

Introduzione

Perchè un incontro di formazione sull'audio?
E' capitato diverse volte negli scorsi mesi che:

1. ci venissero forniti contenuti audio e audiovisivi
2. o che magari questi contenuti li abbiamo prodotti noi registrando un video di documentazione o per una intervista;
3. oppure abbiamo scaricato files da freesound o FMA

E abbiamo dovuti comporli tra loro per ottenere un prodotto finito da integrare all'interno di una installazione, un software oppure da pubblicare su un social per comunicare un messaggio.

In questi casi potrebbe esserci capitato di notare che l'audio file suonasse male per qualche motivo:

1. il volume troppo basso (o troppo elevato);
2. contiene rumori e altri elementi di disturbo che affaticano l'ascolto.
3. la voce poco intellegibile per percepirne il messaggio;
4. non si adatta agli altri contenuti

Quello che vorrei fare oggi con voi è esplorare un poco questi elementi.

Lo faccio su base di quella che è la mia esperienza personale, mostrandovi dei metodi e degli strumenti che ho avuto modo di utilizzare in questi mesi. Non posso certo dire di conoscerli perfettamente ma mi piaceva comunque condividere quello che ho imparato nell'usarli visto che penso possa tornarci utile.

Non si possono fare i ****miracoli****: se il materiale di partenza è molto deturpato, ci si può lavorare sopra una infinità di tempo e spenderci risorse hardware e software di ogni tipo ma il risultato sarà sempre insufficiente.

Per tutti gli altri casi, ovvio, non esiste un'unica possibile soluzione al problema ma quello che vediamo potrebbe essere di spunto e di stimolo per affrontarlo autonomamente.

Con l'uso, l'ascolto e l'esperienza si diventerà sempre più sicuri e autonomi.

Definizione formale di suono (una)

Soundo è la propagazione della perturbazione dello stato di quiete delle particelle di un mezzo elastico (gas, liquido, oppure anche certi tipi di solidi).

Caratteristica 1: ampiezza

Il suono viaggia attraverso il mezzo elastico trasferendo la vibrazione di un corpo (uno strumento musicale ad esempio) al nostro orecchio o alla capsula di un microfono.

Le particelle di questi materiali normalmente si trovano in uno stato di quiete, si muovono randomicamente attorno ad una propria posizione di equilibrio potremmo dire.

Un corpo sollecitato (percosso) comincia a vibrare e le proprie superfici a distorcersi rispetto ai propri punti di equilibrio, trasferendo per contatto, queste vibrazioni alle particelle del mezzo in cui sono immersi.

Queste particelle a loro volta scatenano un effetto domino per cui gli strati di particelle via via più distanti dalla sorgente sonora saranno anch'essi perturbati dalle stesse vibrazioni a distanza di un certo tempo.

Si creano nel mezzo elastico zone di compressione e rarefazione della particelle. Queste zone si susseguono, uno dopo l'altra e macroscopicamente emerge un comportamento che è il fenomeno della propagazione della perturbazione: il suono.

velocità del suono: circa 343 m/s in aria (temperatura 20 °C)

forma sferica in casi ideali, con sorgente puntiforme.

[gif animata]()

Attenzione: è la perturbazione che si muove nello spazio. Anche le particelle lo fanno ma il loro moto è piuttosto di tipo oscillatorio attorno al punto di quiete.

Ed si tratta di un movimento che "ricalca" in qualche modo le vibrazioni della sorgente sonora.

Più esasperate le distorsioni della sorgente sonora, più dense e rarefatte rispettivamente le zone di compressione e rarefazione.

La velocità della perturbazione resta invariata, quello che cambia in relazione all'energia della sorgente sonora è la quantità di particelle che si ammassano sui fronti d'onda (e viceversa per la rarefazione).

Questa caratteristica può essere misurata in modo assoluto (come si farebbe con un metro per le distanze). L'unità di misura della pressione è il Pa.

Quando queste onde di pressione arrivano al nostro orecchio vengono tradotte dal cervello il quale è in grado di riconoscere tutti i vari livelli di densità.

