

SEMPLICI CONSIDERAZIONI SULLE TECNICHE DI “RESTAURO” DEL SEGNALE VOCALE

Piero Cosi

IFD-CNR

Istituto di Fonetica e Dialettologia – Consiglio Nazionale delle Ricerche

e-mail: cosi@csrf.pd.cnr.it

www: <http://nts.csrf.pd.cnr.it/Ifd>

1. SOMMARIO

Le registrazioni audio, soprattutto se “datate”, sono spesso corrotte da vari tipi di rumore o di disturbo. Fra questi i più frequenti sono rumori impulsivi (ad esempio i “click”), rumori a banda larga (il fruscio caratteristico dei nastri magnetici o della superficie dei dischi di vinile) o rumori di tipo armonico (ad esempio i 50Hz della linea).

Esistono numerose tecniche per la “rigenerazione” o “ripulitura” di registrazioni audio corrotte da rumore e lo scopo di questo lavoro è quello di descrivere brevemente alcuni dei principali algoritmi utilizzati in letteratura per questo scopo.

2. INTRODUZIONE

Il “ripristino” delle sorgenti degradate di archivi di materiale sonoro si è negli anni consolidato come un campo di ricerca sempre più importante e necessario. L'introduzione dei nuovi media digitali di alta qualità audio ha inoltre alimentato notevolmente le aspettative sulla qualità del suono, in tutti i tipi di registrazioni audio e determinato progressive modificazioni della sensibilità uditiva. Il “Restauro del Segnale Audio” (RSA) risponde 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. Con “*Digital Audio Restoration*” (DAR) spesso si intendono ambigualmente attività differenti dipendenti dal fine stesso ad esse associato.

Una prima interpretazione è legata direttamente alle esigenze archivistiche. Con tale procedura si vuole, infatti, ristabilire, il più precisamente possibile, “il suono” della registrazione originale. 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). Conseguenza diretta di questa interpretazione è che tutti i rumori, i sibili, i fruscii, i *crackle*, i ronzii presenti sul supporto analogico originale, sono udibili anche nel nuovo supporto digitale.

Una seconda interpretazione 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. Si vuole cioè, mediante 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, trasformare una vecchia registrazione in una registrazione “attuale” eliminando “il suono” corrispondente alla vecchia registrazione e ristabilendo l'“esecuzione” originale.

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. DIGITAL AUDIO RESTORATION

Volendo di proposito semplificare, i tre passi fondamentali per una corretta e funzionale “rimasterizzazione” digitale del segnale audio, inteso come il procedimento per giungere ad una copia ottimale di supporti analogici originali, possono essere così riassunti:

- Utilizzare la miglior sorgente analogica disponibile.
- 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.
- 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.

Mediante il “restauro” si vuole invece ovviare a tutti quei fenomeni di deterioramento che 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.

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 statistico *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.

- **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.
- **Compressione d’ampiezza (AC - “Amplitude Compression”)**
Questa procedura modifica le ampiezze relative di differenti regioni del segnale aumentando l’ampiezza nelle regioni più deboli e lasciandola inalterata in quelle più forti. 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.
- **Eliminazione dei “click” (“DeClick”)**
Con il termine “*DeClick*”, ci si riferisce generalmente ad una procedura per l’eliminazione dei disturbi di carattere impulsivo che, mediante tecniche statistiche, si prefigge di eliminare i fenomeni non “*in linea*” (“*outlier*”) con le normali

caratteristiche del segnale audio in esame. Un esempio dell'applicazione di tale procedura è illustrato in Figura 1.

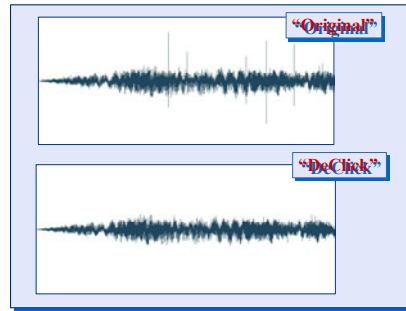


Figura 1. Esempio dell'applicazione della procedura "DeClick" , per l'eliminazione dei click in una registrazione di un brano musicale.

- **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. In una prima fase, si determina il contenuto spettrale del rumore da eliminare e, in una seconda fase, si procede alla sottrazione, da tutto il segnale, dello spettro del rumore precedentemente calcolato. 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 Figura 2 (a destra).

- **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 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.

- **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, che possono essere facilmente eliminati mediante filtri selettivi in grado di sopprimere il disturbo "portante" e le sue armoniche dominanti. Sempre in Figura 2 (a sinistra) è illustrato un esempio dell'applicazione della procedura DeHum su un segnale vocale ('aiuole') a cui è stato sovrapposto un rumore a 50Hz.

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.

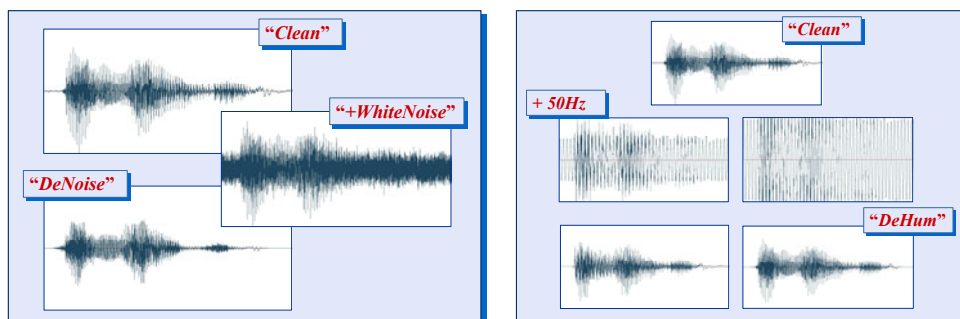


Figura 2 Esempio dell'applicazione della procedura "DeNoise" (a sinistra) per l'eliminazione di un rumore bianco sovrapposto ad un segnale vocale target ('aiuole') e "DeHum" (a destra) per l'eliminazione del disturbo a 50Hz sovrapposto ad un segnale target ('aiuole').

- **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.
- **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 in base al loro specifico contenuto fonetico, aumentandone o diminuendone l'ampiezza, al fine di ottimizzarne la discriminazione.

3. CONSIDERAZIONI FINALI

Mediante il restauro si separa il "suono" dal supporto che lo conserva e lo trasmette. Si vuole, infatti, poter ascoltare il "documento sonoro" senza tutti i disturbi, indotti dal tempo e dall'uso, che corrompono le registrazioni originali. Si può affermare anche che il restauro 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. In alcuni casi però, 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

- [1] Godsill S.J., Rayner P.J.W., Capp O., Digital audio restoration. In *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds., Kluwer, Boston MA, 1998.
- [2] Godsill S.J., Rayner P.J.W., *Digital Audio Restoration*. Springer, Berlin, 1998.
- [3] Adamo G., Debiasi G.B., De Poli G., Giua P., Mian G.A., Re G., Sotgiu M.C., Vidolin A., Problemi di conservazione e restauro di archivi sonori. In *Atti del Congresso Associazione Italiana di Acustica*, Perugia, Maggio 1997, pp. 31-38.
- [4] Mian G.A., Metodologie di restauro di archivi sonori. In *Atti A.E.S. (Italian Chapter)*, Milano, ottobre 1997.
- [5] Niedzwiecki M., *Identification of time-varying processes in signal processin*. John Wiley & Sons, LTD, Chichester, New York, 2000.