

FILTRAGGIO A POLI ACUSTICI COMUNI PER LA LOCALIZZAZIONE BINAURALE DI SORGENTI SONORE

Danilo Comminiello, Raffaele Parisi, Simone Scardapane, Michele Scarpiniti, Aurelio Uncini

Dipartimento di Ingegneria dell'Informazione, Elettronica e Telecomunicazioni (DIET)
"Sapienza" Università di Roma
Via Eudossiana 18, 00184 Roma

Parole chiave: *localizzazione binaurale, riverbero, poli acustici comuni*

La localizzazione di sorgenti sonore per mezzo di una sola coppia di sensori trae ispirazione dal comportamento degli organismi biologici superiori, che sono in grado di percepire suoni e di elaborarli per localizzare le sorgenti sonore attive in un esteso campo di situazioni ambientali. I segnali binaurali possono essere usati secondo una modalità che si ispira al sistema uditivo umano, in modo da imitarne le capacità di localizzazione in presenza di sorgenti sonore multiple e caratteristiche variabili dell'ambiente [1].

In particolare l'ambiente di ascolto è tipicamente caratterizzato dalla presenza di riverbero, che in generale degrada le prestazioni del sistema di localizzazione [1][2]. In tal caso è necessario ricorrere ad un opportuno trattamento preliminare dei segnali [3]. Un esempio è quello del filtraggio cepstrale [4][5].

La localizzazione binaurale può avvenire mediante la stima della *differenza di livello* (*Interaural Level Difference, ILD*) e del *tempo d'arrivo* (*Interaural Time Difference, ITD*) del segnale sui due sensori. Il confronto dei valori di ILD e ITD stimati con valori prefissati di riferimento consente di sviluppare una tecnica di localizzazione piuttosto efficace [6]. Tale tecnica si rivela tuttavia limitata in presenza di livelli di riverbero anche piuttosto basso [2]. Una possibile soluzione consiste nell'utilizzo dei *Poli Acustici Comuni* (*Common Acoustical Poles, CAP*) [7][8].

Il modellamento dei *Poli Acustici Comuni* è stato introdotto come possibile metodo per il filtraggio dei contributi della funzione di trasferimento dell'ambiente (*Room Transfer Function, RTF*) dai segnali acquisiti. L'assunzione di partenza è quella secondo la quale le RTF sono dominate da un insieme di modi comuni caratterizzati da certe *autofrequenze*. Il modello proposto assume che le RTF siano approssimate da funzioni razionali secondo dei modelli ARMA, in cui i poli rappresentano le frequenze di risonanza della stanza mentre gli zeri dipendono dalla posizione relativa di sorgente e sensori. I poli acustici comuni possono essere stimati preliminarmente utilizzando un certo numero di posizioni note fissate all'interno dell'ambiente di ascolto. Successivamente il loro contributo può essere filtrato dai segnali acquisiti, prima di effettuare la stima di ILD e ITD.

Tale tecnica ha consentito un miglioramento delle prestazioni dei sistemi di localizzazione binaurale in ambienti chiusi [9]. A titolo di esempio, la figura seguente riporta il confronto tra le prestazioni di un sistema di localizzazione binaurale senza e con la tecnica di prefiltraggio descritta. In particolare vengono riportati gli istogrammi dell'angolo di azimut stimato per una posizione centrale della sorgente all'aumentare del tempo di riverbero¹. Sono evidenti i miglioramenti che si possono ottenere con la tecnica descritta, soprattutto per valori significativi del tempo di riverbero.

¹ Il tempo di riverbero è definito come il tempo necessario affinché l'energia all'interno dell'ambiente decada di 60 dB a partire dall'istante in cui il segnale ha termine.

