

Capitolo 5 – Applicazioni

Indice

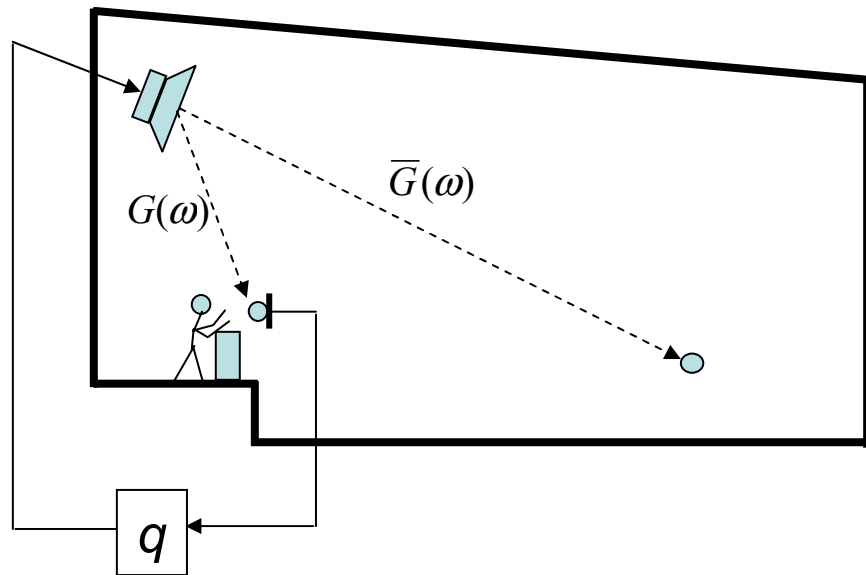
- Cancellazione d'eco
- Audio 3D
- Localizzazione di sorgenti sonore

Cancellazione d'eco (1)

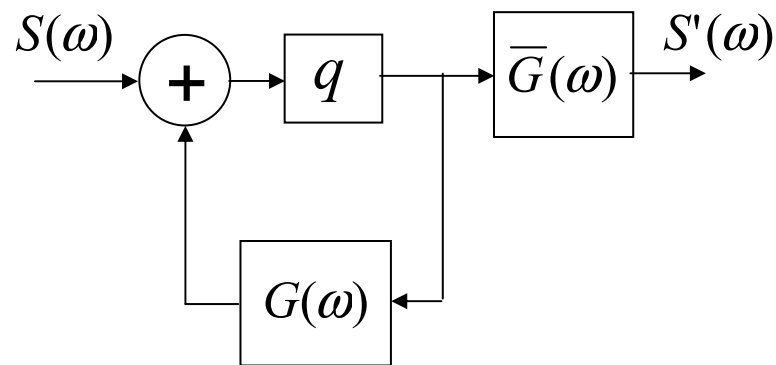
In numerose situazioni di tipo applicativo (per esempio nei sistemi di rinforzo del segnale, nei sistemi viva-voce, nei sistemi per videoconferenza) può accadere che una porzione del segnale emesso da un altoparlante venga prelevato dal microfono presente nello stesso ambiente (*controreazione sonora*).

Se la porzione di segnale prelevata dal microfono non è trascurabile, la controreazione genera effetti di diverso tipo, come distorsioni o addirittura oscillazioni del segnale che si autosostengono, fino ad arrivare all'instabilità dell'intero sistema (nella pratica si ha una dinamica di uscita limitata).

Cancellazione d'eco (2)



Si suppone che lo spettro della sorgente sia $S(\omega)$. Il segnale viene amplificato tramite un guadagno q (idealmente indipendente dalla frequenza).



Il segnale ricevuto è:

$$S'(\omega) = q\bar{G}(\omega) \left[S(\omega) + G(\omega) \frac{S'(\omega)}{\bar{G}(\omega)} \right]$$

Cancellazione d'eco (3)

La funzione di trasferimento complessiva (che tiene conto anche della controreazione) risulta:

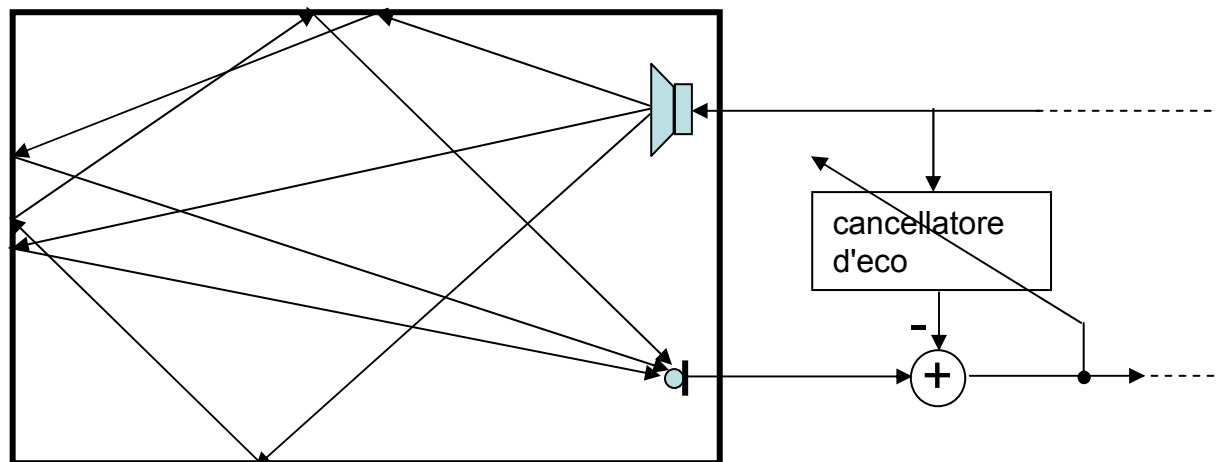
$$G'(\omega) = \frac{S'(\omega)}{S(\omega)} = q\bar{G}(\omega) \sum_{n=0}^{\infty} [qG(\omega)]^n = \frac{q\bar{G}(\omega)}{1 - qG(\omega)}$$

