

# Wyznaczanie przebiegu melodii w sygnale muzycznym

Edward Półrolniczak, dr inż., Katedra Systemów Multimedialnych

# F0

- Jednym z podstawowych elementów systemów jest wyznaczanie częstotliwości podstawowej głosu (ang. fundamental frequency)
- Jest ona utożsamiana z wrażeniem wysokości dźwięku odbieranym przez nasz mózg
- W publikacjach częstotliwość podstawową oznacza się jako F0
- Częstotliwość podstawowa jest możliwa do oszacowania jednym z całego spectrum algorytmów do szacowania F0

# Cel zadania

- Zadanie polega na opracowaniu i implementacji podejścia do wyznaczania melodii w udostępnionych próbkach – uzyskany wektor powinien zawierać częstotliwości kolejnych dźwięków melodii lub ich zapis w notacji z czasem ich trwania (wektor dwuwierszowy z liczbą kolumn odpowiadającą liczbie wykrytych nut (nadmiarowe nuty usuwać)
- Szacowanie proszę wykonać na dostarczonych tutaj próbkach (próbki zawierają śpiew i melodię zagrąną z pianina) i na 2ch próbkach śpiewu z muzyką z internetu ;- ) („Wlaż kotek na płótek i mruga”)
- Opcjonalnie (za to ekstra punktowane) jest opracowanie i zbadanie skuteczności porównywania melodii (czyli wektorów cech F0)

*\*dźwięki można zaśpiewać lub zagrać na wirtualnym pianinie*

*<https://recursivearts.com/virtual-piano/>*

*Przy wciśnięciu klawisza pianina u góry widoczna jest muzyczna notacja dźwięku, np. C4, co może się przydać przy ocenie poprawności szacowania F0.*

*Oznaczenia te mają związek z częstotliwościami – można je znaleźć np. tutaj:*

*<https://pages.mtu.edu/~suits/notefreqs.html>*

*Z tabelki tam zawartej wiadomo np. że C4 to 277,18Hz (czyli jeśli ten dźwięk śpiewaliście to taka powinna być częstotliwość wynikowa*

# Pomoc do zadania

- Zadanie proszę zrealizować w Matlabie (lub Python) – podstawą jest funkcja *Pitch*
- Zwracam tutaj uwagę na fakt, nagrania są próbkowane z określoną częstotliwością
- Zwracam też uwagę na fakt, że w przypadku bardziej nieregularnych sygnałów, jak np. sygnał głosu, możliwe są błędy wskazujące częstotliwości w innych oktavach – można spróbować różnych metod, aby ustalić najbardziej skuteczną

# Pomoc do zadania

pitch (Audio Toolbox) – wyznacza częstotliwość podstawową (F0) sygnału audio

`f0 = pitch(audioln,fs)` returns estimates of the fundamental frequency over time for the audio input. Columns of the input are treated as individual channels.

`f0 = pitch(...,'Range',RANGE)` limits the search range for the pitch between the specified lower and upper band edges, inclusive. Specify range in Hz as a two-element row vector of increasing values. If unspecified, RANGE defaults to [50,400].

# Pomoc do zadania

`f0 = pitch(...,'WindowLength',WINDOWLENGTH)` specifies the analysis window length used to calculate pitch. Specify the window length in samples as a positive scalar. The maximum window length is 192000. If unspecified, WINDOWLENGTH defaults to `round(fs*0.052)`.

`f0 = pitch(...,'OverlapLength',OVERLAPLENGTH)` specifies the number of samples overlap between adjacent windows. Specify the overlap length as a positive scalar smaller than the window length. If unspecified, OVERLAPLENGTH defaults to `round(fs*(0.042))`.

# Pomoc do zadania

`f0 = pitch(...,'WindowLength',WINDOWLENGTH)` specifies the analysis window length used to calculate pitch. Specify the window length in samples as a positive scalar. The maximum window length is 192000. If unspecified, WINDOWLENGTH defaults to `round(fs*0.052)`.

`f0 = pitch(...,'OverlapLength',OVERLAPLENGTH)` specifies the number of samples overlap between adjacent windows. Specify the overlap length as a positive scalar smaller than the window length. If unspecified, OVERLAPLENGTH defaults to `round(fs*(0.042))`.



# Pomoc do zadania

`f0 = pitch(...,'Method',METHOD)` specifies the method used to calculate the pitch.

Valid inputs are:

- PEF - Pitch Estimation Filter

- NCF - Normalized Correlation Function

- CEP - Cepstrum

- LHS - Log-harmonic Summation

- SRH - Summation of Residual Harmonics

If unspecified, METHOD defaults to 'NCF'.



# Pomoc do zadania

`f0 = pitch(...,'MedianFilterLength',MEDIANFILTERLENGTH)` applies a median filter with specified window length. The median filter is a post-processing technique that operates on the estimated pitch values. If unspecified, `MEDIANFILTERLENGTH` defaults to 1 (filter disabled).

`[f0,loc] = pitch(...)` returns the location associated with each pitch decision. The location is the most recent sample used to calculate the pitch.

# Wynik

- Wyniki (oprócz opcji rozszerzającej) powinny dać się osiągnąć na zajęciach, a umieszczony powinien zostać w MTeams w wyznaczonym przez zadanie terminie – skrypt, wavy (całość jako zip)