

SINTESI DI CIRCUITI PER L'INVERSIONE DI SISTEMI DIAFONICI: ANALISI DELLE COMPONENTI INDIPENDENTI

Danilo Comminiello, Raffaele Parisi, Michele Scarpiniti, Aurelio Uncini

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

Negli ultimi anni il problema della separazione e della deconvoluzione di sorgenti tra loro indipendenti ha suscitato un elevato interesse in una vasta area della comunità scientifica testimoniato anche da molte pubblicazioni su tale tematica [1]. Tale metodologia ha assunto un ruolo di primaria importanza alla luce dei molteplici settori applicativi quali, per esempio: le comunicazioni digitali, il miglioramento della qualità dei segnali (immagini, video, audio, ecc), equalizzazione/ricostruzione di segnali, la diagnosi medica come EEG e ECG, sistemi multi sensore, ecc; l'analisi di segnali geofisici, ambientali, serie economiche, ecc.

La ricerca in corso da alcuni anni presso l'Unità di Roma "La Sapienza", riguarda le problematiche relative alla separazione e alla deconvoluzione di sorgenti tra loro statisticamente indipendenti. In particolare lo scopo della presente ricerca consiste nello sviluppo di algoritmi robusti per la separazione di sorgenti vocali che possano funzionare in modo affidabile anche in ambienti altamente riverberanti, dove gli algoritmi classici di separazioni di sorgenti falliscono.

In particolare, sono stati presi in considerazione algoritmi di separazione basati sulle trasformazioni tempo-frequenza in modo da poter convertire un problema convolutivo in n problemi istantanei [2] che lavorano su blocchi più o meno lunghi di segnale a valori complessi. Il numero n di analisi da effettuare è direttamente proporzionale alla banda del segnale preso in considerazione.

Il problema fondamentale è dovuto alla complessità computazionale quando si tiene conto della risposta all'impulso particolarmente lunga che descrive gli ambienti reali caratterizzati da un elevato riverbero. Infatti, la lunghezza dei filtri di deconvoluzione deve essere sufficientemente lunga da coprire la coda di riverberazione dell'ambiente, anche se una finestra particolarmente lunga tende a catturare le non-stazionarietà del segnale e rende difficile la stima delle proprietà statistiche dei segnali, indispensabili per ottenere una buona separazione, dovute all'"effetto somma" delle trasformazioni tempo-frequenza, che producono campioni con statistica più gaussiana [3].

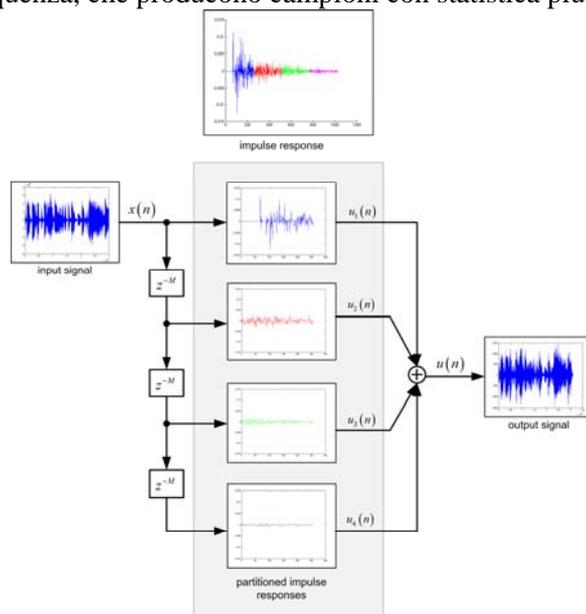


Figura 1 Risposta impulsiva partizionata

E' stata quindi proposta una particolare architettura di de-mixing che è in grado di ridurre il costo computazionale adottando un'elaborazione a blocchi. Un'implementazione efficiente può essere ottenuta dividendo (partizionando) la convoluzione in una somma di convoluzioni più corte (Fig. 1), ognuna delle quali coinvolge un numero minore di termini e risolve oltre al problema del costo computazionale, anche il problema della cattura delle non-stazionarietà [4]. In questo modo si riduce, inoltre, la latenza dell'algoritmo, cioè il forte ritardo tra l'ingresso e l'uscita, rendendo l'applicazione adatta all'impiego in tutte quelle situazioni in cui è necessario che l'elaborazione avvenga con esigenza di *real time*, come ad esempio nei sistemi di comunicazione "hands-free".

