

## Ogg Vorbis

# La compressione audio in Noise & Vibration Works



ARMANI Alberto  
Dicembre 2011

### Premessa

*A seguito dell'esperienza maturata nel supporto dei file audio acquisiti con la strumentazione di tipo fonometrico, si è giunti alla considerazione che data la grande utilità del riascolto, sarà sempre maggiore la richiesta di poter associare alle misure di rumore, una registrazione audio di tipo continuo. Questa richiesta diventerà sempre più frequente sia perché consente all'operatore di aver certezza nell'identificazione di ogni evento sonoro relativo a periodi di misura non presidiati, sia perché gli è poi consentita una documentazione audio di difficile contestazione.*

*Purtroppo sebbene la più recente generazione di strumentazione fonometrica consenta una discreta semplicità nell'acquisizione diretta dei file audio associati alle misure di rumore, rimane aperto il problema relativo alle grandi dimensioni di questo genere di file, che di fatto ne limita molto l'impiego pratico.*

*La soluzione logica porta quindi alla necessità di trovare algoritmi adatti per comprimere adeguatamente il segnale audio.*

### **La registrazione audio nei fonometri.**

Oggi, il problema della registrazione audio nella strumentazione fonometrica è affrontato con modalità molto efficienti e non si ripercuote più come nella precedente generazione, sulle capacità di memorizzazione dello strumento poiché ora tutto il processo è di tipo digitale ed è possibile gestire dimensioni di memoria di parecchi GByte.

L'ostacolo attuale è rappresentato invece dalle capacità di gestione e di playback del software che viene utilizzato per la lettura dei dati e per il corrispondente riascolto.

Un semplice esempio può dare una idea delle dimensioni dei file che sono generati da questo tipo di misure: con un sampling di 51.2kHz a 24 bit avremo necessità per il solo file audio, di 540 MByte per ogni ora di registrazione, a cui andrà aggiunta la dimensione occupata dalle misure fonometriche che saranno almeno di altri 12 MByte nel caso più comune di memorizzare uno spettro in bande di 1/3 d'ottava al secondo.

Facile accorgersi che file con questa dimensione non possono consentire rilievi estesi altro che su brevi periodi e quindi non adatti per misure prolungate su intere giornate come sarebbe desiderabile.

La soluzione di ridurre la frequenza di campionamento ad esempio ad 8 kHz può essere un compromesso accettabile in considerazione al fatto che spesso per le misure di rumore ambientale, si desidera solo avere certezza degli eventi sonori più rilevanti e si è interessati solo ad una qualità audio sufficiente

per la compressione dell'evento ma anche in questo caso si avranno file con dimensioni di circa 1.5 GByte / giorno, quindi ancora una volta non facilmente gestibili.

Per risolvere il problema esiste una sola soluzione : la compressione del segnale audio.

### La scelta dell'algoritmo di compressione.

Numerosi sono infatti gli algoritmi di compressione studiati per comprimere il segnale audio tra cui il più noto è sicuramente l'MP3 ma non è il solo e soprattutto ne esistono alcuni che si adattano meglio allo scopo che noi desideriamo raggiungere.

Dopo varie prove l'algoritmo che si è mostrato più adatto ai nostri scopi, è risultato l' Ogg Vorbis e per avere maggiori dettagli sullo specifico codec, suggeriamo di visitare il sito:

<http://it.wikipedia.org/wiki/Vorbis>

Il formato standard .wav di Windows, memorizza i dati in modo diretto nel formato digitale PCM (Pulse Code Modulation), senza nessun tipo di compressione (lossless) e con conseguenti file di grosse dimensioni che però non richiedono elevata potenza di calcolo per il loro riascolto.

Il formato compresso Ogg Vorbis è invece di tipo Lossy, analogo al più noto formato MP3 ovvero con l'eliminazione di alcune informazioni non uditive, elaborate per ottenere il miglior tasso di compressione possibile, mantenendo però sempre una qualità di riascolto indistinguibile da quella originale.

Un paragone significativo per il formato di compressione lossless è il noto sistema ZIP utilizzato per ridurre la dimensione dei file tradizionali; la dimensione del file viene ridotta ma nessun dato ovviamente viene perduto.

Diversamente il paragone per un compressore di tipo Lossy è il formato JPEG comunemente utilizzato per la compressione delle immagini fotografiche ed è esperienza comune il riscontro visivo dell'efficienza di questo algoritmo.

Formati di compressione di tipo Lossy raggiungono ovviamente capacità di compressione molto superiori a quelle fornite dai compressori Lossless.

La scelta del codec di compressione audio Vorbis è stata condizionata da numerosi fattori tutti scaturiti da prove di confronto e legati alle seguenti ragioni:

- Possibilità di ottenere valori di compressione molto alti ed ottimizzati per ogni tipo di segnale.
- Possibilità di poter comprimere file provenienti da analizzatori multicanali.
- Possibilità di comprimere file con singoli eventi audio registrati in sequenze arbitrarie.
- Massima accuratezza nei risultati delle analisi in 1/3 d'ottava prima e dopo la compressione.
- Massima precisione per il trattamento dei valori di picco.
- Velocità di codifica.
- Formato dati in virgola mobile (floating point) per segnali codificati a 8, 16, 24 o 32 bit.
- Capacità di compressione anche per frequenze di campionamento fino a 200 kHz.

