

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO HỌC PHẦN
XỬ LÍ TÍN HIỆU SỐ

GIẢNG VIÊN HD: TS. NINH KHÁNH DUY

SINH VIÊN THỰC HIỆN:

NGUYỄN ĐẮC ĐỨC

- MSV: 102200166 - 20TCLC_DT4

PHẠM VĂN TIẾN TRƯỞNG

- MSV: 102200198 - 20TCLC_DT4

PHẠM CÔNG HUY

- MSV: 102200174 - 20TCLC_DT4

NGUYỄN ĐẠI

- MSV: 102200164 - 20TCLC_DT4



MỤC TIÊU

1

➔ PHÂN TÍCH ĐẶC TRƯNG PHỔ CỦA CÁC NGUYÊN ÂM NHIỀU
NGƯỜI NÓI ĐƯA RA NHẬN XÉT

2

➔ NHẬN DẠNG NGUYÊN ÂM KHÔNG PHỤ THUỘC NGƯỜI NÓI
DÙNG ĐẶC TRƯNG PHỔ FFT.

3

➔ NHẬN DẠNG NGUYÊN ÂM KHÔNG PHỤ THUỘC NGƯỜI NÓI
DÙNG ĐẶC TRƯNG PHỔ MFCC

PHÂN TÍCH ĐẶC TRƯNG PHỔ CỦA CÁC NGUYÊN ÂM NHIỀU NGƯỜI NÓI ĐƯA RA NHẬN XÉT

Bảng phân công

STT	Nội dung công việc	Sinh viên thực hiện
1	Xuất ảnh phổ băng rộng	Phạm Văn Tiến Trưởng, Nguyễn Đại
2	Xuất bảng tần số format, tính các giá trị để đưa ra nhận xét, kết luận	Phạm Văn Tiến Trưởng, Nguyễn Đại
3	Nhận xét sự khác biệt đặc trưng phổ 5 nguyên âm của 1 người	Phạm Văn Tiến Trưởng
4	Nhận xét sự khác biệt đặc trưng phổ 1 nguyên âm của các người nói khác nhau	Nguyễn Đại
5	Nhận xét chung và kết luận	Phạm Văn Tiến Trưởng, Nguyễn Đại

