

Il restauro digitale del segnale audio

di *Piero Cosi**

I

Sommario

L'introduzione dei nuovi media digitali di alta qualità audio, quali il *Compact Disc* (CD) e il nastro audio digitale (*Digital Audio Tape* – DAT), ha alimentato notevolmente le aspettative sulla qualità del suono, in tutti i tipi di registrazioni audio e determinato progressive modificazioni della sensibilità uditiva. Assieme al sempre più vivo ed attuale interesse per gli archivi di materiale sonoro relativi a registrazioni musicali, discorsi storici o altro materiale audio, il “ripristino” delle sorgenti degradate si è trasformato in un campo di ricerca sempre più attraente ed importante. In questo lavoro si prendono in considerazione le applicazioni delle più moderne tecniche d'elaborazione dei segnali, per il “mantenimento” ed il “ripristino” delle sorgenti audio degradate. Oltre a fornire una descrizione delle attuali metodologie, si è cercato di esplorare gli attuali campi di ricerca e le tendenze future. Queste tecniche, basate essenzialmente sull'elaborazione digitale del segnale audio o su una sua descrizione statistica, sono essenziali per tutti coloro che sono interessati alla registrazione musicale, o in generale alla registrazione audio e all'industria di *remaste-*

* IFD-CNR, Istituto di Fonetica e Dialettologia – Consiglio Nazionale delle Ricerche.

ring e *restoring* rivolta soprattutto alla creazione di nuove librerie multimediali.

2

Introduzione

Due sono essenzialmente i motivi che fanno del “Restauro del Segnale Audio” (RSA) una linea di ricerca sempre più attuale ed attraente. Alle esigenze archivistiche di un patrimonio culturale di vecchie registrazioni, che rischia purtroppo di andare perduto, a causa del continuo e costante degrado dei supporti fisici che contengono il segnale audio, si unisce il sempre più vivo interesse del pubblico per le esecuzioni storiche e i documenti del passato, indotto dall’alta qualità audio di supporti come il CD o il DAT e dalla perfezione raggiunta nelle registrazioni audio. Il senso che viene attribuito al termine *Digital Audio Restoration* (DAR) è, infatti, ambiguo e dipende dal fine stesso che si desidera ottenere mediante questa tecnica d’elaborazione del segnale audio.

Una prima interpretazione, legata direttamente alle esigenze archivistiche, è essenzialmente quella per cui, con tale procedura, si cerca di ristabilire, il più precisamente possibile, “il suono” della registrazione originale. In tal caso, ogni elaborazione viene applicata con lo scopo di ottenere una copia esatta, generalmente digitale, del segnale conservato analogicamente sul supporto originale. In questa situazione non si può parlare di vero e proprio “restauro” o “ripristino”, ma piuttosto di “trasferimento” dal supporto analogico a quello digitale (*Digital Remastering* - DR). Di conseguenza, tutti i rumori, i sibili, i fruscii, i *crackle*, i ronzii presenti sul supporto analogico originale, saranno udibili anche nel nuovo supporto digitale.

Una seconda linea di pensiero è quella per cui, mediante il DAR, si deve cercare di rimanere il più possibile aderenti alle “esecuzioni originali” piuttosto che alla registrazione vera e propria. L’idea quindi è quella di trasformare una vecchia registrazione in una registrazione “attuale”. In altre parole, l’obiettivo è quello di eliminare “il suono” corrispondente alla vecchia registrazione e di ristabilire l’“esecuzione” originale. Per far questo si rende necessaria l’applicazione di alcune procedure per la pulizia di tutti i

rumori sovrapposti all'esecuzione originale, indotti sul supporto, dall'uso e da vari agenti esterni. In altre parole, il desiderio d'alta fedeltà e la nostalgia per le esecuzioni storiche si fondono assieme per superare i limiti di una riproduzione storicamente fedele e per ricostruire il "vero suono di un interprete". Se il valore artistico di una registrazione è fortemente compromesso dal rumore dovuto al deterioramento, il solo mantenimento digitale non è sufficiente ed il "restauro" diviene necessario.

