni di risposte all’impulso utilizzando la strumentazione e la tecnica di elaborazione precedentemente descritte sono state effettuate per la prima volta nel Teatro alla Scala di Milano la notte del 10 maggio 2010.

L’allestimento della sala, visibile in fig. 14, era con il pavimento della fossa orchestrale sollevato, su cui era collocato un pianoforte, con il sipario tagliafuoco ed il relativo paravento rosso entrambi chiusi.



Fig. 14 – allestimento del palcoscenico durante le misure

La foto soprastante mostra il microfono nella posizione del direttore d'orchestra (pos. 7). Le altre posizioni di misura sono indicate nella mappa visibile in fig. 15.

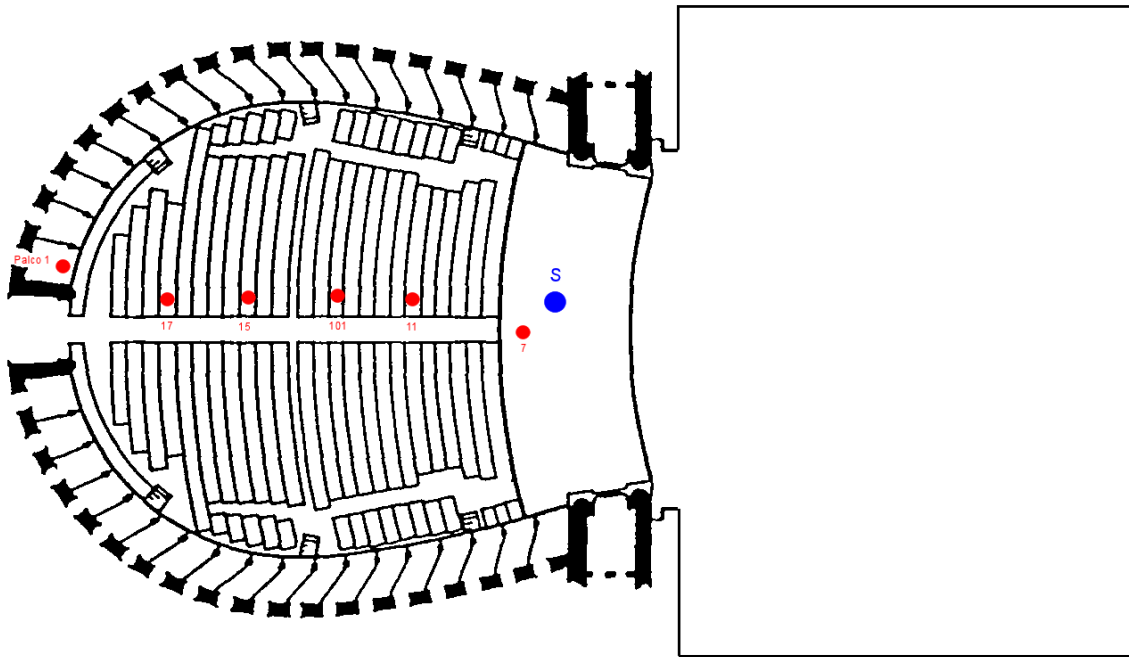


Fig. 15 – posizioni di misura con sonda microfonica Eigenmike

Nel corso della presentazione orale della presente memoria saranno visualizzati alcuni dei filmati AVI, ottenuti con la tecnica di rendering visivo descritta nel precedente capitolo, relativi ad alcuni dei punti di misura suddetti.

Per quanto riguarda i parametri acustici tradizionali misurati secondo ISO3382 dalle risposte all’impulso B-format, il confronto con i valori misurati la stessa giornata con il microfono Soundfield, e con il Bruel & Kjaer 4189 ha mostrato i seguenti risultati:

- I valori del tempo di riverberazione e degli altri parametri acustici monofonici sono risultati in generale molto simili, con la rimarchevole eccezione del parametro EDT, che invece ha mostrato significative differenze fra i tre microfoni, come mostrato dalla fig. 16.
- Si sono invece riscontrate notevoli differenze fra i valori del parametro spaziale LF, a riprova del fatto che i polar pattern del microfono Soundfield tendono ad esagerare la risposta del segnale di *particle velocity* ad alta frequenza, come mostrato in fig. 17.