Il termine $qG(\omega)$ è il *guadagno ad anello aperto* del sistema e determina in modo essenziale lo spettro del segnale ricevuto. Esiste un valore critico q_0 del guadagno per cui ad una certa frequenza si hanno in uscita oscillazioni non attenuate.

La controreazione può essere limitata scegliendo accuratamente posizione e direttività di altoparlanti e microfoni.

Cancellazione d'eco (4)

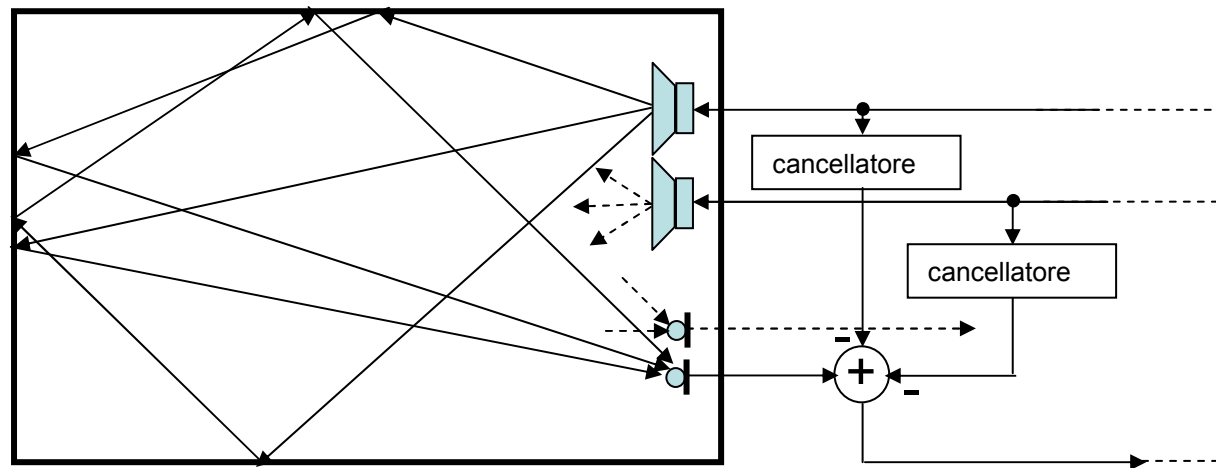
Il cancellatore stima il segnale d'eco e lo sottrae al segnale ricevuto. La stima deve avvenire in modo sufficientemente accurato ed eventualmente seguire le variazioni del canale in modo adattativo.



Si utilizzano dei filtri numerici (FIR o IIR) i cui pesi vengono adattati mediante la minimizzazione di un opportuno funzionale d'errore.

Cancellazione d'eco (5)

In un sistema di tipo stereofonico, l'ascoltatore riceverà anche un'informazione di tipo spaziale e sarà in grado di localizzare la sorgente sonora. In questo caso i percorsi sonori da identificare saranno quattro.



Il problema fondamentale è che in questo caso i due canali trasportano segnali correlati e la soluzione non è unica.

Audio 3D (1)

La conoscenza della funzione di trasferimento di un ambiente chiuso consente in linea di principio di generare anche sorgenti di tipo “virtuale”, tramite opportune schiere di altoparlanti.