³ "Rimasterizzazione" del segnale audio

Nello sviluppo delle tecnologie audio digitali sembra quasi che il processo di riversamento, vale a dire il procedimento per giungere ad una copia ottimale di supporti analogici originali, non sembra essere oggetto di adeguata attenzione. Questo è vero in particolare per la digitalizzazione di contenuti registrati su vecchi supporti analogici quali il nastro magnetico, i dischi di vinile o altri supporti simili. Sembra quasi che si parta dall'assunto che il trasferimento di materiali conservati su questi vecchi supporti analogici, comporti solo la necessità di mettere il supporto "vecchio" in un'opportuna apparecchiatura e premere il bottone giusto. Al contrario, la riproduzione di un nastro, ad esempio, è un processo complesso e gioca un ruolo fondamentale nella riscoperta della qualità del suono originariamente registrato. Specialmente nel campo del restauro è indispensabile prendere in considerazione la riproduzione ottimale, definitiva, dei segnali audio. Ogni errore e ogni omissioni fatti in questo primo passo cruciale non saranno senza conseguenze, perché per lo più non possono essere compensati nel successivo processo di restauro digitale. Particolare attenzione deve essere rivolta alla risoluzione digitale adeguata per un corretto riversamento, senza "perdite" di qualità, del segnale audio originale, sul nuovo supporto digitale. Fondamentale poi risulta essere la questione di quale informazione, contenuta in registrazioni analogiche su supporti magnetici o d'altro tipo, vada perduta nel corso di trasferimenti convenzionali e conseguentemente quali siano le modalità e le elaborazioni (ad esempio le tecniche di equalizzazione) con cui poterla recuperare una volta creato il nuovo segnale digitale.

Semplificando notevolmente, i tre passi fondamentali per una corretta e funzionale rimasterizzazione digitale del segnale audio possono essere così riassunti:

1. Utilizzare la miglior sorgente analogica disponibile.
2. Ottimizzare la riproduzione del suono. Ad esempio, utilizzando opportune equalizzazioni che, senza snaturare la registrazione originale, compensino i disturbi indotti sui supporti analogici dal tempo e dall'usura.
3. Convertire il suono su supporto digitale mediante un buon convertitore analogico/digitale (24 bit) ed assicurarsi che i livelli siano correttamente impostati, al fine di evitare saturazioni e distorsioni sul suono risultante.

4 “Restauro” del segnale audio

Le cause del deterioramento delle registrazioni audio possono avere influenza locale sul segnale, come polvere o graffi, che danno origine prevalentemente a rumori di tipo impulsivo quali sibili, fruscii, *crackle*, ronzii, o influenza globale e influire sull'intero segnale, come ad esempio il rumore bianco, le variazioni dell'altezza dei suoni o le distorsioni ed è in questo scenario che diviene sempre più importante individuare nuove e sempre più raffinate tecniche per il restauro audio digitale. Di seguito sono elencate alcune delle più comuni tecniche di elaborazione applicate sul segnale audio al fine di migliorarne le caratteristiche. Per ogni tecnica esistono innumerevoli varianti in letteratura e questo lavoro si prefigge, esclusivamente, l'obiettivo di illustrare le procedure standard ed il loro funzionamento generale, senza pretendere di esaurire la trattazione del restauro audio digitale in dettaglio. Alcune tecniche sono basate su semplici procedure euristiche, altre su sofisticati modelli probabilistici associati ad opportune procedure di stima e altre ancora sfruttano nuovi modelli ad apprendimento basati su reti neurali artificiali o sull'approccio Bayesiano. Ad ogni procedura, sono associati anche i termini inglesi con cui tali elaborazioni sono comunemente indicate all'interno di testi scientifici di riferimento o di specifici software.

