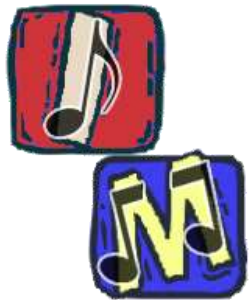


Flac & Mp3: Lossless e Lossy a confronto

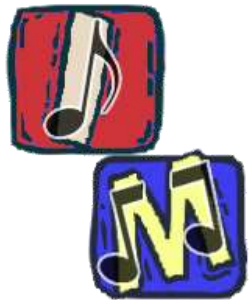


A cura di
Costanzo Giacomo



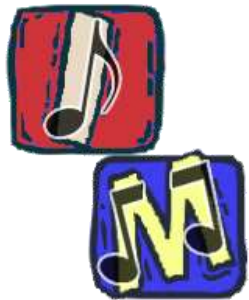
Obiettivi del progetto

- Introduzione: Spiegazione “generica” dei tool di codifica LAME e Flac tool.
- Confronto di una traccia audio con formati di compressione diversi (confronto della differenza di informazione, tramite l'inversione della fase).
- Calcolo del PSNR e MSE delle tracce audio con l'ausilio di MATLAB.



Introduzione

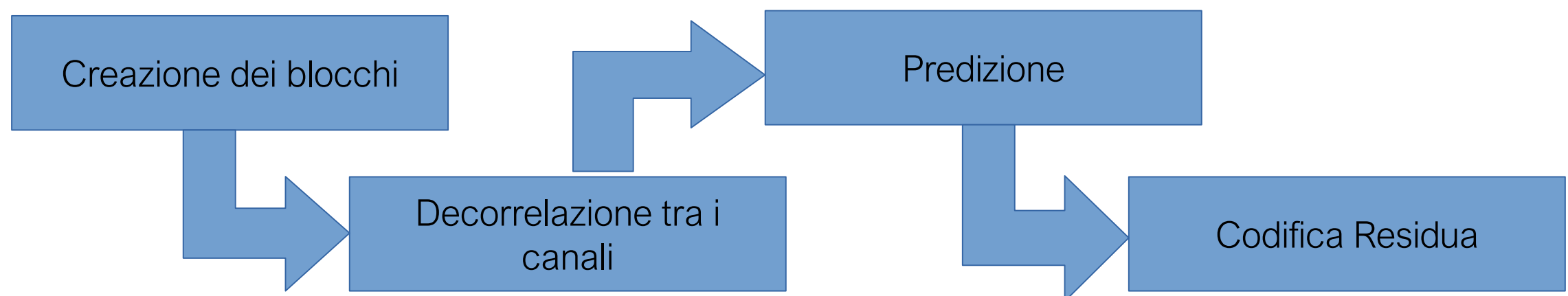
- La compressione è una tecnica utilizzata per ridurre lo spazio occupato da un file, per rendere più agevole l'archiviazione e la trasmissione di quest'ultimo.
- La compressione viene fatta attraverso determinati algoritmi che permettono di dividerla in due Macro gruppi:
 - **Lossless**: compressione senza perdita di informazione, ovvero si dà priorità a mantenere l'informazione integra a discapito dell'efficienza di compressione (es: Cartelle compresse, FLAC)
 - **Lossy**: compressione con perdita di informazione, si dà priorità all'efficienza di compressione rispetto a mantenere l'informazione intatta, spesso utilizzati nel campo multimediale (jpeg, mp3).

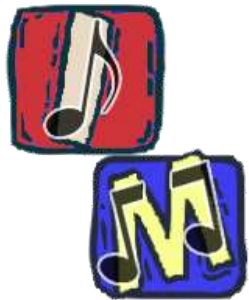


Flac



- Acronimo di Free Lossless Audio Codec
- Ultima versione 1.3.2 (1 Gennaio 2017)
- Nato da un'idea di Josh Coalson del 1999, attualmente si occupa di amministrare e coordinare il progetto FLAC
- Gratis in tutti i sensi: Codice Open Source, con tantissima documentazione disponibile, il formato può essere utilizzato sia in implementazioni gratuite, sia a pagamento.
- I passi fondamentali di questo formato sono i seguenti:

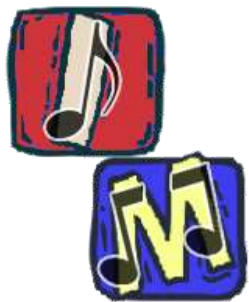




MP3



- Abbreviatio di MPEG1-Layer III (Motion Picture Experts Group - layer III)
- Standard definito nel 1992.
- Massimo numero di canali 2
- BitRate massimo di 320kb/s
- Sfrutta la psicoacustica per rimuovere le informazioni necessarie, che comunque cambiano in base all'encoder utilizzato
- Supporta bitrate fisso e variabile
- Lo standard definisce la qualità CD con un bitrate di 128kb/s
- Brevetti del formato appartenenti a Thomson e Fraunhofer
- Nonostante sia stato sostituito da AAC (brevettato da Dolby), continua ad essere uno dei formati più utilizzati
- Si possono costruire decoder liberamente, ma per encoder a fini commerciali bisogna pagare i diritti.