L'algoritmo di compressione Vorbis è del tipo VBR (*Variable Bit Rate*), ovvero a bit rate variabile in base al tipo di segnale sonoro che è chiamato a codificare; diversamente dalla compressione in MP3 che è del tipo CBR (*Constant Bit Rate*), Vorbis produrrà un volume di dati per la codifica di un segnale sonoro di rumore di fondo o di un rumore continuo molto inferiore al segnale di un pieno di orchestra o di altri

simili suoni di livello maggiore e più variabili nel tempo. Il riascolto di un segnale compresso con Vorbis risulta inoltre a parere di molti esperti di qualità migliore rispetto alla compressione in MP3.

L'operatore che desidera comprimere il segnale con il codec Vorbis ha solo una semplice scelta tra valori compresi tra  $q=0$  e  $q=10$ , che corrispondono alla qualità della codifica; verrà scelto un valore  $q=10$  per ottenere la massima qualità disponibile ed un valore  $q=0$  per ottenere la massima compressione disponibile; ad un indice di qualità 10 corrisponde un bitrate o velocità di flusso molto alta, anche superiore ai 1000 kbit/s che via via si riduce per indici di qualità inferiori sino ad un bitrate variabile tra 64 ed 80 kbit/s per un  $q=0$ .

### **L'impiego pratico della compressione audio.**

Per ottenere un utilizzo pratico orientato al settore delle misure fonometriche, il codice di compressione Ogg Vorbis è stato integrato nel software Noise & Vibration Works in modo tale da rendere possibile la compressione del segnale audio, direttamente in fase di importazione di ogni tipo di misura fonometrica con annessa registrazione audio.

Particolare cura è stata posta nel rendere il più trasparente possibile la gestione dei file compressi; dopo la scelta iniziale, se comprimere o meno il segnale audio durante la fase di importazione, tutti gli altri comandi e le procedure di riascolto, filtraggio, edit sul segnale o esportazione, sono identici per entrambe le condizioni; l'operatore non ha modo di accorgersi delle differenze nell'ascolto altro che constatare la notevole differenza nella dimensione di memorizzazione tra due file.

E' opportuno riportare qualche esempio per evidenziare i vantaggi pratici conseguenti all'uso del compressore dei file audio:

Come primo esempio possiamo considerare un generico monitoraggio del rumore in ambiente di vita dove si desidera registrare il segnale audio in continuità per poter riascoltare oltre agli eventi sonori più importanti anche alcuni periodi dove sono presenti delle sorgenti specifiche di disturbo che però risultano mascherate dal rumore del traffico locale ma pur sempre giudicate disturbanti a giudizio degli esposti.

Il fonometro Larson Davis LD-831 viene impostato con una registrazione audio continua a 16 bit, 8kHz in parallelo ai soliti valori fonometrici ed ad una analisi in frequenza a 1/3 d'ottava acquisiti ogni secondo, con un consumo di memoria pari a 56.8 MByte / ora.

Il file originale di misura per una intera giornata risulta di circa 1.363 GByte ma, se compresso al massimo valore disponibile dal compressore di Vorbis, otteniamo un file di soli 167 MByte con un rapporto di compressione di circa 8 volte!

Il secondo esempio considera una misura fatta con l'analizzatore multicanale SoundBook dove il segnale audio registrato in continuità è campionato a 51.2 kHz, con risoluzione a 24 bit e memorizzato a 32 bit floating point con un consumo di memoria di 703 MByte / ora. Per memorizzare tutti i vari valori fonometrici ogni 100 ms, dovremo aggiungere circa un altro MByte mentre per una analisi in frequenza in 1/3 d'ottava ogni secondo aggiungeremo un altro 1/2 MByte, per un totale di circa 706 MByte /ora ovvero quasi 17 GByte al giorno!!