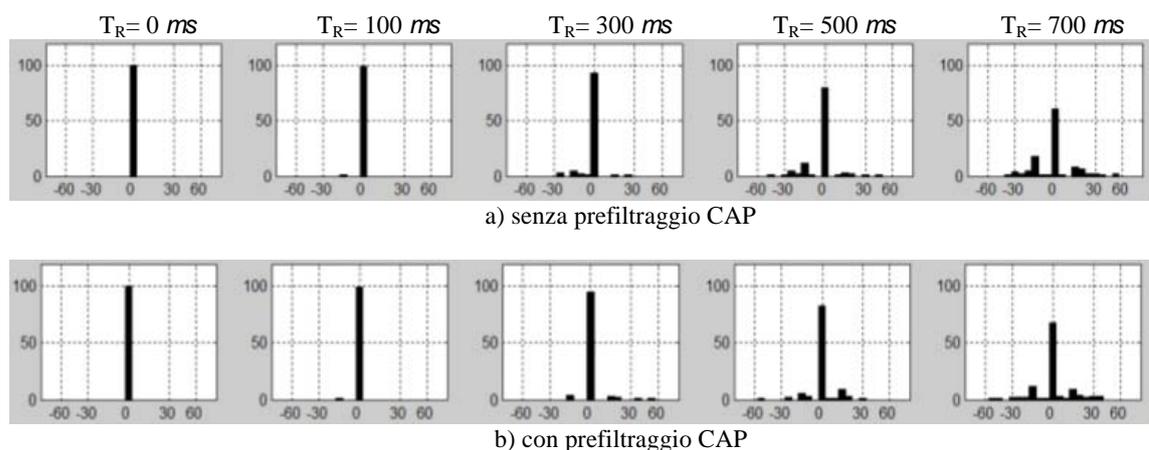


Figura 1. Confronto tra le prestazioni di un sistema di localizzazione binaurale senza e con prefiltraggio CAP all'aumentare del tempo di riverbero T_R .

BIBLIOGRAFIA

- [1] D. L. Wang and Guy J. Brown, “*Computational Auditory Scene Analysis - Principles, Algorithms and Applications*”, IEEE Press, Wiley Interscience, 2006.
- [2] C. M. Zannini, R. Parisi, A. Uncini, “Binaural sound source localization in the presence of reverberation,” *17th International Conference on Digital Signal Processing (DSP 2011)*, Corfù (Grecia), 5-7 July 2011.
- [3] R. Parisi, R. Gazzetta, and E.D. Di Claudio, “Prefiltering approaches for time delay estimation in reverberant environments,” in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP’02)*, 2002, vol. 3, pp. 2997–3000.
- [4] A. Stephenne and B. Champagne, “A new cepstral prefiltering technique for estimating time delay under reverberant conditions,” *Signal Processing*, vol. 59, no. 3, pp. 253–266, 1997.
- [5] R. Parisi, F. Camoes, M. Scarpiniti and A. Uncini, “Cepstrum Prefiltering for Binaural Source Localization in Reverberant Environments,” *IEEE Signal Processing Letters*, Vol. 19, N. 2, pp. 99-102, February 2012.
- [6] M. Raspaud, H. Viste, and G. Evangelista, “Binaural source localization by joint estimation of ILD and ITD,” *IEEE Trans. On Audio, Speech and Language Processing*, vol. 18, no. 1, pp. 68–77, 2010.
- [7] Y. Haneda, S. Makino, and Y. Kaneda, “Common acoustical pole and zero modeling of room transfer functions,” *IEEE Trans. on Speech and Audio Processing*, vol. 2, no. 2, pp. 320–328, April 1994.
- [8] Y. Haneda, S. Makino, and Y. Kaneda, “Multiple-point equalization of room transfer functions by using common acoustical poles,” *IEEE Trans. on Speech and Audio Processing*, vol. 5, no. 4, pp. 325–333, July 1997.
- [9] R. Parisi, D. Comminiello, M. Scarpiniti and A. Uncini, “Common pole prefiltering for binaural sound source localization in reverberant environments,” accettato a *Fifth International Conference on Information, Intelligence, Systems and Applications (IISA 2014)*, July 07 – 09, 2014, Chania Crete, Greece.