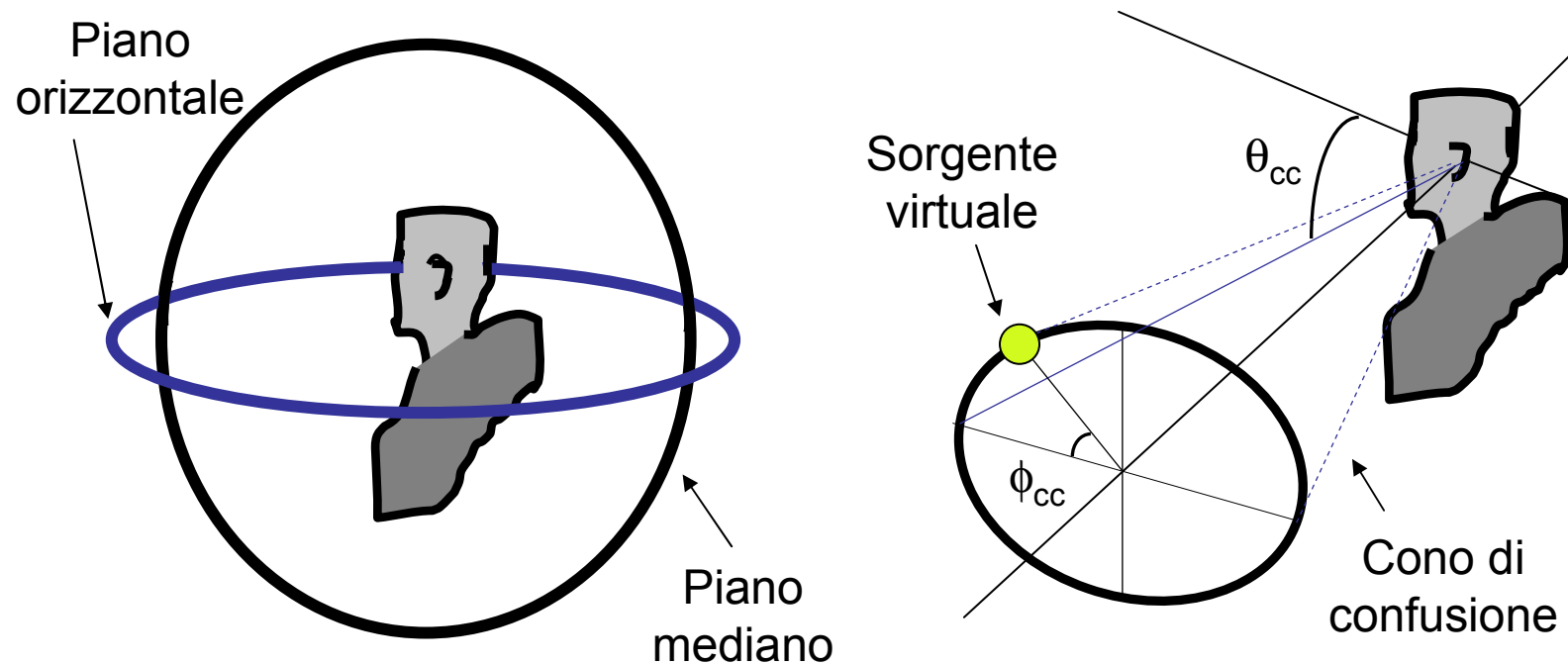
I possibili campi di applicazione sono molteplici:

- Multimedialità e intrattenimento
- Realtà virtuale (simulatori di volo, videogiochi)
- Medicina (riabilitazione post-traumatica)

Il problema fondamentale è quello di rendere la *spazialità* del campo sonoro.

Audio 3D (2)

Localizzazione della sorgente sonora: la percezione della spazialità deriva da fattori diversi (effetto precedenza, parametri monoaurali e binaurali). In particolare i parametri binaurali riguardano le differenze di livello (ILD) e le differenze temporali (ITD) tra i segnali nei condotti uditivi.



Audio 3D (3)

La posizione della sorgente e le riflessioni causate dal torace e dalla testa modificano il segnale ricevuto. La spazialità può essere introdotta mediante le “Head-Related Transfer Functions” (HRTF).



Le HRTF possono venire stimate empiricamente. I problemi principali derivano dal loro carattere fortemente individuale e dall'effetto di localizzazione “dentro alla testa”.

Audio 3D (4)

Tecniche di posizionamento di sorgenti virtuali: il panning di ampiezza

Lo stesso segnale, con ampiezze diverse, viene applicato in fase ai vari diffusori. Il segnale emesso dall'altoparlante i -esimo è dato da:

$$x_i(t) = g_i \cdot x(t)$$

La posizione della sorgente virtuale è determinata dai fattori di guadagno e giace sulla superficie delimitata dagli altoparlanti.

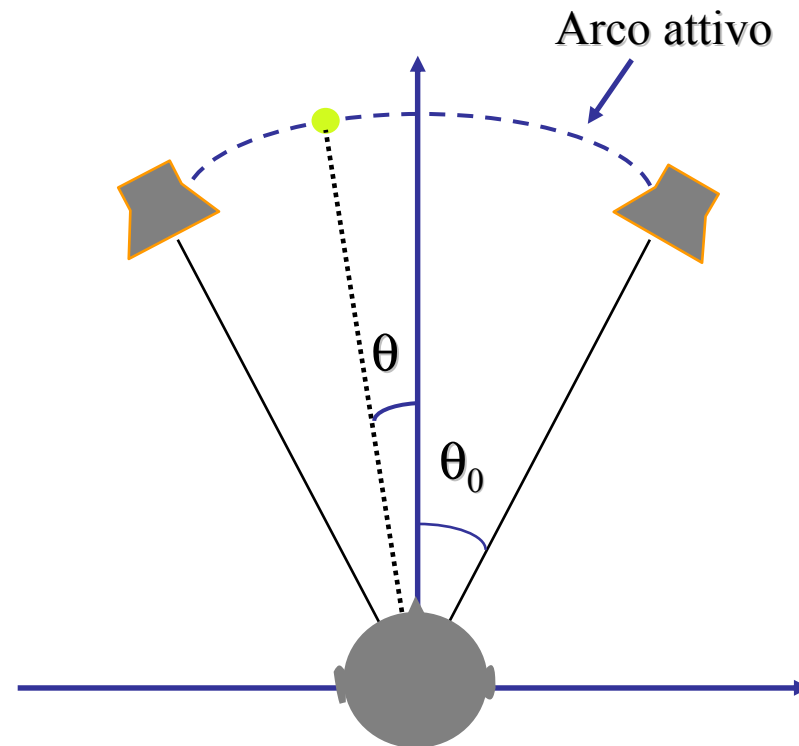
Audio 3D (5)

I fattori di guadagno sono legati alla direzione della sorgente virtuale mediante la legge del seno:

$$\frac{\sin \theta_s}{\sin \theta_0} = \frac{g_1 - g_2}{g_1 + g_2}$$

e da quella della tangente:

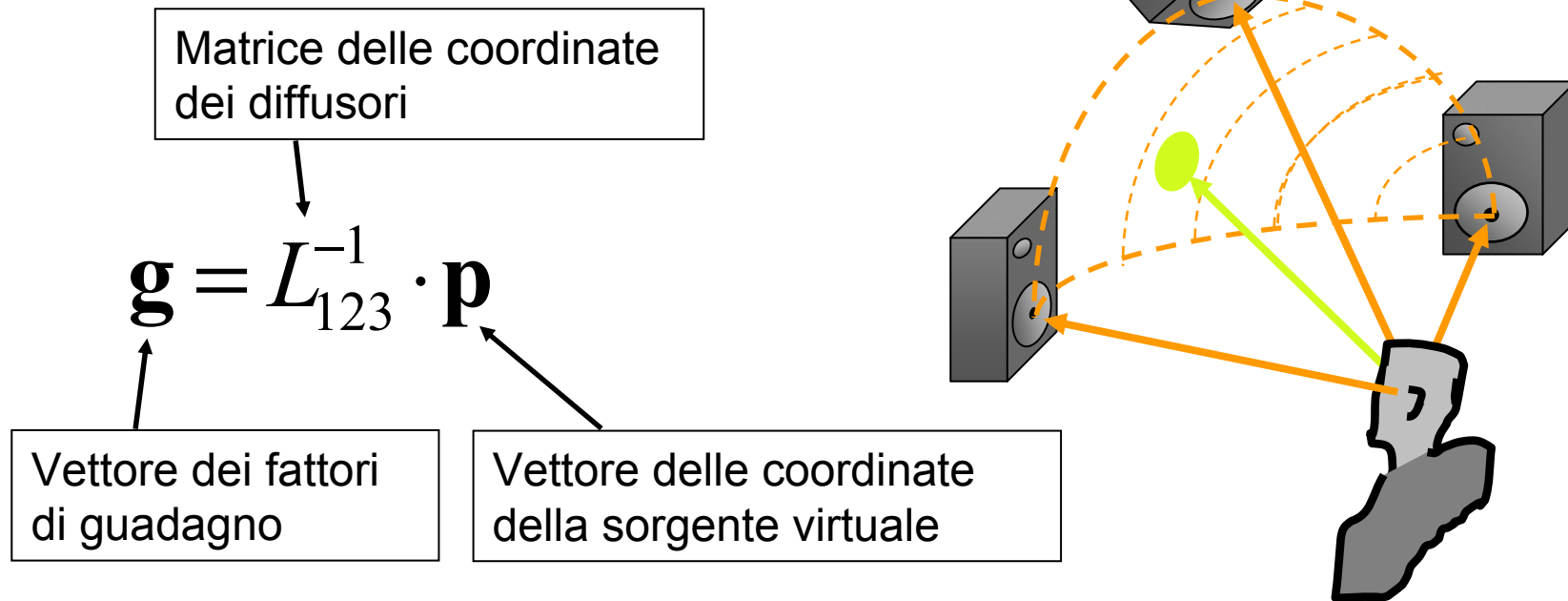
$$\frac{\tan \theta_t}{\tan \theta_0} = \frac{g_1 - g_2}{g_1 + g_2}$$



Audio 3D (6)

Panning di ampiezza vettoriale

Si tratta di una riformulazione del panning tradizionale che, grazie all'uso dei vettori, risulta più efficiente. Il sistema può essere esteso ad un numero maggiore di diffusori, purché i triangoli attivi non si sovrappongano.



Audio 3D (7)

Risultati del panning vettoriale

- Sorgenti definite e stabili all'interno della regione attiva
- Corretta percezione della direzione
- Best Listening Area limitata
- Piccoli spostamenti della sorgente virtuale all'aumentare del tempo di riverbero

Filtraggio numerico con HRTF

- Buona capacità di localizzazione sul piano orizzontale
- Sorgenti definite soprattutto lungo le direzioni laterali
- Inside-head localization per elevazioni non nulle
- Forte carattere di individualità

Localizzazione di sorgenti sonore (1)

Obiettivo

Stimare la posizione di sorgenti sonore attive all'interno di ambienti chiusi (quindi in presenza di riverbero) mediante opportune trasformazioni dei segnali acquisiti da sensori multipli (microfoni) opportunamente posizionati.

Applicazioni

- Teleconferenza
- Sorveglianza
- Telefonia senza fili
- ...

Localizzazione di sorgenti sonore (2)

Due possibili approcci

1. *Array processing* (Schiere di sensori)

La direzione di provenienza del fronte d'onda ed eventualmente la posizione della sorgente vengono stimate elaborando tutti insieme i segnali catturati dai sensori.