La percezione del volume!

L'orecchio non è tuttavia in grado di riconoscere tutte le possibili variazioni.

20μPa minima ampiezza percepibile da un orecchio medio (link)

[https://en.wikipedia.org/wiki/Sound_pressure#Examples_of_sound_pressure]

100 Pa - la pressione sonora oltre alla quale si determinano pesanti danni al sistema uditivo

100 Pa --> 135 dBspl
1 Pa – 94dB SPL
20 µPa --> 0 dBspl

ampiezza di picco, ampiezza picco picco, ampiezza RMS (valore efficace)

la nostra percezione uditiva non è lineare rispetto a variazioni di pressione. Si muove di più con il valore efficace (il quadrato della pressione sonora).

Il valore efficace rappresenta il valore che la grandezza in esame dovrebbe avere se fosse una costante per generare gli stessi effetti energetici.

la risposta del nostro sistema uditivo non è sufficientemente rapida da consentire la percezione completa dell'alternarsi di picchi di compressione e rarefazione.

Questo perchè il nostro orecchio lavora su delle medie (RMS - media quadratica sotto radice) temporali effettuate per altro ad intervalli regolari di tempo (50 ms)

la scala logaritmica si usa per comodità di rappresentazione (il range che ci porta da 20µPa a 100Pa si estende su 8 ordini di grandezza mentre una scala logaritmica comprime il tutto).

il decibel ovvero la decima parte del Bel

il Bel non è una unità di misura come potrebbe esserlo il Volt, il metro o il litro.

Si tratta piuttosto di una misura di un rapporto (relativa).

Se io dico che il valore A è il doppio del valore B ecco che uso il numero 2 non tanto in termini assoluti quanto per descrivere A in relazione a B. Se volessimo dire la stessa cosa sfruttando il concetto di dB allora diremmo che A è dB maggiore di B.

Si basano sui logaritmi che possiamo tralasciare

ma pensiamo a quanto è più semplice usare i dB piuttosto che dire 0.5, 0.24, 10000 volte di più, 100000 volte di più etc

1 Bel rappresenta la differenza in una scala relativa di valori con potenza 10
relativa significa con un riferimento

più vicino al nostro sistema percettivo 1Pa - 2Pa è percepito loud il doppio proprio come 50pa - 100pa (6dB)

$20 \log(\text{valore da misurare} / \text{valore di riferimento})$

Audio digitale accenni

grafico pressione microfono ADC audiofile

audio campionato (frequenza di campionamento / bit depth)

16 bit --> dinamica rappresentabile dal supporto di 98dB $6.02 * 16 * 1.47$ (ogni bit in più sono 6 dB in più)

se anzichè un microfono ne avete due potete raccogliere il doppio delle informazioni da due punti diversi dello spazio e magari ottenere un file che si dice essere stereofonico

* frequenza di campionamento / bit depth

* mono / stereo / multicanale

Pratica

Parentesi: userò due tool in modo intercambiabile, principalmente allo scopo di mostrarvi al meglio le cose
e dal momento che ogni tool ha i propri punti di forza useremo l'uno o l'altro a seconda dei particolari su cui ci focalizzeremo
cercherò di mostrarvi come si fa sia in audacity che in reaper

Esercizio A: dialogo-fireworks.wav

Ascoltiamo il file prima di processarlo in audacity (con un qualunque player)

audacity

carica il file
ascoltiamolo
obiettivo (alzare il volume)
vediamo i samples (super zoom)

la prima cosa che ci viene in mente amplificazione
usiamo un plugin - che cos'è?

amplificazione
moltiplicare significa
mantiene inalterati gli equilibri dinamici tra i suoni presenti nel materiale audio originale.

amplifichiamo di 12dB (moltiplichiamo i sample per un fattore 4)
i fireworks clippano (come si visualizzano i clip?)

annulliamo
meglio (?) un sistema chiamato normalizzazione

(che cosa fa normalize?) effettua una amplificazione del segnale audio
andando a moltiplicare i valori numerici di ciascun sample per una costante impostata

stabilisce lui un livello di amplificazione adeguato basandosi sul sample maggiore
non molto efficace --> la voce è ancora bassa

ci piaceva con l'amplificazione? sì
torniamo ad amplificare allora
dove sono i problemi?

