

# Sonda microfonica Sferica per Surround Sound

Leonardo **Scopece**  
Rai - Centro Ricerche e Innovazione Tecnologica  
Torino

## 1. INTRODUZIONE

Il desiderio di sentirsi, nel proprio salotto, *immerso* in un evento, coinvolto da immagini spettacolari e circondato da musica e suoni realistici, è sempre stata una forte motivazione per cercare di trarre il massimo vantaggio dai continui progressi in campo audiovisivo e multimediale.

L'utente già negli anni '90 era in grado di adottare sistemi di riproduzione audio per ottenere sensazioni uditive prossime a quelle percepite in una sala durante un concerto o uno spettacolo teatrale. L'*home video* nasce nel 1982, con l'introduzione del Dolby Surround. Lo spettatore ha la sensazione di essere al centro della scena, grazie alla presenza di almeno 5 diffusori, alimentati ciascuno da un canale audio, di cui due posizionati alle proprie spalle.

La qualità video disponibile trenta anni fa, alla nascita dell'*home video*, era ancora prossima, in termini di definizione, a quella disponibile alla nascita della televisione. I programmi erano a colori, in formato 4:3, e diffusi (secondo lo standard PAL) o registrati su cassetta (VHS). E' proprio a metà degli anni '80 che si sono avviate le prime sperimentazioni di HDTV: immagini, in formato 16:9, caratterizzate da un numero di elementi (pixel) cinque volte superiore rispetto a quelle della TV. La qualità video a cui si voleva tendere era già quella attuale, solo da pochi anni disponibili a casa dell'utente. La prima sperimentazione di trasmissione HDTV in

### Sommario

*L'obiettivo è fornire al telespettatore, nel suo salotto trasformato in home theatre, il suono surround: offrire un'esperienza coinvolgente anche dal punto di vista uditivo. Per ottenere tale risultato occorre agire nella fase di produzione dei programmi: la ripresa e postproduzione audio devono essere multicanale, di qualità, e contemporaneamente l'impatto sui costi e sulle modalità operative di ripresa deve essere contenuto.*

*Il progetto avviato dal Centro Ricerche della Rai, in collaborazione con l'A.I.D.A., ha consentito la messa a punto di un sistema, recentemente brevettato, che permette la ripresa multicanale utilizzando un sola sonda microfonica dotata di 32 capsule e la sintesi, sulla base della teoria Ambisonic, di un insieme di microfoni virtuali. Grazie all'uso di un joystick è possibile definire in modo semplice sia il polar pattern, sia il puntamento di ciascuno dei sette microfoni virtuali sintetizzabili.*

*E' stata avviata una serie di sperimentazioni per individuare e definire i campi di applicazione del sistema.*

formato digitale fu effettuata dalla Rai in occasione dei mondiali di calcio Italia '90. Si è però dovuto attendere vent'anni prima che la tecnologia degli schermi (Plasma, LCD...) rendesse oggi possibile, in termini di costo e ingombro, la fruizione nelle nostre case delle immagini HD e che i nuovi sistemi di diffusione digitale DVB, via satellite e terrestre, ne consentissero la ricezione domestica, grazie ai progressi in termini di efficienza nello sfruttamento della capacità dei canali.

Da un decennio l'evoluzione tecnologica sembra aver assunto un ritmo sempre più rapido e il prossimo passo verso l'immersione nella realtà virtuale televisiva sarà possibile con la diffusione di programmi tridimensionali: l'avvento della televisione stereoscopica o 3D TV sembra prossima.

Apparentemente tali progressi non sembrano altrettanto rapidi e incisivi in campo audio.

Di fatto l'utente può allestire nel proprio soggiorno un sistema di riproduzione adeguato, sia per quanto riguarda le immagini, sia il suono: un grande schermo e un sistema di diffusione surround. E' l'*home theatre*: il coinvolgimento totale del telespettatore nella scena. Ma spesso alle immagini in alta definizione non si accompagna un audio di qualità altrettanto spettacolare.

Occorre agire su tutta la catena, partendo dal primo stadio, quello della produzione: i sistemi di ripresa video e postproduzione devono operare con segnali video in alta definizione, e i sistemi di ripresa e postproduzione audio devono essere in grado di codificare segnali surround.