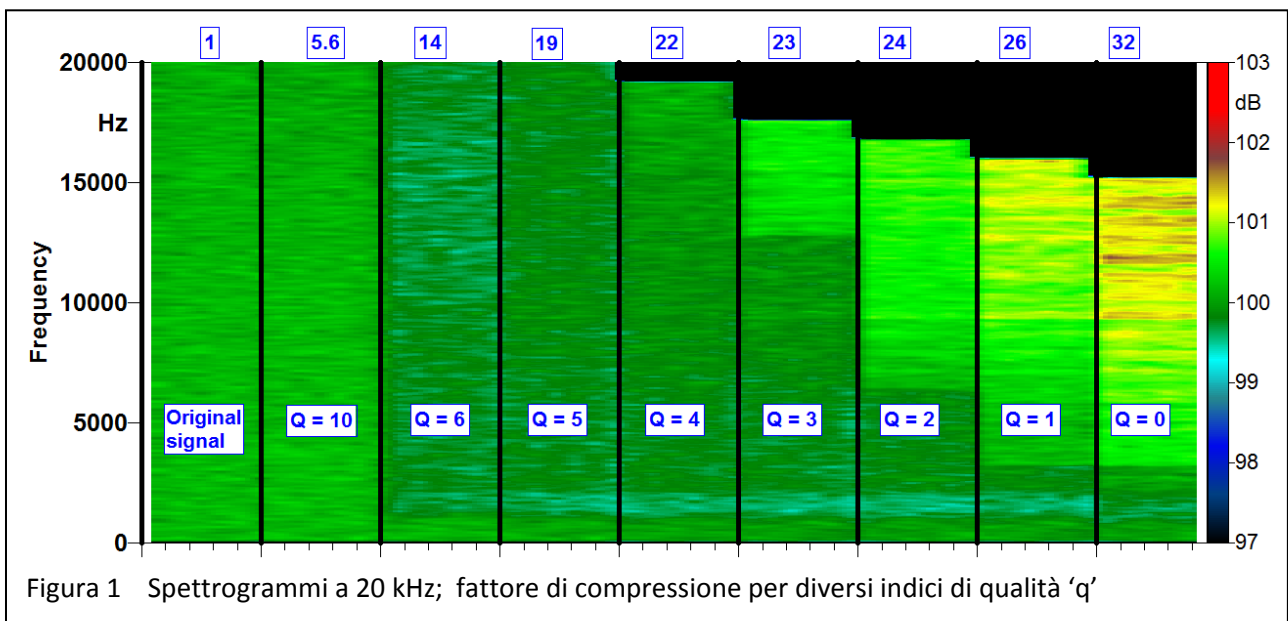
Applicando la compressione massima disponibile da Vorbis, al file di 17 GByte di una giornata di misura acquisita con l'analizzatore SoundBook si può ottenere un rapporto di compressione superiore a 25 volte pari ad una riduzione da 17 GByte a 680 MByte..... e senza una alterazione percepibile della qualità del segnale audio registrato!!

### La rianalisi del segnale audio compresso.

Considerando che stiamo utilizzando un compressore del tipo 'Lossy', qualora fosse necessario comprimere il segnale mantenendo una qualità utile per garantire un possibile riprocessamento del medesimo, per esempio con una analisi in 1/3 d'ottava, allora si dovrà aver cura di mantenere il fattore di compressione minore di 19 scegliendo il parametro di compressione pari ad un indice di qualità compreso tra un  $q=5$  o un  $q=7$ .

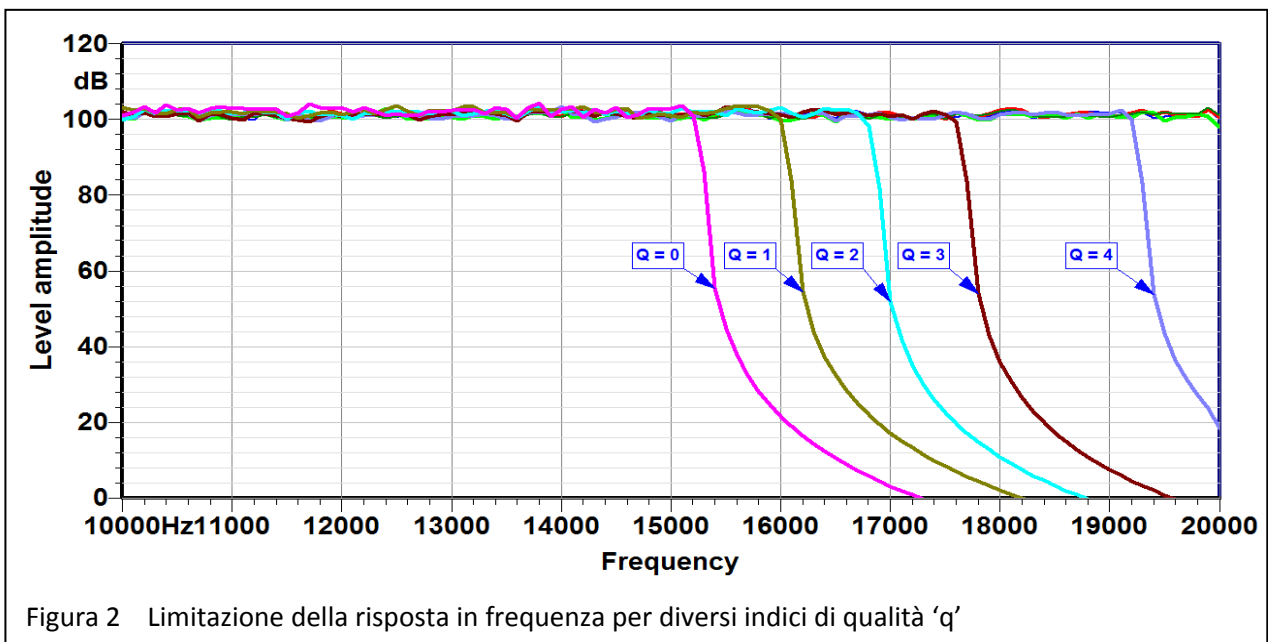
Per illustrare in modo più comprensibile la relazione tra 'fattore di compressione' ed 'indice di qualità', abbiamo riportato in figura 1, con la forma grafica di spettrogrammi, incolonnati in sequenza lungo l'asse delle ascisse, oltre all'analisi del segnale originale corrispondente ad un rumore bianco di qualche secondo di durata, una serie di 9 postelaborazioni eseguite sul medesimo segnale ma compresso con indici di qualità 'q', pari a 10, 6, 5, 4, 3, 2, 1, 0.

