

Pitch detection

attraverso l'autocorrelazione

Armando Boemio
10768150

Music Engineering
Conservatorio di Milano - Politecnico di Milano

Teorie dell'ascolto e della percezione sonora e musicale

- Pitch
- Autocorrelazione
- Algoritmi
- Virtual Pitch
- Risultati

Pitch

Frequenza fondamentale

La frequenza fondamentale F_0 di un segnale periodico è definita come l'inverso del suo periodo, ovvero il minimo tempo dopo il quale il segnale si ripete uguale a sè stesso.

- Definizione valida solo per segnali perfettamente periodici
- I segnali reali non rispondono a questa definizione

La frequenza del pitch è solitamente collegata alla frequenza fondamentale di un segnale, tuttavia si tratta di una grandezza **psicoacustica**. Storicamente sono state fornite diverse definizioni, ad esempio:

- "Suoni aperiodici possono comunque creare la sensazione di un pitch"
- "Risoluzione di un pattern armonico delle frequenze parziali risolte dalla coclea"
- "Corrisponde alla frequenza fondamentale F_0 "

Nell'ambito di pitch detection, quest'ultima definizione è la più utilizzata.

Autocorrelazione

Definizione

La funzione di autocorrelazione (ACF) di un segnale è definita come la correlazione di un segnale con una copia ritardata del segnale stesso. Descrive quanto un segnale è simile a se stesso in funzione di un parametro temporale detto "lag".

Dato un segnale tempo continuo $f(t)$, l'autocorrelazione $R_{ff}(\tau)$ è definita come:

$$R_{ff}(\tau) = \int_{-\infty}^{+\infty} f(t) \overline{f(t - \tau)} dt \quad (1)$$

Autocorrelazione

Proprietà

- Simmetria:

$$R_{ff}(\tau) = R_{ff}(-\tau). \quad (2)$$

- Periodicità:

$$R_{ff}(\tau) = R_{ff}(\tau + T). \quad (3)$$

- Picco nell'origine:

$$|R_{ff}(\tau)| \leq R_{ff}(0). \quad (4)$$

- Convoluzione:

$$R_{ff}(\tau) = (f * g_{-1}(\bar{f}))(\tau). \quad (5)$$

Autocorrelazione

Analisi in short-time

I segnali vocali e sonori non sono perfettamente periodici e non sono stazionari, pertanto è necessario analizzarli operando in short-time, ovvero segmentando il segnale in brevi finestre temporali. In questo caso si definisce un'autocorrelazione in short time:

$$\phi_l(m) = \frac{1}{N} \sum_{n=0}^{N'-1} [x(n+l)w(n)][x(n+l+m)w(n+m)] \quad (6)$$

La correlazione è una funzione periodica con picco nell'origine. Al variare del lag τ l'ACF:

- raggiunge un minimo quando le onde sono in opposizione di fase
- raggiunge un massimo quando le onde sono perfettamente in fase

La difficoltà sta nell'identificare tali massimi e minimi nel caso di segnali non perfettamente periodici, dove le armoniche superiori e spurie generano a loro volta picchi e valli.

Il primo metodo calcola l'autocorrelazione lungo tutto il segnale.

Algorithm 1 simple_acf

- 1: Load audio data
 - 2: compute_acf
 - 3: find_peaks
 - 4: lag = first peak found
 - 5: compute pitch from lag
 - 6: plot power spectrum and spectrogram
-

Algoritmi

Metodo in short-time

Il secondo metodo calcola l'autocorrelazione in short time dopo aver effettuato una pre-elaborazione del segnale.

Algorithm 2 short_time_acf

```
1: Load audio data
2: trim_audio
3: Apply low pass filter
4: Compute clipping threshold
5: Apply non linearities
6: for i in range of #frames do
7:   correlate processed outputs
8:   compute lags in range [acf_min:acf_max]
9:   lag = argmax[x_corr]
10:  compute pitch from lag
11: end for
12: Plot and compare results with YIN
```

Il metodo YIN si compone di sei step:

- 1 Metodo dell'autocorrelazione**
- 2 Funzione alle differenze**
- 3 Funzione cumulativa alle differenze media normalizzata**
- 4 Soglia assoluta**
- 5 Interpolazione parabolica**
- 6 Miglior valore locale stimato**

Ogni step si fonda sul precedente e ne migliora l'errore della stima del pitch.