tagliamo il primo firework sovrascrivendogli il silenzio
eventuale fade in (smoothing)

altrove effettuiamo una automazione del volume

esportazione

Esercizio B: ABC.wav (25minuti)

ascoltiamolo prima di inserirlo

automazione di volume in Reaper

Reaper

carichiamo

possiamo cominciare con il normalizzarlo

ascoltiamolo

elementi estranei

potremmo scrivere il silenzio (in reaper è più un fatto di taglia)

infattibile in termini di sbatti

infattibile perchè può capitare di sbagliarsi e sovrascrivere audio utile

un processo automatico --> gate

Gate

che cosa è e come funziona

catena audio in/out

curva lineare delle ampiezze

[illustrazione 1]

[illustrazione 2]

ora ci serve applicare il gate come si fa?

click su fx, cerchi gate, inserisci gate

avvantaggiamoci dell'interfaccia grafica che è un po' più intuitiva

modificate la th è già solo così dovrete avere l'effetto desiderato, il clock è sparito

Anche la voce potrebbe cancellarsi per via del fatto che ha un inviluppo

![grafico dell'inviluppo]()

Perchè non funziona sempre?

a volte il contributo utile è miscelato al contributo indesiderato

compromesso

esportazione in Reaper

Esercizio Bbis: snare.wav

snare con gate

il rumore è persistente ed è presente assieme al suono utile. Il gate non si dimostra efficace in questo caso.

Forse è più fastidioso udire un rumore che a volte c'è e a volte no: diventa un elemento di interesse per l'orecchio il quale ci presta attenzione.

Un suono, pur fastidioso ma continuo in realtà finirà per annoiarlo e il cervello a lavorare per mandarlo in sottofondo, lontano dalla parte cosciente.

Caratteristica 2: frequenze (contenuto spettrale)

Finora abbiamo analizzato il fenomeno suono basandoci su una delle sue caratteristiche principali: l'ampiezza.

E' anche vero che quando un corpo, sollecitato comincia a vibrare in sé lo fa in molti modi diversi.

le sue superfici si distorcono nel tempo in modo molto particolare.

Esse, spostandosi dalla propria posizione di equilibrio cominciano a spostarsi - in avanti e indietro - con diverse rapidità.

In queste distorsioni sono presenti moltissimi tipi diversi di vibrazione che, trasferite per contatto al mezzo elastico, arrivano al microfono e al nostro orecchio il quale è in grado di percepirle tutte.

Non proprio tutte in realtà: il nostro orecchio è in grado di percepire e discriminare vibrazioni che abbiano una rapidità che varia tra 20 volte al secondo fino a 20 mila volte al secondo.

le interpreta come **Pitch** (intonazione).

L'unità di misura che sta ad indicare la rapidità di queste vibrazioni, che in altre parole è la rapidità nel susseguirsi di zone di compressione e rarefazione è l'**Hertz**.

L'hertz è una unità di misura per descrivere la frequenza.

esempio con Pure Data

Discorso sul **timbro**

Il timbro è ciò che ci permette di distinguere una chitarra elettrica da un pianoforte, anche se questi due strumenti stanno suonando la stessa nota e con lo stesso volume.

esempio con Pure Data

Tutte queste vibrazioni sono simultaneamente presenti nel movimento di un corpo sollecitato da un colpo o uno sfregamento, però sono presenti con energie diverse.

Tutte queste energie, distribuite su tutte queste possibili vibrazioni, concorrono a creare l'ampiezza del suono.

Il suono può essere letto quindi secondo una prospettiva differente:

* se prima interpretavamo il suono esaminando il livello di ampiezza complessivo su base del tempo;

* ora potremmo esplorare cosa succede, sempre su base temporale, all'energia delle vibrazioni individuali presenti in esso.

Costruiamo un altro tipo di rappresentazione dove sull'asse verticale sono indicate tutte le possibili vibrazioni.