BÀI 1

PHÂN TÍCH ĐẶC TRƯNG PHỔ CỦA CÁC NGUYÊN ÂM NHIỀU NGƯỜI NÓI ĐƯA RA NHẬN XÉT

1

CÁC BƯỚC THỰC HIỆN

1

Nguyên âm huấn luyện



2

Xuất ảnh phổ và bộ 3 tần số Formant



3

Nhận xét sự khác biệt đặt trưng về phổ

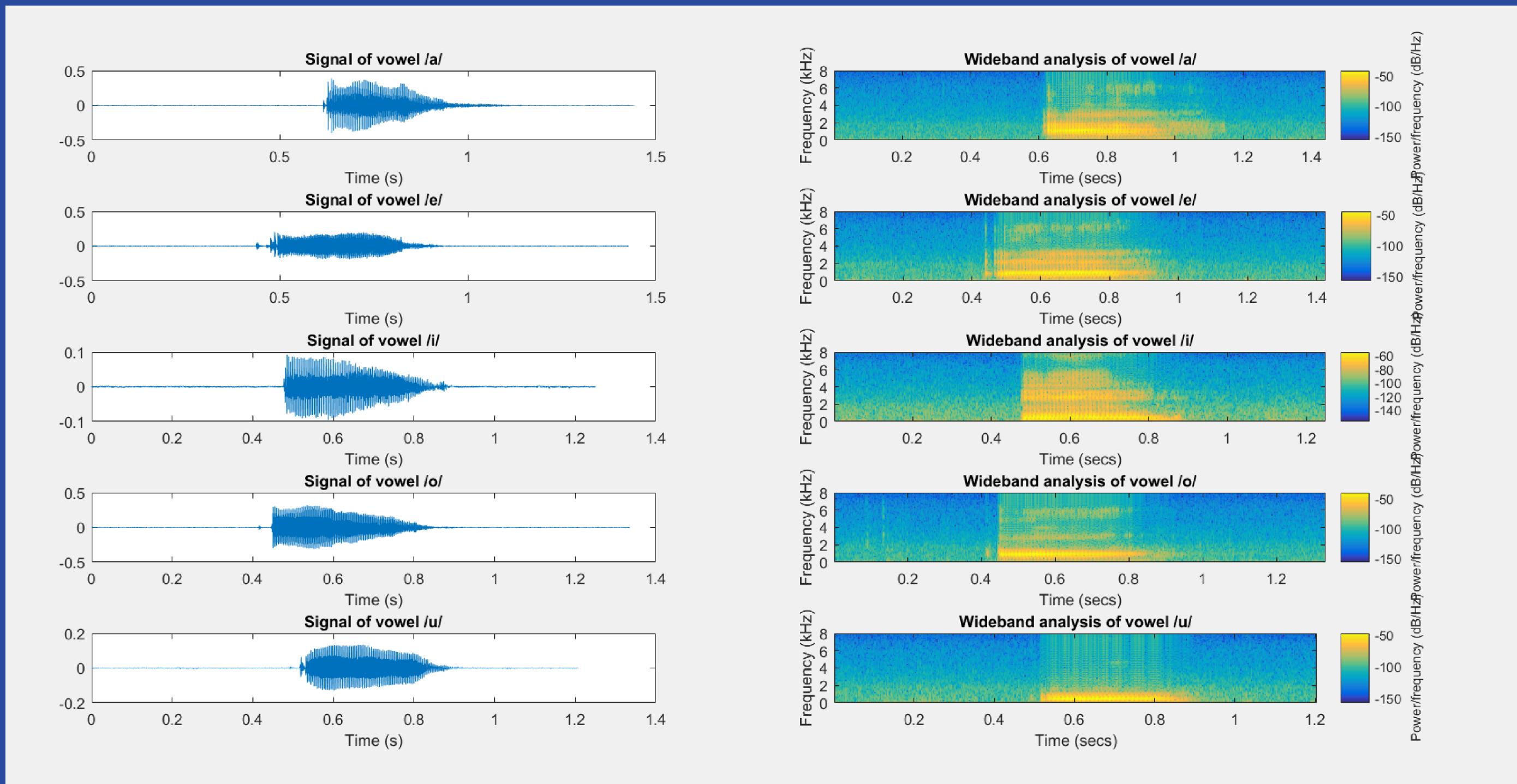


4

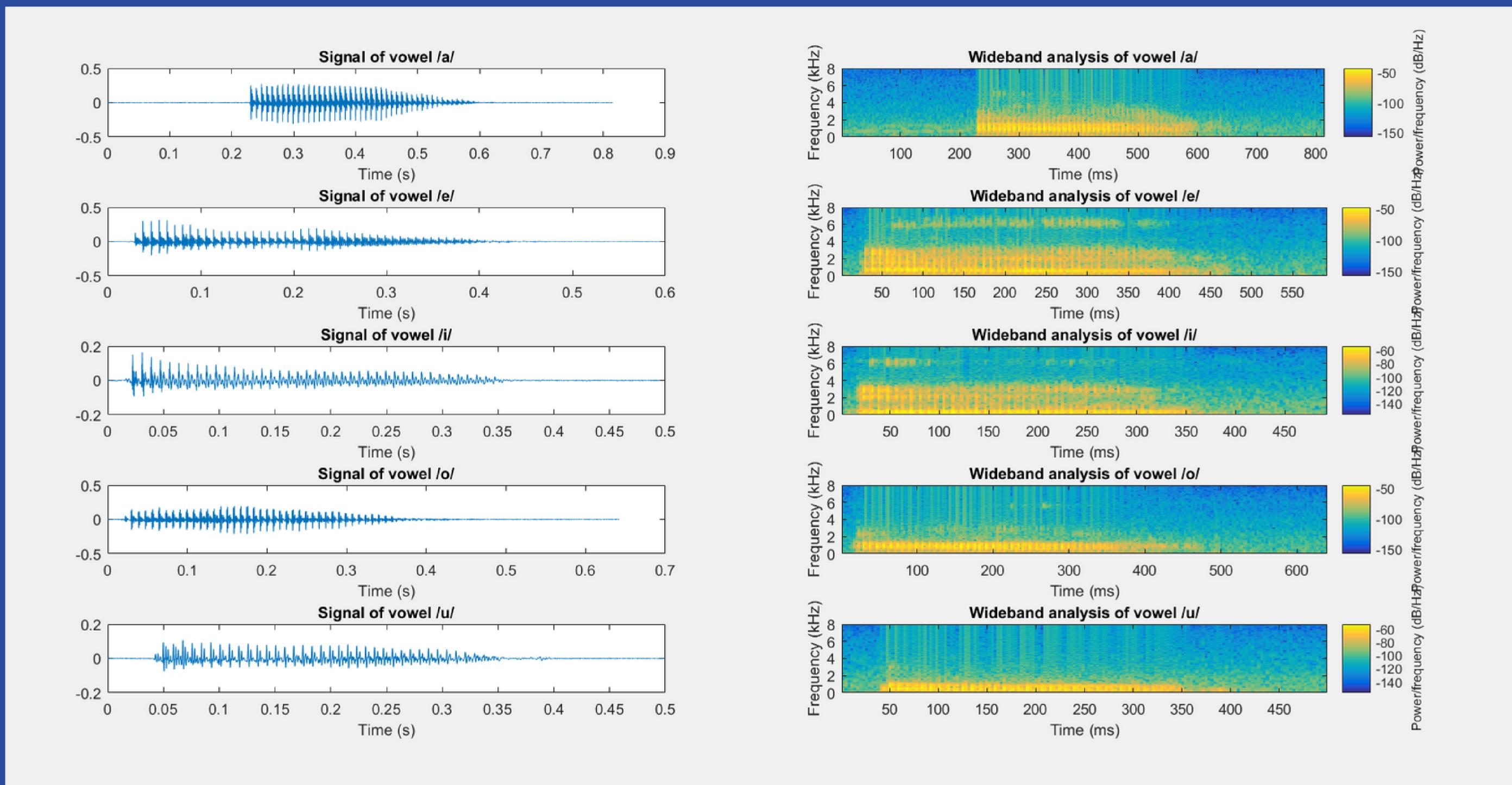
Nhận xét chung và kết luận



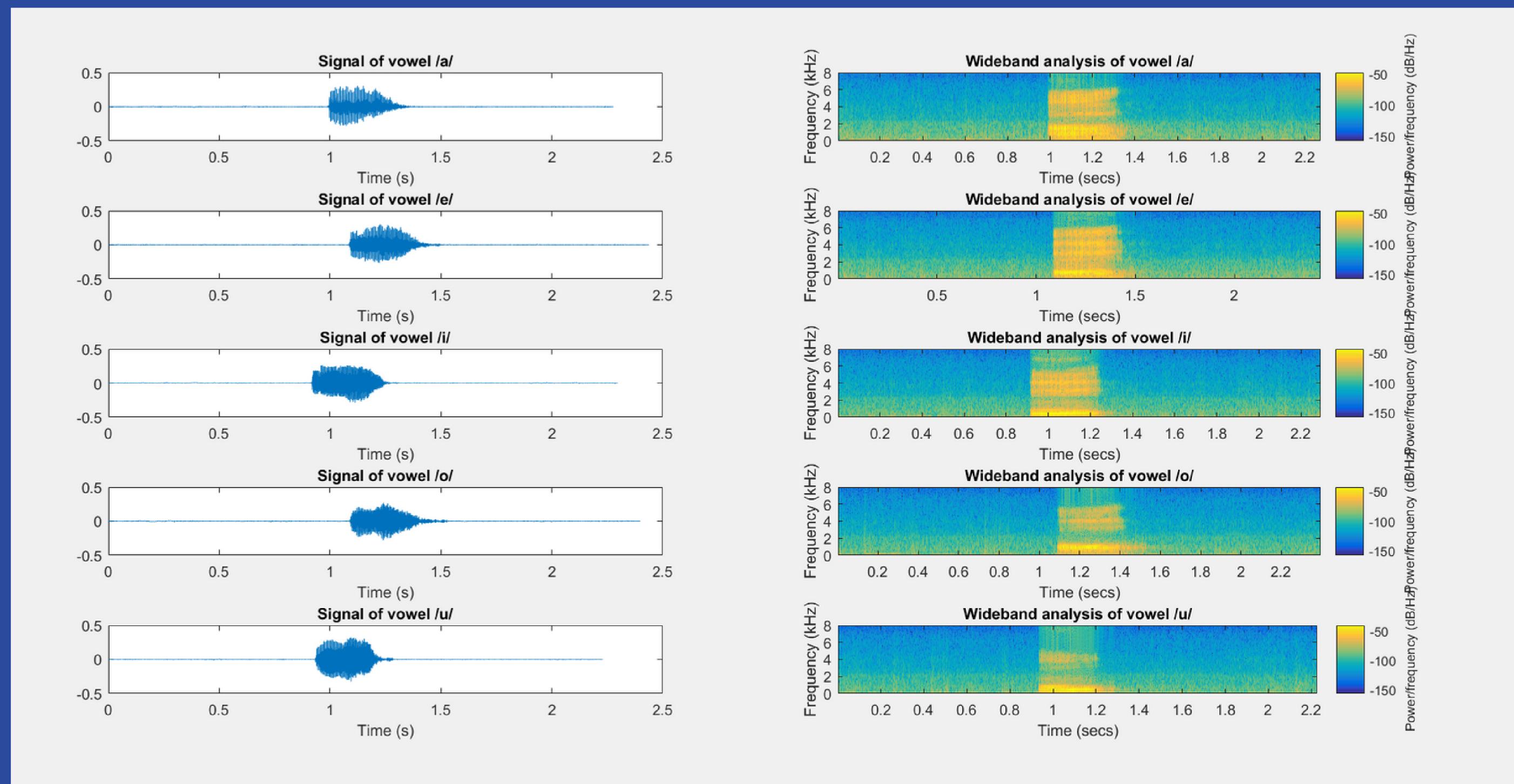
02FVA



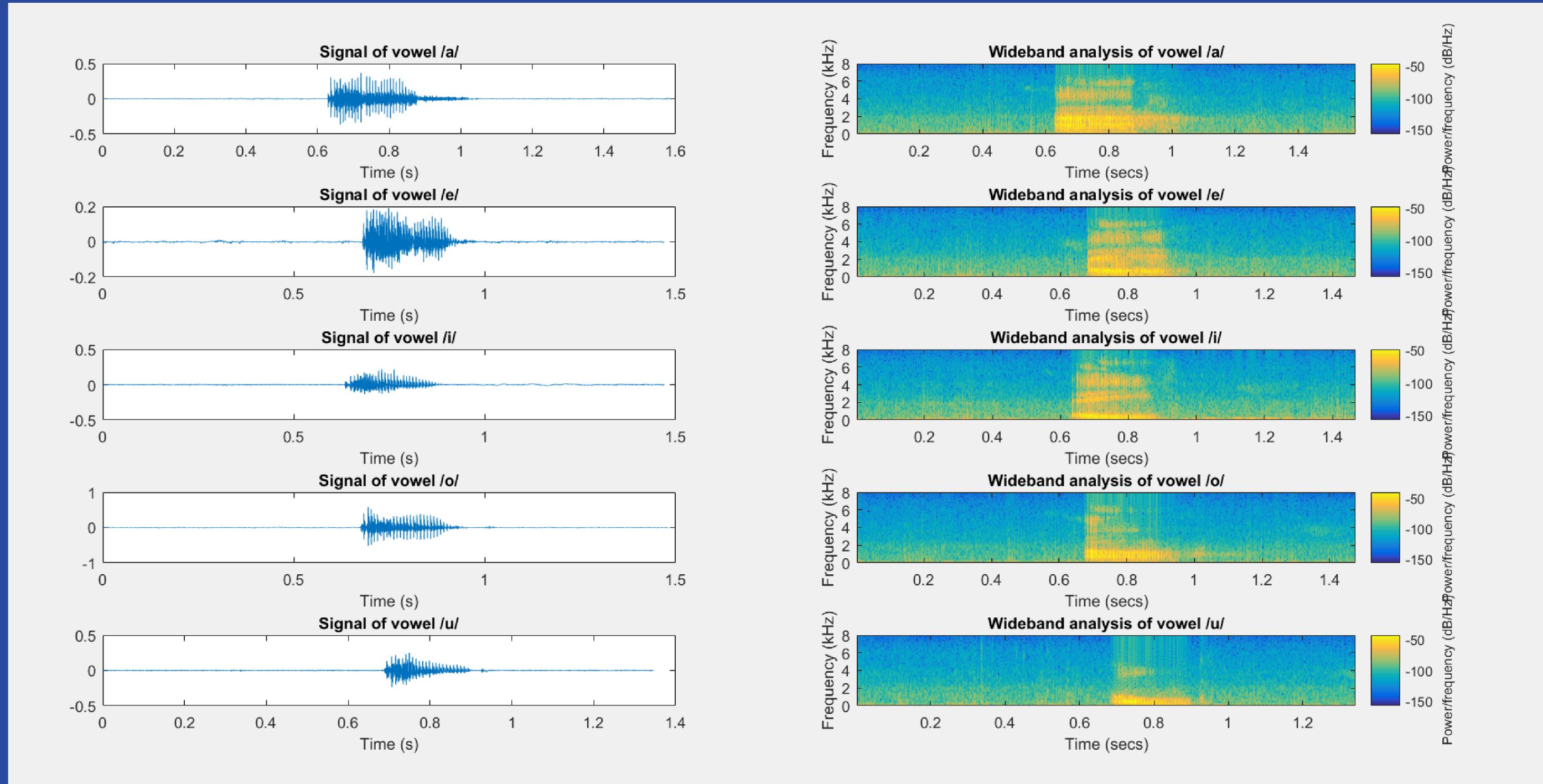
04MHB



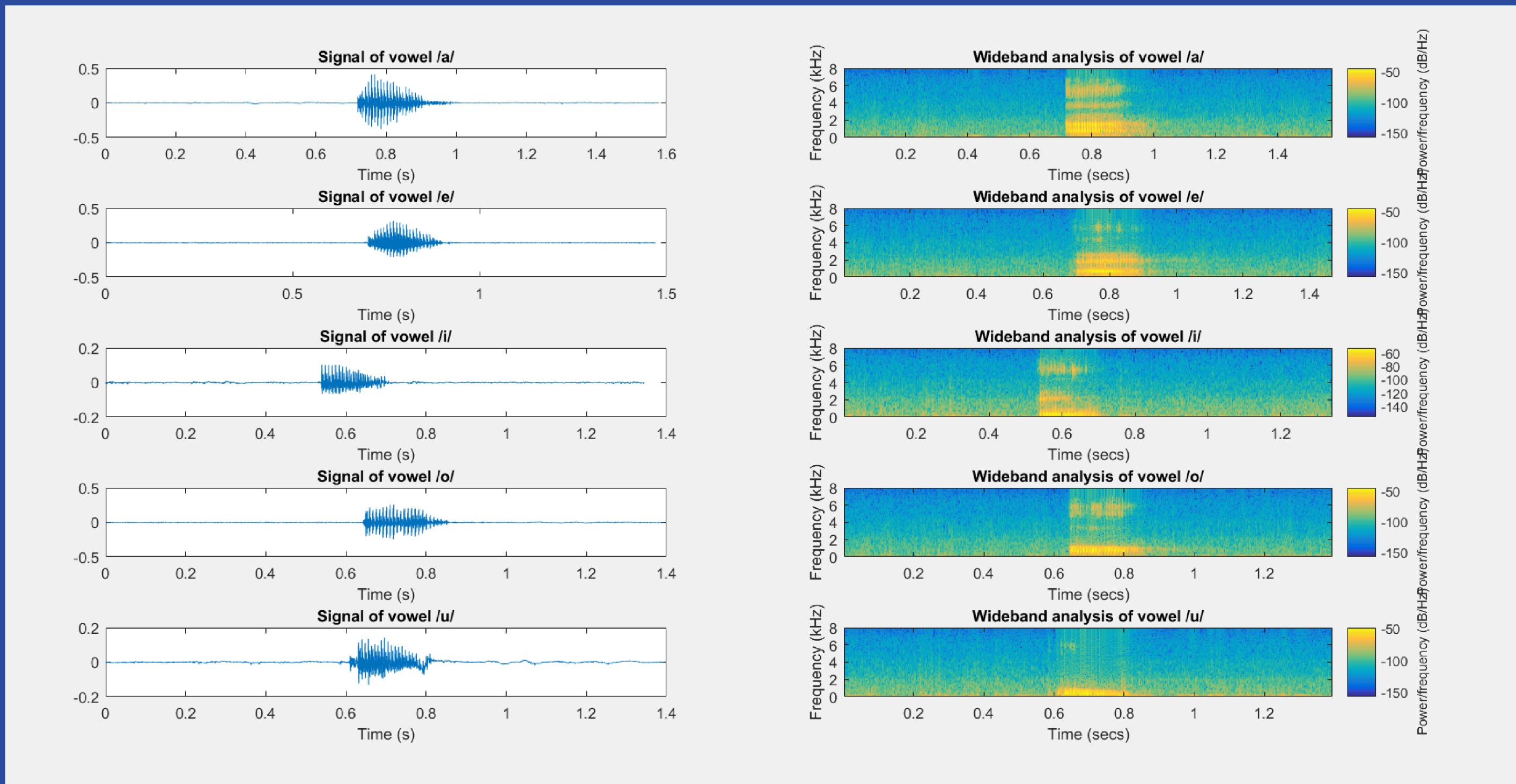
06FTB



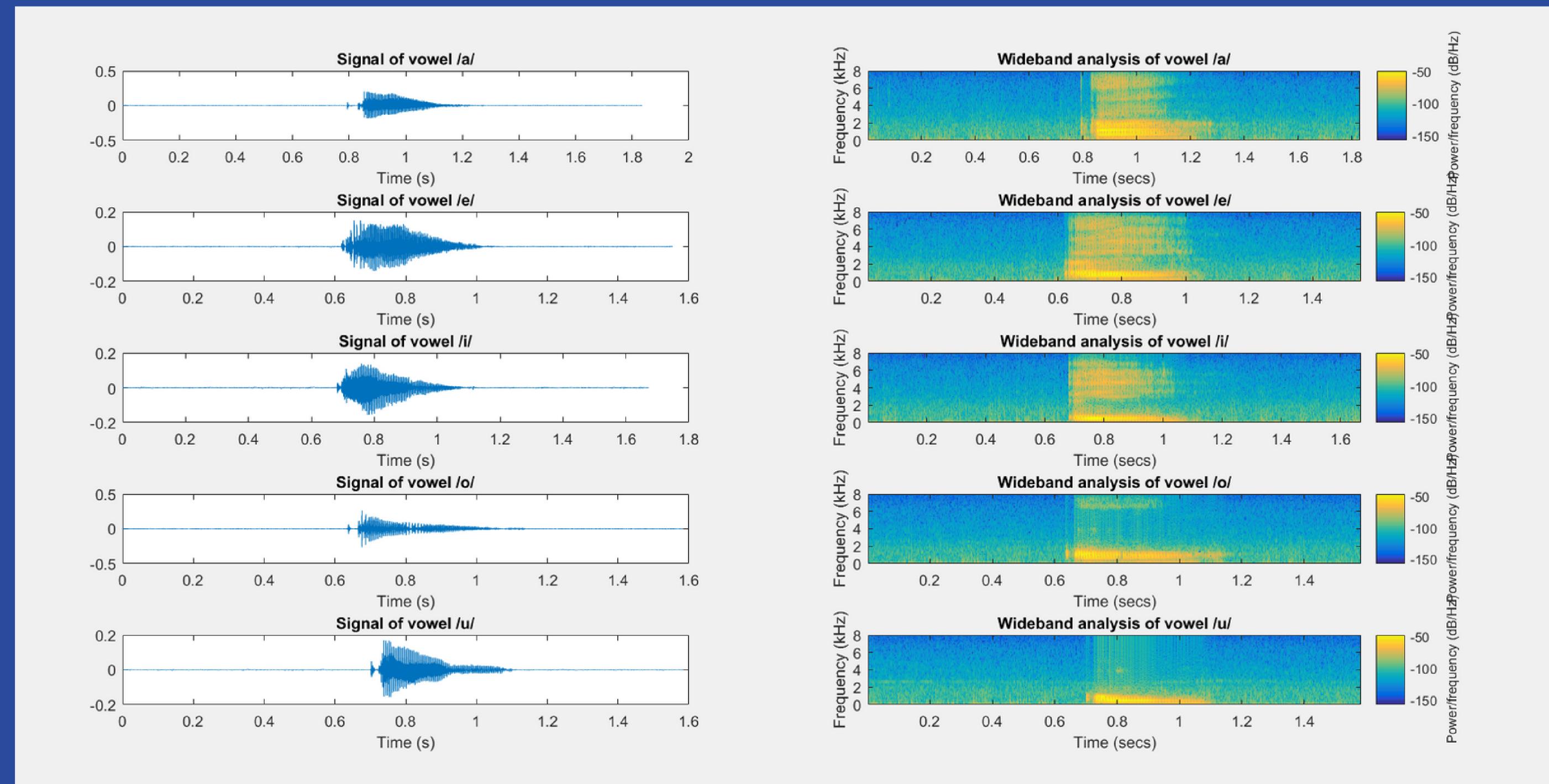
08MLD



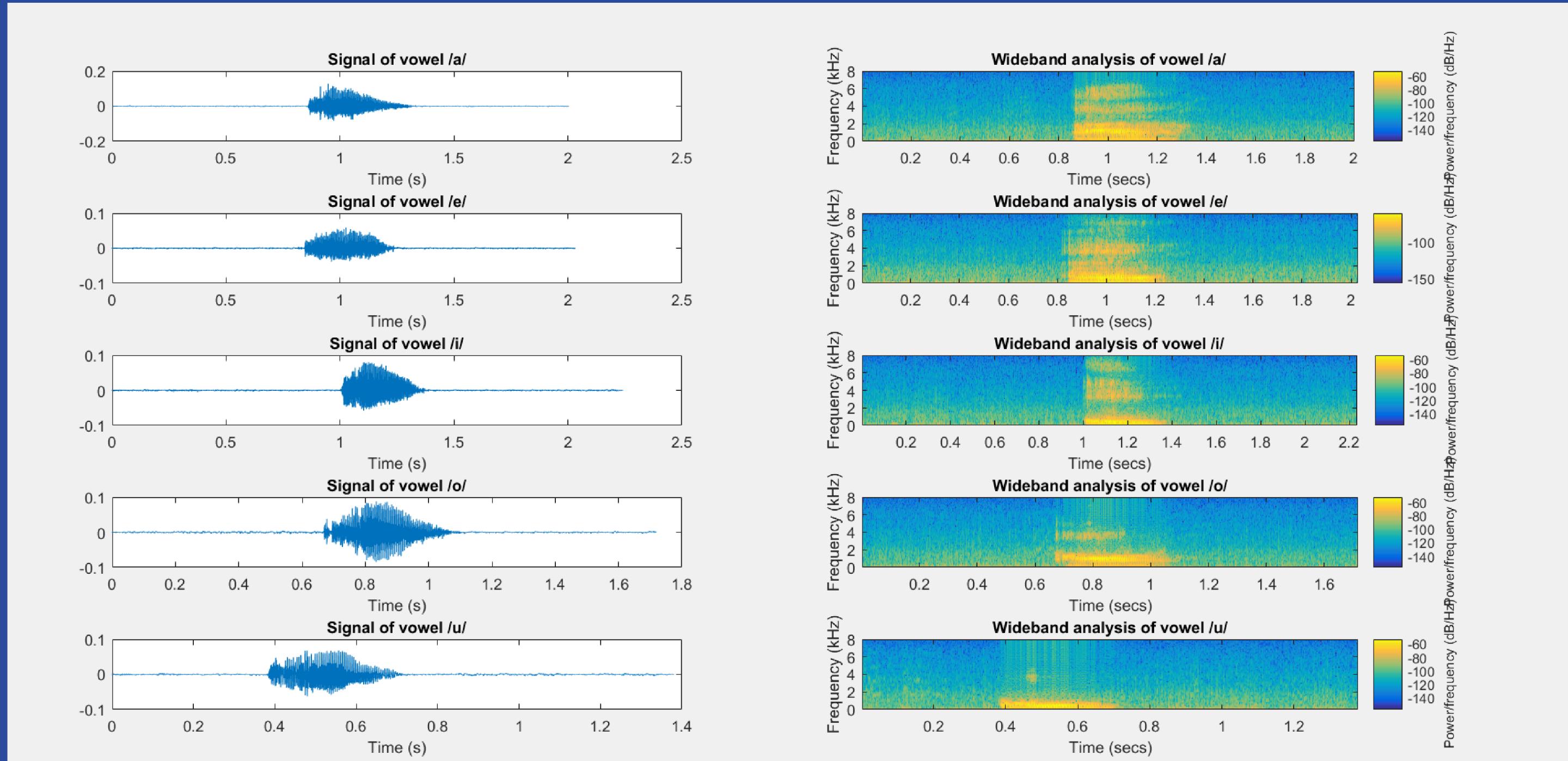
10MSD



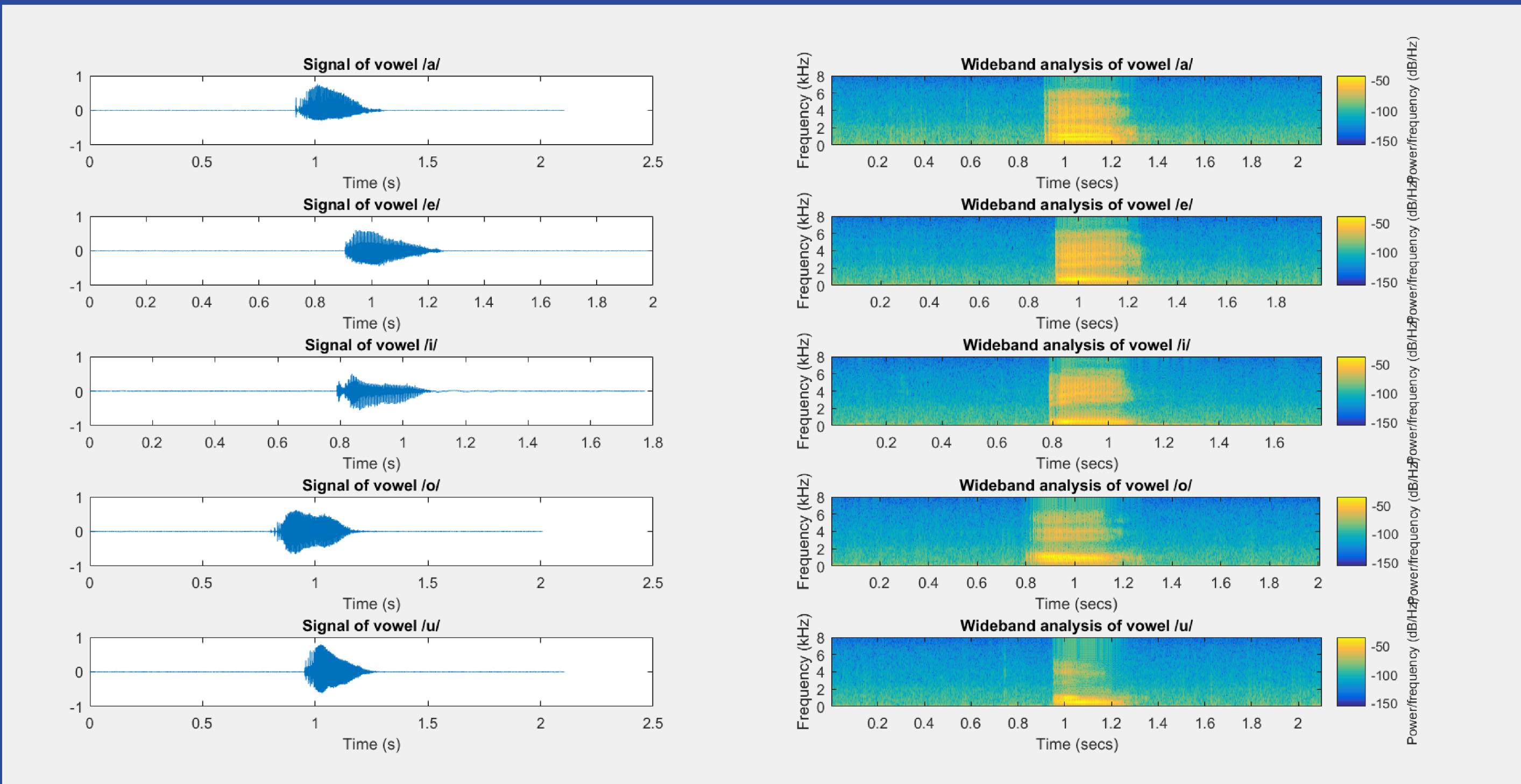
12FTD



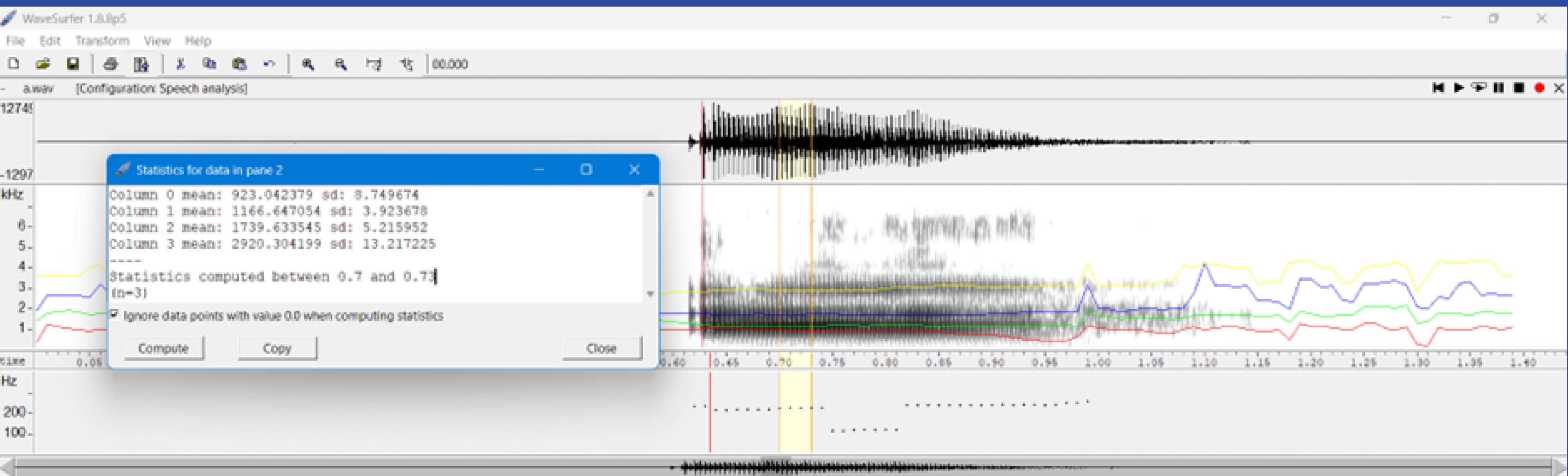
14FHH



16FTH



ĐO BỘ 3 TẦN SỐ FORMANT



- DÙNG WAVESURFER ĐO 3 LẦN TẦN SỐ
- FORMANT TRÊN 3 KHUNG KHÁC NHAU CÓ
- KHOẢNG CÁCH LÀ 30MS

BỘ 3 TẦN SỐ FORMANT

Nguyên âm	02FVA			04MHB			06FTB			08MLD		
	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
/a/	1004.54	1130.29	1736.57	895.21	1313.02	2099.24	776.978	1527.92	3098.63	1161.13	1794.71	2729.42