4.1

Controllo automatico del guadagno
(AGC - *Automatic Gain Control*)

Mediante tale procedura si modifica l'ampiezza complessiva del segnale audio *target* allo scopo di mantenerne ottimale il livello. Vi sono molte varianti e ad esempio mediante la procedura AGC-RMS (*Root Mean Square*) si modifica il livello complessivo del segnale in modo che il valore RMS del segnale stesso risulti approssimativamente di 20dB inferiore al livello massimo di un segnale registrabile digitalmente su 16 bit quantizzati linearmente. Con questa elaborazione in pratica si stima la potenza media del segnale nelle regioni di "non-silenzio" poiché, infatti, le regioni di silenzio sono escluse dal calcolo.

4.2

Compressione d'ampiezza
(AC - *Amplitude Compression*)

Questa procedura modifica le ampiezze relative di differenti regioni del segnale. Si cerca, infatti, di "parificare" il livello del segnale aumentandone l'ampiezza nelle regioni più deboli e lasciandola inalterata in quelle più forti. Esistono anche per questa procedura innumerevoli varianti. Una possibile implementazione, ad esempio, divide il segnale in piccole finestre sovrapposte e applica su ciascuna una mappatura di ampiezza di tipo "μ-law" come illustrato in FIG. 1. Se l'ampiezza di una finestra è al di sotto di 40dB rispetto alla finestra più intensa la sua ampiezza è modificata secondo la funzione di compressione scelta. Valori più elevati di μ aumentano il guadagno e tendono quindi a rendere il segnale più simile in ampiezza. L'ampiezza massima del segnale rimane in pratica inalterata poiché sono le regioni più deboli ad essere maggiormente interessate dal procedimento. Una piccola quantità di compressione aumenta, ad esempio, l'intelligibilità del segnale vocale quando questo viene presentato in situazioni in cui è simultaneamente presente un elevato rumore di sottofondo.

FIGURA 1
Funzione μ -law di compressione d'ampiezza.

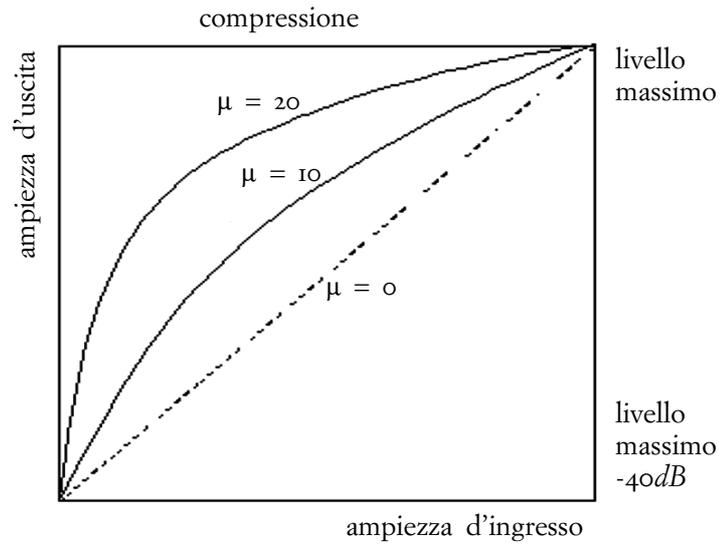
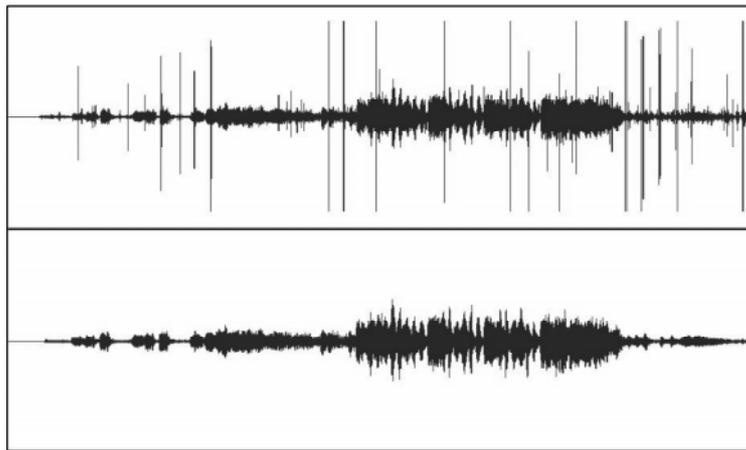