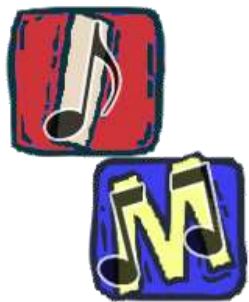


Flac tool



- Disponibile nel sito del formato
- Tool “ufficiale”
- Ultima release v1.3.2
- Al suo interno è presente sia codificatore che decodificatore
- con i tool è possibile modificare i metadati del file direttamente da terminale (cosa fattibile pure con LAME per gli mp3)
- esempio di utilizzo:

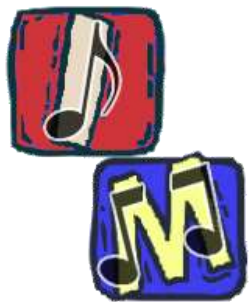
```
PS D:\Users\giaco\Desktop\flac-1.3.2-win\win64> D:\Users\giaco\Desktop\flac-1.3.2-win\win64\flac.exe -5  
"D:\Users\giaco\Desktop\flac-1.3.2-win\win64\01 Grendel.wav"  
flac 1.3.2  
Copyright (C) 2000-2009 Josh Coalson, 2011-2016 Xiph.Org Foundation  
flac comes with ABSOLUTELY NO WARRANTY. This is free software, and you are  
welcome to redistribute it under certain conditions. Type 'flac' for details.  
  
01 Grendel.wav: wrote 104998973 bytes, ratio=0,575  
PS D:\Users\giaco\Desktop\flac-1.3.2-win\win64>
```



Flac tool



- Il comando di default è :
 - `flac "nome_file"`
- che darà in output in file `"nome_file.flac"` con un livello di compressione 5
 - Esistono 9 livelli di compressione da 0 a 8, questi valori migliorano la compressione del file, a discapito della velocità di codifica, posso scegliere utilizzando il flag `-x`, con `x` che va da 0 a 8.
 - 0 è il più veloce, 8 è il migliore
 - Utilizzando il flag `-e`, è possibile migliorare la compressione dello 0.5% con un aumento del tempo di codifica (di base disabilitato)
- Per convertire tutti i file wav nella stessa cartella con un unico comando posso utilizzare
- `flac *.wav`
- Tra le varie cose che cambiano tra i livelli abbiamo:
 - Codifica lineare predittiva `-l` (default `-l5`): va associata con un numero da 0 a 12 (migliore qualità) con 0 si utilizzeranno dei predittori fissi (molto più veloce, file più grande del 10% rispetto allo standard)
 - l'utilizzo del mid size (`-m`, default) o adaptive mid size (`-M`)
 - Le dimensioni dei blocchi utilizzati

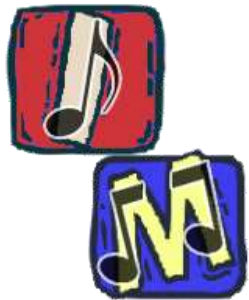


MP3 tool: LAME



- Il nome sta per LAME Ain't an MP3 Encoder (ovvero LAME non è un encoder mp3), nome rimasto quando ancora era solamente una serie di patch, successivamente è diventato un encoder vero e proprio.
- Ultima release v3.1
- Possibilità di codifica a bitrate fisso o variabile
- Decodificatore incorporato
- esempio:

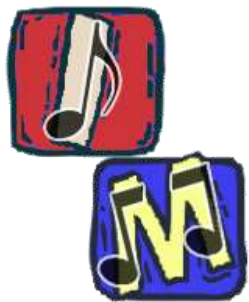
```
PS C:\Users\giaco> D:\Users\giaco\Desktop\lame3.100-64\lame.exe --preset insane "D:\Users\giaco\Desktop\lame3.100-64\01 Grendel.wav"
LAME 3.100 64bits (http://lame.sf.net)
CPU features: SSE (ASM used), SSE2
Using polyphase lowpass filter, transition band: 20094 Hz - 20627 Hz
Encoding D:\Users\giaco\Desktop\lame3.100-64\01 Grendel.wav
to D:\Users\giaco\Desktop\lame3.100-64\01 Grendel.mp3
Encoding as 44.1 kHz j-stereo MPEG-1 Layer III (4.4x) 320 kbps qval=3
  Frame      | CPU time/estim | REAL time/estim | play/CPU |    ETA
39635/39635 (100%)| 0:18/ 0:18| 0:18/ 0:18| 56.300x| 0:00
-----
  kbps      LR  MS %    long switch short %
  320.0     98.8 1.2    95.8 2.4 1.9
Writing LAME Tag...done
ReplayGain: -2.1dB
PS C:\Users\giaco> D:\Users\giaco\Desktop\lame3.100-64\lame.exe --preset 320 "D:\Users\giaco\Desktop\lame3.100-64\01 Grendel.wav"
LAME 3.100 64bits (http://lame.sf.net)
CPU features: SSE (ASM used), SSE2
Using polyphase lowpass filter, transition band: 20094 Hz - 20627 Hz
Encoding D:\Users\giaco\Desktop\lame3.100-64\01 Grendel.wav
to D:\Users\giaco\Desktop\lame3.100-64\01 Grendel.mp3
Encoding as 44.1 kHz j-stereo MPEG-1 Layer III (4.4x) average 320 kbps qval=3
  Frame      | CPU time/estim | REAL time/estim | play/CPU |    ETA
39635/39635 (100%)| 0:18/ 0:18| 0:18/ 0:18| 54.573x| 0:00
32 [ 96] %
40 [ 0]
48 [ 0]
56 [ 0]
64 [ 0]
80 [ 0]
96 [ 0]
112 [ 0]
128 [ 0]
160 [ 1] %
192 [ 0]
224 [ 112] %
256 [11536] %%%%%%%%%%%
320 [27890] %%%%%%%%%%%
-----
  kbps      LR  MS %    long switch short %
  300.4     98.8 1.2    95.8 2.4 1.9
Writing LAME Tag...done
ReplayGain: -2.1dB
```