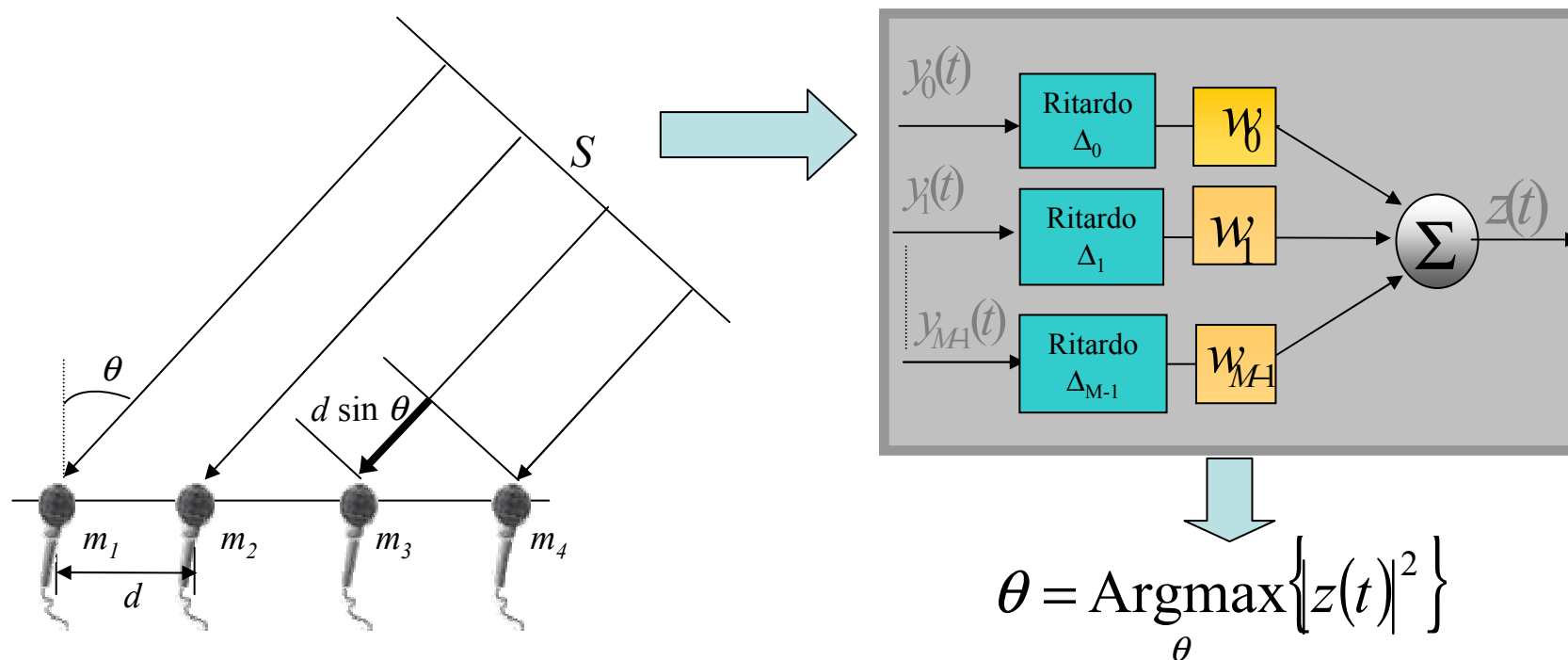
2. **Stima dei tempi di ritardo e triangolazione**

In questo caso la posizione della sorgente viene stimata attraverso due fasi successive. La prima fase consiste nella stima del ritardo relativo tra i segnali acquisiti da *coppie* di microfoni. La seconda fase fornisce la posizione stimata della sorgente a partire dall'insieme dei tempi di ritardo mediante la minimizzazione di un opportuno funzionale di errore.

Localizzazione di sorgenti sonore (3)