Dal grafico possiamo osservare che ad una compressione del segnale con indice di qualità  $q=10$  abbiamo uno spettrogramma dello stesso colore di quello corrispondente al segnale originale che viene riportato all'inizio della sequenza dei 9 spettrogrammi; praticamente identici sono gli spettrogrammi per gli indici  $q=9$ ,  $q=8$ ,  $q=7$ , non riportati; per un indice di qualità  $q=6$  e  $q=5$  si possono invece osservare nella colorazione, alcune sfumature che iniziano ad indicare leggere alterazioni del livello in decibel. Con indici da  $q=4$  a scendere, si nota chiaramente l'azione di un taglio sulle frequenze superiori che giunge per un  $q=0$  ad una limitazione della banda di analisi a 15.2 kHz.



Sulla parte superiore del grafico di figura 1, sono riportati per un segnale campionato a 51.2 kHz, i fattori di compressione corrispondenti agli indici di qualità indicati dalle 9 barre a spettrogramma; per un  $q=10$  che garantisce la massima qualità si può ottenere una compressione di circa 5.6 volte, per un  $q=5$  che ci può fornire comunque una qualità di segnale tale da poter essere nuovamente utilizzata per una postanalisi di tipo fonometrico in classe 1 o in 1/3 d'ottava, la compressione fornita è di quasi 20 volte, mentre con un  $q=0$  la compressione massima ottenibile supera le 30 volte.

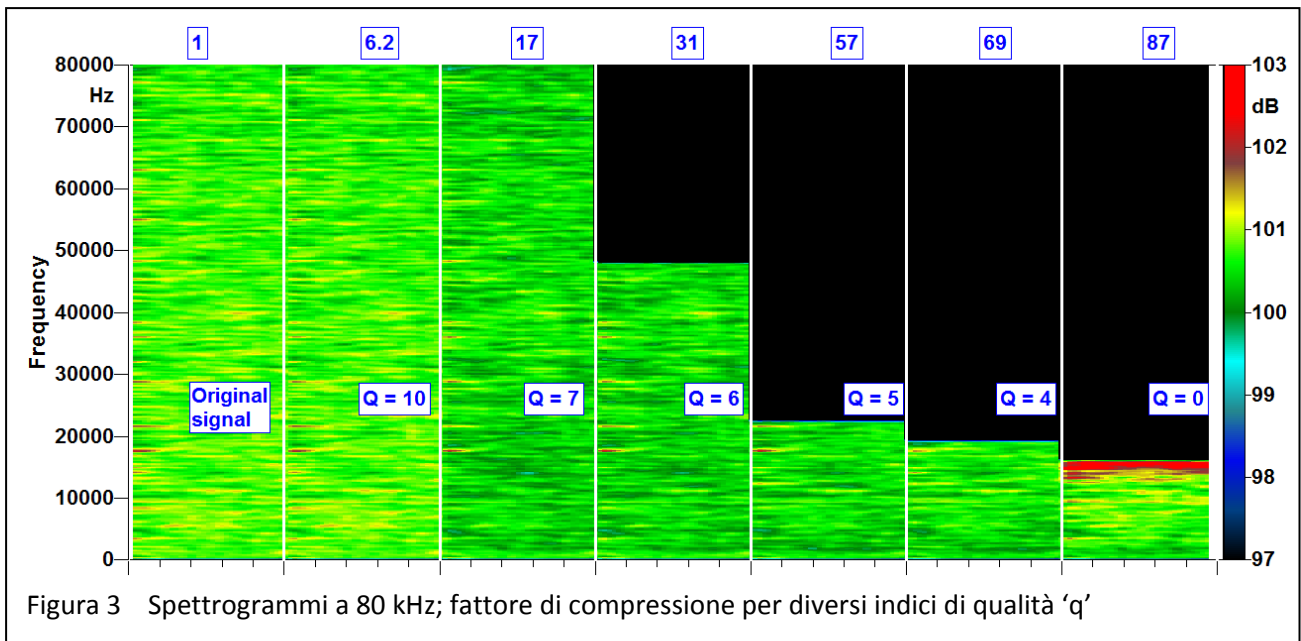
Per avere una analisi in banda sino a 20kHz campioniamo il segnale a 51.2 kHz ed utilizzando convertitori A/D con risoluzione a 24 bit, per non limitare la banda di analisi, dovremo almeno disporre di una velocità di trattamento dati superiore ai 150÷200 kbit/s; questa velocità di flusso dati è consentita con la scelta di un indice di qualità uguale o superiore a 5, mentre la velocità si riduce scegliendo un indice inferiore a 5, con conseguente limitazione nell'estensione di banda alle frequenze superiori come evidenziato dal grafico di figura 2.