FIGURA 2
Esempio dell'applicazione della procedura *DeClick*, per l'eliminazione dei click in una registrazione di un brano musicale



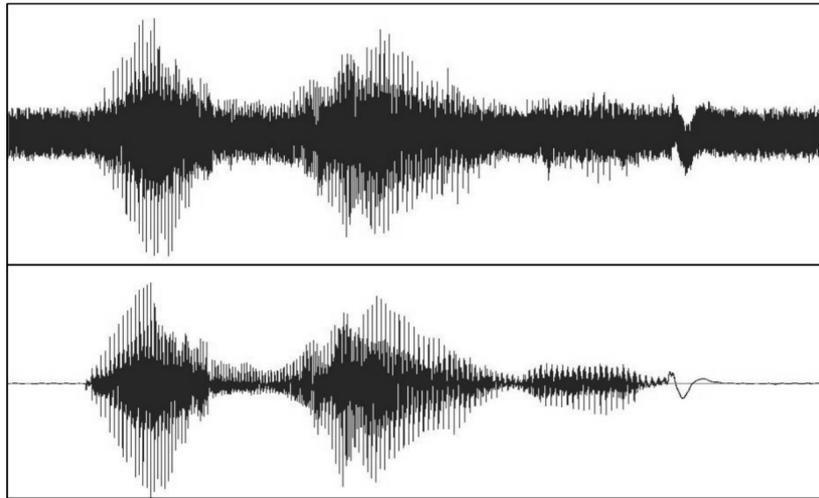
4.3 Eliminazione dei “click” (*DeClick*)

Con il termine “DeClick”, ci si riferisce generalmente ad una procedura per l’eliminazione dei disturbi di carattere impulsivo. Questa procedura si basa essenzialmente sulle indicazioni fornite da un rivelatore statistico di fenomeni “outlier”, cioè di fenomeni non “in linea” con le normali caratteristiche del segnale audio in esame. In altre parole, si ricercano i fenomeni “impulsivi” estranei alle normali caratteristiche del segnale e, una volta individuati, si passa ad una loro progressiva eliminazione per riportare il segnale audio alla sua purezza originale. Un esempio dell’applicazione di tale procedura è illustrato in FIG. 2.

4.4 Eliminazione di “rumore a banda larga” di tipo statico e predefinito (*DeNoise*)

Con il termine *DeNoise*, ci si riferisce generalmente ad una procedura per l’eliminazione di un rumore a banda larga le cui caratteristiche principali siano “statiche” e soprattutto note a priori. La procedura completa è quindi generalmente suddivisa in due fasi. In una prima fase, si determina il contenuto spettrale di un frammento di segnale in cui sia presente solo il rumore in oggetto (ad esempio un piccolo “silenzio” precedente o successivo al segnale audio target in cui è presente solo il rumore di fondo), e, in una seconda fase, si procede alla sottrazione, da tutto il segnale, dello spettro del rumore precedentemente calcolato. Dopo aver opportunamente suddiviso il segnale in finestre sovrapposte di opportune dimensioni si opera la sottrazione nel dominio della frequenza e si riconverte successivamente il segnale nel dominio del tempo. Qualora il rumore di riferimento non fosse facilmente descrivibile, perché ad esempio non esistono frammenti di segnale sufficientemente estesi in cui sia presente solo la componente di rumore, lo spettro da sottrarre può essere ovviamente estratto da segnali simili per natura al segnale in esame. Vi è da sottolineare che, mentre è evidentemente chiaro, a livello percettivo, che il segnale risultante è sicuramente più piacevole all’ascolto, non esi-

FIGURA 3
 Esempio dell'applicazione della procedura *DeNoise*, per l'eliminazione di un rumore bianco sovrapposto ad un segnale vocale target ("aiuole").