Questi risultati confermano quanto l’autore aveva già verificato riguardo l’uso di diversi tipi di microfoni, quali diverse teste binaurali o diversi microfoni WY: questi parametri “oggettivi” dipendono tantissimo dal particolare sistema microfonico impiegato, ed hanno una pessima riproducibilità, anche se le misure vengono eseguite una di seguito all’altra, nella stessa sala, e con la stessa sorgente sonora [13].

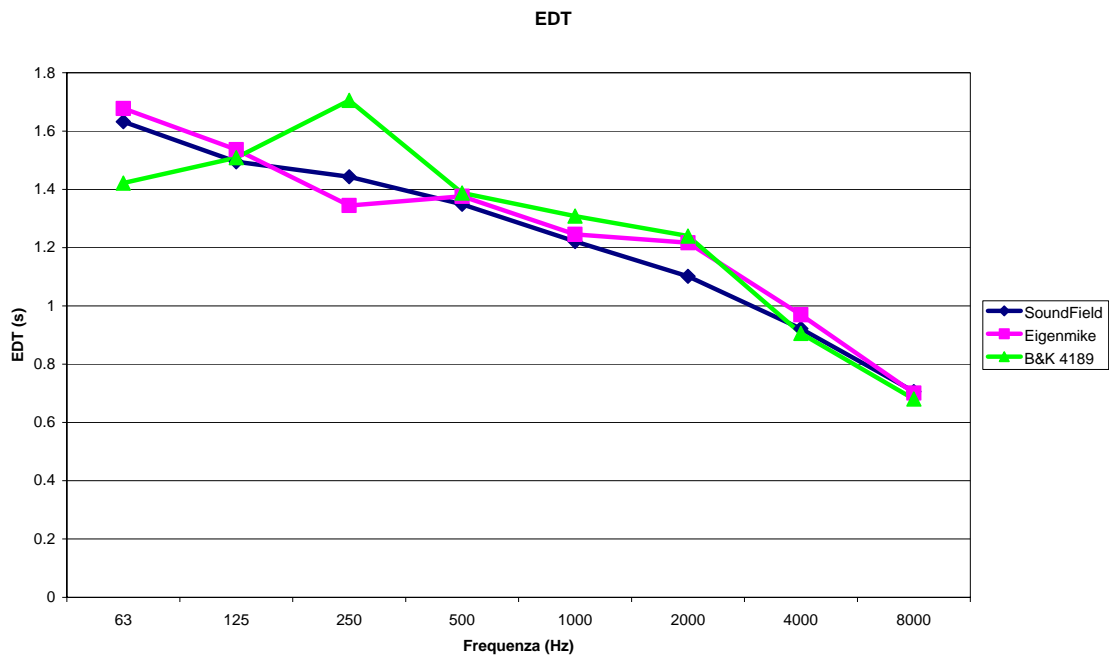


Fig. 16 – confronto valori di EDT misurati dai tre microfoni

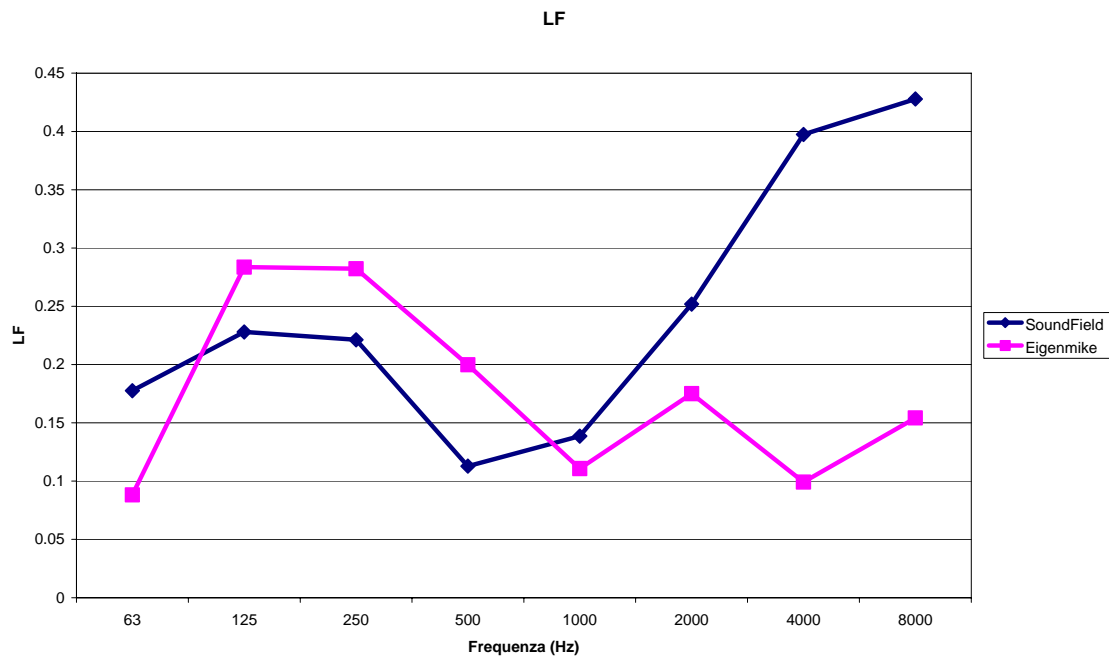


Fig. 17 – confronto valori di LF misurati dai due microfoni

## 8. Rendering acustico delle risposte all'impulso

In perfetta analogia con quanto si fa normalmente con le risposte all'impulso binaurali o B-format, ed ora anche con le risposte Ambisonics di ordine elevato, é possibile ricostruire la sensazione sonora tridimensionale di trovarsi entro il teatro storico, mediante l'operazione matematica di convoluzione.



Il problema, al giorno d’oggi, non consiste certo nel convolvere un segnale anecoico con una risposta all’impulso multicanale, seppur in questo caso dotata di ben 32 canali: il PC la esegue facilmente in tempo reale, impiegando solo una piccola frazione della potenza di calcolo disponibile.

Il fatto é che convolvendo il segnale anecoico con le risposte all’impulso misurate, ritroviamo esattamente gli stessi segnali che i microfoni avrebbero registrato se la musica o il parlato utilizzati fossero stati suonati o pronunciati nella sala. Ma questi segnali non sono direttamente utilizzabili come segnali di alimentazione per gli altoparlanti usati per la riproduzione, dovendo in generale effettuarsi una ulteriore operazione di filtraggio multicanale o matriciale, onde derivare gli opportuni segnali.