Una delle ragioni per cui è stato scelto il compressore Ogg Vorbis rispetto ad MP3, è la sua capacità di poter operare anche su segnali campionati a frequenza maggiore dei tradizionali 44.1 o 48 kHz. La nuova versione di SoundBook MK2, può campionare su tutti i canali alla frequenza di 204.8 kHz con una banda di analisi che si estende quindi fino ad 80 kHz.

Abbiamo quindi provato a comprimere un segnale di rumore bianco campionato a 204.8 kHz con un indice di qualità di  $q=10$ ,  $q=7$ ,  $q=6$ ,  $q=5$ ,  $q=4$  e  $q=0$ . I risultati sono riportati nel grafico di figura 3 con modalità analoghe a quanto fatto in precedenza per le analisi in banda 20 kHz. Nel grafico possiamo osservare che con un  $q=10$  a cui corrisponde un fattore di compressione di 6 volte, il sonogramma del segnale compresso appare identico a quello del sonogramma originale, mentre per un  $q=7$ , con una compressione di 17 volte, sono già visibili alcune alterazioni cromatiche e per un  $q=6$ , la compressione raggiunge un rapporto pari a 31 volte ma già interviene una limitazione di banda che taglia tutte le frequenze superiori ai 48 kHz; con fattori di qualità inferiori a  $q=5$  il segnale viene ulteriormente compresso e la banda passante si riduce sino a frequenze inferiori ai 20 kHz.

Per gestire quindi la compressione di un segnale campionato a 204.8 kHz in modo da poi poter riutilizzare il segnale per una postanalisi, dovremo porre semplicemente attenzione a non scegliere un indice di qualità inferiore a  $q=10$ ; per la medesima ragione, per un segnale campionato a 102.4 kHz non scenderemo sotto ad un  $q=8$  e per uno campionato a 51.2 kHz non sotto  $q=5$ . Se diversamente lo scopo della compressione è solo orientato al riascolto del segnale audio, l'algoritmo Ogg Vorbis ci rende disponibili fattori di compressione anche superiori a 30 volte.



Per una verifica operativa più vicina alle realtà di una misura di tipo fonometrica, è stata fatta una registrazione del segnale in formato .wav campionato alla frequenza di 51.2 kHz con risoluzione a 24 bit. Nei 30 secondi di misura, sono stati registrati sia suoni continui, sia tonali, a bassi ed ad alti livelli oltre a 5 impulsi con levato fattore di cresta e valori di picco superiori ai 137 dB(C).

Nel grafico di figura 4 sono riportati i risultati nella forma di 5 diverse time history sovrapposte, con i profili di short LAeq da 100 ms, relativi alle tracce della misura originale, del segnale wav memorizzato e rianalizzato e delle successive 3 analisi condotte sullo stesso segnale wav compresso con indici di qualità pari ad un  $q=10$ ,  $q=6$  e  $q=1$ .

A prima vista tutte le 5 tracce sembrano perfettamente sovrapposte tanto da apparire come una traccia singola; abbiamo pertanto aggiunto tre tabelle numeriche che riportano i valori dei livelli istantanei rilevati in corrispondenza di due degli eventi impulsivi più importanti e dei valori globali di LAeq dell'intera misura. Prendendo come riferimento i valori in testa alla lista numerica corrispondenti alla misura diretta, possiamo facilmente rilevare che tutti gli altri valori differiscono al massimo per valori inferiori ai 0.2 dB.

Nota: per ottenere il segnale compresso utilizzato per le rianalisi, il segnale .wav originale è stato compresso con l'indice di qualità desiderato ottenendo un file .ogg. Il file .ogg viene poi trasformato nuovamente in formato wav che sarà ora distinguibile dall'originale per la sua dimensione ridotta dal processo di compressione. Il confronto tra il file wav originale ed il nuovo file wav ottenuto dal file ogg consentirà di evidenziare in dettaglio gli eventuali artefatti introdotti dal processo di compressione.

Sia il software N&VW sia SoundBook rendono disponibili queste funzioni di compressione e rielaborazione.