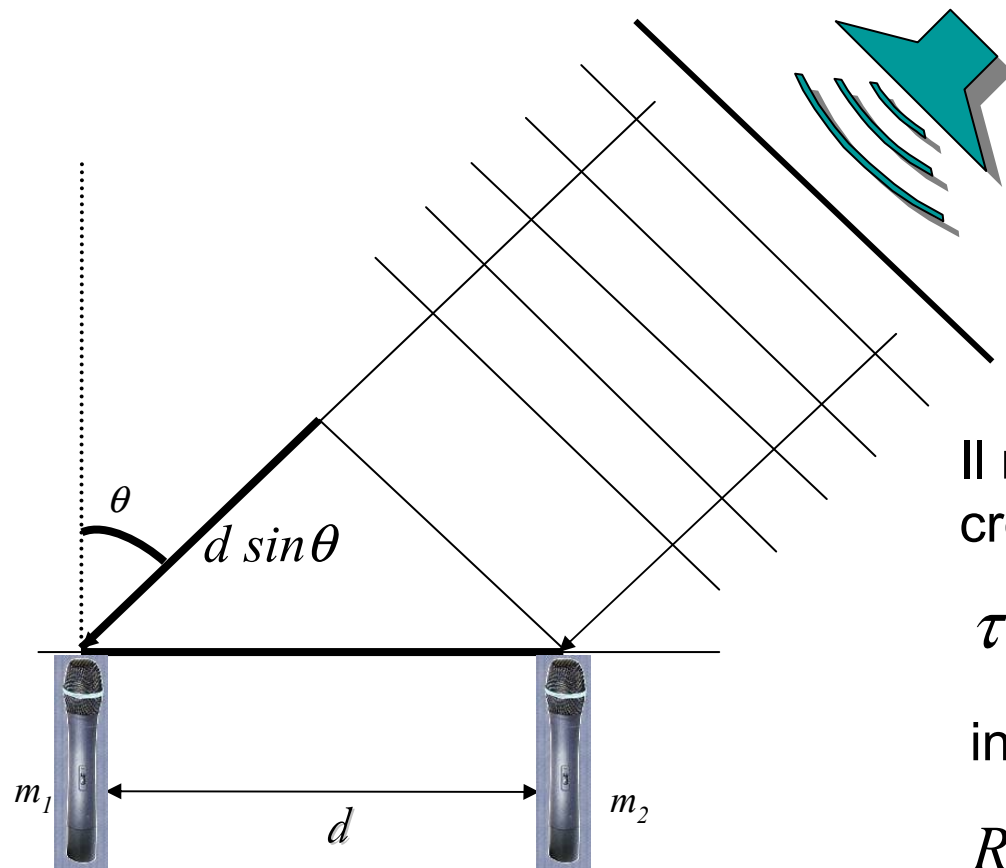
Array processing

Un esempio di array processing è il cosiddetto *beamforming*. Nella sua versione più semplice il beamforming consiste nel trovare lo sfasamento da attribuire a ciascun segnale dell'array in modo da massimizzare l'energia ricevuta (*delay and sum beamformer*).



Localizzazione di sorgenti sonore (4)

Stima dei tempi di ritardo



L'angolo θ può essere stimato a partire dal ritardo τ mediante la relazione:

$$\theta = \arcsin\left(\frac{c\tau}{d}\right)$$

Il ritardo τ può essere stimato dalla cross-correlazione dei due segnali:

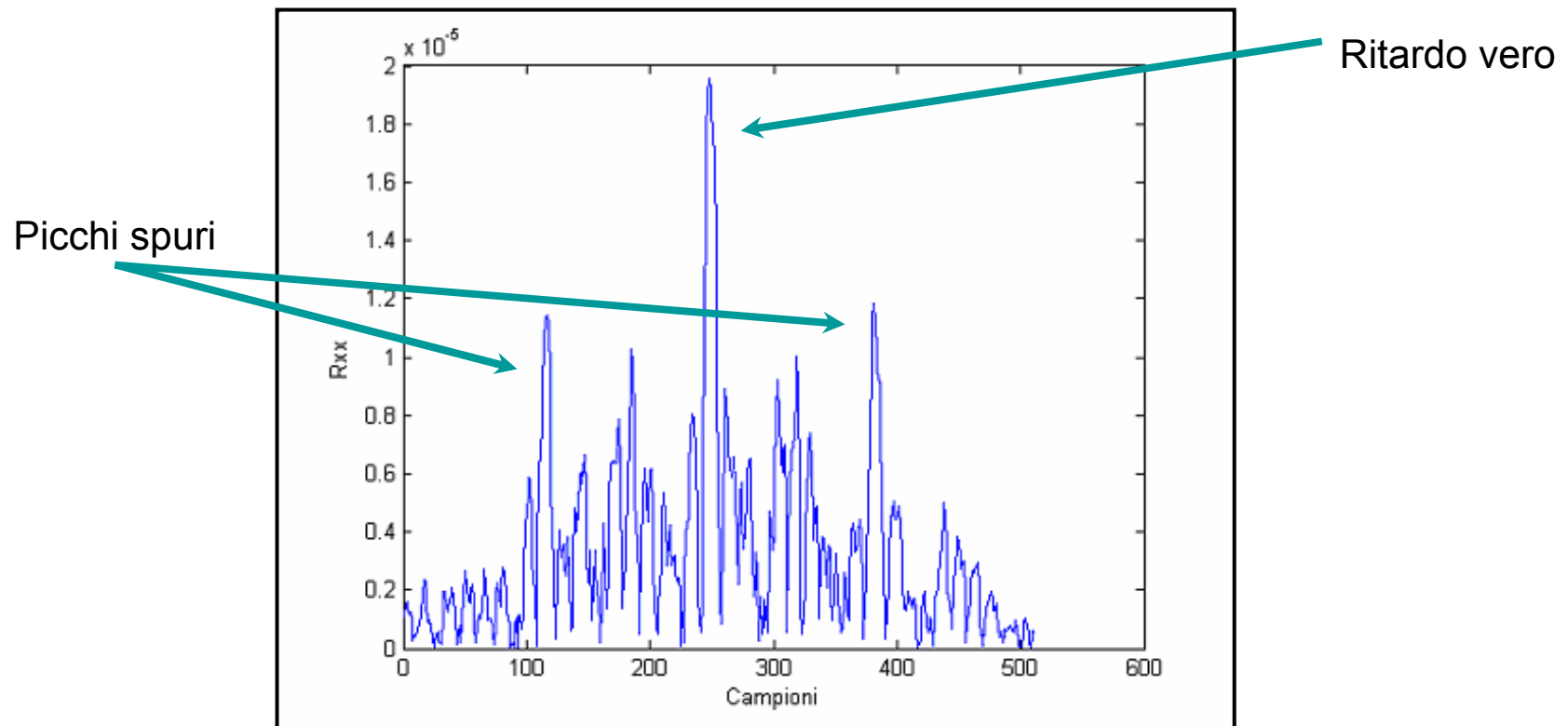
$$\tau = \underset{D}{\operatorname{Argmax}}\{R_{x_1x_2}(D)\}$$

in cui:

$$R_{x_1x_2}(D) = E\{x_1(t) * x_2(t - D)\}$$

Localizzazione di sorgenti sonore (5)

La presenza di riverbero origina picchi spuri nella funzione di cross-correlazione, rendendo meno affidabile la stima del tempo di ritardo.