Ad esempio nel caso di riproduzione di segnali binaurali in cuffia, bastano due semplici filtri FIR per correggere la risposta in frequenza ed in fase delle cuffie impiegate, che possono essere facilmente determinati posizionando le cuffie sulla stessa testa artificiale utilizzata per le registrazioni binaurali.

Nel caso di riproduzione di un segnale binaurale su una coppia di altoparlanti, serve invece una matrice di  $2 \times 2$  filtri, in grado sia di operare la equalizzazione degli altoparlanti, sia di causare la cancellazione del crosstalk, come chiaramente spiegato in [20].

Nel caso di segnali Ambisonics, sia di primo ordine che di ordine piu’ elevato, si utilizza normalmente per la decodifica una semplice matrice di guadagni, in sostanza il segnale di ciascun altoparlante viene ottenuto come una media pesata dei segnali delle singole armoniche sferiche, a ciascuna delle quali viene applicato un opportuno guadagno, dipendente dalla posizione dell’altoparlante considerato [21,22].

In realtà questo ha mostrato gli evidenti limiti della tecnologia Ambisonics, limiti che possono venire in parte mitigati se si sostituisce un vera e propria convoluzione matriciale con opportuni filtri FIR alla semplice media pesata dei segnali di partenza, come mostrato da Nelson e Fazi [23].

Ed in effetti anche quando si parte dai segnali catturati da una sonda Eigenmike, sia “nativi” (A-format), sia già processati con la matrice di filtraggio [h], che li ha trasformati nei segnali PCM spaziali, si può procedere con la stessa metodologia matematica suggerita da Nelson e Fazi.

La matematica é di fatto identica a quella utilizzata per generare la matrice di encoding [h], ora vogliamo invece generare una matrice di decoding [f], tale per cui i segnali provenienti dal microfono (siano essi i segnali a-format, {a}, o i segnali già filtrati in formato SPS, che chiameremo {s}), passando attraverso la matrice di filtraggio [f] si trasformano nei segnali che alimentano N altoparlanti disposti tutt’attorno all’area di ascolto.