Con l'obiettivo di individuare sistemi efficaci per produrre programmi dotati di surround, il Centro Ricerche Rai ha avviato tre anni fa un ampio programma di sperimentazioni. Lo scopo è la messa a punto di un sistema che consenta di effettuare riprese sonore e registrazioni con sistemi multicanale, utilizzando le infrastrutture già esistenti, con costi contenuti, per offrire all'utente, nel modo più semplice dal punto di vista operativo, un segnale surround a complemento di quello video in alta definizione.

## 2. DALLA SPERIMENTAZIONE ALL'INNOVAZIONE

Precedenti articoli hanno illustrato le tecniche di ripresa stereofonica e quella olofonica [1] e le sperimentazioni effettuate [2, 3].

Per effettuare tali sperimentazioni, in ambienti differenti e diverse tipologie di sorgenti sonore, si è utilizzato un microfono già adottato da alcune emittenti del Nord America e sperimentato da celebri artisti: l'Holophone H2Pro.

Il microfono Holophone è dotato di 8 capsule, sette localizzate sulla periferia del supporto ed una all'interno, per l'acquisizione delle basse frequenze. Gli otto segnali in uscita dalla "testa" sono distinti fra loro: gli otto canali possono essere avviati ad un mixer o codificati come audio 5.1, 6.1, 7.1 o semplicemente stereo.

La tecnica di ripresa olofonica si aggiunge alle tradizionali tecniche di ripresa, che presentano specifiche caratteristiche del suono riprodotto:

♪ con la tecnica multimicrofonica il suono è tutto presente, a fuoco, pulito, ma senza profondità, un "muro" di suoni provenienti dagli strumenti disposti sulla scena.

♪ con le tecniche stereofoniche, dall'MS, alla Stereosonic, all'AB, alla XY, il suono comincia ad essere "profondo".

Le nuove tecniche multicanale, come quella adottata da Holophone, offrono un suono avvolgente, più naturale, meno "a fuoco" ma più reale rispetto a quello ottenibile con le tecniche fino ad ora utilizzate.

I risultati ottenuti con le sperimentazioni citate sono molto interessanti e hanno consentito di comprendere che il "modo di sentire" dei prodotti ottenuti con tecniche di ripresa multicanale cambia.

Resta ora da valutare quanto differisca la percezione da parte del pubblico nel passaggio dal modo di sentire fin qui offerto, artificioso, a confronto con il nuovo modo di sentire, più avvolgente e più reale, anche se meno "presente".

### 3. IL SISTEMA BASATO SU SONDA MICROFONICA H.O.A.

In base alle conoscenze acquisite e avendo valutato positivi i vantaggi della ripresa multicanale, il Centro Ricerche ha deciso di avviare un progetto per realizzare un sistema con caratteristiche innovative, basato sulla tecnica Ambisonic (vedere Appendice).

Questo progetto ha portato allo sviluppo a al brevetto di un sistema di ripresa e registrazione basato sulla Sonda Microfonica HOA (High Order Ambisonics) da parte di Rai e A.I.D.A., spin-off dell'Università di Parma.

Lo scopo del sistema è di riprendere la scena sonora ottenendo tre risultati:

- 🎵 una ripresa multimicrofonica o, in alternativa, surround;
- 🎵 poter effettuare uno zoom microfonico, ovvero cambiare la direttività, in modo dinamico e in tempo reale;
- 🎵 poter posizionare i punti di ripresa nello spazio sia sul piano azimutale che mediano sempre in modo dinamico e in tempo reale.

#### 3.1 LA SONDA H.O.A. E L'INTERFACCIA

La sonda microfonica (figura 1) è costituita da una sfera di 8,4 centimetri di diametro ed ha sulla superficie 32 capsule da 1/2" ad elettrete di ottima qualità.

Le capsule hanno una risposta in frequenza che va da 20 Hz a 20 kHz, con sensibilità di 32 mV/Pa e una dinamica prima del clipping di 135 dB.

All'interno della sfera stessa sono presenti i preamplificatori delle 32 capsule e i convertitori A/D.

I 32 segnali digitali sono multiplati e trasferiti, mediante un cavo ethernet di categoria 5 o superiore che può essere lungo fino a 140 metri, ad una interfaccia (figure 2) in cui possono essere memorizzati su un dispositivo USB per elaborazioni successive oppure trasferiti alla unità di elaborazione.

#### 3.2 UNITÀ DI ELABORAZIONE

Costituisce il cuore del sistema. Provvede all'elaborazione dell'audio, effettuando in tempo reale i complessi calcoli per l'emulazione dei filtri che consentono di sintetizzare i segnali corrispondenti a sette microfoni virtuali.

