

2.8 Bibliografia

1. B. Bartlett, "Stereo Microphone Technique," *db*, Vol. 13, No. 12, 1979 (Dec.), pp.34+46.
2. J. Eargle, "Stereo Microphone Techniques," dell'opera *The Microphone Handbook*, Elar Publishing, Plainview, NY, 1981, Chapter 10.
3. R. Streicher and W. Dooley, "Basic Stereo Microphone Perspectives-A Review," *Journal of the Audio Engineering Society*, Vol. 33, No. 7/8, 1984 (July/Aug.), pp.548+556.
4. David Griesinger, "Spaciousness and Localization in Listening Rooms and Their Effects on the Recording Technique," *J. Audio Eng. Soc.*, Vol. 34, No. 4, 1986 (April), pp. 255+268.
5. S. Lipshitz, "Stereo Microphone Techniques: Are the Purists Wrong?" *J. Audio Eng. Soc.*, Vol. 34, No. 9, 1986 (September), pp. 716+744.

3

Teoria dell'immagine stereofonica

Un valido sistema audio con buona definizione stereo è in grado di formare sorgenti sonore apparenti; può ad esempio riprodurre gli strumenti musicali in un posizionamento ben definito, di solito fra una coppia di monitor messi di fronte all'ascoltatore. Queste sorgenti sonore apparenti sono dette *immagini*.

Questo capitolo spiega:

- i termini correlati all'immagine stereofonica;
- come si localizzano le sorgenti sonore reali;
- come si localizzano le immagini fra i monitor;
- in quale modo il posizionamento microfonico controlla la localizzazione delle immagini.

Si possono usare le tecniche stereo anche senza leggere questo capitolo. Comunque, se si vuole comprendere l'evoluzione in corso e sviluppare una propria tecnica stereofonica, è bene studiare la teoria e capire le semplici formule matematiche riportate di seguito.

3.1 Definizioni

Innanzitutto si definiscono alcuni termini correlati con l'immagine stereofonica.

La fusione si riferisce alla sintesi di una singola sorgente sonora apparente (immagini o immagine fantasma) da due o più sorgenti reali (per esempio gli altoparlanti).

La collocazione di una immagine è la sua posizione angolare rispetto ad un punto perfettamente perpendicolare di fronte all'ascoltatore o la sua posizione rispetto ai monitor (fig. 3.1). Scopo dell'alta fedeltà è quello di riprodurre le immagini sonore nella posizione voluta dal tecnico del suono o dal produttore. In alcune produzioni, solitamente le registrazioni di musica classica, un obiettivo del tecnico del suono o del produttore è quello di collocare le immagini sonore nelle stesse posizioni relative in cui erano al momento della registrazione dal vero.

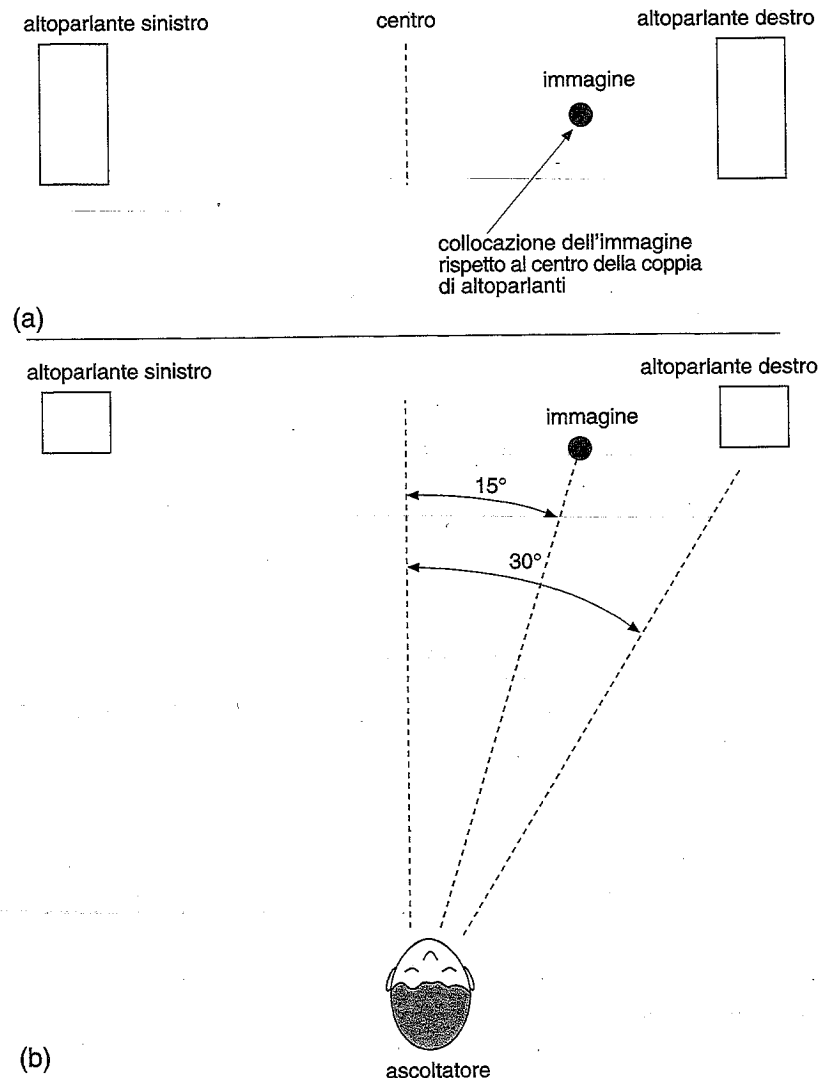


Fig. 3.1 Esempio di collocazione dell'immagine: (a) vista dall'ascoltatore; (b) vista dall'alto.

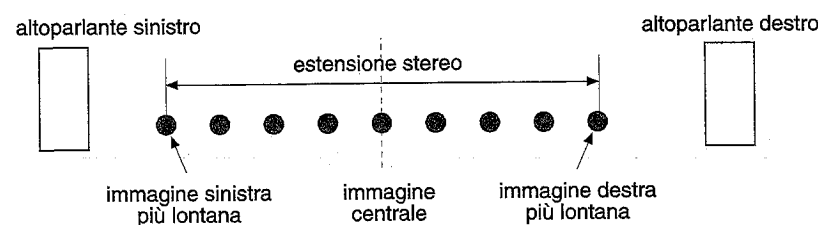


Fig. 3.2 Estensione o ampiezza stereo.

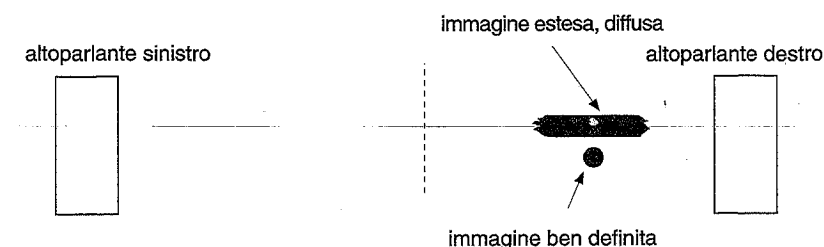


Fig. 3.3 Definizione o dimensione dell'immagine (percezione dell'ascoltatore).

L'estensione stereo, o larghezza del campo (fig. 3.2), è la distanza fra le immagini all'estrema destra e all'estrema sinistra di un insieme di strumenti riprodotto.

L'estensione stereo è grande se l'insieme sembra estendersi completamente fra una coppia di monitor; viceversa l'estensione è limitata se l'insieme occupa solo una parte dello spazio fra i monitor. A volte la riverberazione o l'ambiente riprodotti si estendono da monitor a monitor anche se la larghezza dell'insieme risultante è minore.

La focalizzazione, o dimensione dell'immagine (fig. 3.3), si riferisce al grado di fusione di una immagine, o alla definizione della sua posizione.

Una immagine altamente definita può essere descritta come precisa, definita, ad alta risoluzione e facile da localizzare. Al contrario, quella poco definita è difficile da localizzare, estesa, diffusa e vaga. Una immagine naturale è focalizzata allo stesso modo dello strumento riprodotto.

La profondità è la distanza apparente di una immagine dall'ascoltatore e dà il senso di vicinanza e di lontananza dei vari strumenti.

L'elevazione è un posizionamento dell'immagine in altezza al di sopra del piano dei monitor.

