

Mỗi SV viết báo cáo MS-Word theo trình tự sau:

1. Thu thập tín hiệu (signal acquisition):

Mỗi SV thu âm 05 file tín hiệu chứa 5 nguyên âm /a/, /e/, /i/, /o/, /u/ của chính giọng nói của mình, dùng tính năng “record” của phần mềm xử lý âm thanh (ví dụ, Audacity). Sau đó dùng phần mềm cắt lấy 1 khung tín hiệu ổn định ở phần chính giữa (không cần quá chính xác) của mỗi nguyên âm dài 30 ms (milliseconds) để dùng cho các Bước 3 và 4 tiếp theo.

Các chú ý khi thu âm:

- định dạng file: format .WAV, tần số lấy mẫu (sampling frequency/sample rate) là 16 kHz, chế độ đơn kênh (mono), độ phân giải 16 bit.
- không đọc nhầm /e/ thành /ê/, /o/ thành /ô/.
- thu trong môi trường yên tĩnh nhất có thể để giảm ảnh hưởng của nhiễu môi trường.
- có khoảng lặng dài khoảng 0,3 đến 1 giây ở hai đầu của mỗi file tín hiệu.

2. Tìm tần số cơ bản (fundamental frequency hay F0):

a. Đo thủ công:

Mỗi SV đo chu kỳ cơ bản T_0 (fundamental period, chính xác đến mili giây) của khung tín hiệu đã cắt ra ở Bước 1 của mỗi nguyên âm bằng cách đo thủ công một chu kỳ nào đó của sóng tín hiệu (waveform) hiển thị trên phần mềm xử lý âm thanh (zoom in tín hiệu để đo chính xác hơn). Tính nghịch đảo của T_0 để thu được ước lượng (estimate) của tần số cơ bản F_0 (Hz) đo trên miền thời gian (F_0 được làm tròn, không lấy số lẻ ở phần thập phân). Với mỗi nguyên âm, SV thu được 1 ước lượng F_0 thủ công. SV so sánh và nêu nhận xét về 5 ước lượng F_0 thủ công thu được.

b. Tính tự động:

SV dùng chương trình (mỗi SV tự code) xuất ra đường F_0 của cả file tín hiệu nguyên âm (SV nào chưa thi GK bài này thì tự code thuật toán tính F_0 dùng hàm tự tương quan `xcorr()` và hàm tìm cực đại cục bộ `findpeaks()` của Matlab). Sau đó tính giá trị F_0 trung bình của mỗi file tín hiệu nguyên âm và dùng giá trị này làm ước lượng F_0 tự động. SV tính sai số tương đối (%) của ước lượng F_0 tự động so với giá trị tham chiếu là ước lượng F_0 thủ công. Báo cáo kết quả với 5 nguyên âm.

3. Viết chương trình phân tích phổ của các khung tín hiệu dài 30 ms của các nguyên âm dùng Fast Fourier Transform (FFT) và vẽ các đồ thị phổ biên độ thang logarithmic (log magnitude spectrum, theo dB-decibel) của các nguyên âm. Tham khảo: <https://vlab.amrita.edu/?sub=3&brch=164&sim=908&cnt=2> (Figure 2, chỉ dùng “voiced speech” vì nguyên âm là âm hữu thanh). Đồ thị sẽ có dạng gần giống phổ vạch (line spectrum) do tín hiệu nguyên âm gần như tuần hoàn.

4. Dựa vào 2 đồ thị phổ biên độ và dữ liệu của đồ thị lưu trong Matlab, SV đo các tham số sau của mỗi nguyên âm:

- Tần số cơ bản F_0 : chính là khoảng cách giữa các vạch phổ trên trục tần số (đơn vị Hz). Giá trị này gọi là ước lượng của tần số cơ bản F_0 đo trên miền tần số. Với mỗi nguyên âm, SV tính sai số tương đối (%) của ước lượng F_0 trên miền tần số so với giá trị tham chiếu là ước lượng F_0 thủ công trên miền thời gian (đo ở mục 2).

- 3 tần số cộng hưởng đầu tiên F_1, F_2, F_3 (gọi là formant frequencies hay formants, theo Hz) của mỗi nguyên âm. Các formant chính là các đỉnh (peak) của đường bao (envelope) của đồ thị phổ biên độ. Với mỗi formant F_i , SV đo 3-dB bandwidth B_i (theo Hz) và giá trị phổ biên độ G_i (theo dB) tại tần số formant. Sau bước này, SV nhận được 3 cặp giá trị (F_i, B_i, G_i) ($i=1,2,3$) cho mỗi nguyên âm. (xem [https://en.wikipedia.org/wiki/Bandwidth_\(signal_processing\)](https://en.wikipedia.org/wiki/Bandwidth_(signal_processing)), mục “x dB bandwidth” để hiểu thuật ngữ “3-dB bandwidth”).

5. SV tính giá trị trung bình của 5 ước lượng F_0 thủ công trên miền thời gian để được giá trị F_0 đại diện cho cao độ của giọng nói của chính mình.

Giá trị F_0 đại diện này và 3 cặp giá trị (F_i, B_i, G_i) ($i=1,2,3$) cho mỗi nguyên âm là kết quả phân tích nguyên âm, sẽ được SV dùng làm bài TH 2 về tổng hợp nguyên âm dùng như các tham số đầu vào của chương trình.

Chú ý:

- SV đưa kết quả thực hiện (dưới dạng đồ thị/hình vẽ và số liệu cụ thể) của từng bước thực hiện vào báo cáo.
- Sản phẩm nộp cho GV: thư mục chứa các file tín hiệu và báo cáo (file PDF).
- Hướng dẫn thực hành: Tuần đầu tiên sau khi thi cuối kỳ xong.
- Thời hạn nộp báo cáo: Tuần sau đó.