Ognuno di questi microfoni virtuali ha un tempo di latenza sufficientemente basso da non inficiare la ripresa, ed una buona qualità su tutte le frequenze della banda base audio. Inoltre, ciascun microfono è sintetizzato in base a tutti i 32 segnali componenti, prelevati dalle singole capsule.

Il polar pattern può variare da bidirezionale, a omnidirezionale, a cardioide di vario ordine. Gli algoritmi utilizzati sfruttano il principio Ambisonic: l'informazione sonora in un punto dello spazio è ottenuta conoscendo la pressione e la velocità delle particelle in quel punto.

Attualmente Il sistema realizzato consente, per ognuno dei 7 microfoni sintetizzati, di elaborare in tempo reale le informazioni ottenute dalle 32 capsule e realizzare la sintesi di un microfono virtuale equivalente Ambisonic fino al sesto ordine.



Fig. 1 - Sonda Microfonica H.O.A.

Foto Enrico Cavallini

Fig. 2 - Interfaccia audio



Foto Enrico Cavallini

### 3.3 PUNTAMENTO E JOYSTICK

Obiettivo del sistema sono semplicità operativa e ergonomia.

La semplicità operativa è ottenuta grazie all'interfacciamento del sistema di elaborazione con un joystick (figura 3) e con una telecamera wide-angle che permette di controllare il "puntamento" dei singoli microfoni virtuali.

Grazie a tali dispositivi è possibile definire in modo semplice sia il diagramma polare, sia il puntamento di ciascuno dei microfoni virtuali, indipendentemente l'uno dall'altro.

Con il joystick è possibile selezionare ciascun microfono, "inseguire" dinamicamente il soggetto sulla scena e decidere la direttività del microfono virtuale selezionato, visualizzando il corrispondente beam con un cerchio colorato, le cui dimensioni forniscono in modo intuitivo le sue caratteristiche, relative all'angolo di ripresa.

Le informazioni relative ai singoli microfoni virtuali utilizzati sono visualizzate su un monitor, sovrapposte all'immagine ripresa dalla telecamera (figura 4). E' quindi possibile decidere in tempo reale quale diagramma polare adottare per ciascun microfono, e la loro posizione spaziale.

Nel caso la scena sia statica o quasi, può essere utile usare, invece di una telecamera, una macchina fotografica e realizzare una fotografia panoramica, me-

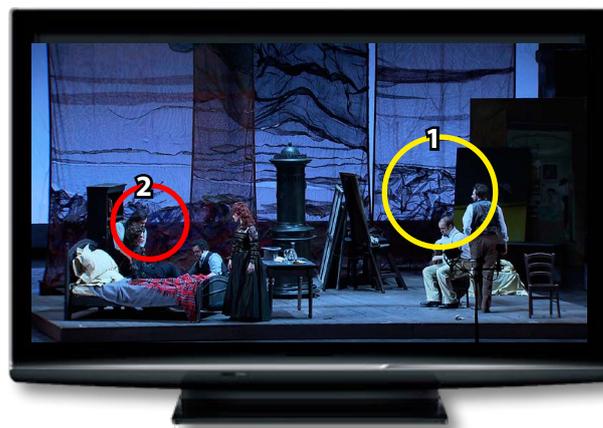


Fig. 5 - Esempio di visualizzazione del diagramma polare relativo a due dei microfoni virtuali sovrapposti all'immagine della scena.

morizzare le posizioni dei soggetti presenti in scena senza "inseguirli" potendoli richiamare facilmente, semplificando così enormemente il lavoro (figura 6).

### 3.4 LA CONSOLE DI CONTROLLO

Per la gestione del sistema si utilizza una console di controllo, un notebook. Sulla sua memoria di massa possono essere inoltre registrati in tempo reale i 32 segnali provenienti dalla sonda e/o i 7 segnali sintetizzati corrispondenti ai microfoni virtuali.

L'applicativo di gestione del sistema consente di avviare i sette segnali in uscita dal sistema di elaborazione allo spazio di ascolto oppure ad un mixer, tramite la porta ADAT, o di memorizzarli su un registratore audio esterno.

I segnali memorizzati possono ovviamente subire un processo di postproduzione. In particolare i segnali corrispondenti ai 7 microfoni virtuali possono essere sintetizzati e modificati in postproduzione, a partire dall'insieme dei 32 segnali catturati grazie alla sonda H.O.A.