MP3 tool: LAME

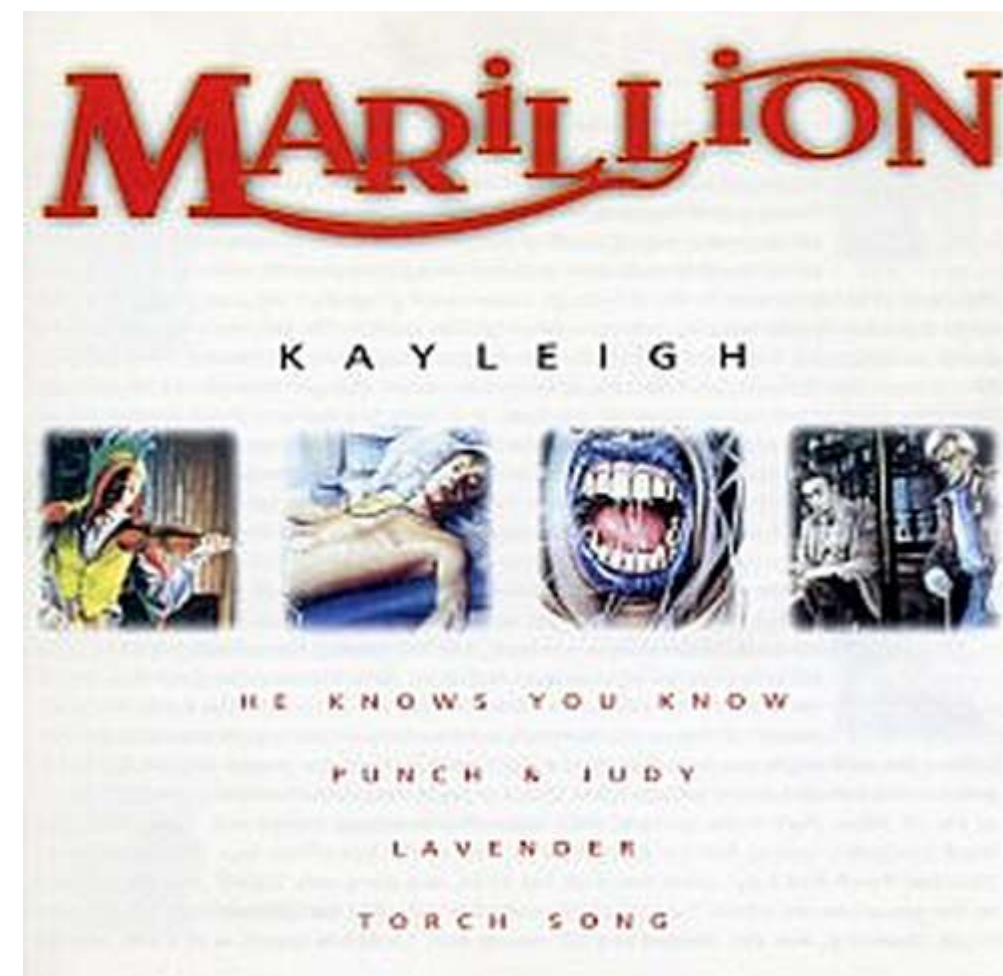


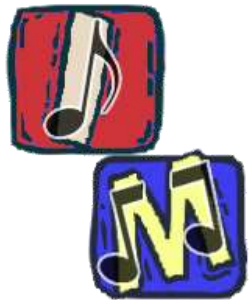
- Il comando di default è :
- lame "nome_file" che
- da in output un file "nome_file.mp3" con un bitrate costante di 128 kbps (qualità cd secondo lo standard) con un valore qual di 3 (è un valore che permette di scegliere la qualità finale del file, in totale 10 livelli). Per impostare la qualità massima bisogna utilizzare il flag "-q=0"
- Il nome del file si può cambiare specificando il nuovo nome a fine comando (nel tool flac bisogna usare il flag -o)
- Per scegliere un bitrate costante si può utilizzare -bx con x massimo di 320kbps
- Per il bitrate variabile sono disponibili due opzioni:
 - VBR: con il flag -V (che va da la qualità più alta 0, alla più bassa 9.999), di default è il 4
 - ABR: con il flag --abr number oppure con --preset number
- I preset standard del encoder sono:
 - preset medium: -V5
 - preset standard: -V2
 - preset extreme: -V0
 - preset insane: -b320



Inizio Test: Brano utilizzato

Questo test è stato eseguito estraendo le tracce audio da CD in formato non compresso .wav,
l'album preso in questione per questo test è la compilation Kayleigh dei Merillion (1996), per la precisione la traccia omonima. Per il confronto della fase le tracce mp3 sono state riconvertite in .wav








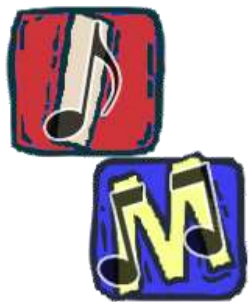
Inizio Test: Software per l'estrazione dei dati

Per l'estrazione delle tracce audio è stato utilizzato il software free:ac su windows, disponibile anche su linux (installando manualmente delle librerie extra).



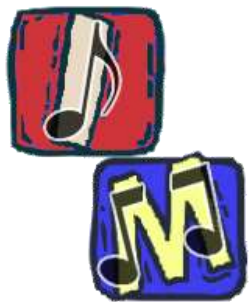
Altre alternative sono:

- Sound Juicer (linux) 
- Exact Audio Copy (windows) 
- KaudioCreator (linux) 



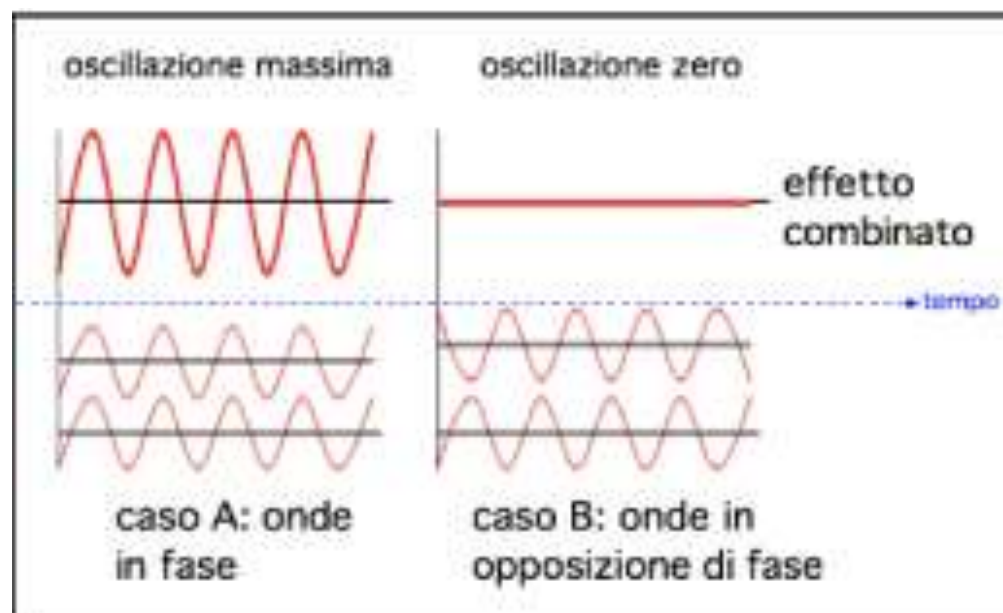
Informazioni sulle tracce audio post conversioni

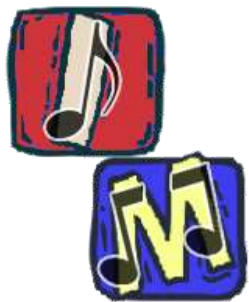
	WAV	MP3 128 Costante	MP3 abs 320	MP3 Preset insano	FLAC preset -5
Dimensione (MiB) (tra parentesi dimensione % rispetto all'originale)	36,1 (100%)	3,27 (9%)	7,61 (21%)	8,19 (22,7%)	20,3 (56,2%)
Bitrate (kb/s)	1411	128	297	320	792
Modalità Bitrate	Costante	Costante	Variabile	Costante	Variabile
Compressione	//	Con Perdita	Con Perdita	Con Perdita	Senza perdita