Quindi è stata inserita l'idea degli algoritmi partizionati nel contesto della separazione di sorgenti. A tal proposito è stata ottenuta un'architettura circuitale che implementa una nuova versione del noto algoritmo di Smaragdis per la separazione in frequenza (Fig. 2) [5], [6]. L'algoritmo proposto è stato abile di ottenere ottime prestazioni in confronto agli algoritmi standard, riducendo notevolmente la latenza dell'architettura. Inoltre tale soluzione riesce a risolvere il problema fondamentale della separazione in frequenza, ovvero svincola le prestazioni dell'algoritmo dalla scelta della lunghezza della finestra di elaborazione, punto critico degli altri algoritmi disponibili in letteratura.

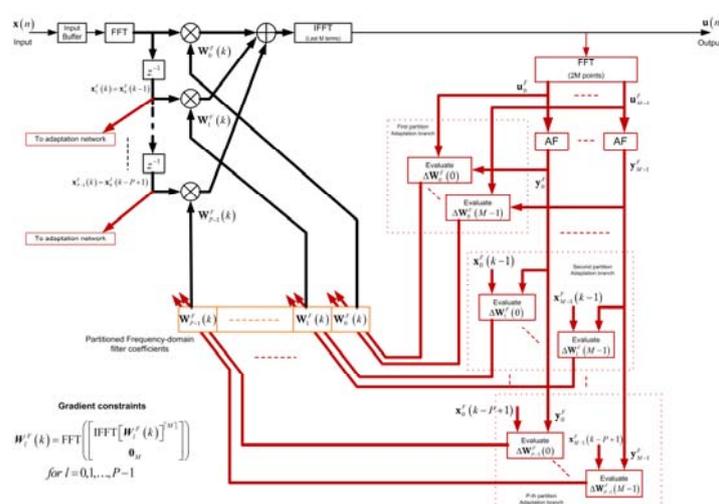


Figura 2 Architettura partizionata proposta (PFBSS)

Referenze

1. A. Cichocki, and S. Amari, *Adaptive Blind Signal and Image Processing*, Wiley, 2002.
2. R. Mukai, H. Sawada, S. Araki and S. Makino, (2006), "Frequency-Domain Blind Source Separation of Many Speech Signals Using Near-Field and Far-Field Models", *EURASIP Journal on Applied Signal Processing* **2006**, 1-13.
3. S. Araki, R. Mukai, S. Makino, T. Nishikawa and H. Saruwatari, (2003), "The fundamental limitation of Frequency Domain Blind Source Separation for convolutive mixtures of speech", *IEEE Transaction on Speech and Audio Processing* **11**(2), 109-116.
4. Farhang-Boroujeny, B. (1996), 'Analysis and efficient implementation of partitioned block LMS adaptive filters', *IEEE Transactions on Signal Processing* **44**(11), 2865-2868.
5. M. Scarpiniti, A. Picaro, R. Parisi and A. Uncini, "A Partitioned Frequency Block Algorithm for Blind Separation in Reverberant Environments" in *Neural Nets WIRN09* (B. Apolloni, S. Bassis, and C. Morabito, eds.), pp. 81-90, IOS Press, Amsterdam, 2009.
6. M. Scarpiniti, A. Picaro, R. Parisi and A. Uncini, "A Partitioned Frequency Domain Algorithm for Convolutive Blind Source Separation", in *Proc. of IEEE International Workshop on Machine Learning for Signal Processing (MLSP'09)*, pp. 1-6, 2009.