Il movimento dell'immagine è la riproduzione del movimento della sorgente sonora, se questo avviene. Le immagini non dovrebbero muoversi senza motivo.

La localizzazione è l'abilità di un ascoltatore nell'individuare la direzione di un suono.

È anche la relazione delle differenze interaurali, o fra i canali, e la posizione dell'immagine percepita.

3.2 Come localizzare le sorgenti sonore reali

Il sistema uditivo umano usa il suono diretto e le prime riflessioni per localizzare le sorgenti sonore.

Il suono diretto e le riflessioni entro 2 millisecondi contribuiscono alla localizzazione. Le riflessioni che si verificano fra 5 e 35 millisecondi dopo il suono diretto influenzano la spazialità dell'immagine.

Gli echi aventi un ritardo superiore a 5 millisecondi e fino a 50 (a seconda dei materiali riflettenti) non si fondono in tempo con le prime riflessioni, ma contribuiscono a far percepire lo spettro tonale.

Immaginiamo una sorgente sonora e un ascoltatore. Supponiamo che la sorgente sia di fronte all'ascoltatore, spostata a sinistra (fig. 3.4). Il suono percorre una distanza maggiore per arrivare all'orecchio destro rispetto a quello sinistro, giungendo dopo al destro; in altre parole, il segnale all'orecchio destro è in ritardo rispetto al segnale all'orecchio sinistro.

Ad ogni localizzazione della sorgente corrisponde un'unica differenza dei tempi di arrivo alle orecchie. Inoltre la testa funziona come un ostacolo per i suoni al di sopra di 1000 Hz circa.

Le frequenze alte sono oscurate dalla testa, per cui a ciascun orecchio si presenta uno spettro diverso (ampiezza in rapporto alla frequenza). Ogni localizzazione di sorgente produce una sola differenza di spettro fra le orecchie (fig. 3.5).

Abbiamo imparato ad associare alcune differenze fra le orecchie (cioè interaurali) con direzioni specifiche delle sorgenti sonore. Quando dobbiamo localizzare una nuova sorgente, confrontiamo ciò che ascoltiamo con un modello memorizzato di una situazione simile per determinare la direzione.

Come abbiamo già visto, un ruolo importante nella localizzazione è affidato alla differenza dei tempi di arrivo interaurali dell'involuppo del segnale. Noi percepiamo questa differenza ad ogni cambiamento del suono: un transiente, una pausa o una variazione nel timbro.

Perciò, localizziamo i passaggi (transienti) più facilmente che il suono continuo. La differenza di tempo fra i segnali alle orecchie può essere considerata anche come una differenza di fase tra le onde sonore in arrivo (fig. 3.6); questa differenza di fase aumenta con la frequenza.

Quando le onde sonore di una sorgente reale raggiungono la testa di un ascoltatore, ad ogni orecchio si presenta un diverso spettro di ampiezza e di fase.

Queste differenze interaurali sono percepite dalla mente come un'apprezzabile direzionalità della sorgente sonora. Ogni direzione è associata con un certo numero di differenze interaurali.

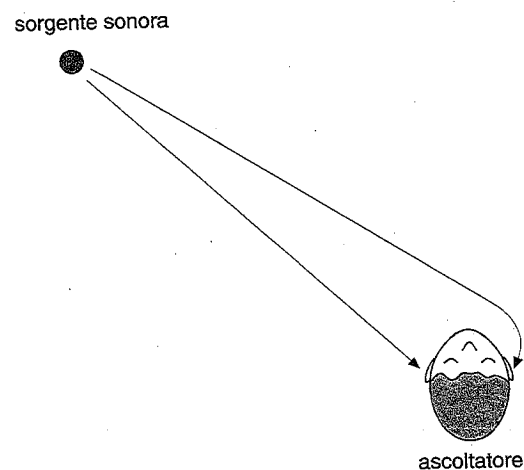


Fig. 3.4 Percorso del suono da una sorgente fino all'orecchio dell'ascoltatore.

Le orecchie usano le differenze interaurali di fase per localizzare i suoni con frequenza comprese fra 100 e 700 Hz. Le frequenze inferiori a 100 Hz non vengono localizzate (rendendo fattibile un sistema di amplificazione "subwoofer/satellitare"). Attorno ai 1500 Hz circa, le differenze di ampiezza interaurali contribuiscono alla localizzazione; fra 700 e 1500 Hz circa, sono usate sia le differenze di fase che quelle di ampiezza per individuare la direzione del suono.

Piccoli movimenti della testa cambiano le differenze dei tempi di arrivo alle due orecchie.

La mente usa queste informazioni come ulteriori suggerimenti per la localizzazione, specialmente per la discriminazione delle distanze e della differenza davanti e dietro.

Anche la parte esterna dell'orecchio, detta pinna, ha la sua importanza. In ogni pinna le riflessioni sonore su ogni rilievo si combinano col suono diretto, provocando cancellazioni di fase e irregolarità nello spettro percepito.

Le posizioni delle irregolarità dello spettro variano con l'altezza del suono sorgente. Noi le percepiamo non come una colorazione tonale, ma come informazioni sull'altezza.

Inoltre possiamo discriminare i suoni anteriori da quelli posteriori in virtù dell'effetto ombra dei padiglioni auricolari alle alte frequenze. Alcune delle informazioni usate dalle orecchie possono essere omesse senza alterare l'accuratezza della localizzazione, se altre sono presenti.

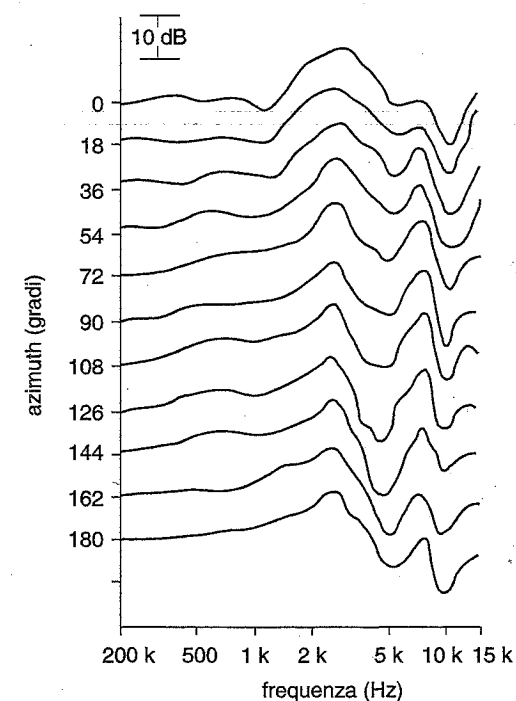


Fig. 3.5 Risposta in frequenza dell'orecchio a diversi angoli di azimuth. 0° individuano una provenienza perpendicolare, di fronte; 90° individuano il lato dell'orecchio preso a riferimento per la misurazione; 180° si riferiscono a una posizione dietro all'ascoltatore.

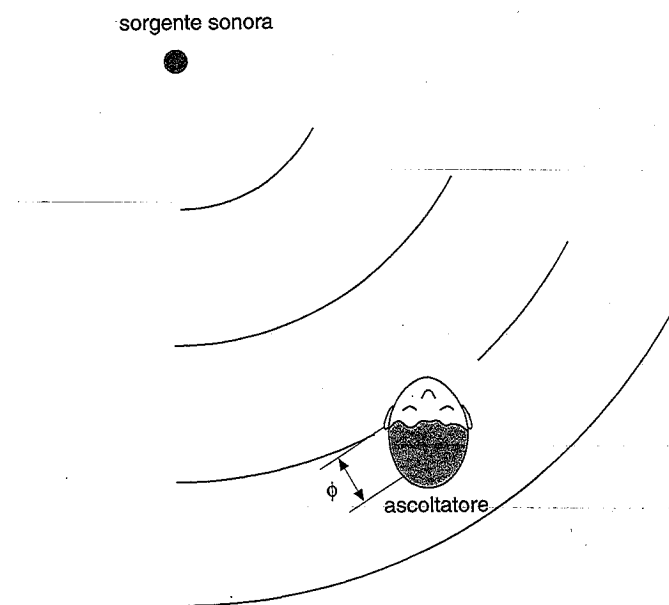


Fig. 3.6 Variazione della fase fra onde che arrivano all'orecchio.