/e/	764.48	916.34	2134.78	901.3	1319.65	1883.17	738.87	2131.22	3080.17	664.45	2230.75	2943.26
/i/	592.84	2754.94	3352.06	379.68	2084.56	2829.03	422.24	1066.16	3100.3	357.72	2423.98	2913.92
/o/	904.056	1161.23	2867.26	796.37	1047.53	2868.42	832.58	1025.37	3270.77	734.13	1198.21	2408.75
/u/	480.96	822.91	1621.23	414.04	822.84	1752.17	420.2	780.94	1782.38	411.5	754.19	3699.34

Nguyên âm	10MSD			12FTD			14FHH			16FTH		
	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
/a/	746.85	1528.77	2348.94	748.54	1210.03	1903.82	1216.48	1853.28	3467.18	738.19	1269.95	2140.38
/e/	690.09	1902.7	2558.16	759.35	997.97	2150.83	648.97	2325.45	3627.41	773.34	2413.03	3623.07
/i/	347.86	2158.83	2977	478.69	3133.56	3498.65	371.78	2772.48	3624.05	492.88	2759.03	3540.21
/o/	712.71	981.4	2108.48	763.49	1131.73	1531.79	838.62	999.27	3365.23	795.07	1165.12	3232.73
/u/	369.76	633.15	1645.2	503.06	843.12	1922.33	393.69	683.8	3742.85	509.65	987.96	2975.01

BỘ 3 TẦN SỐ FORMANT

<i>/a/</i>	F1	F2	F3
Mean(Hz)	910.99	1453.5	2440.52
STD	194.953	268.493	605.822
CV(%)	21.40%	18.47%	24.82%

