DEEE725 음성신호처리실습
Speech Signal Processing Lab
Instructor: Gil-Jin Jang

진행방법: 2시간 강의 + 2시간 실습 (유동적)
Topics:
- Review of digital signal processing
- Fundamentals of speech production and perception
- Speech coding / Audio coding
- Speech recognition
Prerequisite: Probability and Random Process, Digital Signal Processing

주교재
[Edinburgh] https://www.inf.ed.ac.uk/teaching/courses/asr/lectures-2023.html
[Rabiner] Theory and Applications of Digital Speech Processing, Lawrence Rabiner and Ronald Schafer, Prentice Hall (Pearson Education), 2011.
부교재(계속 추가될 예정)
[Kondoz] Digital Speech: Coding for Low Bit Rate Communication Systems, 2nd edition, A. M. Kondoz, John Wiley & Sons, 2004.
[Bosi] Introduction to Digital Audio Coding and Standards, Marina Bosi and Richard E. Goldberg, Springer, 2002.
[HTK] The HTK book:
http://www.ee.columbia.edu/ln/labrosa/doc/HTKBook21/HTKBook.html
http://htk.eng.cam.ac.uk/docs/docs.shtml

Detailed Topics
==========
1. Fundamentals of speech signal processing
. Basics of Digital Signal Processing
. Introduction to Digital Speech Processing
. Fundamentals of Human Speech Production
. Hearing, Auditory Models, and Perception
. Sound Propagation in the Human Vocal Tract
. STFT (short-time Fourier transform)

2. Speech coding
. Sampling and Quantization

. Linear predictive coding: Levinson-Durbin recursion
. V/UV (voiced / unvoiced) detection
. Pitch estimation
. Excitation modeling
. LPC to LSF Conversion
. Efficient LSF Quantization
. Speech codec (coder/decoder) design

3. Noise suppression
. Voice activity detection and endpoint detection
. Noise spectrum estimation
. Spectral subtraction
. Wiener filtering
. Speech reconstruction from spectrum

4. Speech feature extraction
. Linear / Mel Filterbank design
. Log filterbank energy extraction
. MFCC (mel-frequency cepstral coefficients)

5. DTW (dynamic time warping)
. DTW by dynamic programming
. Word recognition using DTW

6. HMM (hidden Markov model)
. First-order Markov model
. Observation probability modelling
. Viterbi decoding
. Forward-backward algorithm
. Segmental k-means
. Baum-Welch Reestimation
. Basic word recognition

7. Speech recognition applications and backend techniques
. GMM (Gaussian mixture model) for observation probability modeling
. NN (Neural Networks) for observation probability modeling
. Connected word recognition
. Phoneme recognition
. Continuous sentence recognition from phoneme recognition results

. Language models

8. Advanced topics
. Deep learning for speech recognition
. Speech recognition tools: HTK and Kaldi

==========
* Programming Assignments
- Speech codec implementation
- Noise suppression
- Speech feature extraction and DTW
- Word recognition using HMM
- HMM implementation using Neural networks
- Connected word recognition

==========
Programming assignments
- ~~Python 3.0~~ Python 3 버전
- 개인 github 만드는 것 강력하게 추천 개인 github에 과제 업로드해서 채점하는 방식으로 할지 생각 중
==========