3.3 Come localizzare le immagini fra i monitor

Ora che abbiamo visto come localizzare le sorgenti sonore reali, cerchiamo di capire come localizzare le loro immagini riprodotte attraverso i monitor. Immaginiamo di essere seduti fra due monitor, come nella fig. 3.7; se si invia un segnale musicale uguale ad entrambi i canali e con la stessa polarità, si percepirà una immagine fra i due monitor. Normalmente si ascolterà una singola sorgente risultante dalla sintesi delle immagini, piuttosto che due sorgenti separate nei monitor.

Ogni orecchio sente entrambi i monitor. Per esempio, l'orecchio sinistro ascolta il segnale del monitor sinistro e poi, dopo un breve ritardo dovuto alla maggiore lunghezza del percorso, sente quello del monitor di destra. Ad ogni orecchio i segnali da entrambi i monitor si sommano vettorialmente per produrre un segnale risultante. Supponiamo di aumentare il segnale in uno dei monitor, creando una differenza di livello tra di essi.

Sorprendentemente si manifesta una differenza di tempi d'arrivo fra le orecchie. Ciò è il risultato della somma di fase dei segnali di entrambi i monitor ad ogni orecchio.

Ricordiamoci di distinguere fra differenza intercanale (fra monitor) e differenze interaurali (fra orecchie). Una differenza di livello fra canali di monitor non si presenta come una differenza di livelli fra le orecchie, ma piuttosto come una differenza di tempi interaurali. Possiamo usare queste differenze di tempi interaurali generate dai monitor per posizionare le immagini. Ecco come: supponiamo di voler posizionare una immagine a 15° da un lato. Una sorgente sonora reale a 15° da un lato produce una differenza di tempi interaurali di 0,13 millisecondi. Se facciamo in modo che i monitor producano una differenza di tempi interaurali di 0,13 millisecondi, ascolteremo una immagine spostata

di 15° da un lato. Possiamo ingannare il sistema uditivo facendogli credere che ci sia una reale sorgente sonora a quella angolazione; ciò avviene quando i monitor si differenziano di livello di un certo ammontare. Anche la polarità dei due canali influenza la localizzazione. Per capire la polarità si tenga presente che se i segnali inviati ai due monitor sono in polarità, significa che essi sono in fase a tutte le frequenze, ovvero hanno una tensione positiva allo stesso istante di tempo. Se i segnali sono fuori polarità, significa che essi sono ruotati di fase a tutte le frequenze (di 180°). Il segnale di un canale è positivo quando l'altro è negativo e i segnali con polarità opposte sono a volte definiti, anche se non del tutto correttamente, come "fuori fase".

Se i segnali hanno polarità opposte sui due canali e uguale livello, l'immagine risultante è diffusa e senza direzione e non può essere localizzata. Se i segnali hanno polarità opposte e un canale ha il livello più alto rispetto all'altro, l'immagine spesso appare al di fuori dei limiti dei monitor. Si ascolterebbe un'immagine a sinistra del monitor di sinistra e a destra del monitor di destra. Polarità opposte si possono verificare in modi diversi. Due microfoni hanno polarità opposte se i fili ai piedini di collegamento 2 e 3 sono invertiti in uno di essi. Due monitor hanno polarità opposte se i conduttori del cavo sono invertiti in uno dei due. Un singolo microfono può avere parti diverse del suo diagramma polare in polarità opposte. Per esempio la parte posteriore dell'impronta di un bidirezionale è in opposizione a quella anteriore. Se un suono proveniente da una data direzione raggiunge il lobo anteriore in uscita al canale sinistro e contemporaneamente il lobo posteriore in uscita al canale destro, i due canali avranno polarità opposte. L'immagine risultante di quella sorgente sonora sarà diffusa, o sarà al di fuori dei monitor.

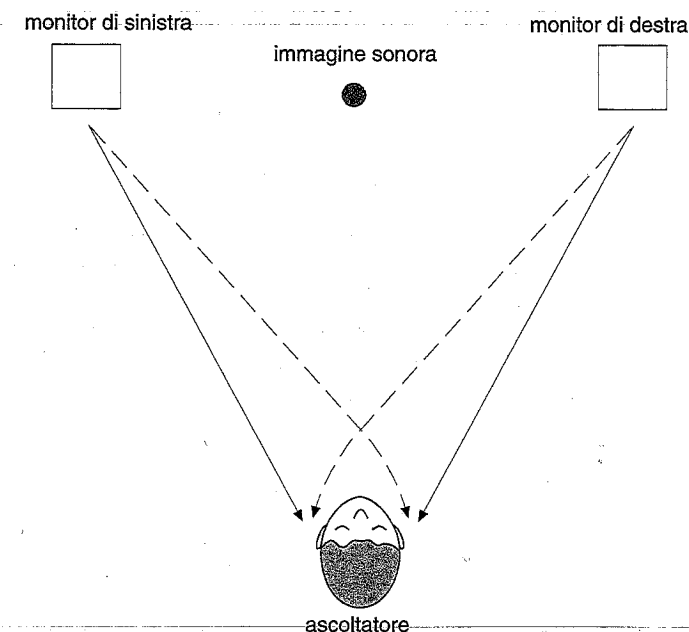


Fig. 3.7 Due orecchie che ricevono segnali da due monitor.

3.4 Requisiti per un posizionamento naturale dell'immagine fra i monitor

Grazie alla vastità con cui un sistema di registrazione e di riproduzione sonora può ricreare le differenze interaurali prodotte da una sorgente sonora reale, l'immagine risultante sarà accuratamente localizzata.

In altre parole, quando suoni riprodotti raggiungono le orecchie dell'ascoltatore e hanno l'ampiezza e la differenza di fase corrispondenti a quelle di una sorgente sonora reale posizionata ad un determinato angolo, l'ascoltatore percepisce una immagine con la stessa angolazione, ben amalgamata e naturalmente definita. Al contrario, quando si producono relazioni di fase ed ampiezza innaturali, l'immagine sembra relativamente diffusa piuttosto che definita, ed è più difficile da localizzare.

Le differenze interaurali richieste per una immagine realistica possono essere prodotte da differenze fra canali.

Collocare una immagine con precisione richiede particolari differenze di ampiezza e di fase rispetto alle frequenze, fra i canali.

Queste sono state calcolate da Cooper e Bauck relativamente a molte angolazioni di immagini. Gerzon, Nakabayashi e Koshigoe hanno calcolato le differenze intercanali e interaurali richieste per produrre qualunque direzione di immagine ad una singola frequenza.

La fig. 3.8 mostra le differenze fra canali richieste per posizionare una immagine a 15° verso sinistra, rispetto al centro, quando i monitor sono posizionati a $\pm 30^\circ$ davanti all'ascoltatore.

Come si vede nella fig. 3.8, le differenze intercanale richieste per una immagine naturale variano con la frequenza. Cooper e Bauck indicano che le differenze di ampiezza fra canali sono richieste sotto i 1700 Hz circa, mentre le differenze di tempi (cioè di fase) sopra tale frequenza. Più precisamente:

- alle basse frequenze la differenza di ampiezza richiesta per posizionare una immagine ad un angolo di 15° rispetto al centro, è di circa 10 dB;
- fra 1,7 kHz e 5 kHz arriva a circa 0 dB;
- sopra 1,7 kHz la differenza di fase corrisponde ad un ritardo (differenza di tempo fra canali) di circa 0,547 millisecondi oppure 20 cm di distanza ipotetica fra i microfoni usati per la registrazione stereo.

Questa teoria è basata sull'oscuramento di un suono che si muove attorno ad una sfera. La descrizione data qui, per quanto semplificata, esprime il concetto di base.

Cooper fa notare che piccoli allontanamenti da queste specifiche possono non portare a turbamenti percepibili dagli ascoltatori o a immagini ingannevoli. Si possono verificare i criteri di Cooper e Bauck registrando con una testa artificiale i cui segnali vengono processati in modo particolare. Una testa artificiale usata per una registrazione binaurale è una struttura normalmente conformata con un microfono in ciascun orecchio.

Le differenze di tempo e di spettro fra i canali creano l'immagine stereo (le differenze di spettro sono le differenze di ampiezza in funzione della frequenza).

Anche se una registrazione binaurale con testa artificiale può dare un eccellente posizionamento in cuffia, essa dà una localizzazione scarsa nei monitor alle basse frequenze, a meno di usare un'equalizzazione spaziale (con un circuito *shuffler*).