Passi eseguiti per il test

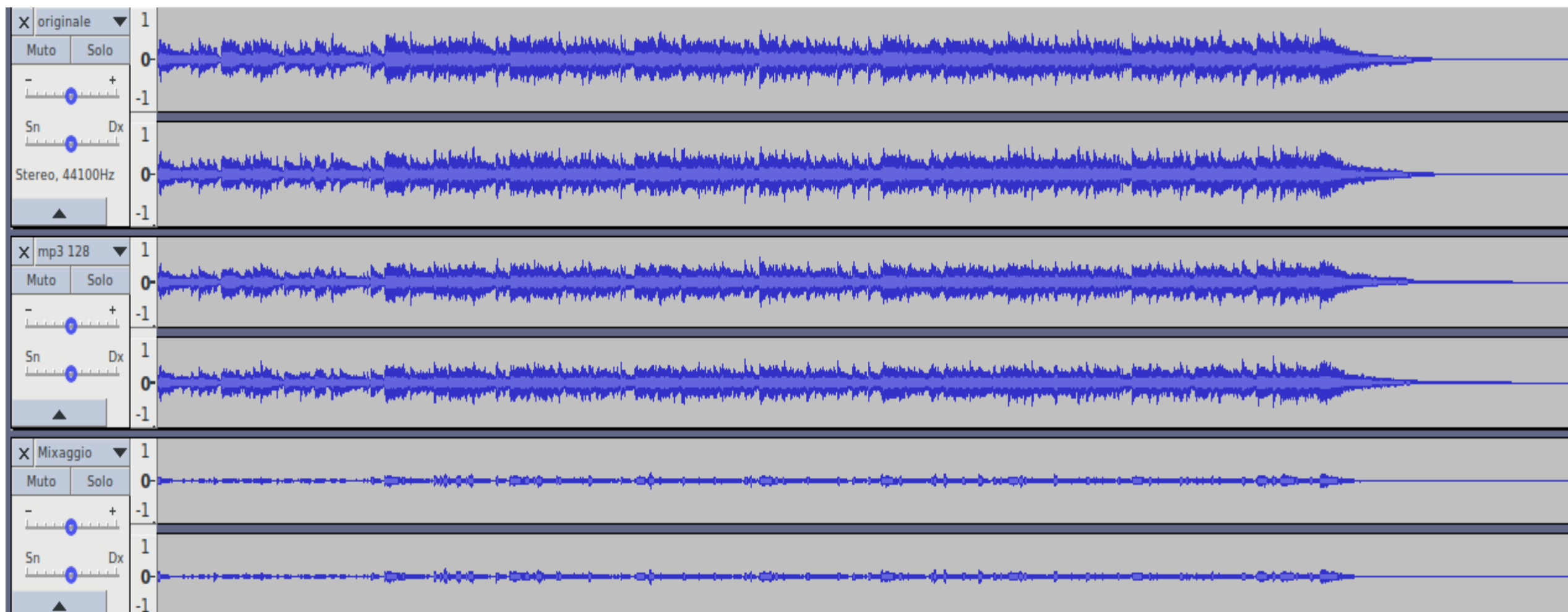
- Importare il file originale e il file convertito su Audacity
- Inversione della fase tramite il comando Inverti
- Mixaggio delle due tracce in una sola traccia