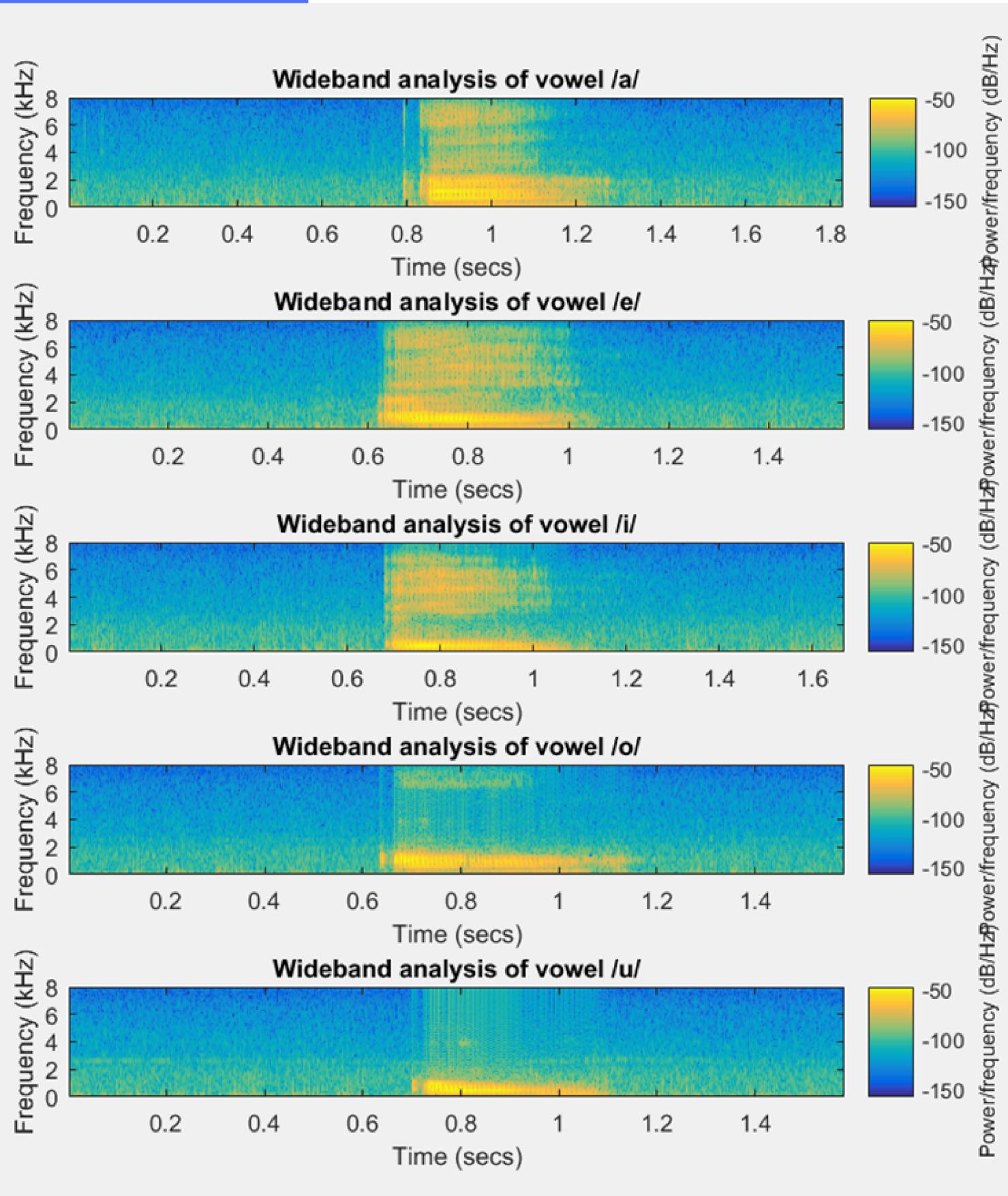
<i>/e/</i>	F1	F2	F3
Mean(Hz)	742.606	1779.64	2750.11
STD	79.7632	610.53	676.563
CV(%)	10.74%	34.31%	24.60%

<i>/o/</i>	F1	F2	F3
Mean(Hz)	797.128	1088.73	2706.68
STD	61.6945	84.6516	644.486
CV(%)	7.74%	7.78%	23.81%

<i>/u/</i>	F1	F2	F3
Mean(Hz)	437.858	791.114	2392.56
STD	52.6785	107.84	927.919
CV(%)	12.03%	13.63%	38.78%

<i>/i/</i>	F1	F2	F3
Mean(Hz)	430.461	2394.19	3250.71
STD	85.183	640.183	311.698
CV(%)	19.79%	26.74%	9.59%

Nhận xét đặc trưng phổ giữa các nguyên âm khác nhau của cùng một người nói 12FTD



Về đặc trưng phổ:

- Đối với nguyên âm /a/ thì năng lượng sẽ được tập trung nhiều ở tần số từ 0-2200 Hz, âm /e/ và âm /i/ thì ở tần số 0-1200 Hz, Ở các tầng số cao hơn năng lượng được trải đều.
- Đối với nguyên âm /o/ và /u/ thì năng lượng tập trung nhiều ở tần số 0-1800 Hz và ở các tần số cao hơn có rất ít năng lượng tập trung.

Nhận xét đặc trưng phổ giữa các nguyên âm khác nhau của cùng một người nói 12FTD

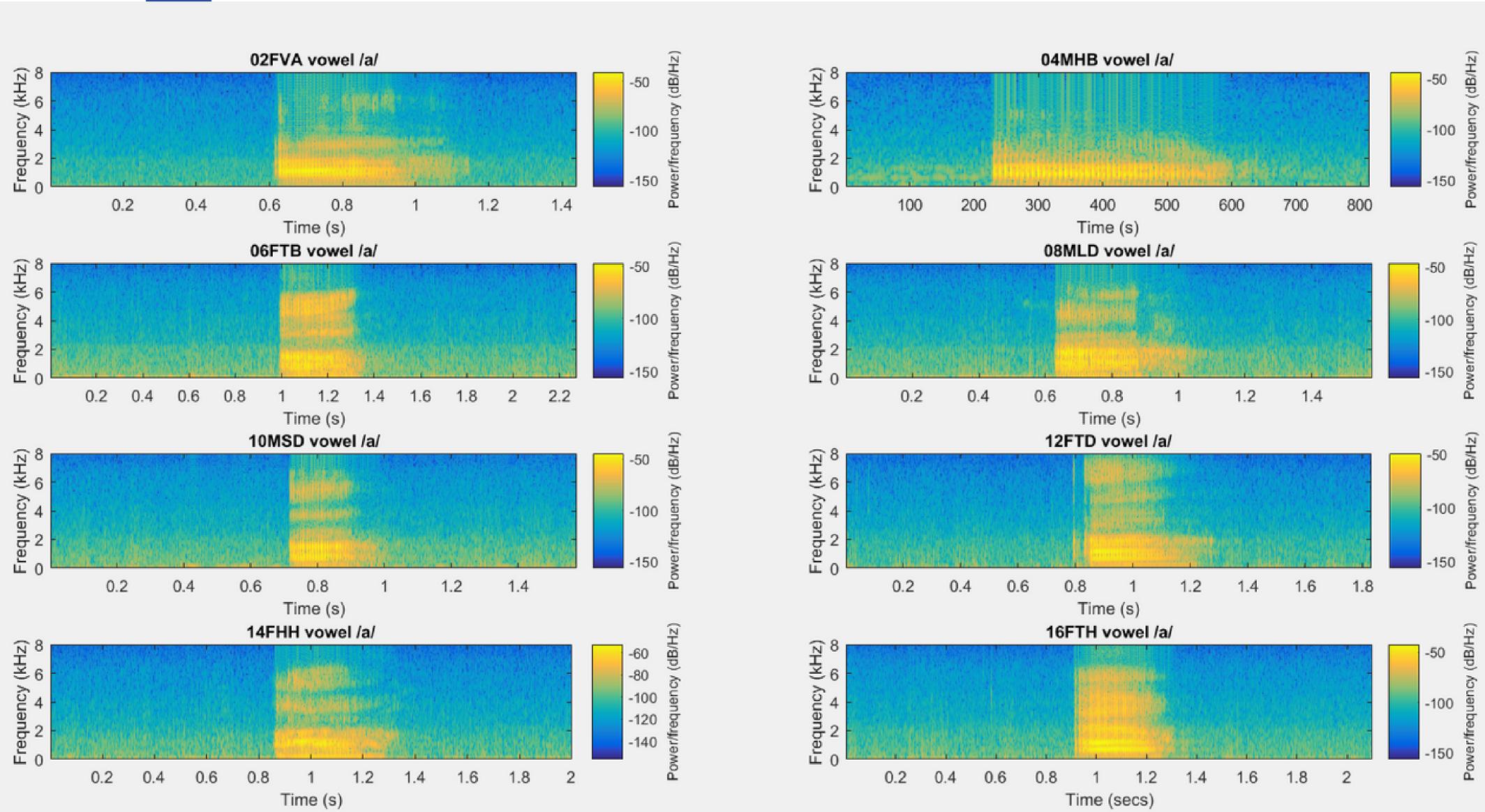
12FTD

12FTD		
F1	F2	F3
748.54	1210.03	1903.82
759.35	997.97	2150.83
478.69	3133.56	3498.65
763.49	1131.73	1531.79
503.06	843.12	1922.33

Về tần số formant:

- Nguyên âm /o/ có mean F1 cao nhất
- Nguyên âm /i/ có mean F2, F3 cao nhất .
- Độ rộng của vùng tần số formant đều tăng dần từ formant bậc thấp đến cao.

Nhận xét đặc trưng phổ giữa nguyên âm /a/ của 8 người nói khác nhau.

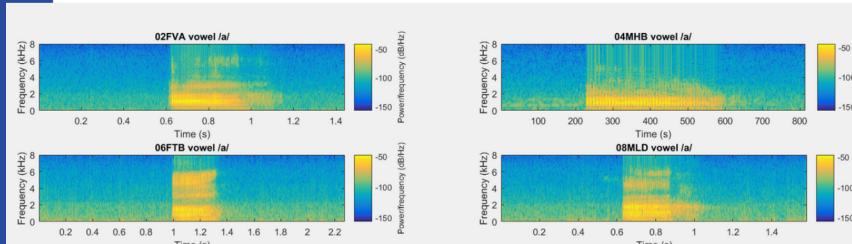


/a/	F1	F2	F3
Mean(Hz)	910.99	1453.5	2440.52
STD	194.953	268.493	605.822
CV(%)	21.40%	18.47%	24.82%

Về đặc trưng phổ:

- Nhìn chung, năng lượng của tín hiệu của cả 8 người đều tập trung nhiều ở tần số từ 0- 2000Hz và trải đều ở tần số cao hơn, riêng ở người nói 02FVA và 04MHB thì năng lượng mờ nhạt ít đi.
- Vì năng lượng tín hiệu nguyên âm /a/ giữa các người nói không có khác biệt quá lớn nên khó có thể chỉ dựa vào đặc trưng phổ của nguyên âm /a/ để nhận diện được người nói.

Nhận xét đặc trưng phổ giữa nguyên âm /a/ của 8 người nói khác nhau.



/a/	F1	F2	F3
Mean(Hz)	910.99	1453.5	2440.52
STD	194.953	268.493	605.822
CV(%)	21.40%	18.47%	24.82%

Về tần số formant

Nguyên âm /a/ có tần số formant F1 cao nhất trong các nguyên âm. Độ rộng vùng tần số formant của nguyên âm /a/ tăng dần từ formant bậc thấp đến bậc cao. Độ biến thiên tần số formant F1, F2, F3 khá thấp (< 25%) và khá ổn định. F1 dao động 738Hz - 1217Hz. F2 dao động 1130Hz - 1854Hz. F3 dao động 1736Hz - 3468Hz.

BÀI 2

NHẬN DẠNG NGUYÊN ÂM
KHÔNG PHỤ THUỘC NGƯỜI
NÓI DÙNG ĐẶC TRƯNG PHỔ
FFT.