L'equalizzazione spaziale incrementa le basse frequenze nel segnale differenza (L - R).

Una registrazione binaurale può produrre immagini avvolgenti in cuffia, ma solo imma-

gini frontali nei monitor, a meno di non usare un convertitore transaurale, che è un dispositivo elettrico che converte il segnale sdoppiato per le cuffie in un segnale stereo per i monitor.

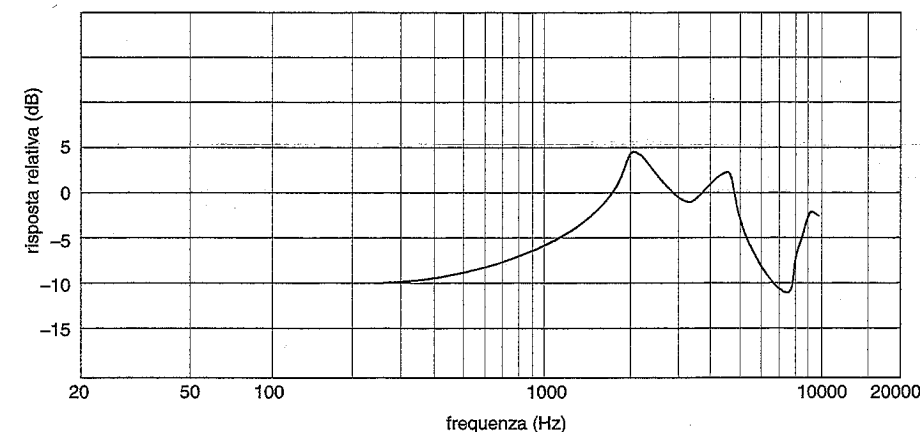


Fig. 3.8a Ampiezza del canale di destra rispetto al canale di sinistra, per un posizionamento dell'immagine a 15° a sinistra rispetto al centro quando i monitor sono a $\pm 30^\circ$ rispetto all'ascoltatore.

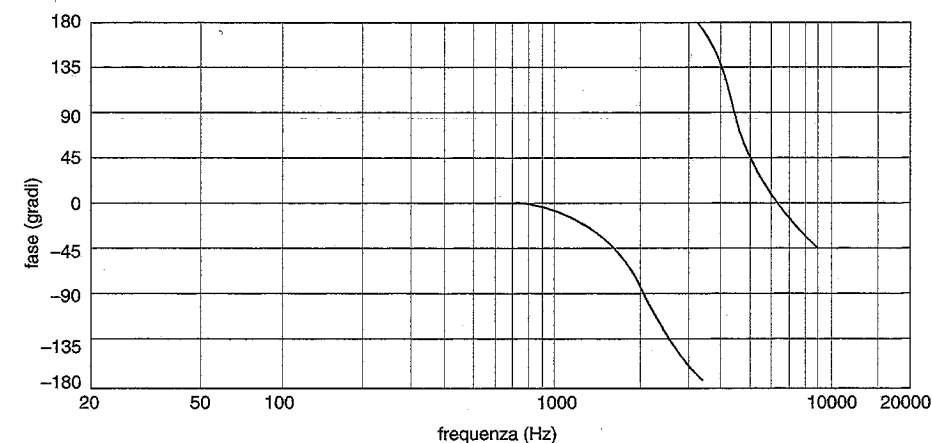


Fig. 3.8b Fase del canale di destra rispetto al canale di sinistra, per un posizionamento dell'immagine 15° a sinistra rispetto al centro.

La stereofonia transaurale è una tecnica di riproduzione sonora "avvolgente" che usa una testa microfonica per la registrazione binaurale, che viene processata elettronicamente per eliminare gli scambi di informazioni relativi alla testa stessa quando la registrazione è ascoltata nei monitor (ulteriori note si trovano nel Capitolo 6). Cooper raccomanda, per una immagine naturale, che le differenze fra canali nei monitor siano controllate in modo che i loro segnali somma, alle orecchie, producano le corrette differenze interaurali. Secondo Theile la teoria di Cooper (basata sulla localizzazione per somma) è errata perché si applica solo alle sinusoidi e non può essere applicata ai segnali con spettro a larga ban-

da. Egli propone una teoria diversa, chiamata modello associativo. Questa teoria dice che quando si ascoltano due monitor stereo, si ignorano le differenze interaurali, e si usano invece le differenze fra canali dei monitor per localizzare l'immagine. Secondo Theile, le differenze fra canali richieste per una migliore stereofonia sono collegate alla testa. La tecnica di microfonaggio stereo ideale dovrebbe usare due microfoni distanziati come la testa, montati allo stesso livello in una sfera delle dimensioni della testa ed equalizzati per una risposta piatta e soggettiva. Ciò produrrebbe differenze di tempo e di spettro fra canali, che, sostiene Theile, sono ottime per la ripresa stereo. Le differenze fra canali (differenze di tempo alle basse frequenze e di ampiezza alle alte) sono opposte ai requisiti di Cooper per una naturale immagine stereo. Solo il tempo dirà quale teoria è più vicina alla realtà.

3.5 Meccanismi per la localizzazione dell'immagine

L'orecchio può essere tratto in inganno nell'ascoltare immagini ragionevolmente ben definite fra i due monitor mediante un procedimento sul segnale che sia meno sofisticato.

La pura e semplice ampiezza e/o le differenze di tempi fra i canali, a parità di frequenza, possono produrre immagini localizzabili. Bartlett, Jordan, Dutton, Cabot, Williams, Blauert e Rumsey danno risultati di test che mostrano il posizionamento dell'immagine come dipendente da differenze di tempi o di ampiezza fra canali. I risultati di Bartlett sono mostrati più avanti in questo capitolo.

Per esempio, dato un segnale vocale, se il canale di sinistra ha un volume maggiore di 7,5 dB rispetto al canale di destra, l'immagine apparirà circa 15° verso sinistra rispetto al centro quando i monitor sono messi a $\pm 30^\circ$ di inclinazione rispetto all'ascoltatore. Un ritardo nel canale di destra di circa 0,5 millisecondi porterà allo stesso risultato, anche se le immagini prodotte unicamente da differenze di tempo sono relativamente vaghe e difficili da localizzare.

Griesinger fa notare che la pura e semplice differenza di tempi fra canali non fa localizzare nei monitor una voce maschile filtrata con un passa basso al di sotto di 500 Hz.

Le differenze di ampiezza (livello) sono necessarie per localizzare i suoni a bassa frequenza. Nè le differenze di ampiezza nè quelle di tempo possono localizzare i suoni ad alta frequenza. Comunque, Blauert mette in evidenza, tramite Wendt, che le differenze di tempi fra canali consentono la localizzazione a 327 Hz.

Le differenze fra canali risultanti nelle usuali registrazioni stereo sono soltanto una semplice approssimazione di ciò che si richiede. La pratica corrente fa uso delle differenze di ampiezza fra canali e/o delle differenze di tempi per localizzare le immagini; nonostante queste differenze siano costanti rispetto alla frequenza, si producono immagini ragionevolmente definite. Vediamo esattamente come queste differenze localizzano le immagini.

3.5.1 Localizzazione tramite differenze di ampiezza

La localizzazione di una immagine fra i monitor dipende in parte dalle differenze di ampiezza tra i rispettivi segnali.

Supponiamo che venga inviato un segnale vocale a due monitor stereo e che il segnale sia uguale nei due monitor, tranne che per la differenza di ampiezza, cioè di livello (fig. 3.9). Si crea una differenza di ampiezza inserendo un attenuatore in un canale.

La figura 3.10 mostra la localizzazione approssimativa dell'immagine sonora fra i monitor rispetto alle differenze di ampiezza nei canali, in decibel.

Una differenza di 0 dB (cioè un segnale uguale in entrambi i monitor) fa apparire l'immagine della sorgente sonora nel centro, a metà strada fra i due monitor.

Aumentando la differenza, l'immagine si allontana dal centro. Una differenza di 15 o 20 dB fa apparire l'immagine solo in un monitor.