Virtual Pitch

Definizione

Il virtual pitch è il pitch di un tono complesso che non ha necessariamente un corrispettivo acustico nella frequenza alla quale viene percepito.

Quando la frequenza percepita non è fisicamente nel segnale, essa è ricostruita tramite riconoscimento di pattern acustici dalla coclea e dal cervello.

Virtual Pitch

Teoria della fondamentale mancante

Un suono si dice avere una fondamentale mancante, soppressa o fantasma quando *"i suoi overtoni suggeriscono una frequenza fondamentale ma il suono è privo di tale componente"*.

- Fenomeno psicoacustico il cui meccanismo non è ancora completamente chiaro
- Diretta conseguenza del virtual pitch
- Potrebbe rispondere ad un meccanismo simile a quello dell'autocorrelazione

Virtual Pitch

Teoria della fondamentale mancante

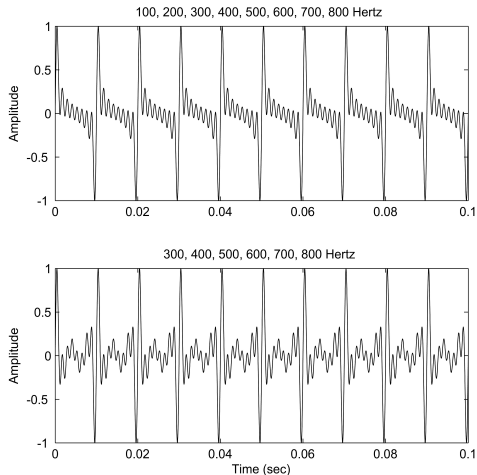


Figura: Effetto della mancanza di fondamentale su una serie armonica

Gli algoritmi sono stati testati utilizzando in ingresso diversi segnali. Gli esempi riportati sono relativi a:

- Un diapason, come segnale stazionario semplice
- Una linea vocale, come segnale armonicamente complesso e non stazionario