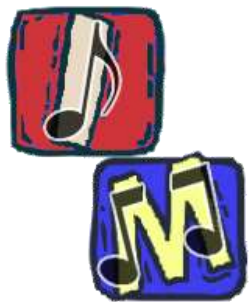




Primo Test mp3 128 bit

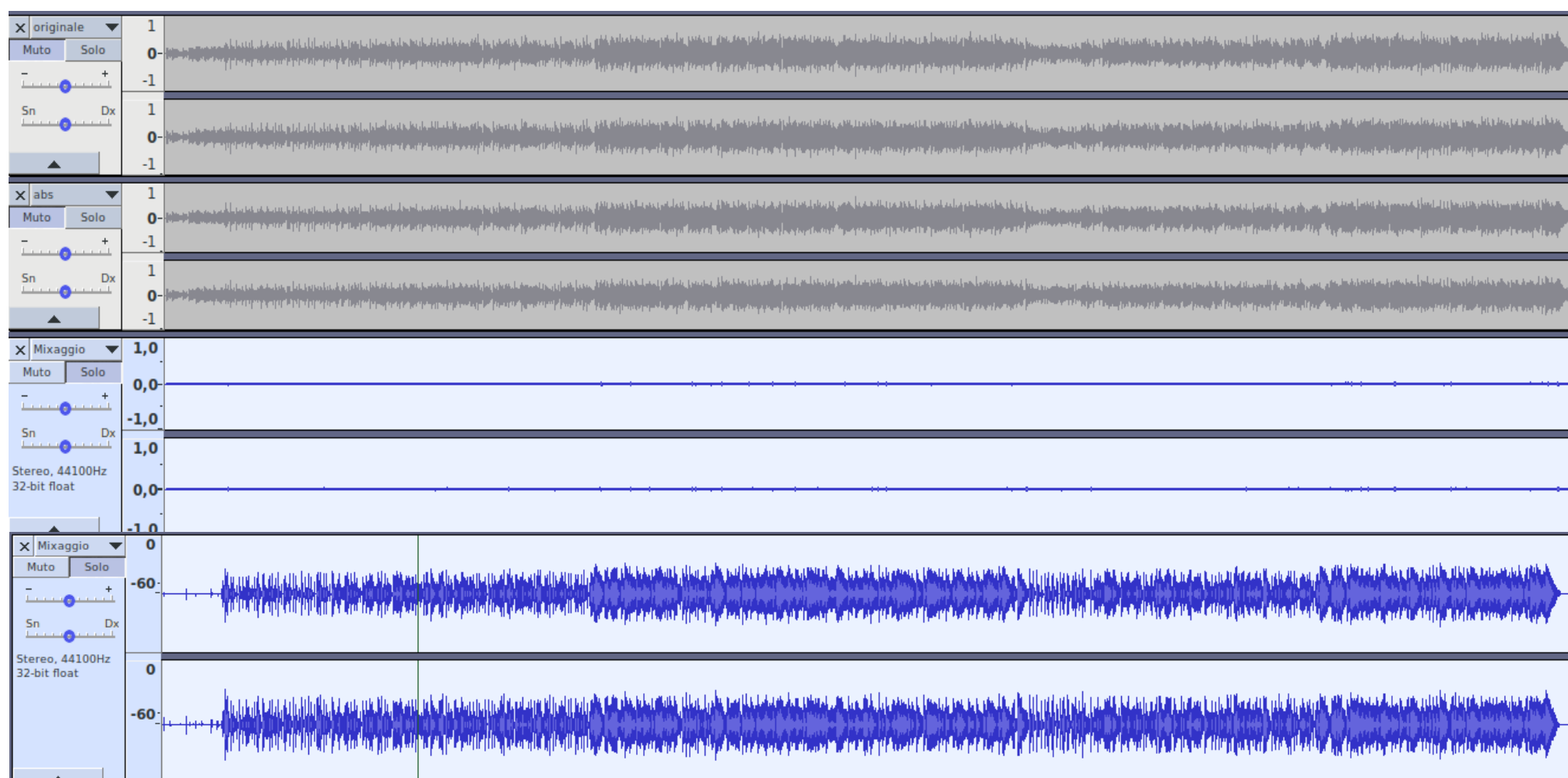
- Tra le perdite d'informazione si individua pure il parlato, oltre che a parti strumentali

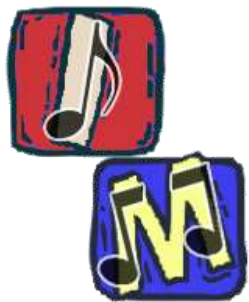




Secondo test mp3 abs (bitrate variabile)

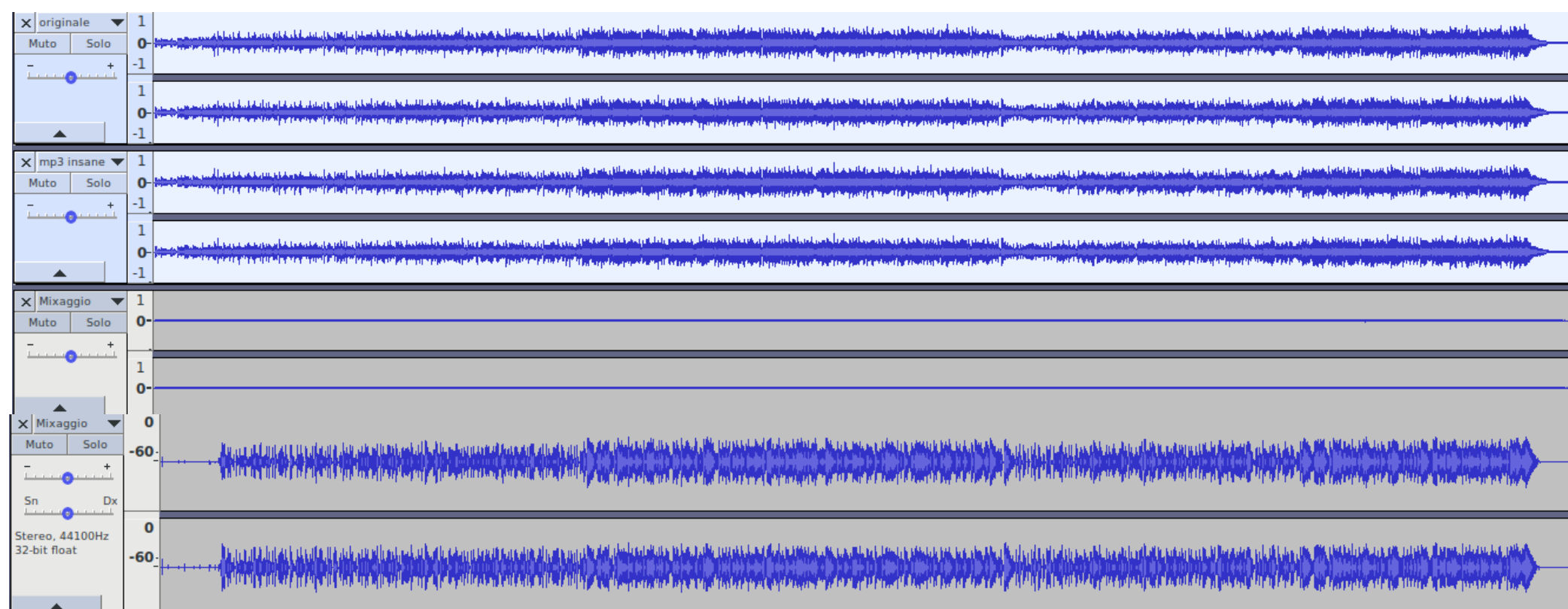
- Come per il test precedente, ma con meno perdita di informazioni, nella parte iniziale sembra quasi nulla, la traccia mixata sembra quasi silenziosa (da un punto di vista grafico) la forma d'onda db...