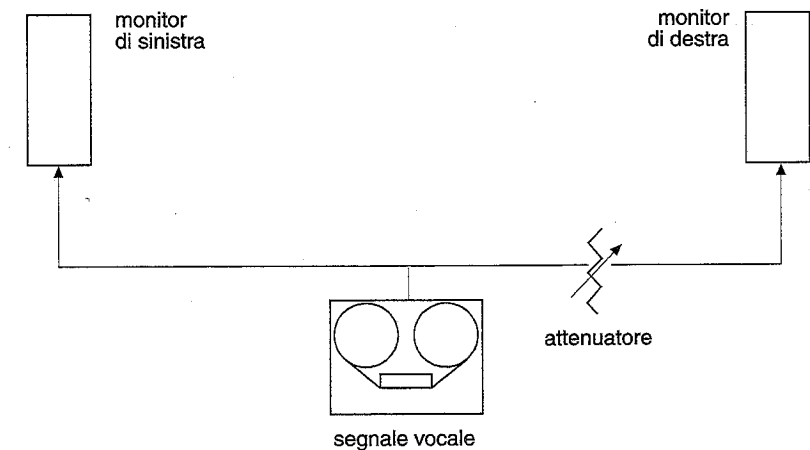


Fig. 3.9 Invio di un segnale vocale a due monitor stereo con attenuazione in un canale.

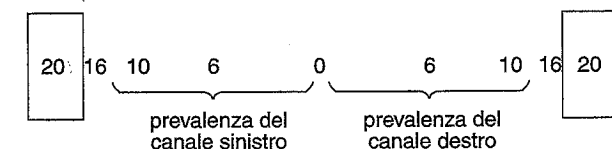


Fig. 3.10 Localizzazione dell'immagine stereofonica rispetto alle differenze di ampiezza fra canali, in decibel (percezione dell'ascoltatore).

Le informazioni in questa figura sono basate su test di ascolto molto ben controllati. I dati sono la media delle risposte di 10 ascoltatori ben allenati. Hanno ascoltato monitor di alta qualità, allineati e posti ad alcuni metri dalla parete in una tipica sala d'ascolto, seduti centralmente fra i monitor, inclinati di 60° .

Come è possibile creare questo effetto con una tecnica stereofonica di microfonaggio? Supponiamo di incrociare due cardioidi a 90° , con le membrane esattamente sovrapposte (fig. 3.11). I microfoni sono quindi angolati di 45° verso destra e verso sinistra rispetto al centro dell'orchestra. I suoni provenienti dal centro dell'orchestra saranno ripresi in maniera eguale da entrambi i microfoni. Durante il riascolto ci sarà un livello uguale su tutti e due i monitor e quindi si avrà una immagine centrale.

Supponiamo che l'estrema destra dell'orchestra sia fuori asse di 45° rispetto all'angolo di ripresa dei microfoni. I suoni che arrivano dall'estrema destra dell'orchestra rag-

giungono il microfono che guarda a destra perfettamente in asse, ma raggiungono l'altro di 90° fuori asse (fig. 3.11). Un diagramma polare cardioidi ci dà, a 90° fuori asse, un livello inferiore di 6 dB rispetto a quello dato in asse. Quindi le sorgenti sonore all'estrema destra produrranno un livello di uscita di 6 dB inferiore dal microfono rivolto a sinistra rispetto al microfono rivolto a destra.

Perciò abbiamo una differenza in ampiezza fra canali di 6 dB. Come si vede nella fig. 3.10, l'immagine della sorgente sonora all'estrema destra sarà posta verso destra rispetto al centro. Gli strumenti a metà strada fra il centro e la destra dell'orchestra saranno riprodotti in qualche parte fra il punto a 0 dB e quello a 6 dB.

Se angoliamo ulteriormente i microfoni, ad es. 135° , la differenza risultante ai canali per la stessa sorgente sarà circa 10 dB. Perciò l'immagine stereo destra sarà ancora più spostata verso destra che in precedenza (cioè con l'angolo di 90°). (Da notare inoltre che non è necessario rivolgere i microfoni esattamente verso la destra e la sinistra dell'insieme).

Più una sorgente sonora è posizionata da una parte, maggiore è la differenza di ampiezza fra canali che si produce, quindi l'immagine risultante sarà più lontana dal centro.

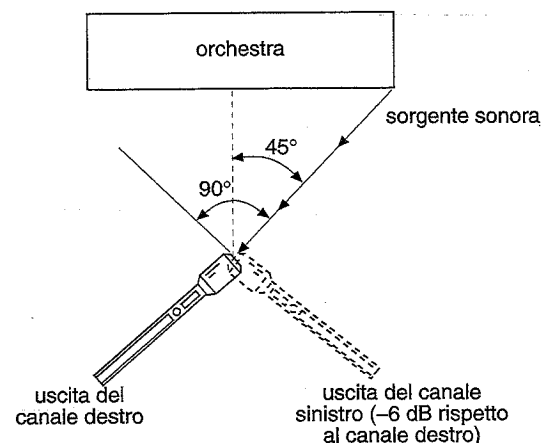


Fig. 3.11 Cardioidi incrociati a 90° , che riprendono una sorgente ad una estremità di un'orchestra.

3.5.2 Localizzazione tramite differenze di tempo

La localizzazione dell'immagine dipende inoltre dalle differenze di tempi fra i canali. Supponiamo di mandare lo stesso segnale vocale a due monitor ad uguale livello ma con un canale in ritardo, come si vede nella fig. 3.12.

La fig. 3.13 mostra la localizzazione approssimativa dell'immagine sonora fra i monitor, con diverse differenze di tempo fra i canali, in millisecondi. A 0 millisecondi di differenza (cioè nessuna differenza di tempo fra i canali) l'immagine appare centralmente. All'aumento della differenza l'immagine si sposta fuori centro. Una differenza o un ritardo di 1,5 millisecondi è sufficiente per posizionare l'immagine in uno solo dei due monitor.

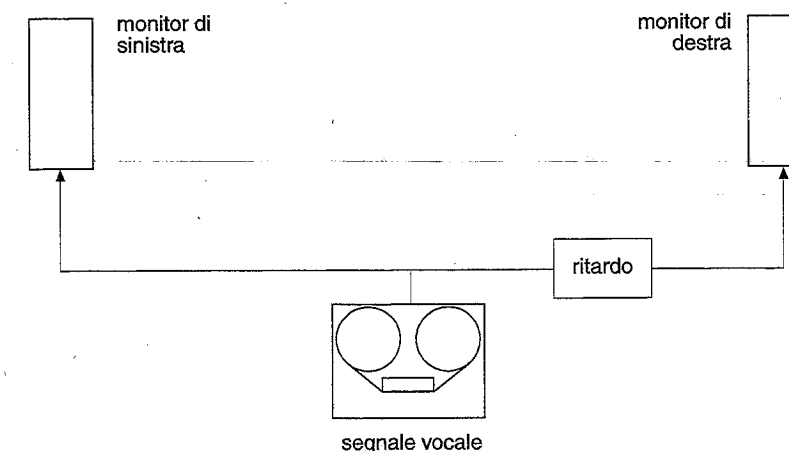


Fig. 3.12 Invio di un segnale vocale a due monitor con un canale in ritardo.

Distanziando orizzontalmente i microfoni di pochi centimetri, si producono differenze di tempo fra canali per sorgenti fuori centro.

Un suono che arriva dalla parte destra dell'orchestra raggiungerà il microfono destro prima del sinistro unicamente perché è più vicino (fig. 3.14).

Per esempio, se la sorgente sonora è a 45° verso destra ed i microfoni sono distanziati di 20 centimetri, la differenza di tempo risultante ai canali per questa sorgente è di circa 0,4 millisecondi.

Per la stessa sorgente una distanza di 50 centimetri fra i microfoni produce una differenza di tempo di 1,5 millisecondi fra i canali, posizionando completamente l'immagine sonora riprodotta in uno dei due monitor.

Con microfoni distanziati, maggiore è la distanza di una sorgente sonora dal centro dell'orchestra, maggiore è la differenza di tempo fra i canali e quindi maggiore la distanza a cui verrà posizionata l'immagine sonora riprodotta.



Fig. 3.13 Posizionamento approssimato dell'immagine rispetto alle differenze di tempo fra canali, in millisecondi (secondo la percezione dell'ascoltatore).

