Plan pracy:

1. Preliminaria
   1. Definicja problemu rozpoznawania
   2. Funkcja dyskryminująca
   3. Modele parametryczne i nieparametryczne
2. Ukryte modele Markowa
   1. Definicja
   2. Algorytm forward
   3. Algorytm Viterbi’ego
   4. Algortym EM
   5. Algorytm Baum’a-Welch’a
   6. HMM w systemach ASR
      1. Embedded training, beam search, tied states
      2. Viterbi Training
      3. Skala logarytmiczna
3. Ekstrakcja cech
   1. Wstęp
   2. Algorytm MFCC
4. Jednostki fonetyczne
   1. Fonemy
   2. STU
   3. Alfabety fonetyczne
5. Hybryda HMM-MLP
   1. Wady GMM-HMM
   2. Trening dyskryminujący
   3. Architektura hybrydy
   4. Algorytm uczenia
6. Specyfikacja techniczna
   1. Diagram klas
   2. Karty CRC
   3. Scenariusze uczenia
   4. Skalowalność
   5. Szczegóły implementacyjne
7. Eksperymenty
   1. Środowisko testowe i metodyka pomiarów
   2. Korpus Językowy AN4
   3. Rezultaty
   4. Omówienie wyników
8. Literatura

|  |  |
| --- | --- |
| [1] | J.B. Hampshire, B.A. Pearlmutter “Equivalence Proofs for Multi-Layer Perceptron Classifiers and the Bayesian Discriminant Function” |
| [2] | D. Jurafsky, J.H. Martin “Speech and language processing” |
| [3] | L. E. Baum, T. Petrie, G. Soules, and N. Weiss "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains" |
| [4] | Sakoe, H. and Chiba, S., “Dynamic programming algorithm optimization for spoken word recognition” |
| [5] | C. S. Myers and L. R. Rabiner. “A comparative study of several dynamic time-warping algorithms for connected word recognition.” |
| [6] | Aldrich, John "R.A. Fisher and the making of maximum likelihood 1912-1922” |
| [7] | H. Bourlard, N. Morgan “Connectionist Speech Recognition A Hybrid Approach” |
| [8] | R. Gemello, D. Albesano, F. Mana, “Continuous Speech Recognition with Neural Networks and Stationary-Transitional Acoustic Units” |
| [9] | R. Gemello, D. Albesano, F. Mana, "Multi-source neural networks for speech recognition |
| [10] | R. Gemello, D. Albesano, F. Mana “CSELT Hybrid HMM/Neural Networks Technology for Continuos Speech Recognition” |
| [11] | M.A. Franzini, K.F. Lee and A. Waibel, "Connectionist Viterbi Training: A new hybrid method for continuous speech recognition" |
| [12] | S. Renals , N. Morgan, H. Bourlard, M. Cohen, H. Franco “Connectionist Probability Estimators in HMM Speech Recognition” |
| [13] | H. Hermansky, “Perceptual linear predictive analysis of speech” |
| [14] | H. Hermansky, N. Morgan, “RASTA Processing of Speech” |
| [15] | R. Gemello, D. Albesano, F. Mana, “Multi-source neural networks for speech recognition” |
| [16] | R. De Mori, R. Gemello, D. Albesano, F. Mana, “Ear-model derived features for automatic speech recognition” |
| [17] | R. Gemello, D. Albesano, F. Mana, P. Pegoraro, “Multi Source Neural Networks based on fixed and Multiple Resolution Analysis for Speech Recognition” |
| [18] | R. Gemello, D. Albesano, F. Mana, “Hybrid HMM-NN for speech recognition and prior class probabilities” |
| [19] | L. Fissore, F. Ravera, P. Laface, “Acoustic-Phonetic Modeling For Flexible Vocabulary Speech Recognition” |
| [20] | Otwarty słownik transkrypcji fonetycznej j. Polskiego V.11.2007 <http://www.elektronikjk.republika.pl/y11.html> |
| [21] | Andrew J. Viterbi. “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm” |
| [22] | <http://www.cslu.ogi.edu/people/hosom/cs552/> |
| [23] | <http://www.stanford.edu/class/cs224s/> |
| [24] | R. Gemello, D. Albesano, F. Mana, “Hybrid HMM-NN for speech recognition and prior class probabilities” |