

ZAAWANSOWANE TECHNIKI PRZETWARZANIA SYGNAŁÓW W TELEKOMUNIKACJI

Ćwiczenie 6b

Rozpoznawanie słów z wykorzystaniem Ukrytych Modeli Markowa

Przebieg ćwiczenia

Program symulacyjny nosi nazwę HMM.m. Działa pod Matlab'em. Uruchamia się po wpisaniu instrukcji HMM lub kliknięciu w HMM.fig. W "Pomocy" przedstawiono 3 etapy ćwiczenia. Jest do nich dostęp po wybraniu "Rozpocznij ćwiczenie".

1. Trening HMM

Modele HMM są już zaprojektowane dla bazy danych zawierającej 20 słów (cyfry i komendy dla robota). Każdy z 20 HMM jest projektowany dla 90 wypowiedzi tego samego słowa. W tej sytuacji **nie korzystamy z opcji "Uczenie algorytmu"**, gdyż otrzymalibyśmy, po długim okresie oczekiwania, te same HMM.

Zaprojektowano HMM o strukturze:

- ergodycznej
- Bakisa z dojściem do ostatniego węzła
- szeregowej z dojściem do ostatniego węzła

Dla każdego HMM można obejrzeć macierz przejść A (a_{ij} - prawdopodobieństwo przejścia z węzła i do węzła j) i wektor prawdopodobieństw startu ścieżki π (π_j - prawdopodobieństwo startu z j -go węzła).

ZADANIE 1. Dla wybranego słowa narysuj (na podstawie macierzy A i wektora π) struktury trzech HMM: ergodycznego, Bakisa i szeregowego.

2. Analiza algorytmu

Testowanie przeprowadza się na podstawie 53 wypowiedzi każdego z 20 słów. Artykułowane są one przez mówców, których wypowiedzi nie należały do bazy uczącej. Po wybraniu opcji

¹ programy przygotowała Joanna Grabowska

"Weryfikacja algorytmu" odczytujemy średnią stopę błędów i liczbę błędów (w 53 testach) rozpoznawania każdego z 20 słów.

ZADANIE 2. Dla 3 struktur HMM porównaj stopy błędów rozpoznawania. Dla których słów stopa błędów jest największa? (postaraj się uzasadnić dlaczego). Która struktura oferuje najmniejszą stopę błędów i dlaczego?

3. Test algorytmu

Po wybraniu opcji "Testuj algorytm" można otworzyć plik (wav) zawierający pojedynczą wypowiedź. Można też nagrać i zapisać (do pliku o nazwie test.wav) własną wypowiedź. W tym drugim przypadku należy sprawdzić, czy plik zawiera całe wypowiedziane słowo. Należy też sprawdzić czy część słowa nie została usunięta w procesie eliminacji ciszy poprzedzającej słowo i następującej po nim.

ZADANIE 3. Wybierz opcję "Oblicz parametry sygnału" i obserwuj zmiany w czasie energii oraz kilku współczynników mel-cepstrum. Jaki sens fizyczny ma pierwszy współczynnik cepstrum?

ZADANIE 4. Wybieraj opcję "Testuj" dla kilku słów z plików .wav (np. szesc_1.wav, chwytak_1.wav, jeden_1.wav) i dla własnych nagrań (plik test.wav). Przetwarzanie każdego pliku należy rozpocząć od "Otwórz plik" i "Oblicz parametry sygnału". Zbadaj 3 struktury HMM. Obserwuj zlogarytmowane gęstości prawdopodobieństwa ("loglik") dla 20 modeli HMM i badanej wypowiedzi. Czy rozpoznanie było poprawne? Jeśli nie, to jaki wystąpił błąd? Jeśli tak, to jaki był margines? (różnica między "loglik'iem" dla właściwego HMM i największym z "loglik'ów" dla niewłaściwych HMM). Która struktura HMM zapewnia największy margines?