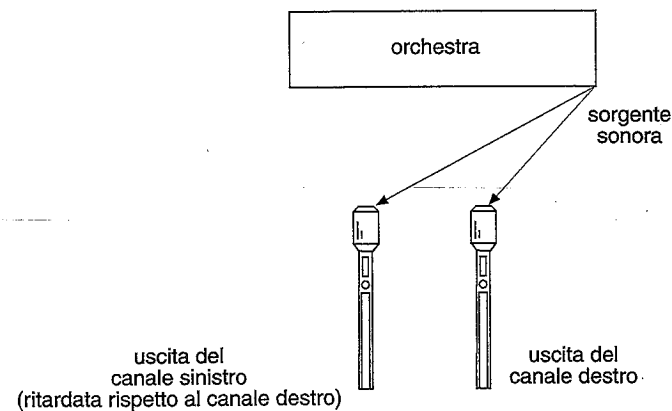


Fig. 3.14 Microfoni distanziati che riprendono una sorgente ad una estremità di un'orchestra.

3.5.3 Localizzazione tramite differenze di tempo e di ampiezza

Le immagini possono essere localizzate anche tramite una combinazione di differenze di ampiezza e di tempo. Supponiamo di avere due microfoni cardioidi inclinati di 90° e distanziati di 20 centimetri (fig. 3.15). Una sorgente sonora inclinata di 45° verso destra produrrà una differenza di livello di 6 dB fra i canali ed una differenza di tempo di 0,4 millisecondi. Lo spostamento nell'immagine dovuto ai 6 dB si aggiunge a quello dovuto alla differenza di 0,4 millisecondi e posiziona l'immagine sonora nel monitor di destra. Tale risultato è ottenibile anche tramite altre combinazioni.

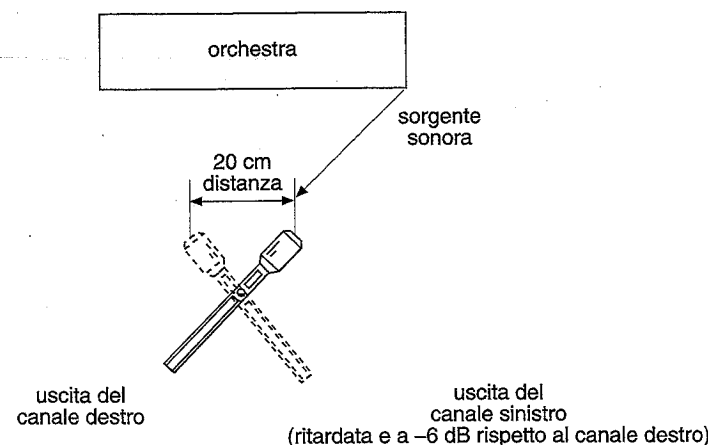


Fig. 3.15 Cardioidi angolati a 90° e distanziati di 20 cm, che riprendono una sorgente ad una estremità dell'orchestra.

3.5.4 Ricapitolando

Se un segnale vocale è registrato su due canali, la sua immagine sonora riprodotta apparirà in uno solo dei due monitor quando:

- il segnale in un canale è inferiore di $15 + 20$ dB rispetto all'altro;
- il segnale in un canale è in ritardo di 1,5 millisecondi rispetto al segnale dell'altro;
- il segnale in un canale è più basso in livello e in ritardo di un certo ammontare rispetto al segnale dell'altro.

Quando si combinano differenze di tempo e di livello per posizionare le immagini, si ha l'immagine più definita se il canale più basso in livello è anche il canale in ritardo. Se è in ritardo il canale col più alto livello, l'immagine risulterà confusa per le differenze di tempo e di livello in conflitto tra loro.

Abbiamo visto che l'angolazione di microfoni direzionali (con capsule coincidenti) produce differenze di ampiezza fra i canali. Il distanziamento di microfoni ("microfoni distanziati") introduce differenze di tempi fra i canali. L'angolazione e il distanziamento di microfoni direzionali (quasi coincidenti) introduce differenze fra i canali, sia in tempo che in ampiezza. Tali differenze localizzano l'immagine sonora riprodotta fra una coppia di monitor.

3.6 Previsione del posizionamento dell'immagine

Supponiamo di avere due microfoni per la registrazione stereo. Dati i loro diagrammi polari, le loro angolazioni e distanze, si possono prevedere le differenze di ampiezze e di tempo fra canali per qualunque angolazione della sorgente. Quindi, si può prevedere la localizzazione per ogni tecnica stereo. Questa previsione considera che i microfoni abbiano diagrammi polari ideali e non variabili a seconda della frequenza. Ciò può essere irrealistico, ma la previsione ben si sposa con i test di ascolto.

La differenza di ampiezza in decibel fra i canali è data dalla seguente formula:

$$(4.1) \quad \Delta \text{dB} = 20 \log \left[\frac{a + b \cos((\vartheta_m/2) - \vartheta_s)}{a + b \cos((\vartheta_m/2) + \vartheta_s)} \right]$$

dove ΔdB è la differenza in ampiezza fra canali, in decibel; $a + b \cos(\vartheta)$ è l'equazione polare del microfono.

Le costanti a e b dipendono dal tipo di microfono; si ha:

- | | | |
|-------------------|-------------|-------------|
| • omnidirezionale | $a = 1$ | $b = 0$ |
| • bidirezionale | $a = 0$ | $b = 1$ |
| • cardiode | $a = 0,5$ | $b = 0,5$ |
| • supercardioide | $a = 0,366$ | $b = 0,634$ |
| • ipercardiode | $a = 0,25$ | $b = 0,75$ |

ϑ_m è l'angolo fra gli assi dei microfoni, in gradi;

ϑ_s è l'angolazione della sorgente rispetto alla perpendicolare, in gradi.

Si veda nella fig. 3.16.

La differenza di tempo fra i canali è data da :

$$(4.2) \quad \Delta T = \frac{\sqrt{D^2 + [(S/2) + D \tan \vartheta_s]^2} - \sqrt{D^2 + [(S/2) - D \tan \vartheta_s]^2}}{C}$$

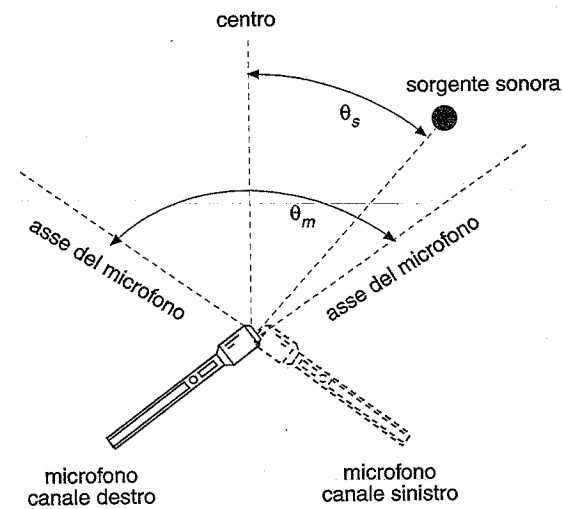


Fig. 3.16. Angolo fra i microfoni ϑ_m , angolo della sorgente ϑ_s .

dove ΔT è la differenza di tempo fra i canali, in secondi;
 D è la distanza dalla sorgente alla linea di connessione dei microfoni, in piedi;
 S è la distanza fra i microfoni, in piedi;
 ϑ_s è l'angolazione della sorgente rispetto alla perpendicolare, in gradi;
 C la velocità del suono (1130 piedi al secondo).
 Si veda la fig. 3.17.

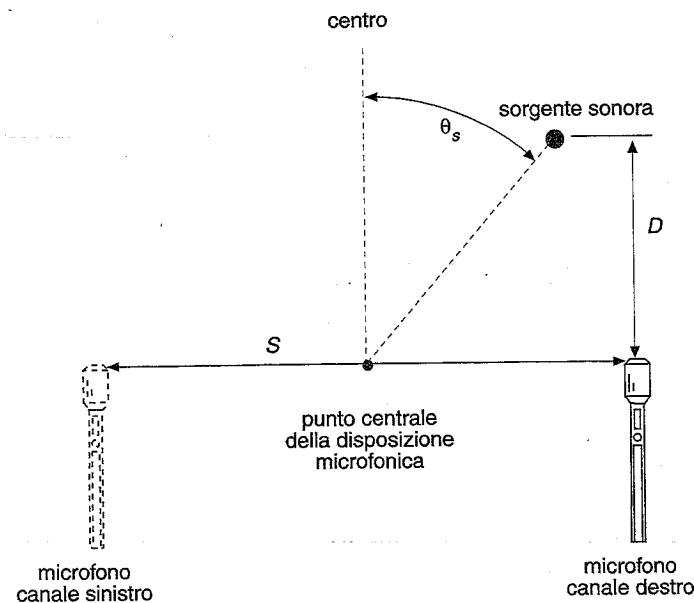


Fig. 3.17. Angolo della sorgente ϑ_s , distanza microfono-sorgente D e distanza fra microfoni S .