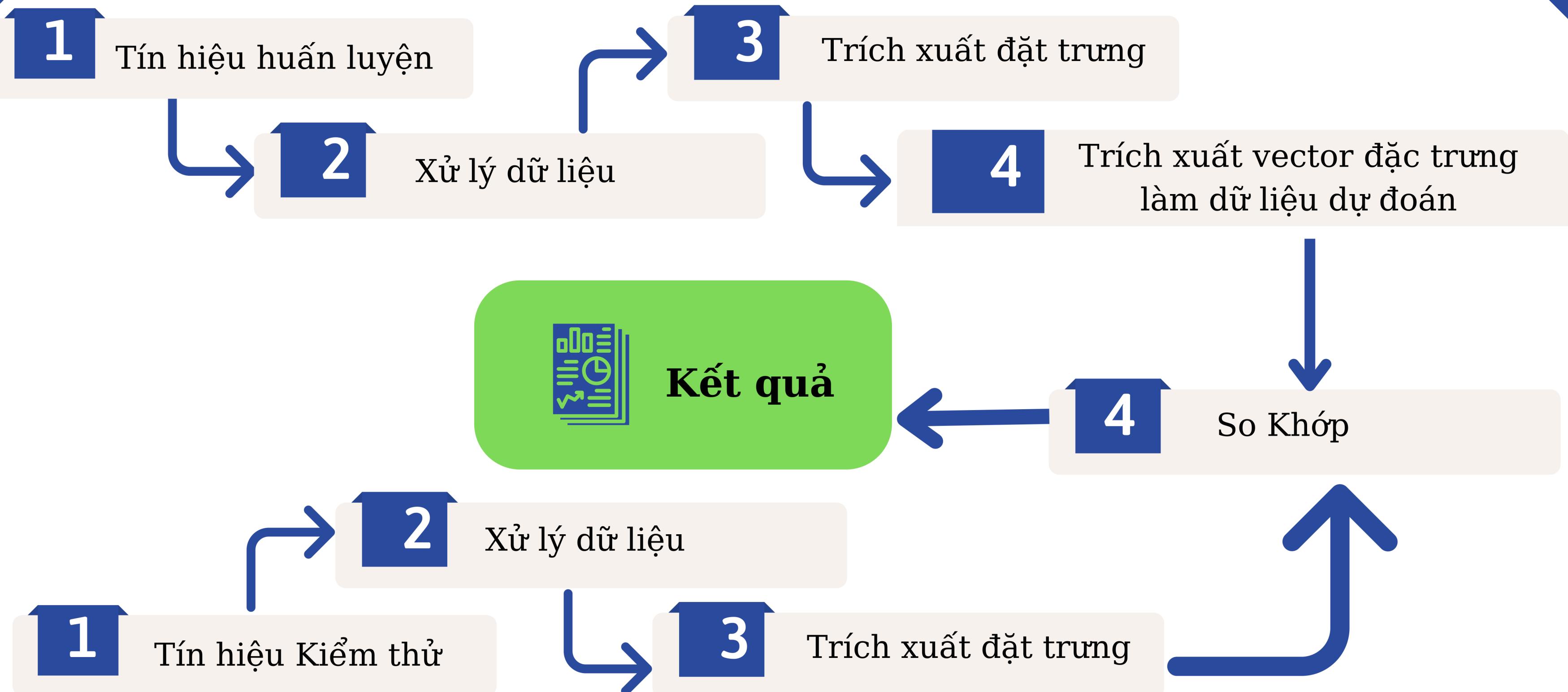
2

NHẬN DẠNG NGUYÊN ÂM KHÔNG PHỤ THUỘC NGƯỜI NÓI

Bảng phân công

STT	Nội dung công việc	Sinh viên thực hiện
1	Phân đoạn tín hiệu thành nguyên âm và khoảng lặng, chọn vùng ổn định	Nguyễn Đắc Đức, Phạm Công Huy
2	Trích xuất vector đặc trưng cho 1 nguyên âm (dựa vào mẫu dữ liệu huấn luyện)	Nguyễn Đắc Đức, Phạm Công Huy
3	Tính vector đặc trưng cho 1 nguyên âm của 1 người nói và của nhiều người nói	Nguyễn Đắc Đức, Phạm Công Huy
4	So khớp vector của tín hiệu kiểm thử với vector đặc trưng của nguyên âm.	Nguyễn Đắc Đức, Phạm Công Huy
5	Nhận xét chung và kết luận	Nguyễn Đắc Đức, Phạm Công Huy

