

ZAAWANSOWANE TECHNIKI PRZETWARZANIA SYGNAŁÓW W TELEKOMUNIKACJI

Ćwiczenie 6a Dyskryminatory liniowe

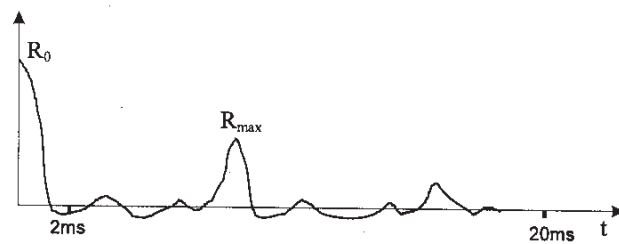
Uwagi ogólne

- Programy symulacyjne noszą nazwę `tester.m` i `dyskryminator.m`. Działają pod Matlab'em. Dyskryminatory w tym ćwiczeniu są zbudowane w celu rozróżnienia mowy dźwięcznej i bezdźwięcznej.
- Program `tester.m` przetwarza nagrania 5 mówców, z których można wybrać od 1 do 4 do stworzenia bazy treningowej a 1 (nie należący do bazy treningowej) tworzy bazę do testowania dyskryminatora.
- W obu programach wykorzystywane są parametry sygnału mowy pokazane na rysunku. Są to:
 - względna wielkość autokorelacji $r_{max} = R_{max}/R_0$, gdzie R_{max} - maksimum autokorelacji sygnału mowy w zakresie charakterystycznym dla tonu krtaniowego (2-20 ms), R_0 - autokorelacja przy zerowym przesunięciu, czyli energia sygnału mowy
 - dolnopasmowość $L_{lp} = E_{lp}/E_{sig}$, gdzie E_{lp} energia zawarta w zakresie niskich częstotliwości, E_{sig} - całkowita energia sygnału
 - predykcyjność (właściwie niepredykcyjność): $L_{res} = E_{res}/E_{sig}$, gdzie E_{res} - energia błędu predykcji.
- Można obserwować wartości tych parametrów dla bazy treningowej i testowej. Na wykresie 2-dim (jeśli wybrano 2 z 3 parametrów) lub 3-dim (jeśli wybrano 3 parametry) ukazuje się ponadto hiperpłaszczyzna podziału przestrzeni na dwa zbiory decyzyjne.
- Można też obserwować histogramy wartości funkcji dyskryminującej (Fishera) i próg decyzyjny. Listowane są też współczynniki funkcji dyskryminującej otrzymane dla bazy uczącej.

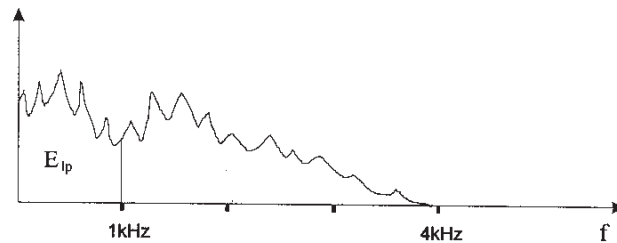
¹programy przygotował Paweł Konopko

- Najistotniejszą informacją jest stopa błędów w bazie uczącej i dla mówcy testowego.
- **Program dyskryminator.m** pozwala na dokonywanie nagrań mowy, ma też dostęp do szeregu fraz testowych w plikach .wav.
- Pozwala na obserwacje w funkcji czasu:
 - decyzji V/UV (dźwięczny - bezdźwięczny),
 - parametrów r_{max} , L_{lp} , L_{res} , mocy sygnału oraz wartości funkcji dyskryminującej w kolejnych oknach

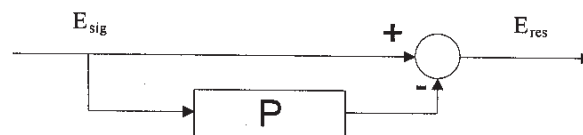
KLASYFIKACJA MOWA DŹWIĘCZNA/BEZDŹWIĘCZNA PARAMETRY



max znormal. autokorelacji: $r_{max} = R_{max} / R_0$



dolnopasmowość: $L_{lp} = E_{lp} / E_{sig}$



predykcyjność: $L_{res} = E_{res} / E_{sig}$

Zadania do wykonania

1. W programie **tester.m** ustaw 4 mówców w bazie uczącej i 1 w testowej. U szereguj 3 parametry pod względem "mocy dystynktywnej", tzn zbadaj ich wpływ na stopę błędów dyskryminacji. Badania przeprowadź dwiema metodami:
 - usuwając 1 parametr i sprawdzając stopę błędów w obu bazach
 - uruchamiając dyskryminator z jednym parametrem.
2. Wymień mówcę "testowego" i powtórz badania. Czy wnioski co do mocy dystynktywnej są identyczne?
3. Przyjrzyj się znakom współczynników funkcji dyskryminującej. Które z nich powinny być dodatnie, a które ujemne?
4. W jaki sposób można określić moc dystynktywną parametrów mowy na podstawie 2-dim lub 3-dim wykresów ich wartości dla mowy dźwięcznej i bezdźwięcznej oraz położenia hiperpłaszczyzny podziału?
5. W programie **dyskryminator.m** zarejestruj kilka słów lub fraz mowy, zawierających głoski dźwięczne i bezdźwięczne. Uruchom dyskryminator i obserwuj wyniki (r_{max} , L_{lp} , L_{res} , moc sygnału oraz wartości funkcji dyskryminującej w kolejnych oknach). Jakie błędy występują? Czy dochodzi do "udźwięczniania" głosek bezdźwięcznych i odwrotnie?