Per microfoni quasi coincidenti distanziati di pochi centimetri, l'equazione può essere semplificata nel seguente modo:

$$(4.3) \quad \Delta T = \frac{S \sin \vartheta_s}{C}$$

dove ΔT è la differenza di tempo fra i canali, in secondi;

S è la distanza fra i microfoni in pollici;

ϑ_s è l'angolazione della sorgente, in gradi;

C la velocità del suono (13.560 pollici al secondo).

Si veda la fig. 3.18.

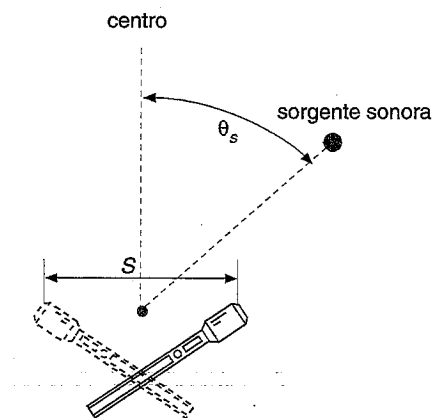


Fig. 3.18. Distanza fra microfoni S e angolo della sorgente ϑ_s .

Per esempio, se si angolano due cardioidi di 135° e la sorgente è angolata di 60° (come nella fig. 3.19), la differenza in decibel fra i canali per tale sorgente è calcolata come segue:

per un cardioide $a = 0,5$ e $b = 0,5$ (dalla tabella successiva alla equazione 4.1);

ϑ_m , l'angolo fra gli assi dei microfoni, è 135° ;

ϑ_s , l'angolazione della sorgente, è 60° , cioè la sorgente sonora è spostata di 60° e quindi la differenza di ampiezza fra i canali, usando l'equazione 4.1, è data da:

$$\Delta \text{dB} = 20 \log \left[\frac{0,5 + 0,5 \cos((135^\circ/2) - 60^\circ)}{0,5 + 0,5 \cos((135^\circ/2) + 60^\circ)} \right]$$

$$= 14 \text{ dB di differenza di ampiezza fra i canali.}$$

Quindi, facendo riferimento alla fig. 3.10, la sorgente sonora sarà riprodotta vicinissima ad uno dei due monitor.

Altro esempio. Se si posizionano due microfoni omnidirezionali a 25 cm di distanza e la sorgente sonora è spostata di 45° fuori asse, qual è la differenza di tempo fra i due canali? (Riferirsi alla fig. 3.20).

$S = 10$ (pollici);

$\vartheta_s = 45^\circ$;

dall'equazione 4.3 si ha:

$$\Delta T = \frac{10 \sin 45^\circ}{13.560}$$

$$= 0,52 \text{ millisecondi di differenza fra i canali.}$$

Quindi, facendo riferimento alla fig. 3.10, tale sorgente sonora sarà posizionata circa a metà strada tra i monitor.

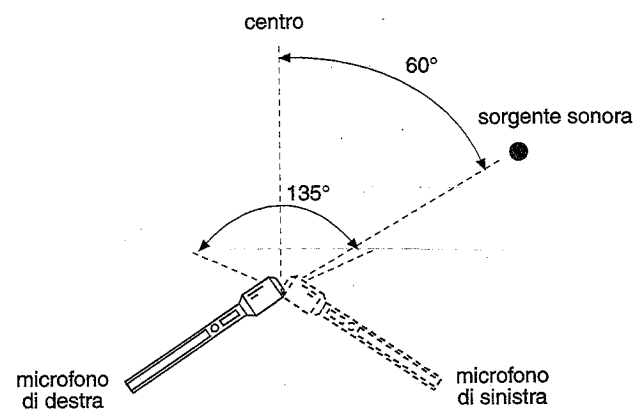


Fig. 3.19 Cardioidi angolati di 135° con un'angolazione della sorgente di 60°.

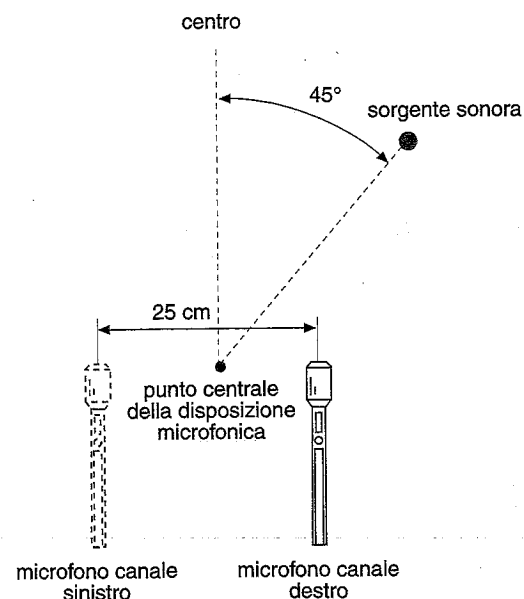


Fig. 3.20 Omnidirezionali distanziati di 25 cm con un'angolazione della sorgente di 45°.

3.7 Scelta degli angoli e delle distanze

Ci sono molte combinazioni di angoli e distanze fra microfoni che posizionano le immagini degli strumenti agli estremi di un'orchestra nel monitor di destra o di sinistra. In altre parole, ci sono molte maniere per ottenere una estensione stereo completa. Si può usare una distanza contenuta ed un angolo ampio, o viceversa. Ad esempio, l'ORTF ha angolo di 110° e distanza di 17 cm, ma se la larghezza del fronte riprodotta è troppo piccola, non c'è modo di aumentarla. Se gli strumenti centrali risultano troppo alti in volume, si possono angolare maggiormente i microfoni e diminuire la loro distanza; così rimane invariata la larghezza del fronte sonoro ripreso e quindi si può controllare il volume degli strumenti centrali per migliorare il bilanciamento.

Per ridurre la ripresa delle prime riflessioni dal pavimento e dalle pareti del palco, bisogna (1) aumentare l'angolazione, (2) diminuire gli spazi, (3) mettere i microfoni più vicini all'insieme. Si ottengono i seguenti risultati:

1. angolare maggiormente i microfoni "ammorbisce" gli strumenti centrali;
2. diminuire la distanza fra i microfoni mantiene la originale larghezza della ripresa;
3. dato che gli strumenti centrali hanno un impatto sonoro minore, si possono mettere i microfoni più vicini all'insieme degli strumenti ed avere comunque un buon bilanciamento;
4. dato che i microfoni sono più vicini, diminuisce il rapporto fra suono riflesso e suono diretto. Si possono aggiungere microfoni distanti o riverberazioni artificiali per l'apporto di ambiente desiderato.

In generale, una combinazione di angolazione e di distanza (differenze di intensità e di tempi) dà una localizzazione più accurata e una immagine più definita rispetto all'uso separato di angolazione e di distanza.

Una maggiore angolazione dei microfoni aumenta il rapporto di riverberazione nella registrazione, e ciò dà l'impressione che l'orchestra suoni molto più da lontano. Invece un eccessivo aumento della separazione non cambia il senso delle distanze, ma peggiora la definizione dell'immagine.

3.8 Spazialità ed equalizzazione spaziale

Le notizie riportate in questo paragrafo sono tratte dalla rivista di Griesinger *Nuovi sviluppi nelle tecniche microfoniche coincidenti e semi coincidenti*.

La spazialità di una tecnica microfonica è il rapporto dell'energia L - R rispetto alla energia L + R del suono riflesso. Idealmente questo rapporto dovrebbe essere uguale o maggiore di 1. In altre parole, c'è una somma e una differenza uguale di energia. La spazialità implica una minima correlazione fra canali per quel che riguarda il suono riflesso. Alcune tecniche che danno buona spazialità (valore pari a 1) sono:

- microfoni distanziati;
- Blumlein (microfoni con diagramma a otto, incrociati di 90°);
- tecnica M-S con un cardiode come trasduttore centrale e un rapporto di 1:1 fra M-S;
- ipercardioidi coincidenti con angolazione di 109°.