Localizzazione di sorgenti sonore (6)


Per ridurre gli effetti del riverbero si usa la *cross-correlazione generalizzata* (GCC):

$$R_{x_1x_2}^{(g)}(D) = F^{-1} \left\{ \Psi_g(f) G_{x_1x_2}(f) \right\}$$

in cui F^{-1} è l'antitrasformata di Fourier, Ψ_g è una opportuna funzione di pesatura e $G_{x_1x_2}$ è lo spettro incrociato (o cross-spettro):

$$G_{x_1x_2}(f) = X_1(f) X_2^*(f)$$

Trasformate di
Fourier dei segnali



Una scelta comune è la cosiddetta “Phase Transform” (PHAT):

$$\Psi_{PHAT}(f) = \frac{1}{|G_{x_1x_2}(f)|}$$

Localizzazione di sorgenti sonore (7)

Prefiltraggio cepstrale

In presenza di tempi di riverbero elevati (in pratica superiori a circa 0,3 s) la GCC non è più sufficiente. In tal caso si ricorre ad ulteriori prefiltraggi. Un esempio è il *filtraggio cepstrale*.

$$\hat{x}(k) = F^{-1} \{ \text{Log } F \{ x(t) \} \} \quad \text{Cepstrum di } x(t)$$

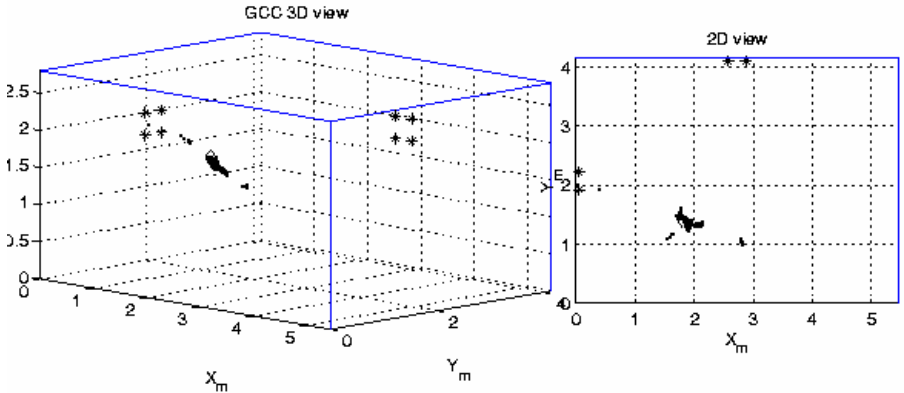
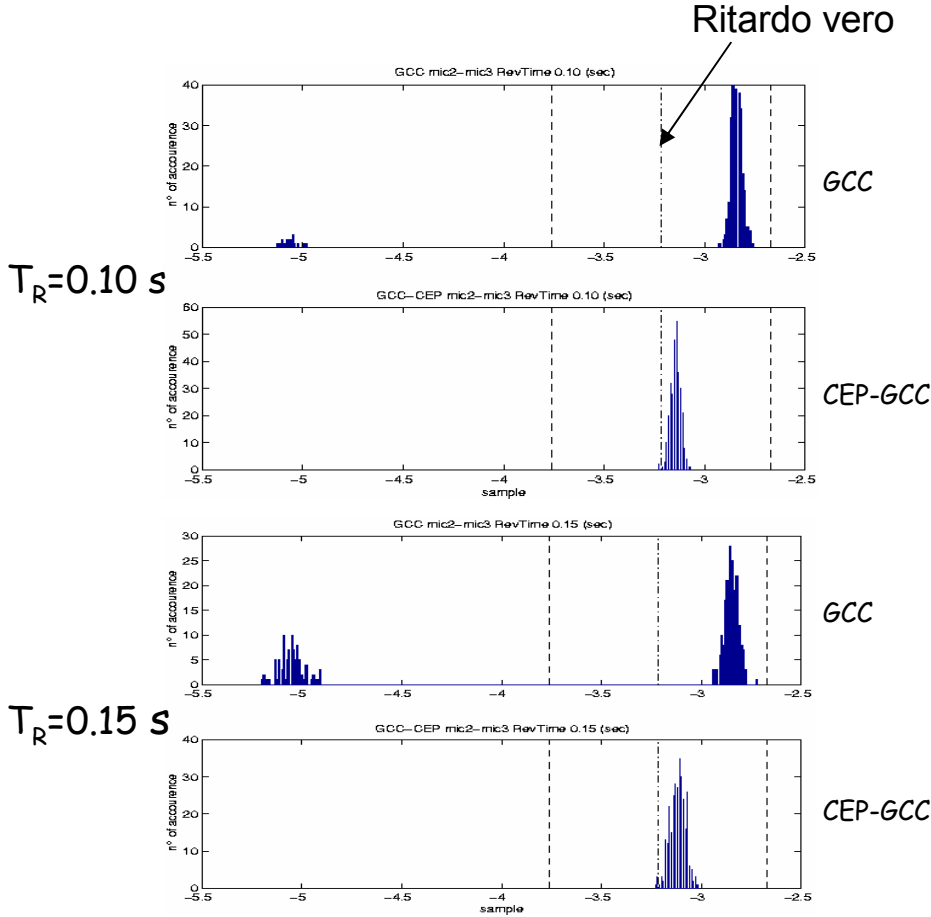
Il vantaggio è che con il logaritmo la convoluzione diventa una moltiplicazione:

$$x(t) = h(t) * s(t) \quad \longrightarrow \quad \hat{x}(k) = \hat{h}(k) + \hat{s}(k)$$