Per determinare la matrice di filtraggio [f], andremo ad imporre che i segnali emessi dagli altoparlanti, ricombinandosi al centro della zona di riproduzione, ricostruiscano esattamente la stessa distribuzione spaziale di energia sonora che era stata rilevata dalla sonda microfonica nel vero teatro in cui erano state fatte le misure.

Come mostrato in fig. 18, supponiamo di avere un sistema di riproduzione sonora dotato, ad esempio, di  $N=16$  altoparlanti, disposti come in figura.

Supponiamo di collocare al centro dello stesso, al posto della testa dell’ascoltatore la nostra sonda microfonica Eigenmike, e misurare le funzioni di trasferimento (risposte all’impulso) fra ciascuno degli altoparlanti e la sonda stessa, che chiameremo [k].

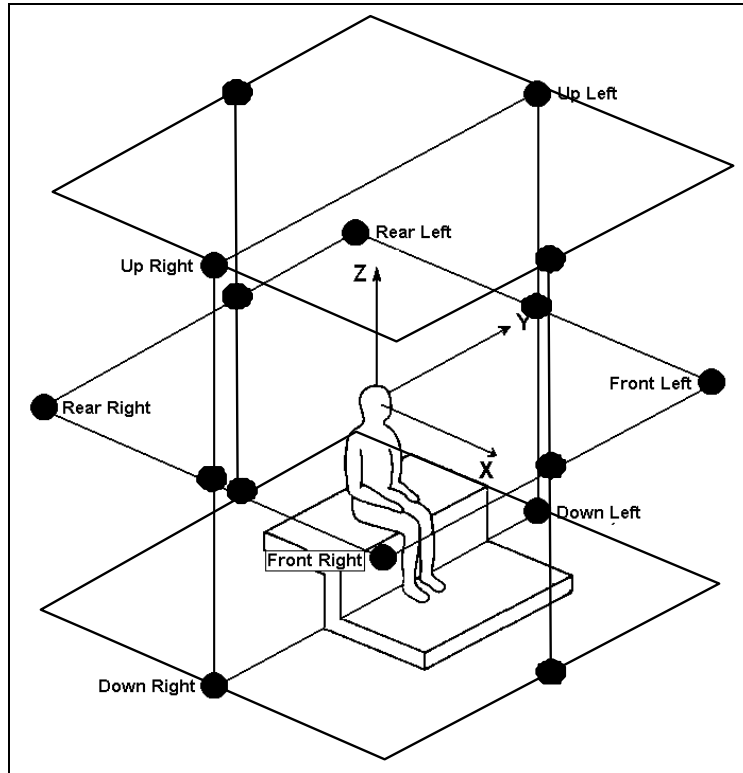


Fig. 18 – array di riproduzione sonora a 16 casse con ascoltatore al centro

Se il sistema di riproduzione sonora é perfettamente trasparente, i segnali registrati dalla sonda all'interno del sistema di riproduzione saranno identici a quelli registrati nell'ambiente di partenza, cioè nel nostro teatro storico.

Supponiamo di lavorare con i segnali SPS, anche se lavorare con i segnali A-format porta di fatto ad un trattamento matematico identico. Noi vogliamo che, qualunque fosse la distribuzione spaziale del suono campionato nel teatro, esso risulti riprodotto identico nell'ambiente di ascolto, a meno di un ritardo di riproduzione necessario per rendere causali i filtri di riproduzione.

Possiamo dunque imporre che:

$$(3) \quad \{s_{out}\} = \{s_{in}\} * [f] * [k] \quad [Hz]$$

Se vogliamo che i segnali in uscita  $s_{out}$  siano identici ad una copia ritardata dei segnali in ingresso  $s_{in}$ , occorre che il prodotto matriciale fra  $[f]$  e  $[k]$  produca una matrice identica (per meglio dire, una matrice diagonale avente come elementi della diagonale delle funzioni Delta di Dirac, che implementano il ritardo di causalità). In sostanza  $[f]$  non é altro che la matrice inversa di  $[k]$ .

Ovviamente, questa identità non potrà mai essere esattamente vera, in quanto abbiamo a disposizione solo 16 altoparlanti, dunque la matrice  $[f]$  é  $32 \times 16$ , mentre la matrice  $[k]$  é  $16 \times 32$ , e l'inversione sarà dunque sempre approssimata.

La soluzione approssimata ci viene fornita dal metodo di Kirkeby & Nelson, nella versione modificata dall'autore, che vi ha introdotto un parametro di regolarizzazione variabile con la frequenza [24,25].