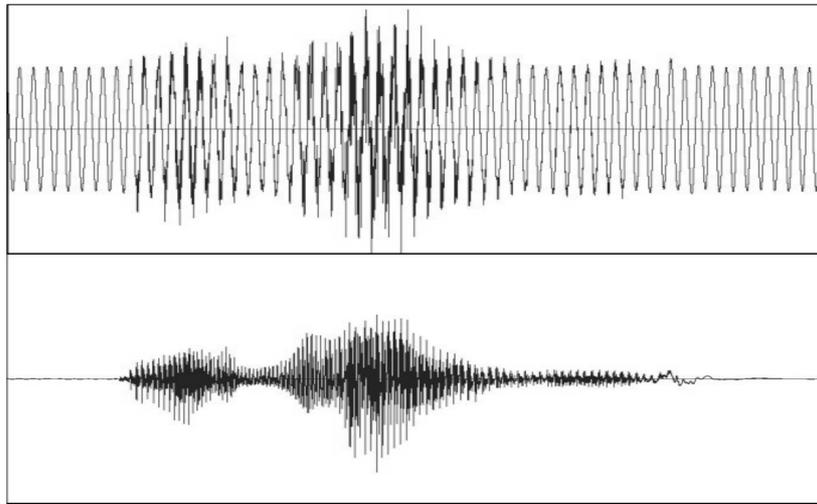
stono riscontri effettivi di un suo aumentato livello di intelligibilità. Questa procedura introduce a volte un fastidioso effetto di disturbo di tipo "musicale" (*musical noise*) che deteriora la qualità del segnale prodotto, per cui in letteratura sono apparse numerose modifiche al procedimento base di sottrazione spettrale tutte rivolte all'eliminazione di questo fenomeno. Un esempio dell'applicazione di questa procedura è illustrato in FIG. 3.

4.5

Eliminazione di "rumore a banda larga" (*DeHiss*)¹

Con il termine *DeHiss*, ci si riferisce generalmente ad una procedura per l'eliminazione di un rumore a banda larga basato su un modello standard del rumore stesso. Questa elaborazione può essere ovviamente applicata quando "l'impronta del rumore" (*noiseprint*) non è disponibile per un qualsivoglia motivo. Qualora, ad

FIGURA 4
Esempio dell'applicazione della procedura *DeHum*, per l'eliminazione del disturbo a 50Hz sovrapposto al segnale target ("aiuole").



esempio, le caratteristiche del rumore sovrapposto varino nel tempo, questa procedura può portare a risultati migliori della tecnica di *DeNoise* precedentemente illustrata, essendo meno selettiva e anche meno sensibile alle variazioni del profilo del rumore. Anche per questo metodo, come per il precedente, sono apparse molte varianti volte principalmente alla risoluzione del problema relativo al *musical noise*. Questa procedura, come d'altronde la precedente, nonostante non sia stata progettata per trattare esplicitamente i rumori di tipo impulsivo, è a volte sufficiente per ridurre un certo tipo di click, specialmente se di piccola intensità.

4.6 Eliminazione di "disturbi armonici" (*DeHum*)

Con il termine *DeHum*, ci si riferisce generalmente ad una procedura per l'eliminazione di disturbi o rumori di tipo armonico quali ad esempio il classico disturbo della portante a 50Hz o altri di-

sturbi di caratteristiche simili. Questo tipo di disturbi può facilmente essere eliminato mediante un filtro *comb* che è un insieme di filtri *notch* (a picco) progettati al fine di sopprimere il disturbo “portante” e le sue armoniche dominanti. In FIG. 4 è illustrato un esempio dell’applicazione della procedura *DeHum* su un segnale vocale (“aiuole”) a cui è stato sovrapposto un rumore a 50Hz.