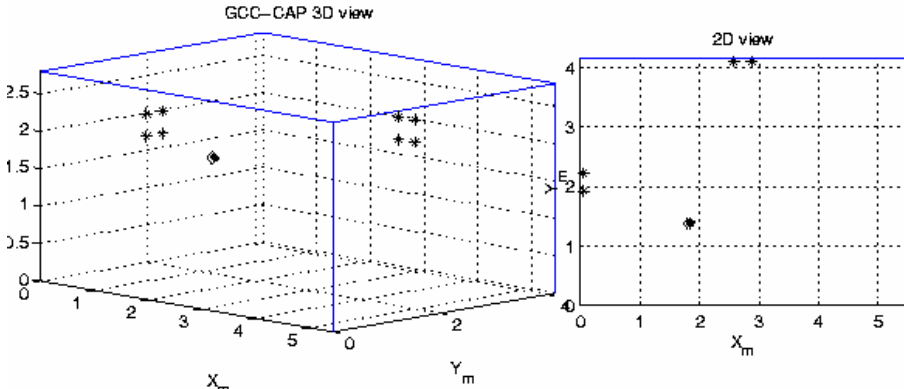
Il contributo del riverbero è contenuto nel cepstrum della risposta impulsiva. E' possibile realizzare un filtraggio nel dominio cepstrale (che diventa una sottrazione) in modo da mantenere solo l'informazione di fase.

Localizzazione di sorgenti sonore (8)

Esempi di localizzazione



GCC

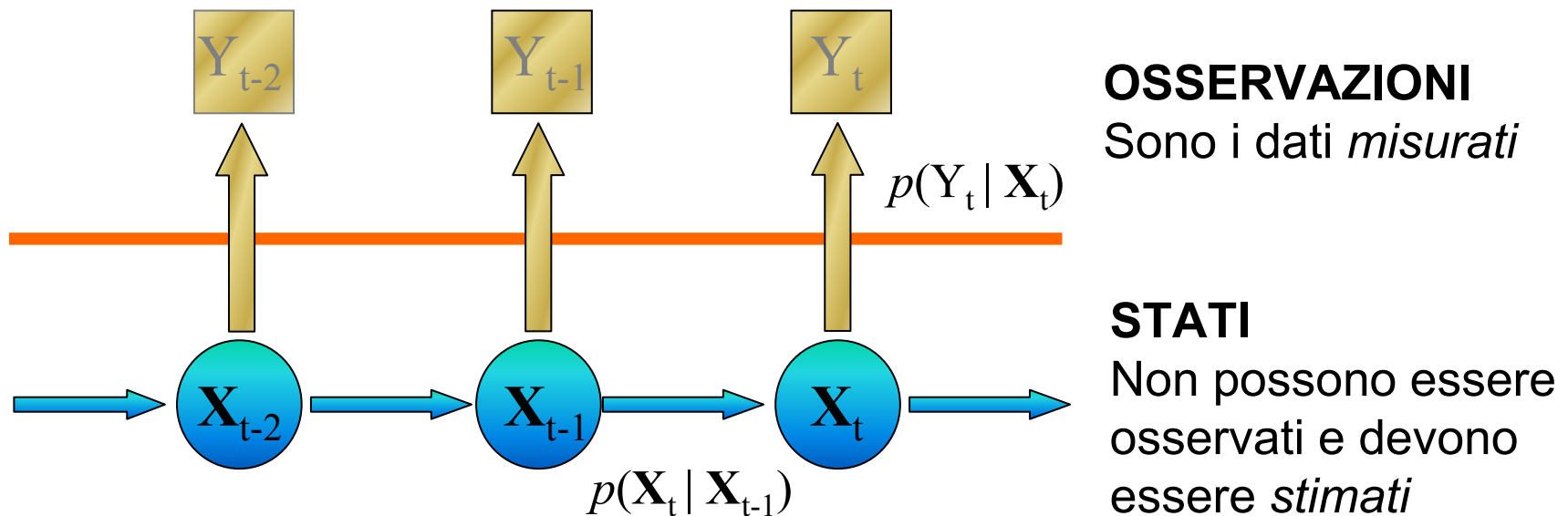


CAP-GCC

Localizzazione di sorgenti sonore (9)

Inseguimento

Se la sorgente sonora è in movimento, si deve considerare il problema del suo inseguimento (*tracking*), che consiste essenzialmente nell'identificazione dello stato di un sistema non lineare a partire da alcune osservazioni. Una classe di metodi che possono essere applicati con successo in questo caso è quella dei *metodi Montecarlo sequenziali*.



Localizzazione di sorgenti sonore (10)

Metodi Montecarlo sequenziali

- Sono utilizzati in presenza di modelli non lineari
- Consentono la stima sequenziale dell'evoluzione di un sistema dinamico (sorgente in movimento)

Particle filtering

E' un metodo Montecarlo sequenziale basato sull'uso di un insieme di entità (o *particelle*) la cui importanza relativa dipende dalla probabilità di rappresentare il vero stato del sistema.

In particolare:

1. ogni particella è una “candidata” sorgente;
2. a ciascuna particella viene associato un peso proporzionale alla probabilità di rappresentare la posizione vera della sorgente;
3. la funzione peso è ottenuta dai tempi di ritardo stimati.

Localizzazione di sorgenti sonore (11)

Il particle filtering può essere interpretata come una tecnica di ottimizzazione basata sull'esplorazione pilotata dello spazio delle variabili.