La fig. 19 mostra lo schema a blocchi del processi di inversione, corrispondente alla equazione n. 3 suddetta.

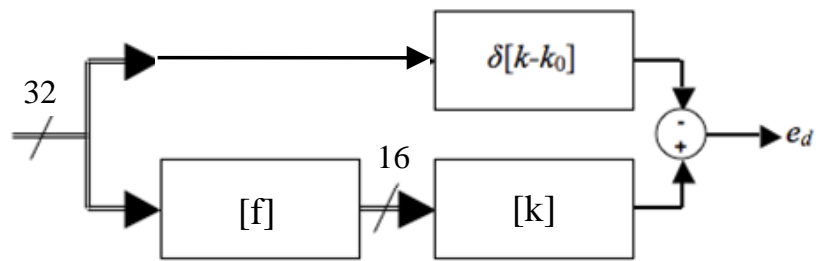


Fig. 19 – schema a blocchi della minimizzazione ai minimi quadrati

La matrice di filtraggio [f] risulterà sempre imperfetta, e quindi in uscita i segnali riprodotti (ramo inferiore del diagramma a blocchi) differiranno dai segnali di partenza ritardati (ramo superiore del diagramma a blocchi) di un errore  $e_d$ , che viene minimizzato con la tecnica dei minimi quadrati, “aggiustando” il valore di un piccolo parametro di regolarizzazione  $\beta$ , al fine di controllare la tendenza dei filtri inversi [f] a diventare molto lunghi, se si vuole un errore molto piccolo:

$$(4) \quad [F]_{16 \times 32} = \frac{[K]_{16 \times 32}^* \cdot e^{-j2\pi f}}{[K]_{16 \times 32}^* \times [K]_{32 \times 16} + \beta_f \cdot [I]_{32 \times 32}}$$

Come si nota nell’equazione 4 le variabili sono scritte con la lettera maiuscola, ad indicare che siamo passati nel dominio della frequenza tramite una trasformata discreta di Fourier.

Il parametro di regolarizzazione  $\beta$  é dipendente dalla frequenza, solitamente con un andamento simile a quello indicato in fig. 20, in modo da garantire una ricostruzione con piccolo errore nel campo di frequenze centrali, e viceversa ammettere una ricostruzione meno accurata a frequenze molto basse o molto alte, ove tale errore diventa poco rivelante dal punto di vista della percezione umana.

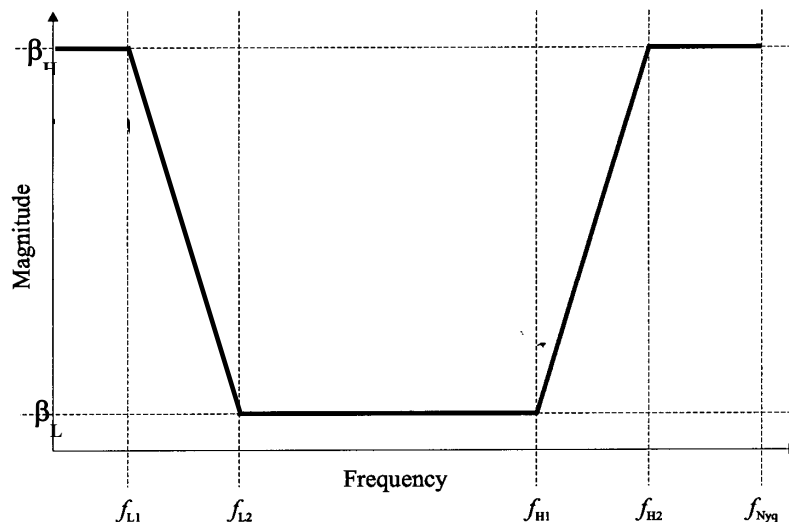


Fig. 20 – tipico andamento del parametro  $\beta$  con la frequenza

Ovviamente la effettiva invertibilità della matrice delle funzioni di trasferimento  $[k]$  è tanto maggiore quanto più la sala di ascolto è anecoica, e la disposizione degli altoparlanti è regolare e in grado di circondare completamente l'ascoltatore da tutte le possibili direzioni. In questo senso, la situazione ideale è quella realizzata da Nelson e Fazi all'ISVR (Southampton, UK), mostrata in fig. 21.