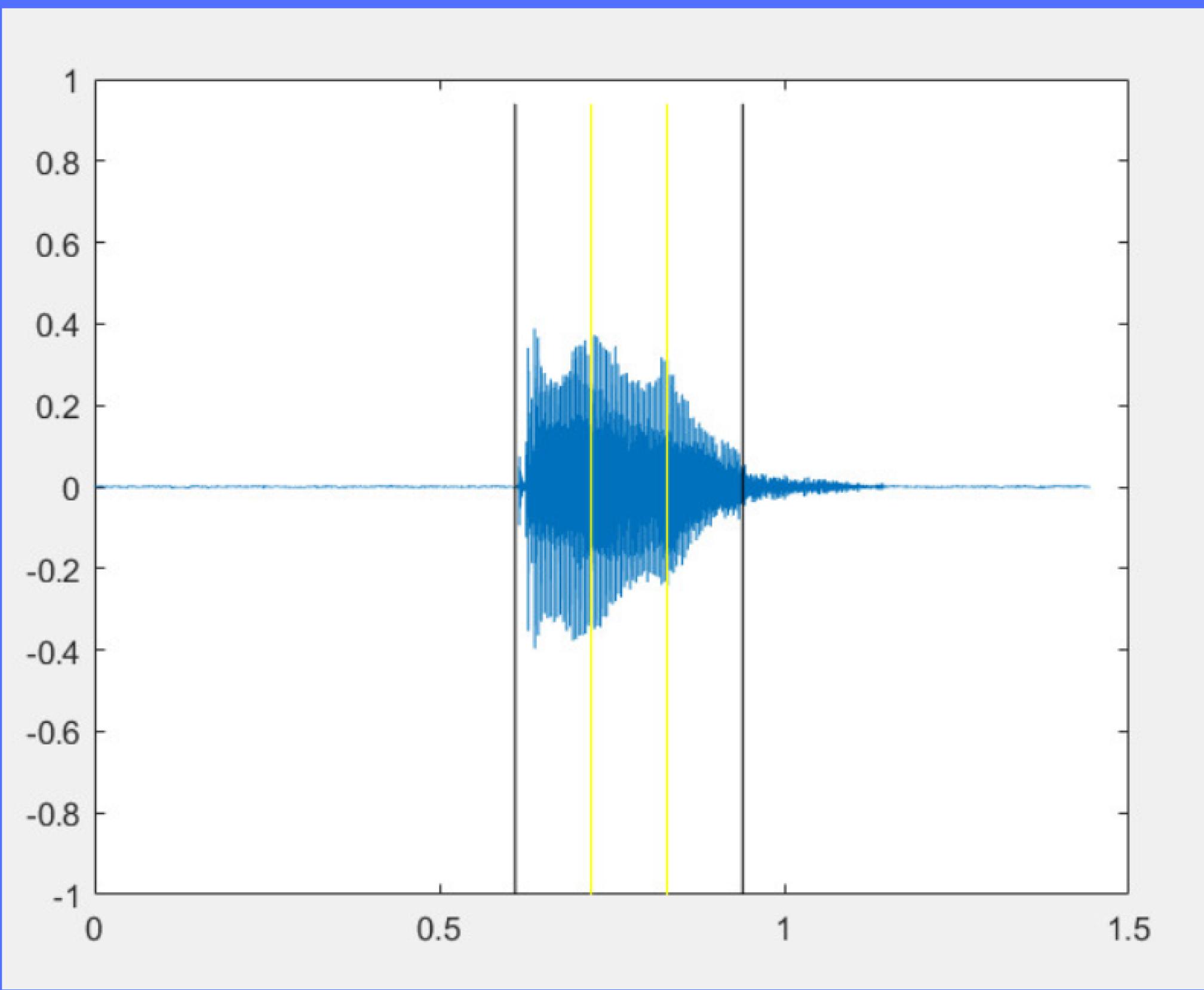
CÁC BƯỚC THỰC HIỆN



XỬ LÝ DỮ LIỆU

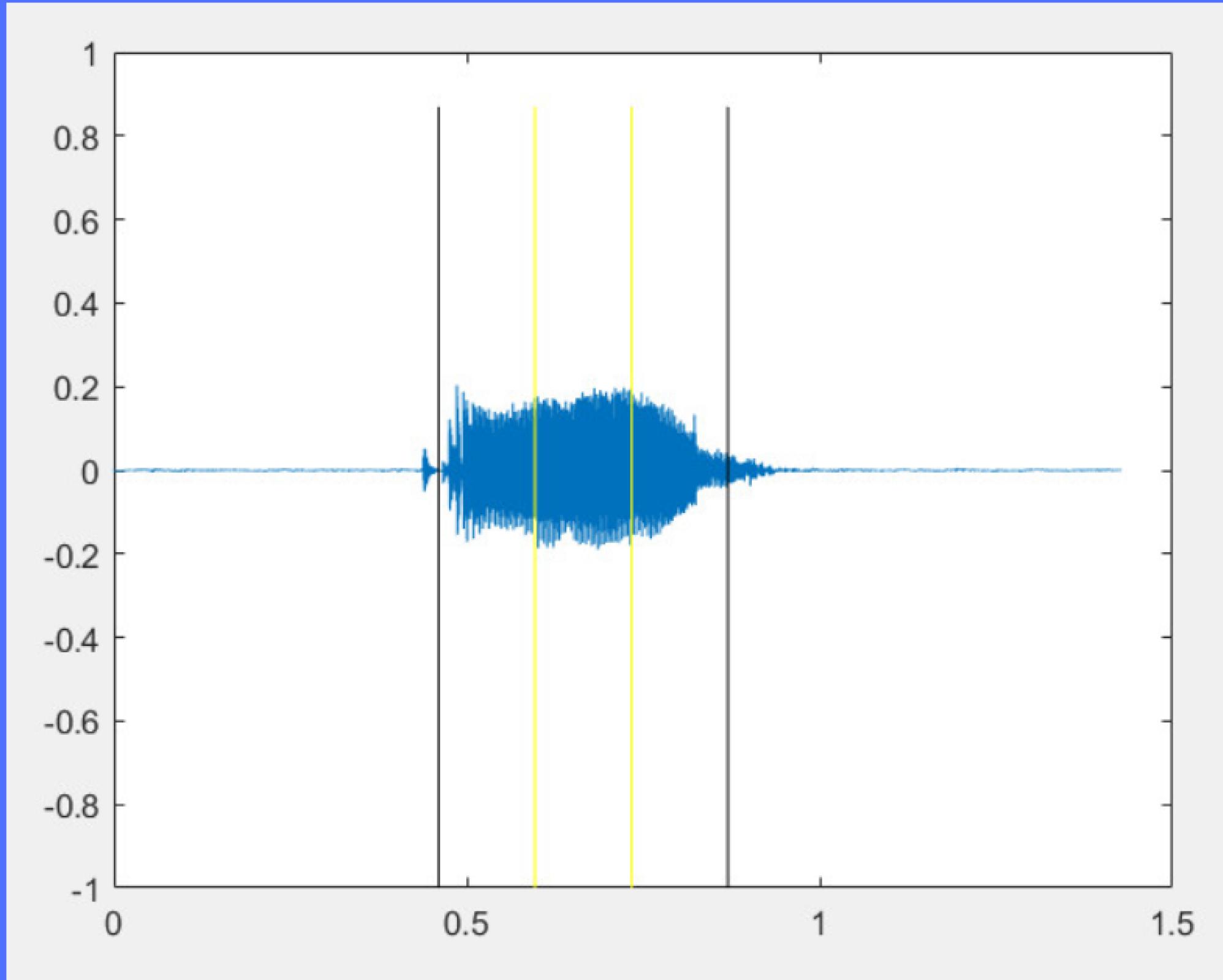


PHÂN ĐOẠN TÍN HIỆU VÀ LẤY PHẦN ỔN ĐỊNH CỦA NGUYÊN ÂM */a/*



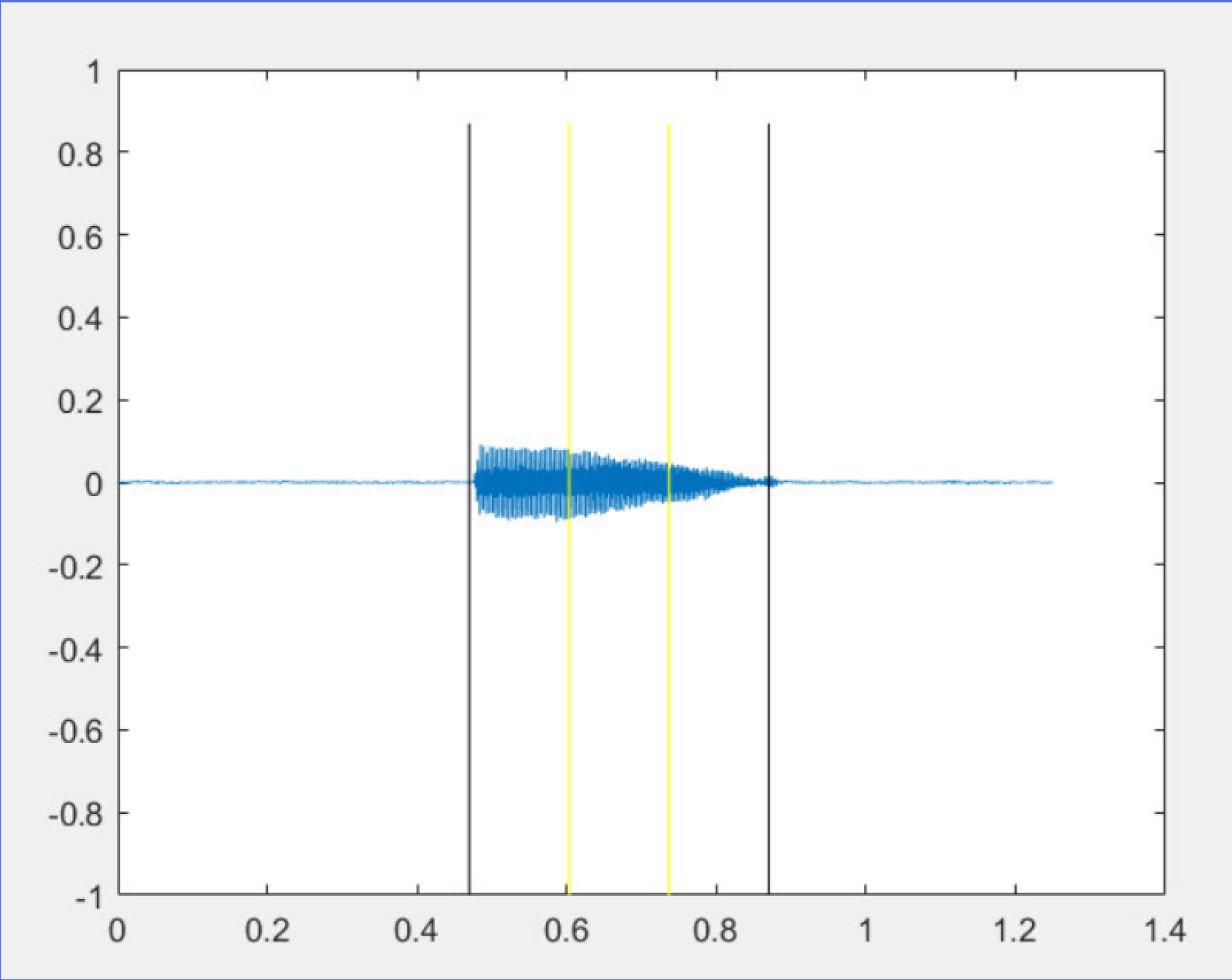
- BIÊN
- BIÊN ỔN ĐỊNH

PHÂN ĐOẠN TÍN HIỆU VÀ LẤY PHẦN ỔN ĐỊNH CỦA NGUYÊN ÂM */e/*



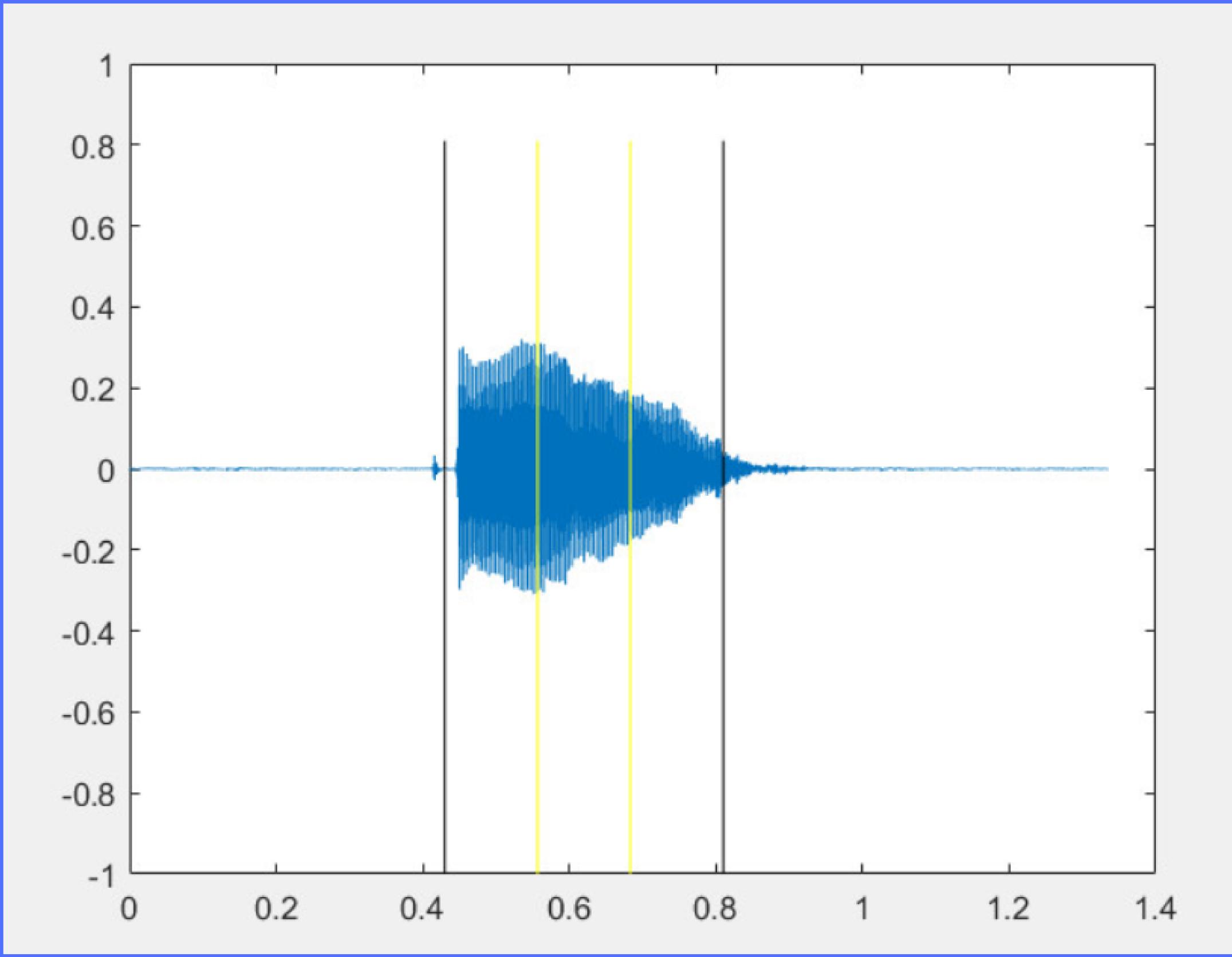
- BIÊN
- BIÊN ỔN ĐỊNH

PHÂN ĐOẠN TÍN HIỆU VÀ LẤY PHẦN ỔN ĐỊNH CỦA NGUYÊN ÂM */i/*



