ECE874 Introduction to Neural Networks Curse Presentation

Speaker Verification Using Series of LVQ Networks

Pálinkó Oszkár

Overview

- Automated biometric authentication
- Speaker recognition – in general
- Speech modeling
- Speaker verifier structure
- LVQ training
- Decision making
- Results and conclusions

Automated biometric authentication

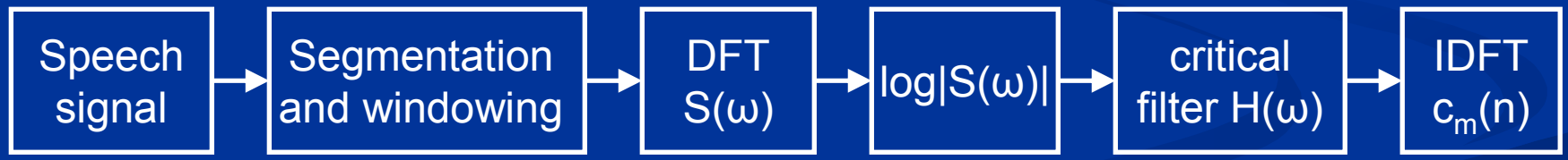
- Non-biometric authentication
- Biometrics – measurement of physiological and behavioral characteristics for authentication
- Authentication based on pattern recognition
- Recognition methods: fingerprint, hand geometry, retina, iris, face, signature, speaker

Speaker recognition

- Based on the speech and way of speaking
- A natural, non-intrusive method
- Both physiological and behavioral characteristic
- Verification vs. Identification
- Main steps: speech modeling (feature extraction), training (classification), recognition

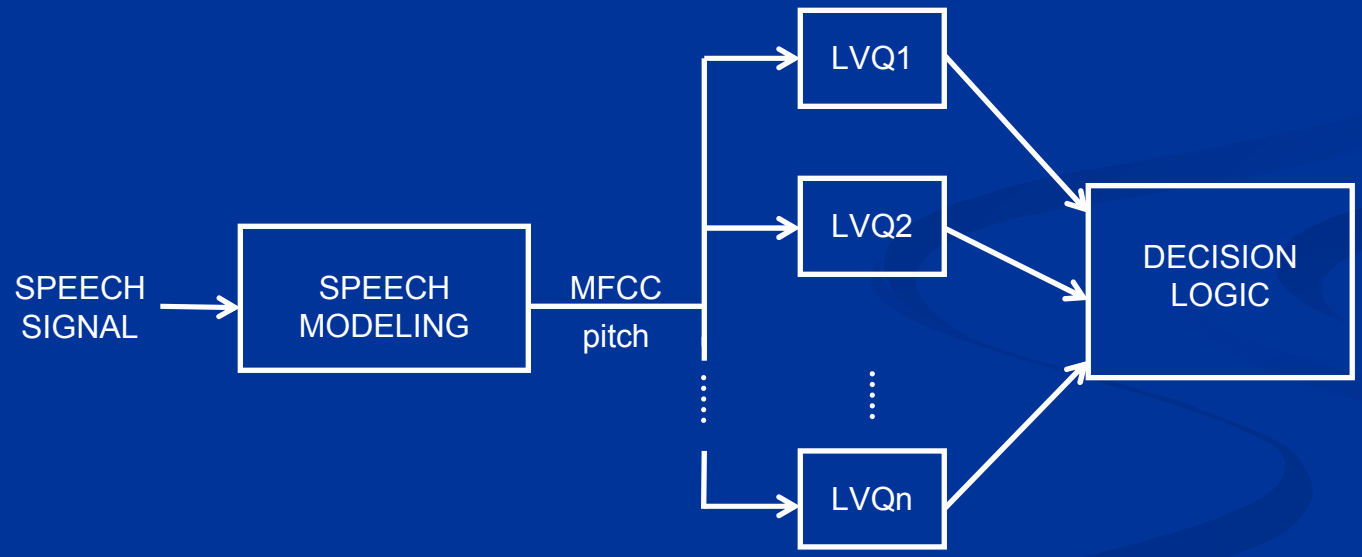
Speech modeling

- Feature extraction
- Appropriate features for authentication: mel-frequency cepstral coefficients (MFCC) and the pitch
- Generating the MFCCs:



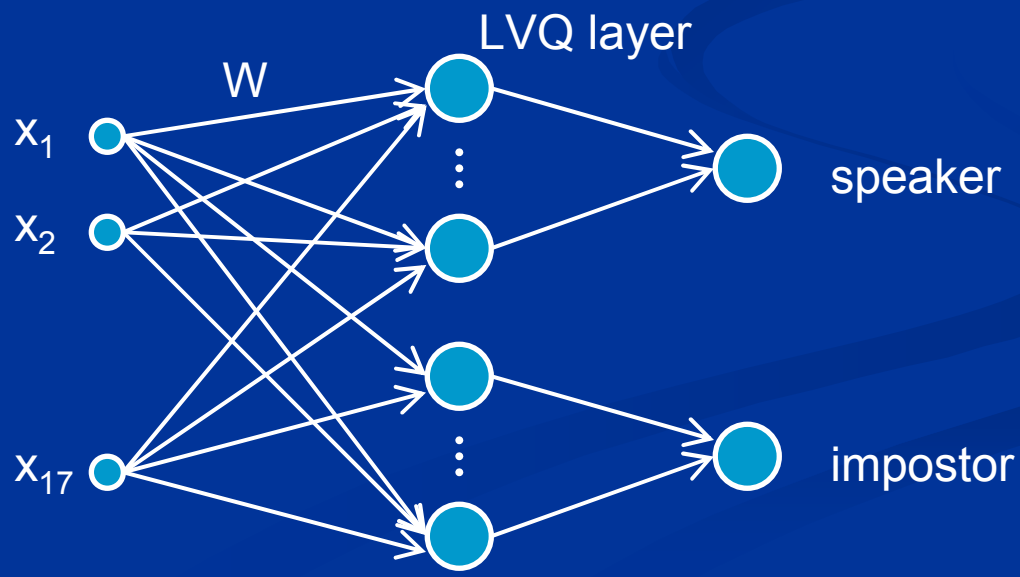
Speaker Verifier Structure

- The verification system is based on a sequence of n LVQ networks



LVQ Training

- Speaker vs. one of the impostors
- The input vector - 16th order MFCC and the pitch
- 18 codewords assigned to one of the final classes
- 45 seconds of training data



Decision making

- LVQ - highly discriminative
- Decision making – based on 3 second segments, every network gives a probability value in this interval
- Every ANN has to perform over 50%, half of those over 60%

Results and Conclusions

- Eight subjects, 8kHz data, 45s training, 3s recognition
- 96% final recognition accuracy
- Further development: higher quality input data (16kHz), use of delta mel coefficients, Markov-modeling

Thank you for your attention.