Fig. 21 – il sistema di riproduzione sonora sviluppato all'ISVR

All'Università di Parma non è disponibile una simile struttura per la ricostruzione sonora tridimensionale, tuttavia presso i laboratori di elettroacustica della Casa della Musica è stato possibile realizzare una piccola sala di ascolto dotata di 16 altoparlanti, visibile in fig. 22.

A parte questa “insignificante” differenza, il metodo sviluppato differisce da quello di Nelson e Fazi solo perché all'ISVR si utilizzano ancora, come “strato” intermedio di memorizzazione e processamento dei segnali, le armoniche sferiche di ordine elevato (sino al quarto ordine), mentre nella variante italiana i segnali vengono mantenuti nel formato SPS (Spatial PCM Sampling) descritto nel presente articolo.

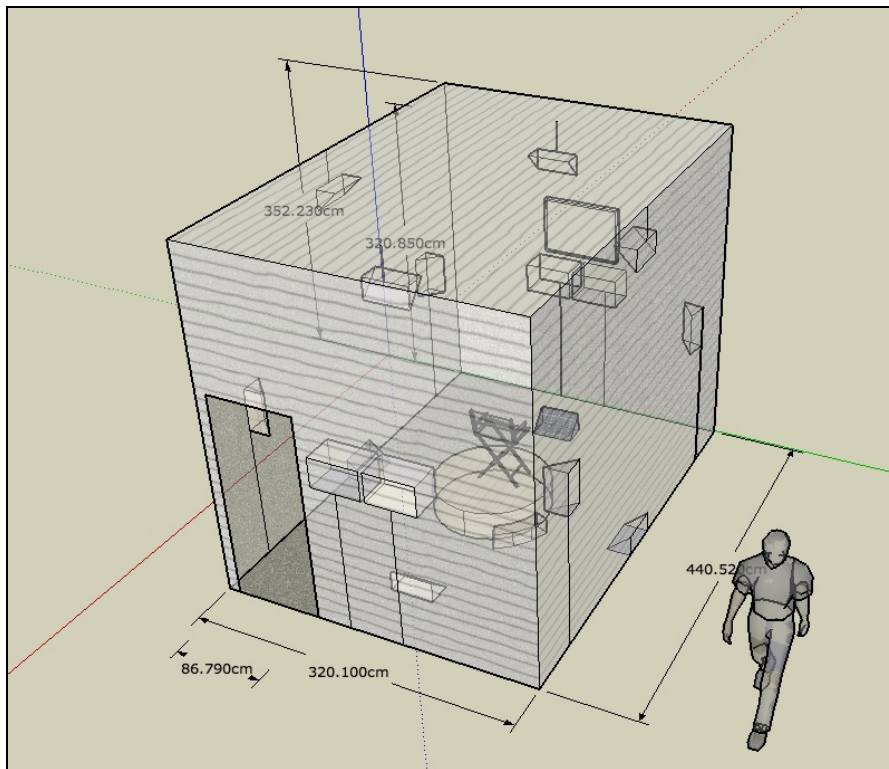


Fig. 22 – il sistema di riproduzione sonora della Casa della Musica di Parma

## 9. Conclusioni

Il presente articolo illustra una nuova metodica di campionamento del campo sonoro all'interno di un teatro mediante una rappresentazione dell'informazione spaziale denominata "Spatial PCM Sampling".

In questo articolo è stato spiegato come sia possibile operare la registrazione di riposte all'impulso multicanale, e processarle al fine di ottenere la visualizzazione grafica della distribuzione spaziale dell'energia sonora lungo l'asse del tempo, cosa che consente di identificare con accuratezza la traiettoria dei raggi sonori riflessi che giungono all'ascoltatore.

È stato infine spiegato come sia possibile ricostruire il campo sonoro originario all'interno di un ambiente di ascolto dotato di un congruo numero di altoparlanti ed adeguato trattamento fonoassorbente.

Tutte le elaborazioni numeriche del segnale necessarie si riconducono a matrici di filtri FIR, che possono venire efficacemente implementati mediante il plugin gratuito Xvolver.

La tecnica proposta si pone dunque come possibile alternativa per la codifica dell'informazione spaziale nei teatri, con significativi vantaggi rispetto alle metodiche sinora in uso (binaurale, Ambisonics del primo ordine, Ambisonics di ordine elevato).