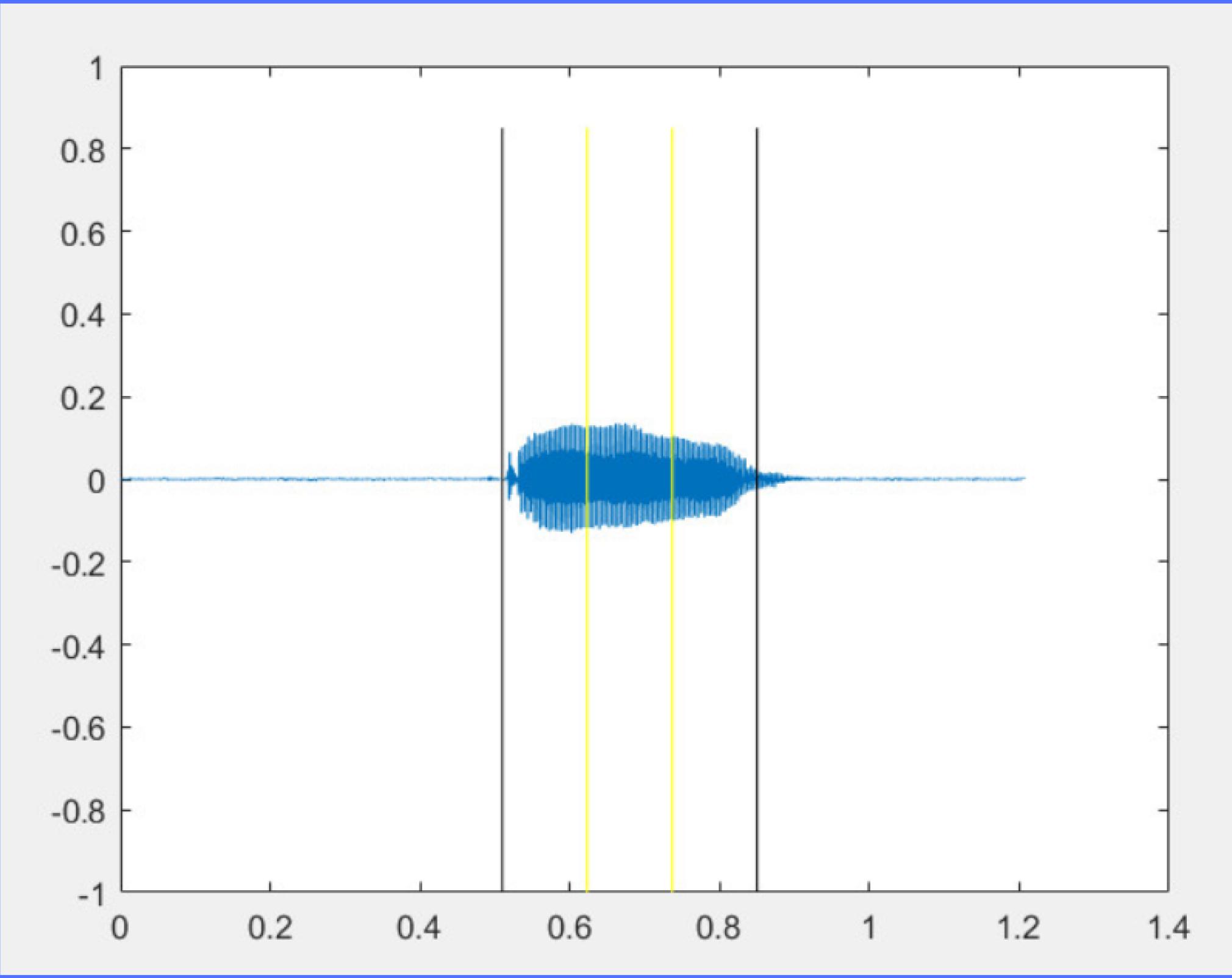
- BIÊN
- BIÊN ỔN ĐỊNH

PHÂN ĐOẠN TÍN HIỆU VÀ LẤY PHẦN ÔN ĐỊNH CỦA NGUYÊN ÂM */o/*



- BIÊN
- BIÊN ỔN ĐỊNH

PHÂN ĐOẠN TÍN HIỆU VÀ LẤY PHẦN ÔN ĐỊNH CỦA NGUYÊN ÂM */u/*

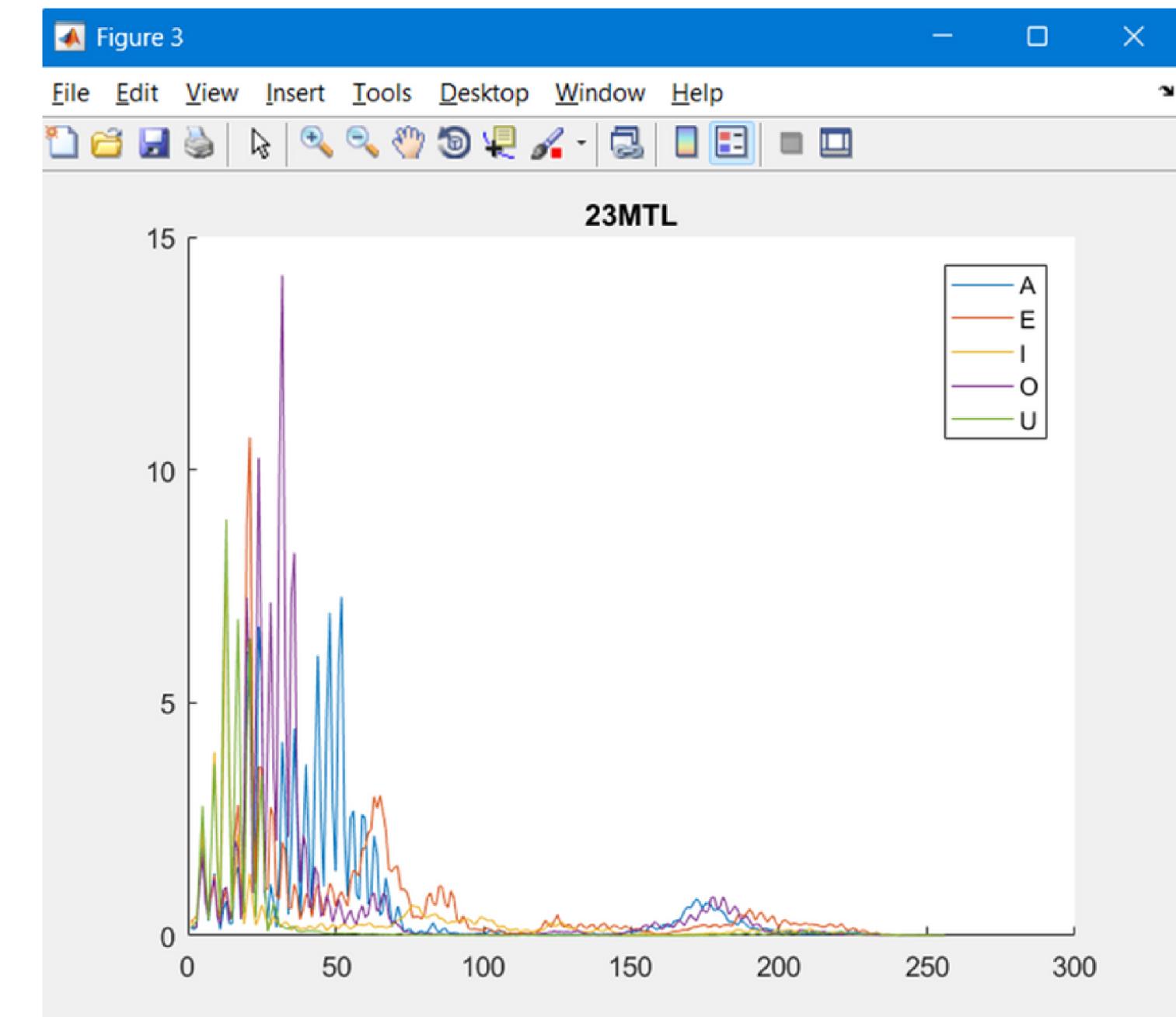
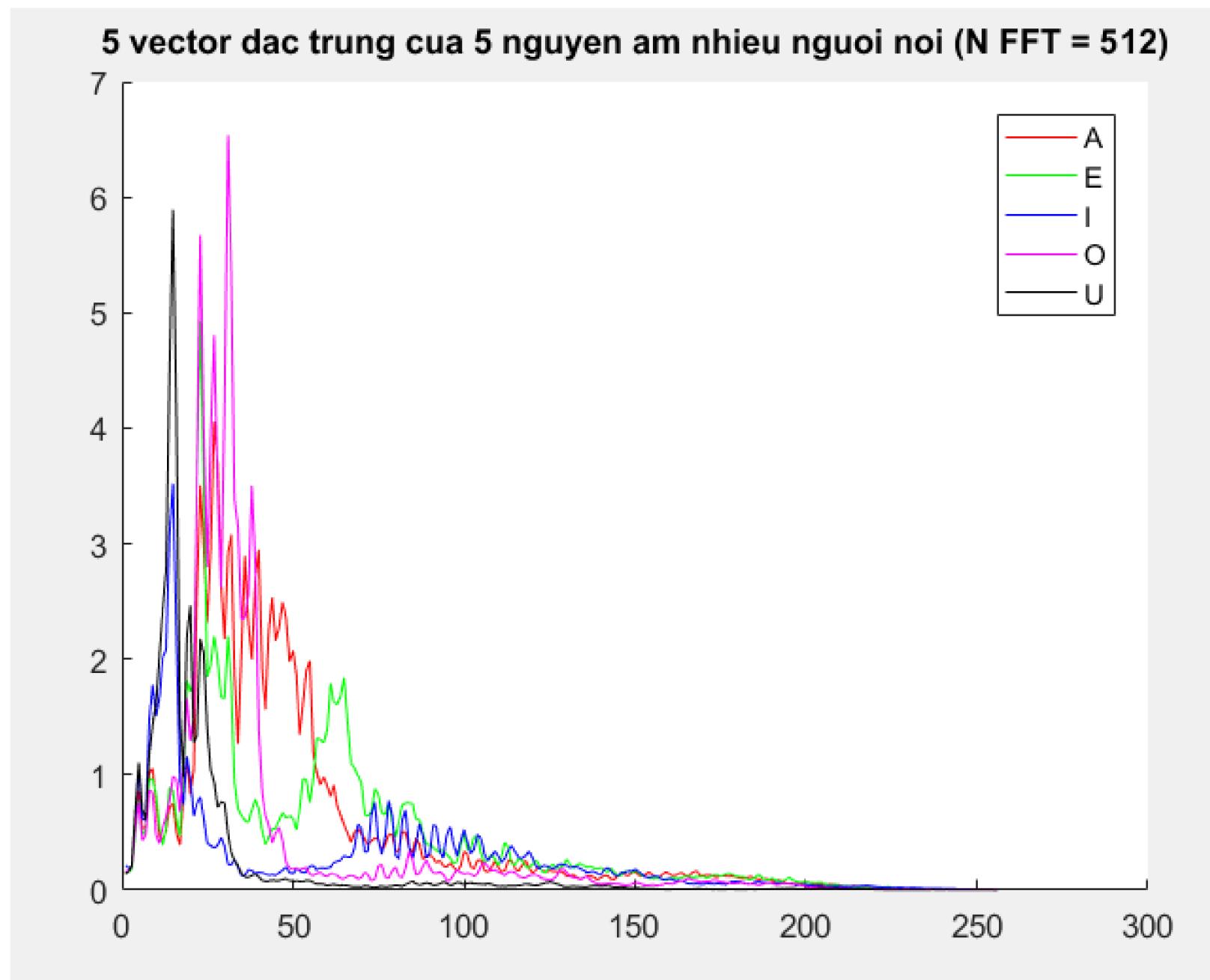


- BIÊN
- BIÊN ÔN ĐỊNH

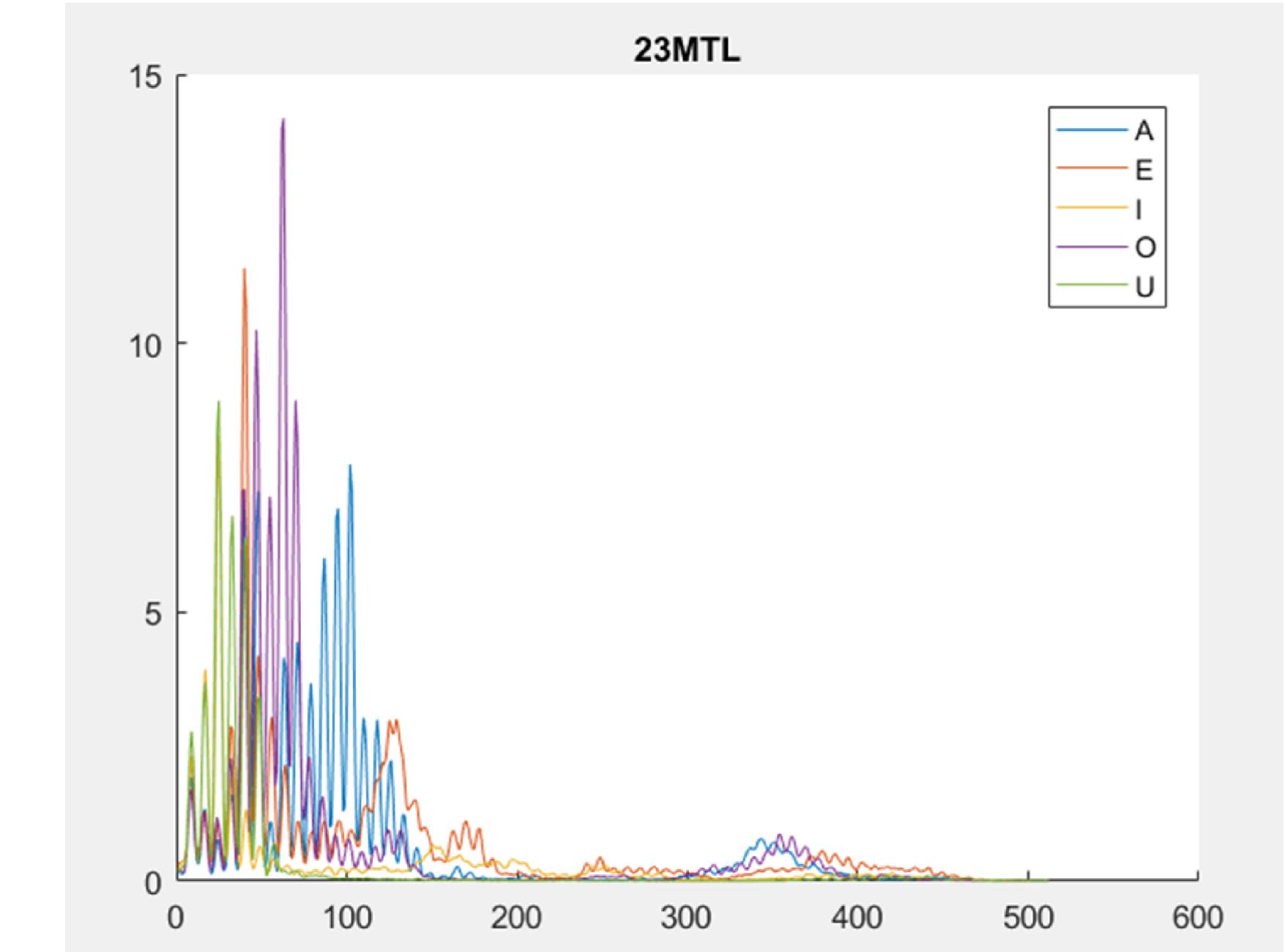
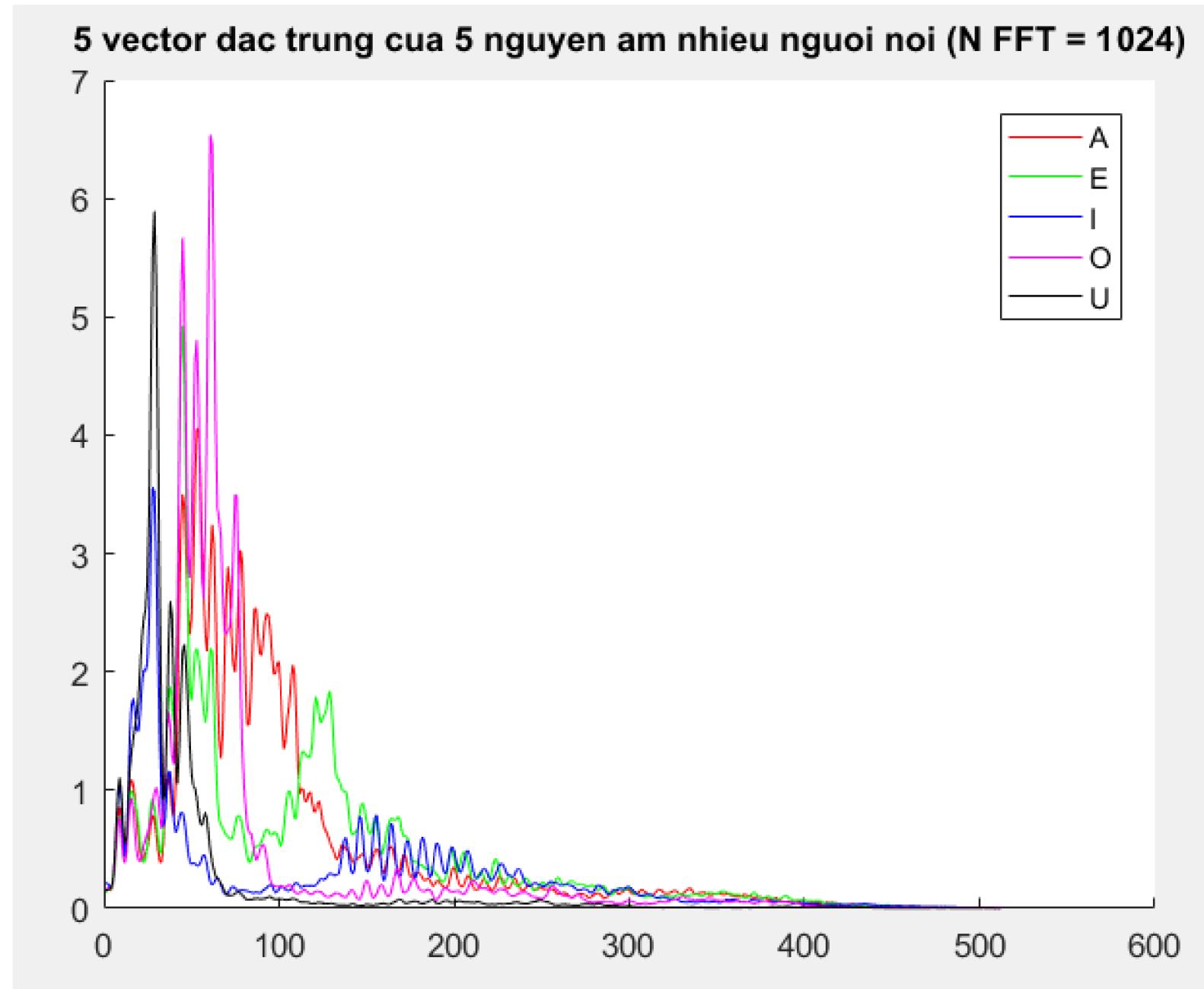
PHÂN ĐOẠN TÍN HIỆU VÀ LẤY PHẦN ỔN ĐỊNH CỦA NGUYÊN ÂM