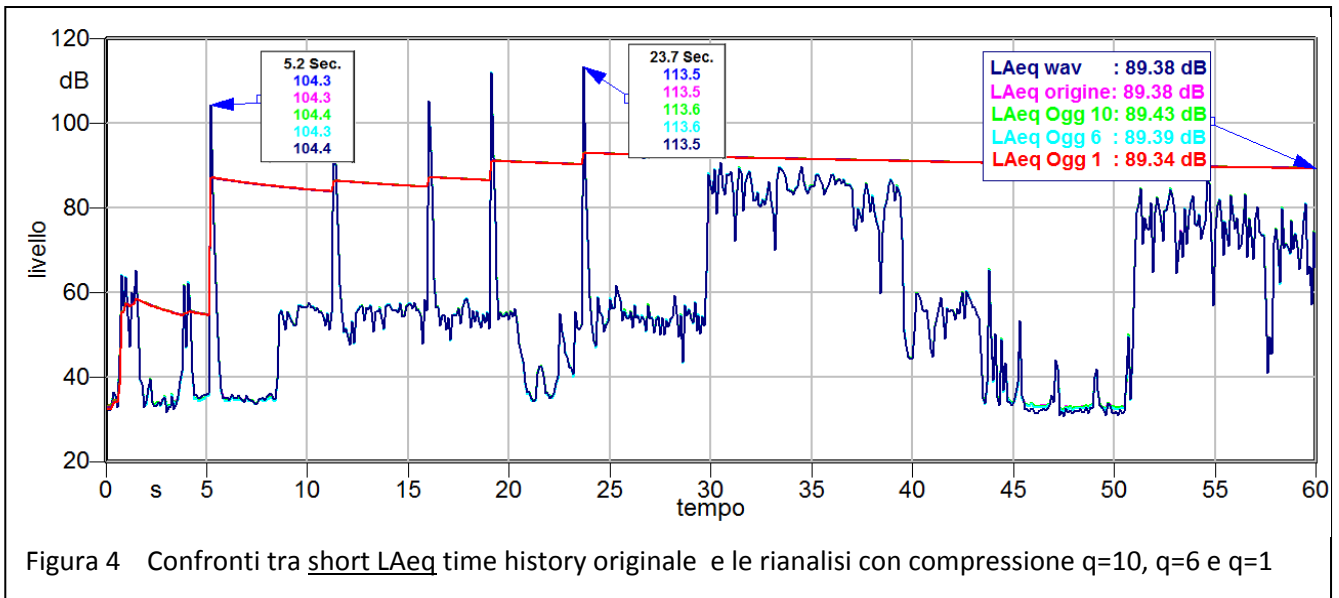


Figura 4 Confronti tra short LAeq time history originale e le rianalisi con compressione q=10, q=6 e q=1

In figura 5 sono riportate sovrapposte, le time history da 100 ms, dei livelli di picco massimo con pesatura (C) relative sempre alla misura diretta, a quella eseguita sul file wav originale ed in successione sui file compressi con qualità q=10, q=6 e q=1. Tutte le 5 tracce si sovrappongono con differenze massime di livello inferiori ai 0.2 dB come rilevabile dalle due tabelline numeriche poste come esempio in prossimità di due impulsi caratteristici della misura presa in esame. Qualche differenza importante è rilevabile solo a livelli di segnale molto bassi ma la funzione del rivelatore di picco è giusto quella di misurare i valori massimi. Molta attenzione è stata posta sui rilievi fonometrici per la misura del picco massimo poiché i vari algoritmi di compressione tendono ad alterare in modo sensibile proprio questi valori. Nell'algoritmo di Ogg Vorbis c'è invece la possibilità di non comprimere i

