

Wyznaczanie częstotliwości podstawowej (F0) głosu.

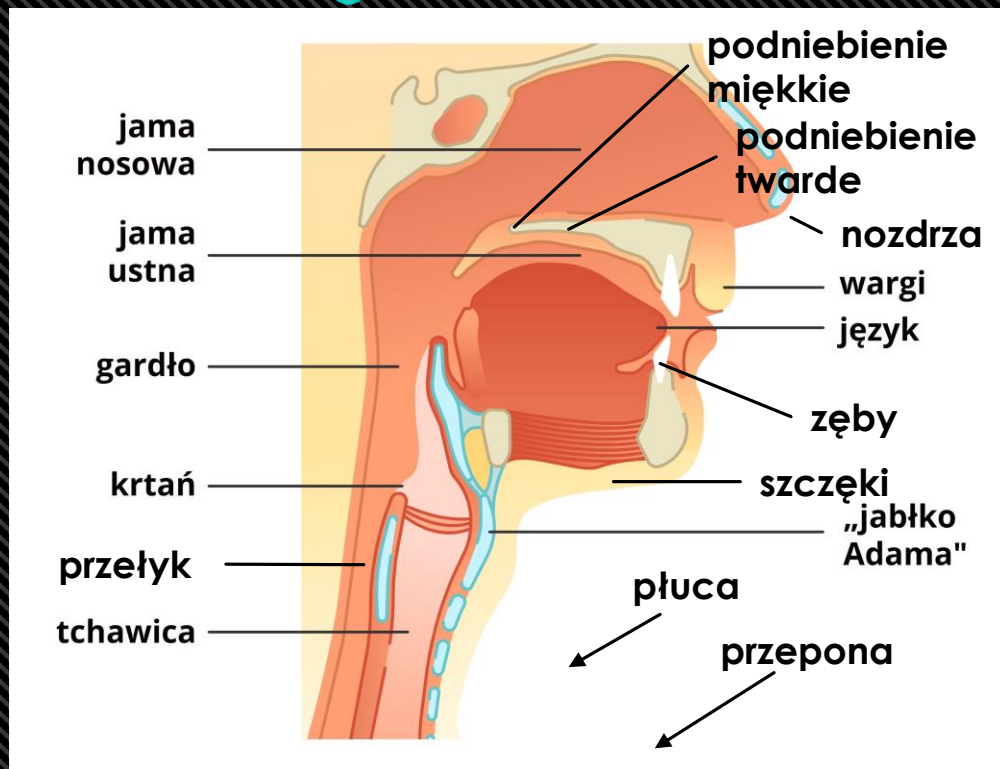
Edward Półrolniczak, dr inż., Katedra Systemów Multimedialnych

Rozpoznawanie mowy

- Jednym z podstawowych elementów systemów jest wyznaczanie częstotliwości podstawowej głosu (ang. fundamental frequency) a medycznie nazywanej tonem krtaniowym
- Jest ona utożsamiana z wrażeniem wysokości dźwięku odbieranym przez nasz mózg
- W publikacjach częstotliwość podstawową oznacza się jako F_0
- Częstotliwość podstawowa jest możliwa do oszacowania ponieważ fałdy głosowe drgają cyklicznie – powoduje to, że w przebiegu sygnału akustycznego głosu można wykryć regularności, na których podstawie można oszacować ton podstawowy
- Proszę zwrócić uwagę, że niektóre dźwięki wydawane przez nas nie muszą angażować fałd głosowych – wtedy jednak są to głoski bezdźwięczne – trudno wyznaczyć jego wysokość, czyli F_0

**ton krtaniowy generowany jest przez drgające cyklicznie fałdy głosowe*

Trakt głosowy



- Trakt głosowy jest konstrukcją złożoną z wielu „elementów” wyszczególnionych na rysunku obok
- F0 wytwarzane jest w fałdach głosowych w krtani