## 4. APPLICAZIONE DEL SISTEMA BASATO SU SONDA H.O.A.

La gamma di applicazioni che possono trarre vantaggio dall'uso del sistema sviluppato e brevettato è molto ampia: si ipotizzano qui alcuni esempi,



Fig. 3 - Joystick

Foto Enrico Cavallini

corrispondenti a diverse condizioni di ripresa. E' in corso la sperimentazione, che deve coinvolgere i possibili utilizzatori, per verificarne la fattibilità e la flessibilità di impiego del sistema. E' così possibile trarre indicazioni utili per il passaggio dalla realizzazione prototipale a quella utilizzabile in produzione.

#### 4.1 RIPRESA DI ORCHESTRA

Nel caso di una orchestra sinfonica, è possibile simulare una ripresa multimicrofonica, puntando, ad esempio, 6 dei 7 microfoni virtuali verso punti stabiliti dell'ensemble e controllando dinamicamente un settimo soggetto, canoro o musicale, che non è statico sulla scena.

Un altro esempio è la realizzazione di una ripresa codificata fino a 7.0 puntando 5 dei 7 microfoni virtuali frontalmente ed i rimanenti 2 posteriormente.

Segnali codificati 5.1, 6.1 o 7.1 possono essere ottenuti elaborando i 7 segnali a disposizione e ricavando da essi il contributo a bassissime frequenze LFE (Low Frequency Effect), oppure utilizzare un microfono aggiuntivo adibito alla ripresa di frequenze da 20 a 110 Hz.

#### 4.2 RIPRESA DI EVENTO SPORTIVO

Allo stadio di calcio, si può ottenere un effetto ambiente in surround. E' pensabile che sia possibile, posizionando la sonda sul terreno di gioco a bordo campo, riprendere anche l'impatto del colpo sul pallone o di questo contro un "legno".

Durante una gara ciclistica, riprendendo la folla che "corre" ai lati degli atleti, si può ottenere un effetto sonoro che "viene incontro" e poi si "smorza" alle spalle.

#### 4.3 RIPRESA DI EVENTO TEATRALE

Analogamente al caso di ripresa di orchestra, anche in teatro è pensabile di puntare 5 o 6 microfoni virtuali in modo fisso e controllare i restanti con uno o due joystick: è quindi realizzabile una ripresa surround, codificata da 5.0 a 7.0.

#### 4.4 RIPRESA DI EVENTO TELEVISIVO

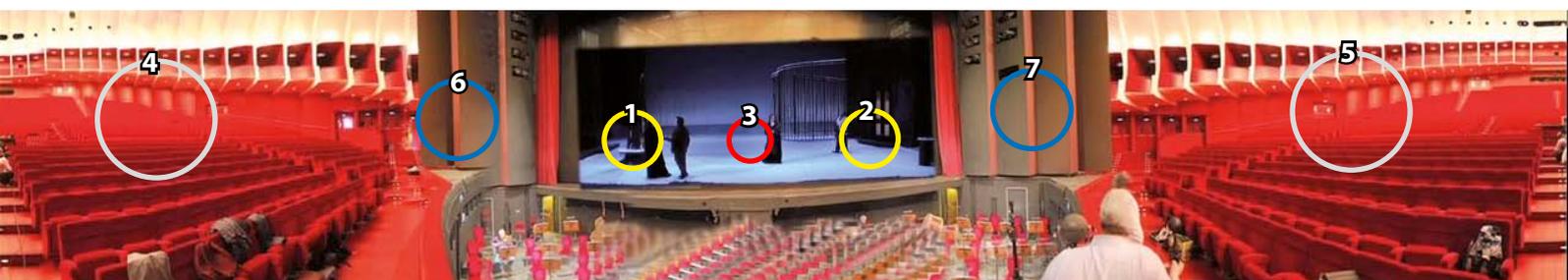
In un programma televisivo in studio, si possono porre una o due sonde in alto indirizzate verso il basso.

"Puntando" i microfoni virtuali corrispondenti a ciascuna sonda con una direttività del quinto o del sesto ordine si possono attenuare notevolmente i rumori ambientali normalmente presenti nello studio televisivo e provenienti dall'alto, quali quelli dovuti agli impianti di condizionamento e di movimento motorizzato delle luci.