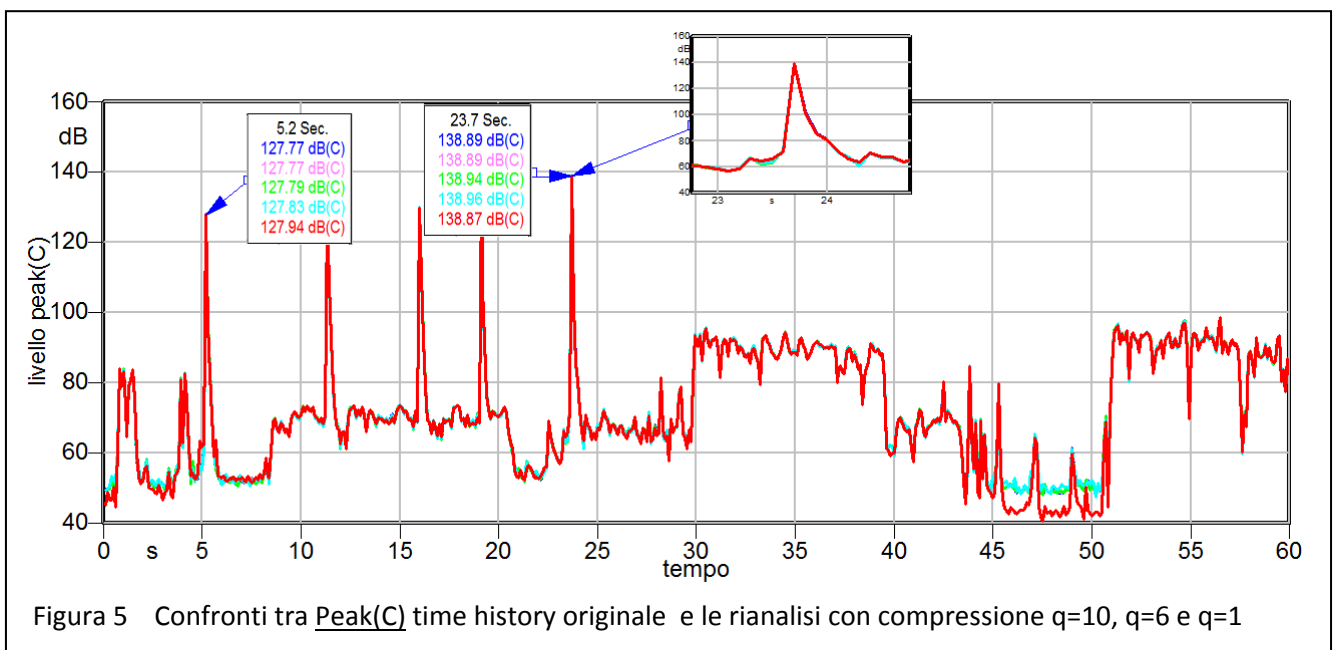
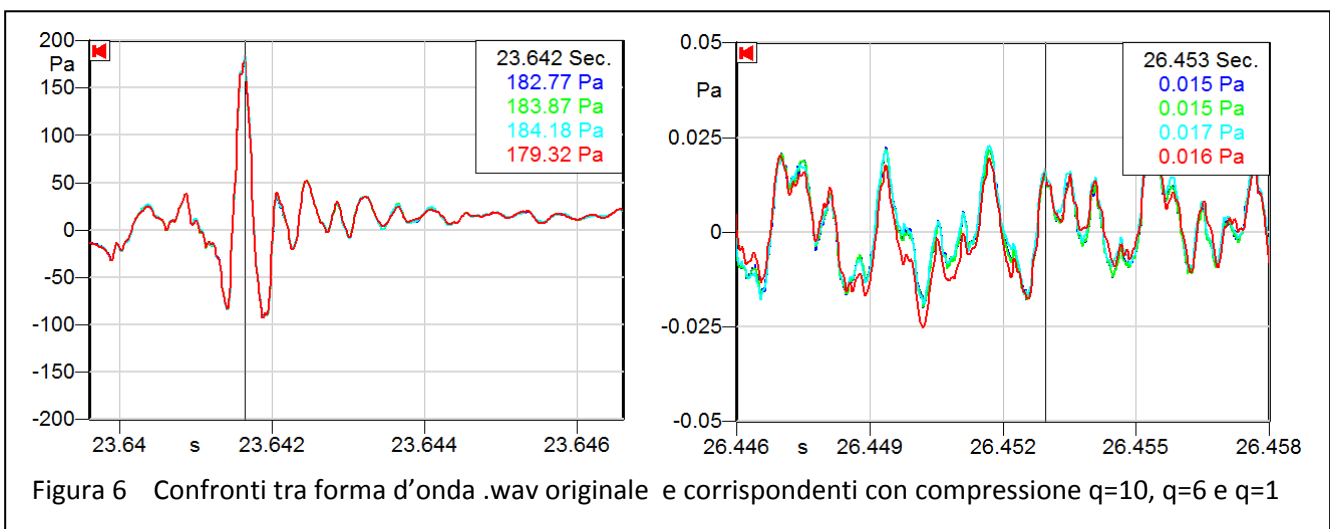


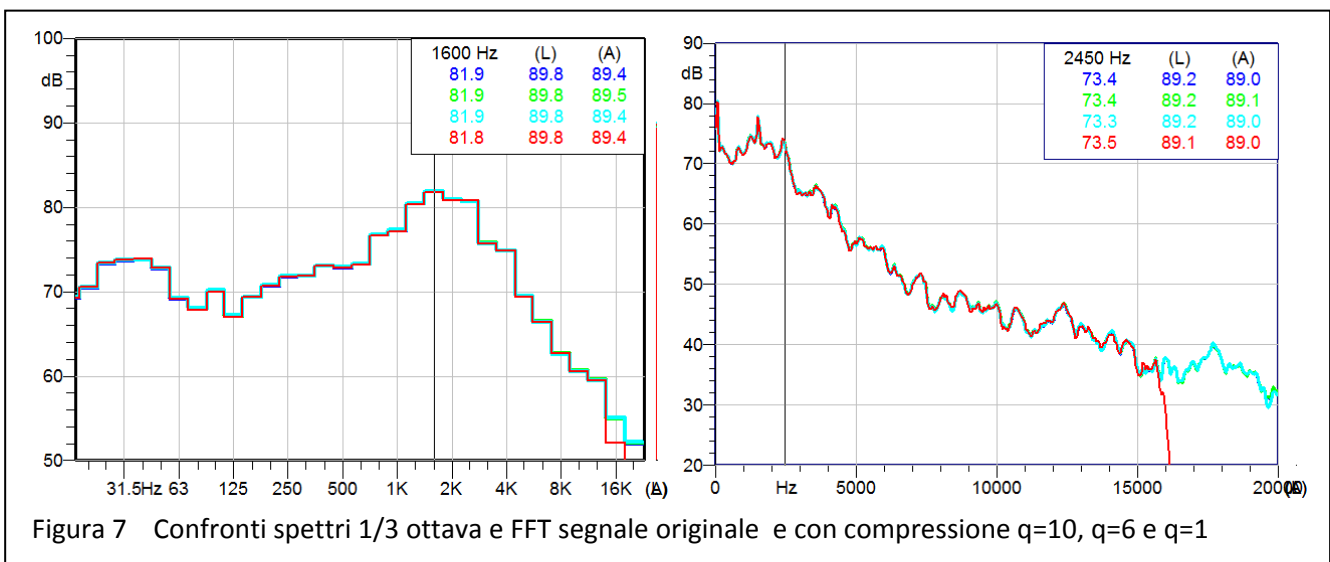
Figura 5 Confronti tra Peak(C) time history originale e le rianalisi con compressione q=10, q=6 e q=1