La disponibilità di apparecchiature di costo ragionevole ed eccellenti prestazioni hanno infine reso possibile effettuare queste misurazioni anche in ambienti non particolarmente attrezzati (ad esempio operando a batteria) ed in tempi molto ridotti.

## 10. Bibliografia

- [1] L. Tronchin, A. Farina, *The acoustics of the former Teatro "La Fenice", Venice*, Journal of the Audio Engineering Society, New York, December 1997
- [2] Michael Gerzon, *Recording Concert Hall Acoustics for Posterity*, JAES Vol. 23, Number 7 p. 569 (1975)
- [3] A. Farina, R. Ayalon, *Recording concert hall acoustics for posterity*, 24th AES Conference on Multichannel Audio, Banff, Canada, 26-28 June 2003
- [4] <http://www.acoustics.net>
- [5] V. Pulkki and J. Merimaa, *Spatial Impulse Response Rendering II: Reproduction of Diffuse Sound and Listening Tests*, Journal of the Audio Engineering Society, vol 54, no. 1, 2006
- [6] Svein Berge, Natasha Barrett, *A new method for B-format to binaural transcoding*, AES 40th international conference, Tokyo, Japan, October 8–10, 2010
- [7] A. Farina; A. Capra; L. Conti; P. Martignon; F. Fazi, *Measuring spatial impulse responses in concert halls and opera houses employing a spherical microphone array*, 19th International Congress on Acoustics (ICA) - Madrid (Spain), 2-7 September 2007
- [8] A. Laborie, R. Bruno, S. Montoya, *A new comprehensive approach of surround sound recording*, 114<sup>th</sup> AES Convention, Amsterdam, 22-25 March 2003
- [9] Angelo Farina, Andrea Capra, Lorenzo Chiesi, Leonardo Scopece, *A Spherical Microphone Array For Synthesizing Virtual Directive Microphones In Live Broadcasting And In Post Production*, 40th AES Conference "Spatial Audio - Sense the Sound of Space", Tokyo, Japan, 8-10 October 2010
- [10] Angelo Farina, Alberto Amendola, Andrea Capra, Christian Varani *Spatial analysis of room impulse responses captured with a 32-capsules microphone array*, 130th AES Conference, London, 13-16 May 2011
- [11] 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 October 21-24, 2001
- [12] A. Farina *Simultaneous measurement of impulse response and distortion with a swept-sine technique*, 108th AES Convention, Paris 18-22 February 2000
- [13] A. Farina *Advancements in impulse response measurements by sine sweeps*, 122<sup>th</sup> AES Convention, Vienna, Austria, 5-8 May 2007
- [14] <http://www.adobe.com/audition>
- [15] <http://www.aurora-plugins.com>
- [16] <http://pcfarina.eng.unipr.it/Public/Xvolver/>
- [17] <http://www.audiomulch.com>
- [18] <http://research.microsoft.com/en-us/um/redmond/groups/ivm/ice/>
- [19] Angelo Farina, Lamberto Tronchin, *New Measurement Technique For 3d Sound Characterization In Theatres*, The Acoustics of Ancient Theatres Conference, Patras, Greece, September 18-21, 2011
- [20] A. Farina, E. Ugolotti, *Subjective comparison between Stereo Dipole and 3D Ambisonics surround systems for automotive applications*, 16th AES Conference, Rovaniemi (Finland), 10-12 April 1999
- [21] A. Farina, E. Ugolotti, *Software Implementation of B-Format Encoding and De-*

- coding*, Pre-prints of the 104th AES Convention, Amsterdam, 1998 May 16-19
- [22] A. Capra, S. Fontana, F. Adriaensen, A. Farina, Y. Grenier, *Listening Tests of the Localization Performance of Stereodipole and Ambisonic Systems*, 123th AES Convention - New York (USA), 5-8 October 2007
- [23] Poletti, M., Fazi, F.M. and Nelson, P.A. *Sound-field reproduction systems using fixed-directivity loudspeakers*, Journal of the Acoustical Society of America, 127, (6), 3590-3601, (2010).
- [24] Kirkeby, O., Nelson, P.A., Hamada, H., Orduna-Bustamante, F., *Fast deconvolution of multichannel systems using regularization*, IEEE Transactions on Speech and Audio, **6**, (1998).
- [25] 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 may 1999.