

Bài 1a: Nhận dạng giới tính giọng nói

Input: tín hiệu tiếng nói

Output: F0 trung bình và giới tính (Nam/Nữ)

Yêu cầu:

- Xác định thuật toán tính F0 nào (ACF/AMDF/FFT/HPS/...) cho ĐCX cao nhất (sai số tương đối giữa F0tb tính được và F0tb tham chiếu là nhỏ nhất) → SV xuất Bảng TK dữ liệu sai số của ít nhất 2 thuật toán dựa trên tập DL huấn luyện **trong thư mục TinHieuHuanLuyen-44k**.
- Báo cáo ĐCX nhận dạng giới tính (Nam/Nữ) (%) của thuật toán tính F0 được lựa chọn trên tập DL kiểm thử **trong thư mục TinHieuKiemThu-44k (GV upload sau)**.
- Chú ý: Cần hậu xử lý (ví dụ: lọc trung vị) để loại bỏ các F0 lệch nhiều (khoảng 20%) so với các giá trị F0 của các khung lân cận.

TLTK: Tinh F0 dung ham tu tuong quan_Luận văn_TV Tam_2019.pdf

Bài 1b: Phân tích đặc trưng phổ các nguyên âm của nhiều người nói

Input: tín hiệu tiếng nói

Output:

- ảnh phổ (spectrogram) và bộ 3 tần số formant (F1, F2, F3) của mỗi nguyên âm được đo trực tiếp trên ảnh phổ
- nhận xét về sự khác biệt đặc trưng phổ giữa:
 - các nguyên âm khác nhau của cùng một người nói.
 - một nguyên âm của các người nói khác nhau.

Yêu cầu:

- Xuất ảnh phổ băng rộng (wideband spectrogram) của các file DL huấn luyện (của 08 người nói **chọn ngẫu nhiên từ 21 người trong thư mục NguyenAmHuanLuyen-16k**).
- Xuất bảng dữ liệu bộ 3 tần số formant của mỗi nguyên âm và đưa ra nhận xét về sự khác biệt giữa đặc trưng phổ giữa 05 nguyên âm của một người nói **trong 08 người đã chọn**.
- So sánh bộ 3 tần số formant của một nguyên âm của các người nói khác nhau và đưa ra nhận xét về sự khác biệt giữa đặc trưng phổ của **08 người nói đã chọn**.
- Dữ liệu 3 tần số formant chỉ cần đo thủ công (cách đo xem phần 3.7.2 trong [3]).

TLTK:

- [1] Bài TH 1-Frequency-domain processing.pdf (phần 1 về spectrogram).
- [2] Spectrogram-Cepstrum-and Mel-Frequency Analysis_CMU_PPT (slide 11-18 về ý tưởng phân biệt các nguyên âm với nhau bằng bộ 3 tần số formant và cách đo tần số formant trên spectrogram).
- [3] Phân tích formant các nguyên âm của nhiều người nói_Luận văn_PĐThiện_2021 (phần 2.2.1 về khái niệm formant, phần 3.5 và 3.6.2 về một số kết quả đo để tham khảo về dải giá trị của F1, F2 và F3 của các nguyên âm).

Kiểm tra và đánh giá (trọng số điểm 40%):

GV kiểm tra chương trình và báo cáo của mỗi nhóm trong tuần sau. Các SV trong nhóm tự phân công nhiệm vụ. Khi GV chấm bài, mỗi SV trình bày riêng phần code và báo cáo (bằng slide) kết quả phần việc mà SV đó làm.

Chú ý:

- SV chỉ báo cáo cách làm step-by-step và kết quả (bằng hình ảnh và số liệu cụ thể), KHÔNG báo cáo lý thuyết.
- SV viết RIÊNG báo cáo và code, làm nhóm chỉ để chia sẻ kiến thức và kỹ năng. Bảng dữ liệu kết quả thì gộp chung để đưa ra nhận xét chung.

Thời gian kiểm tra trên Team riêng theo lớp HP:

- 1910A: 8h00 thứ năm 6/1.
- 1911A: 13h30 thứ năm 6/1.
- 1912A: 8h00 thứ sáu 7/1.
- 1913A: 13h30 thứ sáu 7/1.
- 1914A: 8h00 thứ bảy 8/1.
- 1915A: 13h30 thứ bảy 8/1.