picchi del segnale ottenendo così i buoni risultati evidenziati nell'esempio di figura 5.

A conferma delle ottime capacità di trattamento del segnale negli istanti corrispondenti agli impulsi con alto contenuto energetico, in figura 6 sono riportati due grafici, sempre con le 4 tracce sovrapposte relative al segnale wav originale ed a quello nelle condizioni con compressione  $q=10$ ,  $q=6$  e  $q=1$ ; nel grafico di sinistra, è riportata la forma d'onda della risposta ad un impulso di 140 dB di picco (C), poco più di 180 Pascal e tutte e tre le tracce del segnale compresso seguono fedelmente quella del segnale originale con una perfetta sovrapposizione. Nel grafico di destra riferito invece ad un livello di segnale prossimo ai 60 dB, si può notare qualche differenza tra le tracce, soprattutto per la traccia rossa corrispondente all'indice di qualità più basso  $q=1$ , ovvero quello con la compressione più elevata.



Come indagine finale abbiamo poi eseguito sugli stessi segnali, le analisi in bande di 1/3 d'ottava ed in FFT, sovrapponendo i risultati ed ottenendo i due grafici posti in figura 7. Gli spettri risultano ben sovrapposti, con deviazioni in banda di 1/3 d'ottava contenute in 0.2 dB ed in banda FFT entro i 0.5 dB ad eccezione del segnale compresso con indice di qualità  $q=1$ , che evidenzia il taglio alla frequenza di 16 kHz e che mostra deviazioni prossime al decibel.





## **Conclusioni.**

Nell'ottica di trovare una soluzione pratica ed affidabile per comprimere i dati dei segnali audio acquisiti con la strumentazione di misura di tipo fonometrico, sono state esaminate tutte le possibili soluzioni ad oggi disponibili trovando come migliore compromesso l'algoritmo Ogg Vorbis.

E' stata così integrata la funzione di compressione Ogg Vorbis sia nell'analizzatore SoundBook, con capacità di compressione diretta anche in fase di acquisizione, sia nel software Noise & Vibration Works per poter comprimere in postprocesso anche i file audio acquisiti con i fonometri Larson Davis LD-831.

Come verifica delle prestazioni di compressione e del mantenimento della qualità della riproduzione audio oltrechè dell'integrità del segnale originale per possibili operazioni di postanalisi, sono state eseguite le prove riportate in questa nota tecnica.

I risultati dimostrano la qualità dell'algoritmo scelto e la sua completa integrazione nella gestione generale degli applicativi in cui viene utilizzato; i segnali audio compressi possono essere gestiti in fase di riproduzione, con velocità di accesso assolutamente impensabili con la gestione dei file originali e diventa ora possibile correlare un file audio con registrazione continua a delle time history o a spettrogrammi di durata anche di più giornate ovvero con possibilità di portare il mouse in un qualunque istante, dove lo spettrogramma mostra un evento sonoro importante e con un click avviare il riascolto per il riconoscimento di quel particolare evento.

Trasformare la dimensione di qualche giga byte del file audio di una misura fonometrica, in uno con dimensioni decine o al più centinaia di mega byte, senza perdere in qualità di ascolto e mantenendo anche una sufficiente integrità per consentirne la postanalisi, rende quasi superflua ogni altra argomentazione a sostegno della validità di questa scelta; questa capacità, sarà molto apprezzata soprattutto da tutti gli operatori che si confrontano con le attuali problematiche dei monitoraggi non presidiati.

Si apre e concretizza la prospettiva di rendere disponibile l'ascolto diretto dell'evento sonoro oggetto di valutazione anche in sede di vertenza legale. Anche l'Autorità giudicante potrà disporre di snelli file audio dai quale trarre le informazioni utili per il giudizio sul disturbo da rumore.