Risultati

STFT - Diapason

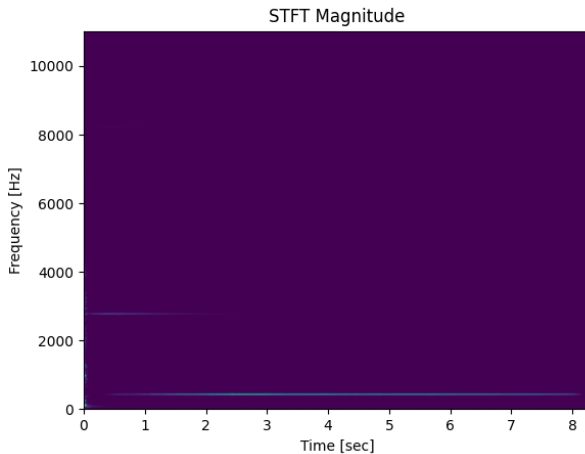


Figura: Spettrogramma del diapason

Risultati

STFT - Linea Vocale

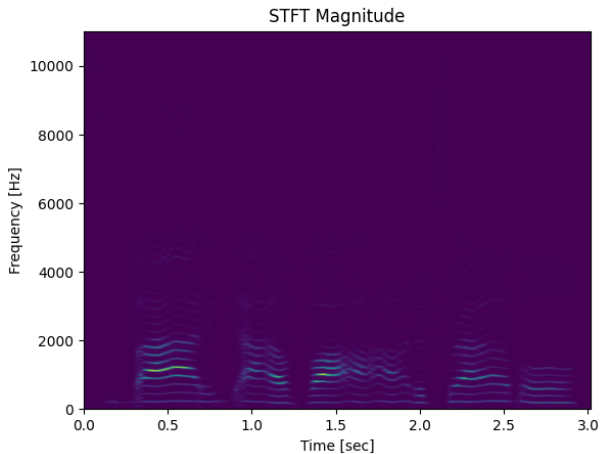


Figura: Spettrogramma della linea vocale

Risultati

Algoritmo 1 - Diapason

Il pitch del diapason rilevato dall'algoritmo 1 viene stampato a console e risulta essere corretto.

```
Sampling frequency: 22050 Hz  
Duration in samples 182711 samples  
Duration in seconds = 8.286213151927438 s  
pitch = 450.0  
Plotting spectrogram...  
Plotting spectrum in the range 0 to 1000 Hz
```

Per la natura stessa dell'algoritmo, non ha rilevanza il pitch trovato in segnali complessi, a causa della non stazionarietà degli stessi.

Risultati

Algoritmo 2 - Diapason

Il pitch rilevato dall'algoritmo 2 viene graficato e confrontato con i risultati dell'algoritmo YIN.

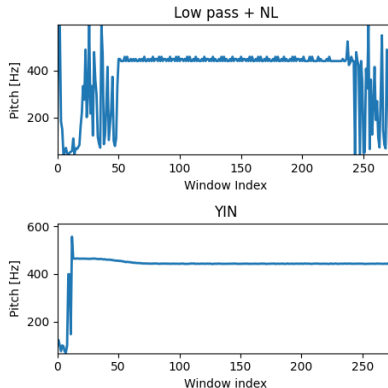


Figura: Pitch tracking del diapason. Sopra: algoritmo presentato. Sotto: algoritmo YIN.

Risultati

Algoritmo 2 - Linea vocale

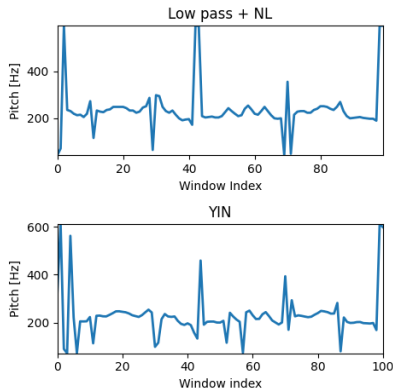


Figura: Pitch tracking della linea vocale. Sopra: algoritmo presentato. Sotto: algoritmo YIN.

L'algoritmo 2 è stato testato anche in caso di fondamentale mancante, utilizzando un filtro passa-alto con frequenza di taglio ben oltre la fondamentale rilevata originariamente dall'algoritmo. Il pitch rilevato in questo caso risulta avere componenti al di sotto di tale frequenza di taglio.

Risultati

Virtual Pitch

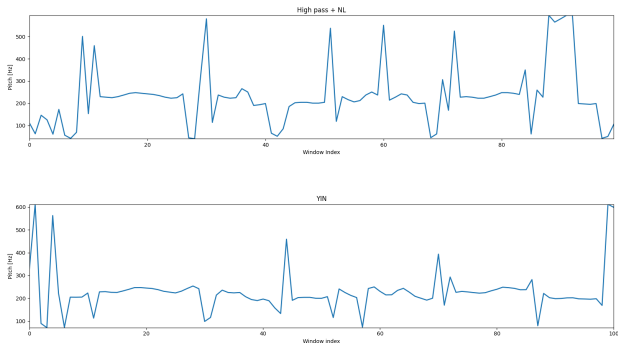


Figura: Virtual pitch detection con l'autocorrelazione.

Dagli esperimenti condotti si possono trarre le seguenti conclusioni:

- I risultati ottenuti dagli algoritmi sono soddisfacenti e coerenti con la teoria
- Il pre-processing del segnale è uno strumento estremamente utile per migliorare la precisione del pitch tracking
- Gli algoritmi sono robusti al fenomeno della fondamentale mancante

Gli algoritmi e l'elaborato associato sono visualizzabili e scaricabili dalla repository di Github

<https://github.com/armandoboemio98/PitchDetector>.