Nhận xét: Các file huấn luyện đầu vào đều xử lý phân đoạn khá chính xác và tìm được phần ổn định của nguyên âm

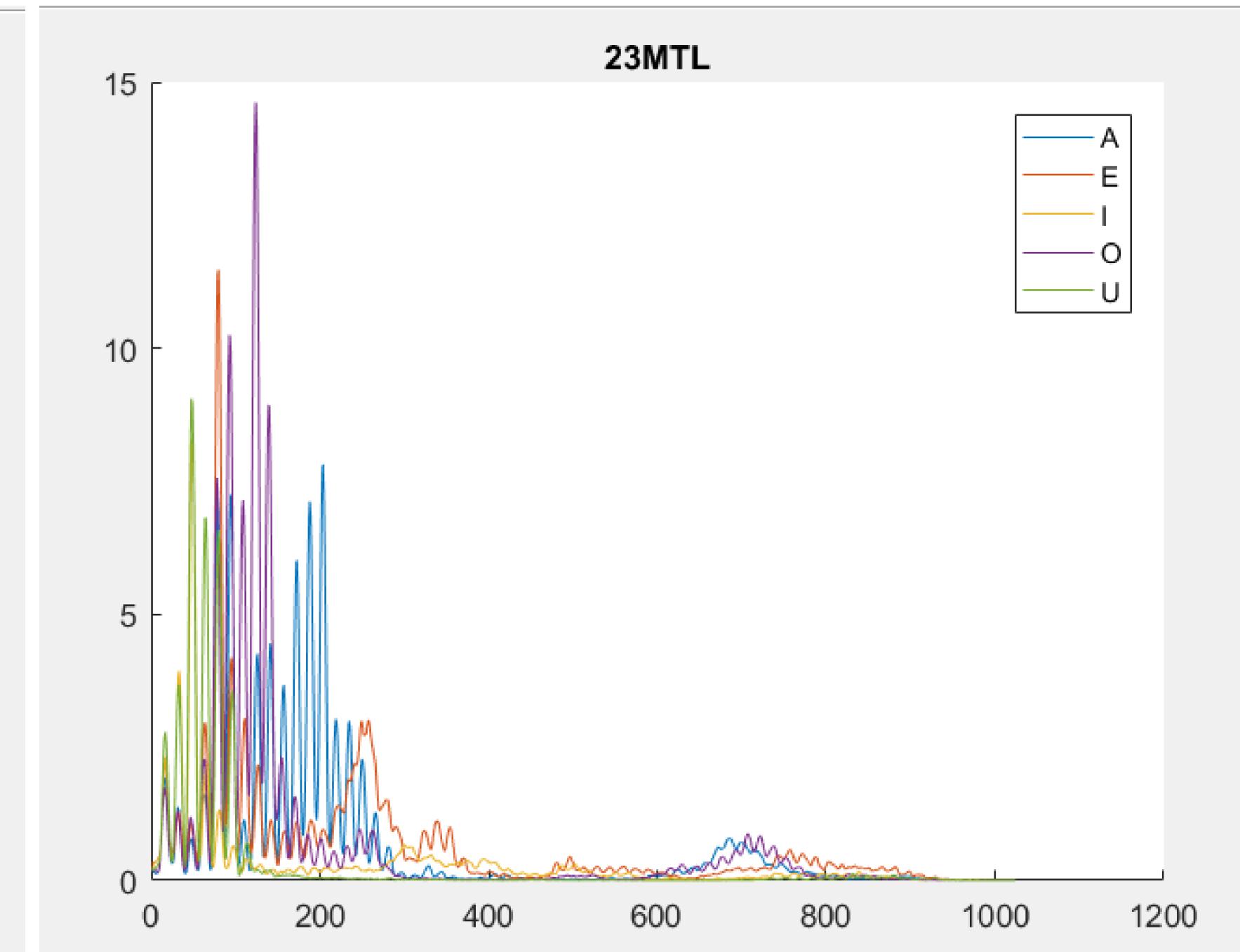
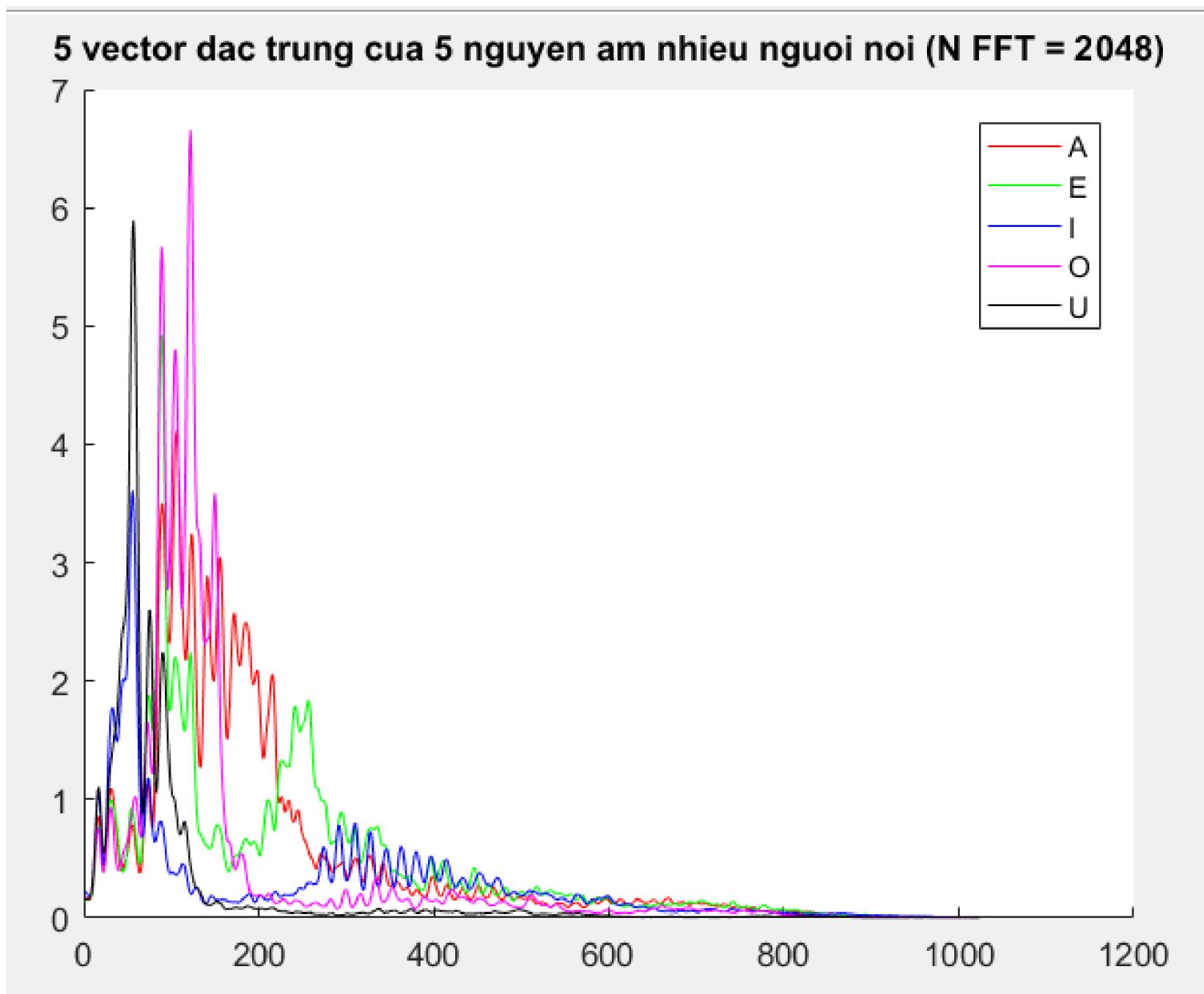
ĐỒ THỊ VECTOR ĐẶC TRƯNG FFT CỦA CÁC NGUYÊN ÂM



ĐỒ THỊ VECTOR ĐẶC TRƯNG FFT CỦA CÁC NGUYÊN ÂM



ĐỒ THỊ VECTOR ĐẶC TRƯNG FFT CỦA CÁC NGUYÊN ÂM



ĐỒ THỊ VECTOR ĐẶC TRƯNG FFT CỦA CÁC NGUYÊN ÂM

Nhận xét: Đặc trưng của nguyên âm tập trung ở vùng có tần số thấp, khi thay đổi số chiều N_{FFT} thì hình dạng đồ thị hầu như không thay đổi.

KẾT QUẢ NHẬN DẠNG NGUYÊN ÂM

File 20MVN: Muc tieu: i - Du doan: i
File 29MHN: Muc tieu: i - Du doan: i
File 30FTN: Muc tieu: i - Du doan: i
File 32MTP: Muc tieu: i - Du doan: i
File 33MHP: Muc tieu: i - Du doan: i
File 34MQP: Muc tieu: i - Du doan: i
File 35MMQ: Muc tieu: i - Du doan: i
File 36MAQ: Muc tieu: i - Du doan: i
File 37MDS: Muc tieu: i - Du doan: i
File 38MDS: Muc tieu: i - Du doan: i
File 39MTS: Muc tieu: i - Du doan: i
File 40MHS: Muc tieu: i - Du doan: i
File 41MVS: Muc tieu: i - Du doan: i
File 42FQT: Muc tieu: i - Du doan: i
File 43MNT: Muc tieu: i - Du doan: i
File 44MTT: Muc tieu: i - Du doan: i
File 45MDV: Muc tieu: i - Du doan: i
File 23MTL: Muc tieu: o - Du doan: o
File 24FTL: Muc tieu: o - Du doan: i
File 25MLM: Muc tieu: o - Du doan: o
File 27MCM: Muc tieu: o - Du doan: o

File 25MLM: Muc tieu: u - Du doan: u
File 27MCM: Muc tieu: u - Du doan: u
File 28MVN: Muc tieu: u - Du doan: i
File 29MHN: Muc tieu: u - Du doan: u
File 30FTN: Muc tieu: u - Du doan: i
File 32MTP: Muc tieu: u - Du doan: i
File 33MHP: Muc tieu: u - Du doan: u
File 34MQP: Muc tieu: u - Du doan: i
File 35MMQ: Muc tieu: u - Du doan: u
File 36MAQ: Muc tieu: u - Du doan: i
File 37MDS: Muc tieu: u - Du doan: u
File 38MDS: Muc tieu: u - Du doan: u
File 39MTS: Muc tieu: u - Du doan: u
File 40MHS: Muc tieu: u - Du doan: u
File 41MVS: Muc tieu: u - Du doan: i
File 42FQT: Muc tieu: u - Du doan: i
File 43MNT: Muc tieu: u - Du doan: i
File 44MTT: Muc tieu: u - Du doan: u
File 45MDV: Muc tieu: u - Du doan: u



MA TRẬN NHẦM LÃN (FFT)

	/a/	/e/	/i/	/o/	/u/
/a/	17	0	3	1	0
/e/	0	16	4	1	0
/i/	0	0	21	0	0
/o/	2	3	2	14	0
/u/	0	0	9	0	12
Tỉ lệ chính xác	76.19%				

NHẬN XÉT VÀ KẾT LUẬN

- Kết quả nhận dạng đúng ở mức tương đối cao
- Độ chính xác trung bình cho 105 file tín hiệu kiểm thử khoảng $>76\%$
- Độ chính xác không thay đổi với số chiều FFT

BÀI 3

NHẬN DẠNG NGUYÊN ÂM
KHÔNG PHỤ THUỘC NGƯỜI
NÓI DÙNG ĐẶC TRƯNG PHỔ
MFCC.
(ĐANG TRONG QUÁ TRÌNH
THỰC HIỆN)

2

THANKS !

Cảm ơn thầy và các bạn đã
lắng nghe