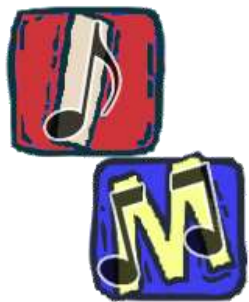




Terzo test mp3 insane (bitrate fisso)

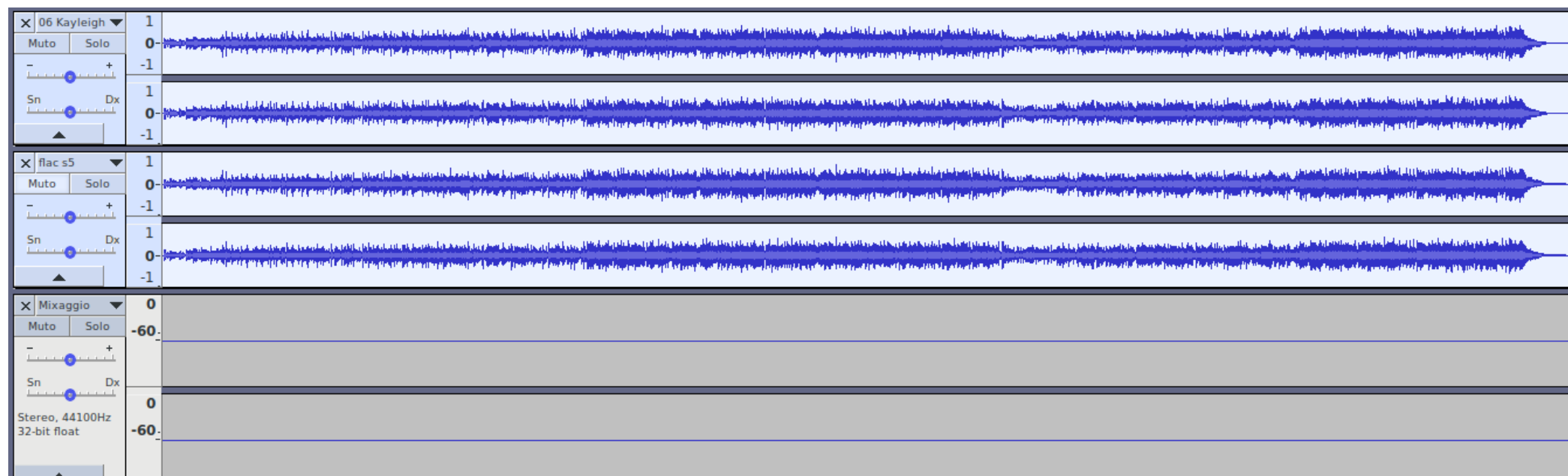
- Perdita di informazione quasi irrisoria in alcuni punti, la traccia mixata spesso sembra quasi silenziosa (in base al dispositivo di ascolto), ma la mancanza di informazione si nota durante le parti “movimentate” della canzone.

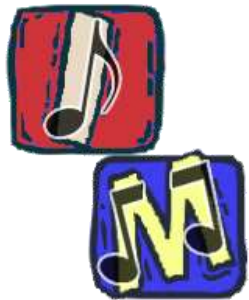




Quarto test flac

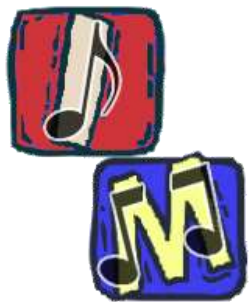
- Perdita di informazione nulla (non per niente comprime senza perdita d'informazioni, la traccia generata è vuota. Testato pure senza riconversione in wave (quindi direttamente con il formato flac) e il risultato non cambia.