#### 4.5 RIPRESA DI EVENTO RADIOFONICO

Negli studi radiofonici si potrebbero effettuare riprese di spazialità che renderebbe l'utente più partecipe di quello che sta ascoltando, avendo la sensazione di "vedere" in uno spazio a 360° la posizione degli attori, speaker e musicisti presenti.

Fig. 6 - La scelta dei diagrammi polari e il puntamento per ciascuno dei 7 microfoni virtuali può essere effettuata e memorizzata con l'ausilio di una foto panoramica realizzata quando la sala è vuota o durante le prove. Nell'esempio un microfono del sesto ordine punta al centro della scena (3), due microfoni del 3° ordine (1 e 2) a  $\pm 30^\circ$ , del 2° ordine (6 e 7) a  $\pm 60^\circ$  e del 1° ordine (4 e 5) a  $\pm 110^\circ$ . Tale scelta può essere modificata nel corso della registrazione, o in postproduzione, selezionando ciascun microfono virtuale mediante il joystick.



#### 4.6 RIPRESA DI TALK-SHOW

L'inseguimento dinamico reso possibile dal joystick può consentire di riprendere, ad esempio nel corso di un talk-show, il giornalista che si muove nello studio e si avvicina agli ospiti, eventualmente "puntati" in modo statico dai restanti 6 microfoni, per interloquire con loro.

Ponendo la sonda in alto, sulla platea, è possibile "puntare" con il joystick i singoli componenti del pubblico a cui si desidera dare la parola.

#### 4.7 RIPRESA DI UNA SCENA E RIPRODUZIONE DEL RISULTATO IN UN ALTRO AMBIENTE

Vi possono essere limitazioni nell'uso del sistema nel caso in cui si voglia riprendere ed irradiare il suono nello stesso ambiente: infatti, a causa dell'alta sensibilità dei microfoni della sonda, è alta la probabilità di innesco (effetto Larsen).

Può invece trovare ampie applicazioni l'uso del sistema per la ripresa dell'evento e la sua riproduzione in tempo reale in uno o più ambienti, anche distanti, garantendo l'effetto presenza.

#### 4.8 POSTPRODUZIONE

Si è detto che le informazioni sonore delle 32 capsule possono essere registrate ed essere quindi disponibili per nuove rielaborazioni. Ciò apre possibilità fino ad ora impensabili a quanto realizzabile in fase di postproduzione.

Sono ridefinibili, in postproduzione, le scelte che normalmente devono essere attuate in fase di ripresa. E' infatti possibile riposizionare i microfoni virtuali nello spazio scenico e ridefinirne la direzionalità, fino a raggiungere il risultato atteso.

La limitazione a 7 microfoni virtuali è legata alla attuale capacità di elaborazione in tempo reale e in futuro potrà aumentare; è già oggi possibile incrementare il numero di microfoni virtuali nella fase di postproduzione, ovviamente rinunciando alla possibilità di operare in tempo reale.

### 5. I PROSSIMI PASSI

I risultati fino ad ora ottenuti sono molto promettenti. Le sperimentazioni in atto sono effettuate con un sistema prototipale: alcuni degli elementi componenti (interfaccia, unità di elaborazione, console) potranno essere integrati favorendo una maggiore trasportabilità e facilità d'uso.

La verifica, coinvolgendo gli utilizzatori, delle varie ipotesi di impiego, in situazioni ambientali differenti e con varie tipologie di generazione sonora, è indispensabile per individuare un sistema utilizzabile nella produzione radiotelevisiva, obiettivo della attività di ricerca e sviluppo avviato dal Centro Ricerche Rai.

Il coinvolgimento dell'industria è essenziale per sviluppare un prodotto che, in base alle indicazioni fino qui ottenute, sembra avere notevoli potenzialità di impiego e diffusione.

#### BIBLIOGRAFIA

1. L. Scopece: "Olofonia - una ripresa sonora di tutto ciò ci circonda", Elettronica e Telecomunicazioni, Agosto 2007.
2. L. Scopece: "OLOFONIA - Sperimentazioni con microfono Holophone H2-PRO", Elettronica e Telecomunicazioni, Agosto 2008.
3. L. Scopece: "Ripresa Audio Multicanale", Elettronica e Telecomunicazioni, Agosto 2009.

#### APPENDICE

##### LA TEORIA AMBISONIC

L'Ambisonic è un metodo di registrazione, o campionamento, e riproduzione delle informazioni inerenti ad un determinato campo sonoro.