Cel zadania – wyznacz f_0 autokorelacją w dziedzinie czasu

- Zadanie polega na zastosowaniu metody autokorelacyjnej do oszacowania F_0
- Szacowanie proszę wykonać na dostarczonych tutaj próbkach i na 2ch próbkach swojego głosu (głoska ,a' zaśpiewana na konkretnych wysokościach dźwięków*) oraz 2ch wygenerowanych za pomocą narzędzia online: <http://onlinetonegenerator.com/> (wygenerowane próbki powinny się zapisać do wav za pomocą przycisku SAVE)

*dźwięki można zaśpiewać do wybranych dźwięków z klawiatury wirtualnego pianina

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

Przy wciśnięciu klawisza pianina u góry widoczna jest muzyczna notacja dźwięku, np. C4

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 będzie względnie łatwe do zrealizowania jeśli zostanie zrozumiany proces szacowania $F0$ metodą autokorelacyjną
- Metoda ta polega na wymnażaniu wartości cyklicznie przesuniętych funkcji (sygnałów) i ich sumowaniu
- Wykres funkcji autokorelacji na osi x prezentuje wartości przesunięcia sygnału (wartość początkowa ,0' sygnały nie są wzajemnie przesunięte, wartość końcowa – sygnały przesunięte w całym cyklu)
- Zagadnienie zostało dobrze omówione np. tutaj:
<http://www.speech.zone/autocorrelation/> (obejrzenie tego jest wręcz konieczne do wykonania zadania) (dobrze jest też ściągnąć stamtąd przykład i go obejrzyć w Excelu)

Pomoc do zadania

- Zadanie proszę zrealizować w Matlabie
- Zwracam tutaj uwagę na fakt, nagrania są próbkowane z określoną częstotliwością i mają one fundamentalne znaczenie w przeliczeniu położenia pików w funkcji autokorelacji na częstotliwość podstawową – przykładowo w materiale instruktażowym analizowany był sygnał próbkowany z częstotliwością 4kHz (4000Hz), pik został znaleziony w przesunięciu na odległość 40 próbek = $4000\text{Hz}/40 = 100\text{Hz}$ (F0)
- Proszę zwrócić uwagę na szerokość okna, które będzie używane do autokorelacji – powinno ono zawierać kilka okresów analizowanego sygnału
- Zwracam też uwagę na fakt, że w przypadku bardziej zakłóconych sygnałów, jak np. jednak rzeczywisty sygnał głosu metoda popełnia błędy wskazując częstotliwości w innych oktavach – jeśli błąd będzie tego rodzaju to należy uznać wynik za dopuszczalny na tym etapie (metody rozwojowe stosują analizę prawdopodobieństwa szacowania i wybierają wtedy właściwy pik, nawet jeśli mniejszy).
- Domyślnie F0 powinno być wskazane przez pierwszy pik lokalny (pomijamy pik w 0)

**Cel zadania – wyznaczyć f_0 autokorelacją
w dziedzinie częstotliwości**

Cel zadania – wyznaczn f_0 autokorelacją w dziedzinie częstotliwości

- Metoda autokorelacyjna wyznaczania f_0 w dziedzinie częstotliwości nazywana jest skrótem ACOLS (ang. Autocorrelation Of Log Spectrum) – ma to związek z przekształceniami jakie należy wykonać w celu osiągnięcia wyniku
- W celu uistalenia wartości f_0 należy dokonać analizy sygnału autokorelacji obliczonego na podstawie zlogarytmowanego widma sygnału wejściowego. Analiza polega na określeniu współrzędnej piku reprezentującego częstotliwość podstawową zlokalizowanego w uzyskanym wyniku w dziedzinie częstotliwości
- Kolejność działań:
 - Pobranie fragmentu nagrania (zakresu próbek)
 - Wyznaczenie widma pobranego fragmentu (w Matlabie `fft(s)`, s – sygnał)
 - Obliczenie wartości bezwzględnej widma (w Matlabie funkcja `abs`)
 - Logarytmowanie widma bezwzględnego (`abs`)
 - Autokorelacja (`xcorr`)

Cel zadania – wyznaczyć f_0 autokorelacją w dziedzinie częstotliwości

- Obliczenia zrealizować na tych samych próbkach, porównać wyniki (można dodać komentarz w kodzie źródłowym)

Wynik

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