Esempio merlo

mostra lo spettrogramma
fai vedere le frequenze

esempio shemekia

complessità dello spettro.
rumori (energia distribuita su tutte le frequenze)
suoni musicali (solo le parziali armoniche)

Esercizio C: female_hum_high_tremolo.wav

indagine sui diversi tipi di filtro
* **Filtri**: passa basso, passa alto, passa banda;
* **Equalizzatori**: peaking e shelving

per enfatizzare o attenuare l'energia di certe vibrazioni
mantendo le altre inalterate (o quasi :)

hum ronza 50Hz --> equalizzatore
suono fastidioso a 10K --> notch o peaking

Esercizio D: noise_snare.wav (20 minuti)

proviamo con un equalizzatore, però finisce che andiamo a modificare anche il suono utile

reaper FIR con
noise remover (learn - apply)
è sempre un compromesso perchè si perde definizione sulle alte frequenze del rullante.

vediamo lo stesso esempio su Audacity e facciamo un confronto. [Audacity - Noise Reduction]
(https://manual.audacityteam.org/man/noise_reduction.html)

Esercizio E: voice_with_noise.wav

Provate voi :)

potremmo scrivere silenzio --> infattibile
abbiamo visto in reaper il gate
vediamo in audacity anche se è più complesso
appliciamo
dove l'audio ha una ampiezza sotto la soglia scrive zero
altrove lascia il file originale --> abbiamo ancora il rumore
più fastidioso che altro --> è come se ce ne accorgessimo di più ora perchè non è più parte di un
fondo costante
al quale il nostro cervello può abituarsi e cancellarlo per noi
il gate non va bene
usiamo il noise reduction

eventuale problema dello zero crossing

step1 get noise profile
spiegazione
step2 remove noise
spiegazione:
è meglio farne di più e meno incisivi
artifatti

normalizzazione
non funziona

compressione

(rivedi)

[Leveler](<https://ttmanual.audacityteam.org/man/Leveller>) --> è una combinazione di Compressione
e Limiter

ulteriore passaggio di noise reduction

confronto

ancora qualcosa che non va in alta frequenza, un altro passaggio di noise removal

ci sono basse frequenze --> usiano un EQ e filtriamole

spettro della voce (lo abbiamo fatto prima?)

Leveler --> se usato con settaggio di tipo hard/heavy può causare distorsioni (mostra cosa è
successo sui picchi problematici)

l'expander lo fai dopo

Approfondimenti

curve isofoniche (Fletcher-Munson)

Il condotto uditivo del nostro orecchio è lungo 3 cm e ha una risonanza pari a 3KHz - 4KHz (amplificazione 10dB)

Compressione (th, ratio, attacco, rilascio)

Limiter: limiter dipende dal software a volte sfuggono

quindi brick wall limiter (event horizon)

fallo vedere con fireworks

[Leveler](<https://ttmanual.audacityteam.org/man/Leveller>) --> è una combinazione di Compressione e Limiter

sistema percettivo (differenze tra individui) per questo si crea uno standard di riferimento

standard ITU

[loudness war](https://en.wikipedia.org/wiki/Loudness_war)

[alignment level](https://en.wikipedia.org/wiki/Alignment_level) = -18dBFS (RMS)

Articolo interessante a riguardo sul sito di T.C. Electronics

[https://www.tcelectronic.com/brand/tcelectronic/loudness-explained#googtrans\(en|en\)](https://www.tcelectronic.com/brand/tcelectronic/loudness-explained#googtrans(en|en))

aliasing - rolling shutter effect/wagon wheel effect

Audacity plugins

Scritti in [Nyquist]([https://en.wikipedia.org/wiki/Nyquist_\(programming_language\)](https://en.wikipedia.org/wiki/Nyquist_(programming_language)))

* [installing](https://wiki.audacityteam.org/wiki/Download_Nyquist_Plug-ins#install)

* [gate audacity](https://wiki.audacityteam.org/wiki/Nyquist_Effect_Plug-ins#Noise_Gate)