Conoscendo infatti pressione e velocità delle particelle d'aria in un determinato punto dello spazio, si posseggono tutte le informazioni necessarie per ricreare in modo fedele il campo sonoro. E' possi-

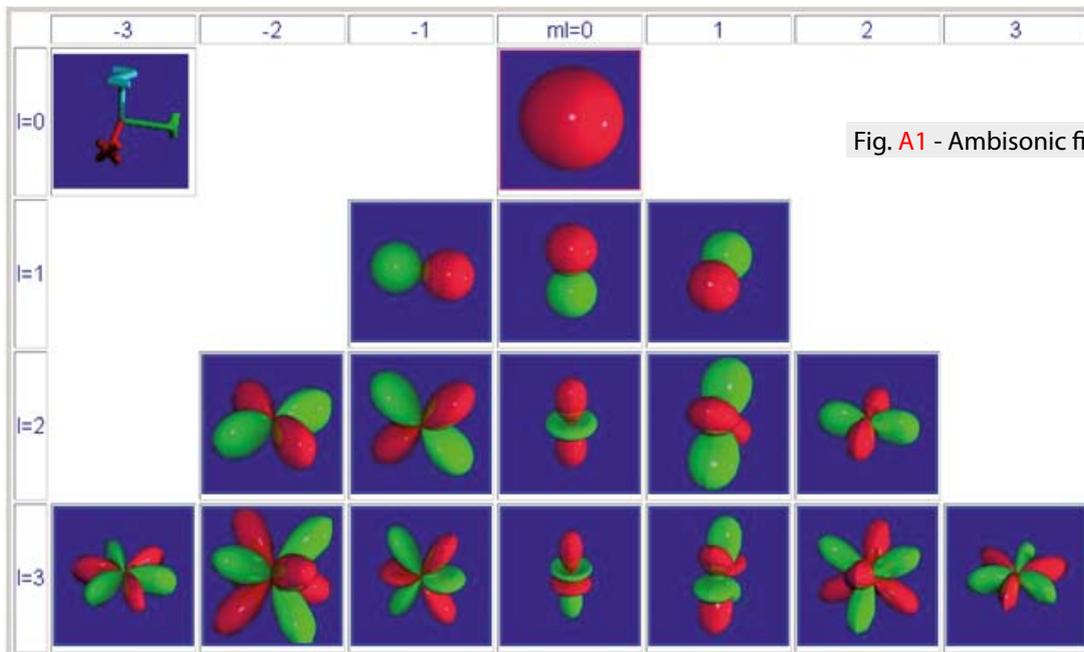


Fig. A1 - Ambisonic fino al 3°ordine

bile sfruttare questi due parametri in due principali direzioni: sintesi di microfoni o riproduzione audio surround.

Negli anni '70 Michael Gerzon realizzò la prima sonda in grado di campionare pressione e velocità in un punto, sfruttando una disposizione tetraedrica di capsule microfoniche con direttività a cardioide.

Sfruttando tre microfoni a figura di otto e un microfono omnidirezionale, teoricamente coincidenti, si potrebbe pensare di campionare le armoniche sferiche (funzioni direzionali) di ordine 0 e ordine 1: il segnale omnidirezionale dà le informazioni sulla pressione (W), i tre segnali rimanenti, orientati secondo gli assi cartesiani, sono proporzionali alla velocità delle particelle d'aria attorno al punto d'origine

Per ogni ordine "m" ambisonic ci sono (2m+1) componenti (figura A1).

Nell'impossibilità di avere a disposizione 4 microfoni coincidenti, Gerzon pensò di utilizzare 4 capsule con direttività a cardioide poste sulle facce di un tetraedro. La conversione dei segnali delle capsule (A-format) in quelli Ambisonics del 1° ordine (B-format) avviene tramite una semplice serie di somme e sot-

trazioni ottenendo una serie di segnali virtualmente provenienti da altre direzioni (figura A2):

$$\begin{aligned} W &= C_1 + C_2 + C_3 + C_4 \\ X &= C_1 + C_2 - C_3 - C_4 \\ Y &= C_1 - C_2 + C_3 - C_4 \\ Z &= C_1 - C_2 - C_3 + C_4 \end{aligned}$$

Le capsule, anche in questo caso, non sono coincidenti: l'approssimazione introdotta si traduce nel disallineamento tra le capsule in termini di tempo/fase che porta alla "colorazione" (filtraggio) dello spettro dei segnali B-format.

Fig. A2 - Microfono tetraedrico