4.7
Eliminazione dei disturbi con filtraggio complesso
(*Complex-Filtering*)

Questa procedura si riferisce a tutti quei metodi in cui il “filtro” utilizzato per la ricostruzione del segnale è progettato e personalizzato a piacere dall’utente.

4.8
Restauro fonetico (*Phonetic Enhancement*)

Con questo termine ci si riferisce generalmente ad una procedura per la modifica selettiva dell’ampiezza di porzioni differenti del segnale vocale. Diversamente da tutti gli altri metodi, questa procedura è stata progettata specificatamente per il segnale vocale, e le sue funzioni differiscono a seconda del contenuto fonetico del segnale elaborato. In fatti, con questo metodo si possono amplificare selettivamente le regioni del segnale che possono migliorare la differenziazione delle varie categorie fonetiche aumentando quindi l’intelligibilità complessiva del messaggio verbale. Ad esempio è noto che le consonanti fricative /f, s, v, z, ʃ/ o nasali /m, n, ŋ/ sono molto più deboli in ampiezza di molte altre consonanti e vocali e possono essere facilmente mascherate, ad esempio, dal rumore di fondo. L’implementazione di un “restauro fonetico” consente, quindi, di agire in modo selettivo sulle differenti classi fonetiche, aumentandone o diminuendone l’ampiezza, al fine di ottimizzare la discriminazione fonetica. Vari esperimenti hanno dimostrato che con tale tecnica si può effettivamente aumentare l’intelligibilità del messaggio verbale.

5 Considerazioni finali

Riassumendo, si può affermare che il restauro audio digitale viene progettato ed applicato al fine di “ripulire” una registrazione audio, allo stesso modo in cui il restauro pittorico è progettato per “ripulire” i quadri antichi dalla patina del tempo. L’obiettivo finale è sempre quello di separare il “suono” dal supporto che lo conserva e lo trasmette, per consentire, ad esempio, di poter ascoltare la musica senza tutti i disturbi, indotti dal tempo e dall’uso, che contagiano e corrompono le registrazioni originali. Bisogna però attentamente considerare che, in alcuni casi, il disturbo che si desidera eliminare è così simile, nelle sue caratteristiche, al suono da restaurare, che deve essere mantenuto per non sacrificare, distorcendole e snaturandole, le parti del suono che desideriamo mantenere.

Bibliografia

- S. J. GODSILL, P. J. W. RAYNER, O. CAPP (1998), *Digital audio restoration*, in M. Kahrs, K. Brandenburg (eds.), *Applications of Digital Signal Processing to Audio and Acoustics*, Kluwer, Boston (MA).
- S. J. GODSILL, P. J. W. RAYNER (1998), *Digital Audio Restoration*, Springer, Berlin.
- G. ADAMO, G. B. DEBIASI, G. DE POLI, P. GIUA, G. A. MIAN, G. RE, M. C. SOTGIU, A. VIDOLIN (1997), *Problemi di conservazione e restauro di archivi sonori*, in *Atti del Congresso Associazione Italiana di Acustica*, maggio 1997, Perugia, pp. 31-8.
- G. A. MIAN (1997), *Metodologie di restauro di archivi sonori*, *Atti AES (Italian Chapter)*, Milano, ottobre.
- M. NIEDZWIECKI (2000), *Identification of time-varying processes in signal processing*, John Wiley & Sons, LTD, Chichester-New York.

Note

1. I termini *DeNoise* e *DeHiss* sono spesso confusi ed il loro significato risulta a volte invertito sia nei testi scientifici di riferimento sia nei programmi software.