L'equalizzazione spaziale, o *shuffling* è un incremento alle basse frequenze del segnale differenza L-R e un taglio complementare, alle basse frequenze del segnale somma L + R. In entrambi i casi si usa un'equalizzazione di tipo *shelving*. Ciò porta due benefici:

- aumenta la spazialità, in modo che le tecniche coincidenti e quasi coincidenti possano dare una spazialità simile a quella della tecnica dei microfoni distanziati;
- allinea le componenti, cioè le basse e le alte frequenze dell'immagine sonora, che risulta più definita.

Si può costruire un equalizzatore spaziale (*shuffler*) seguendo le indicazioni della rivista di Griesinger, oppure usando una tecnica M-S: si incrementano le basse frequenze nel segnale L-R, detto laterale, e si tagliano le basse frequenze nel segnale L + R, detto centrale. Il taglio, o l'incremento necessari, dipendono dalla tecnica microfonica, ma un valore usuale è da 4 a 6 dB al di sotto di 400 Hz tramite un equalizzatore di tipo shelving. Un eccessivo incremento può addirittura suddividere le immagini fuori centro, mettendo i bassi e gli alti in posizioni diverse. La correzione andrebbe fatta sul segnale proveniente dalla tecnica M-S prima di mixarla con altri microfoni. Gerzon fa notare che le fasi dei canali somma e differenza dovrebbero essere compensate (con risposte di fase dei due filtri accoppiate e non minime) come suggerisce Vanderlyn. Gerzon fa notare che l'equalizzazione spaziale si applica meglio alle tecniche stereo che non hanno una grande componente di riverberazione in controfase alle basse frequenze, come ad esempio nei cardioidi coincidenti o quasi coincidenti. Con la tecnica Stereosonic (otto incrociati), le componenti in controfase diventano eccessive. Si suggerisce un taglio di 2,4 dB nel segnale L + R e un incremento di 5,6 dB nel segnale L - R per una migliore risposta alle basse frequenze. Griesinger afferma: "L'equalizzazione spaziale può essere molto utile con le tecniche coincidenti e semicoincidenti, soprattutto se l'ascolto è fatto in piccole stanze. Dato che le maggiori informazioni per la localizzazione vengono dalle alte frequenze, bisogna scegliere impronte microfoniche e angolazioni che diano un'accurata estensione delle immagini alle alte frequenze. L'equalizzazione spaziale può poi essere usata per aumentare la spazialità alle basse frequenze". Alan Blumlein inventò il primo shuffler nel 1933. Lo usò assieme a due capsule omnidirezionali separate da una distanza pari alla testa umana. Lo shuffler separa i due canali (li somma in polarità opposte). Quando due omnidirezionali sono sommati con polarità opposte, il risultato è un'unica impronta bidirezionale che si estende a destra e a sinistra. Blumlein usò questa impronta come parte anteriore in un M-S. La risposta in frequenza dell'impronta bidirezionale risultante è debole alle basse frequenze: diminuisce di 6 dB/ottava al diminuire della frequenza. Così, per compensare, il circuito shuffler, Blumlein include un incremento del primo ordine alle basse frequenze (al di sotto di 700 Hz). Lo shuffler trasforma le differenze di fase in differenze di intensità. Più lontana è la sorgente sonora dal centro, maggiore è la differenza di fase tra i due microfoni distanziati; maggiore è la differenza di fase, maggiore è la differenza di intensità creata dallo shuffler nei due canali.

3.9 Bibliografia

1. H. Wallach, E. Newman, and M. Rozenzweig, "The Precedence Effect in Sound Localization," *Journal of the Audio Engineering Society*, Vol. 21, No. 10, 1973 (Dec.), pp. 817-826.
2. B. Bartlett, "Stereo Microphone Technique," *db*, Vol. 13, No. 12, 1979 (Dec.), pp. 34-46.
3. M. Gardner, "Some Single and Multiple Source Localization Effects," *J. Audio Eng. Soc.*, Vol. 21, No. 6, 1973 (July/Aug.), pp. 430-437.
4. E. Carterette e M. Friedman, *Handbook of Perception*, Academic Press, New York, 1978, Vol. 4 (Hearing), pp. 62, 210.

5. P. Vanderlyn, "Auditory Cues in Stereophony," *Wireless World*, 1979 (Sept.), pp. 55-60.
6. E. Shaw, "Transformation of Sound Pressure Levels From the Free Field to the Eardrum in the Horizontal Plane," *Journal of the Acoustical Society of America*, Vol. 56, No. 6, 1974 (Dec.), pp. 1848-1861.
7. S. Mehrgardt e V. Mellert, "Transformation Characteristics of the External Human Ear," *J. Acoust. Soc. Amer.*, Vol. 61, No. 6, p. 1567.
8. F. Rumsey, *Stereo Sound for Television*, Focal Press, Boston, 1989, p. 6.
9. F. Harvey e M. Schroeder, "Subjective Evaluation of Factors Affecting Two-Channel Stereophony," *J. Audio Eng. Soc.*, Vol. 9, No. 1, 1961 (Jan.), pp. 19-28.
10. J. Eargle, *Sound Recording*, Van Nostrand Reinhold Company, New York, 1976, Chapters 2 and 3.
11. M. Gardner e R. Gardner, "Problems of Localization in the Median Plane - Effect of Pinnae Cavity Occlusion," *J. Acoust. Soc. Amer.*, Vol. 53, 1973 (Feb.), pp. 400-408.
12. K. Nakabayashi, "A Method of Analyzing the Quadraphonic Sound Field," *J. Audio Eng. Soc.*, Vol. 23, No. 3, 1975 (Apr.), pp. 187-193.
13. D.H. Cooper, "Problems with Shadowless Stereo Theory: Asymptotic Spectral Status," *J. Audio Eng. Soc.*, Vol. 35, No. 9, 1987 (Sept.), p. 638.
14. B. Bauer, "Stereophonic Earphones and Binaural Loudspeakers," *J. Audio Eng. Soc.*, Vol. 9, No. 2, 1961 (Apr.), pp. 148-151.
15. M. Schroeder e B. Atal, "Computer Simulation of Sound Transmission in Rooms," *IEEE Convention Record*, Part 7, 1963, pp. 150-155.
16. P. Damaske, "Head-Related Two-Channel Stereophony with Loudspeaker Reproduction," *J. Acoust. Soc. Amer.*, Vol. 50, No. 4, 1971, pp. 1109-1115.
17. T. Mori, G. Fujiki, N. Takahashi, and E. Maruyama, "Precision Sound-Image Localization Technique Utilizing Multi-Track Tape Masters," *J. Audio Eng. Soc.*, Vol. 27, No. 1/2, 1979 (Jan./Feb.), pp. 32-38.
18. N. Sakamoto et. al., "Controlling Sound-Image Localization in Stereophonic Reproduction" (Parts 1 and II), *J. Audio Eng. Soc.*, Vol. 29, No. 11, 1981 (Nov.), pp. 794-799, and Vol. 30, No. 10, 1982 (Oct.), pp. 719-722.
19. H. Moller, "Reproduction of Artificial-Head Recordings through Loudspeakers," *J. Audio Eng. Soc.*, Vol. 37, No. 1/2, 1989 (Jan./Feb.), pp. 30-33.
20. D. Cooper e J. Bauck, "Prospects for Binaural Recording," *J. Audio Eng. Soc.*, Vol. 37, No. 1/2, 1989 (Jan./Feb.), pp. 3-19.
21. E. R. Madsen, "The Application of Velocity Microphones to Stereophonic Recording," *J. Audio Eng. Soc.*, Vol. 5, No. 2, 1957 (Apr.), p. 80.
22. M. Williams, "Unified Theory of Microphone Systems for Stereophonic Sound Recording," *Audio Engineering Society Preprint No. 2466 (H-6)*, presented at 82nd Convention, March 10-13, 1987, London.
23. M. Gerzon, Letter to the Editor, Reply to comments on "Spaciousness and Localization in Listening Rooms and Their Effects on the Recording Technique," *J. Audio Eng. Soc.*, Vol. 35, No. 12, 1987 (Dec.), pp. 1013-1014.