Calcolo PSNR e MSE

- PSNR () e MSE sono due modi per controllare la qualità di un file multimediale rispetto all'originale.
- Comunemente utilizzati su immagini, posso essere usare pure su altri file multimediali
- MSE: (Mean Square Error) serve a stimare l'errore quadratico tra due file multimediali; più basso è l'indice, più sono simili i file che si confrontano (per file uguali 0)
- PSNR: (Peak Signal To Noise Ratio) parametro per confrontare un file compresso rispetto all'originale, più il parametro è alto più sono simili (per file uguali ∞)



Calcolo PSNR e MSE

formule

- MSE: (Mean Square Error)

$$MSE = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [I'(x, y) - I(x, y)]^2$$

- nel nostro caso MN=numero canali* numero campioni
- PSNR: (Peak Signal To Noise Ratio) esistono tre formule equivalenti:

$$PSNR = -10 \log_{10} \frac{MSE}{S^2}$$

$$PSNR = 20 \log_{10} \frac{S}{\sqrt{MSE}}$$

$$PSNR = 10 \log_{10} \frac{S^2}{MSE}$$

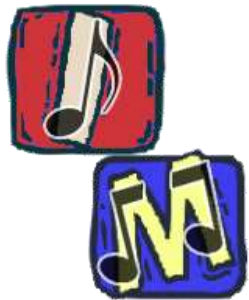
- dove $S=2^n-1$ con n uguale alla profondità dei bit (nel caso dei cd n=16 S=65535)



Calcolo PSNR con Matlab

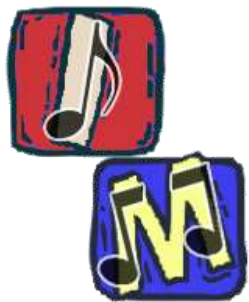
Importare file audio

- Prima di tutto importare le traccie, nel nostro caso
- `[y,fs]=audioread("Nome_Traccia_con_directory")`
- Su fs verrà inserita le frequenza di campionamento (nel nostro caso 44100 Hz)
- Su y verranno memorizzati su una matrice le informazioni del nostro file audio dove le righe corrisponderanno al numero di campioni della traccia, mentre le colonne corrispondono al numero di canali (mono 1, stereo 2...)
- Curiosita: con `"player=audioplayer(y,fs)"` posso riprodurre il brano con matlab usando i comandi `"play(player)"`, `"stop(player)"`, `"pause(player)"` e `resume(player)` per gestire il mio audio.



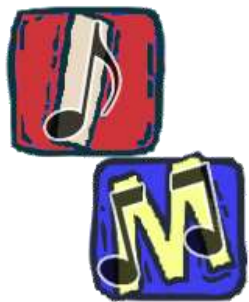
Calcolo PSNR con Matlab Script

- % test di controllo delle dimensioni della mia matrice
- [c1x,c1y]=size(y_originale);
- [c2x,c2y]=size(y_da_confrontare);
- if c1x ~= c2x && c1y ~= c2y
 - disp('dimensioni differenti');
 - exit();
- else
 - MSE = sum(sum((y_originale-y_confrontare).^2)/(c1x*c1y));
 - % alternativa di matlab per calcolare MSE di una matrice a due dimensioni
 - % MSE1=immse(y_originale,y_confrontare);
 - S=65535;
 - PSNR=-10*log10(MSE/S.^2);
 - disp(['mse=' num2str(MSE) ' PSNR= ' num2str(PSNR)]);
- end



Informazioni sulle tracce audio con MSE e PSNR

	WAV	MP3 128 Costante	MP3 abs 320	MP3 Preset insano	FLAC preset -5
Dimensione (MiB) (% rispetto all'originale)	36,1 (100%)	3,27 (9%)	7,61 (21%)	8,19 (22,7%)	20,3 (56,2%)
Bitrate (kb/s)	1411	128	297	320	792
Modalità Bitrate	Costante	Costante	Variabile	Costante	Variabile
Compressione	//	Con Perdita	Con Perdita	Con Perdita	Senza perdita
MSE	0	$1,3038 \times 10^{-4}$	4.4083×10^{-6}	3.6757×10^{-6}	0
PSNR	∞	135.1774	149.8867	150.6761	∞



Documentazione e software

- LAME: <http://lame.sourceforge.net/>
- Guida lame: <https://svn.code.sf.net/p/lame/svn/trunk/lame/USAGE>
- Flac (documentazione): <https://xiph.org/flac/documentation.html>
- Flac Tool: https://xiph.org/flac/documentation_tools.html
- Free:ac: <https://www.freac.org/>
- MATLAB: <https://it.mathworks.com/products/matlab.html>
- Vincenzo Lombardo, Andrea Valle. *Audio e Multimedia*. Maggioli editore. 2014. 4 edizione.
- David Salomon. *Data Compression, The Complete Reference*. Springer. 2007. 4 edizione.