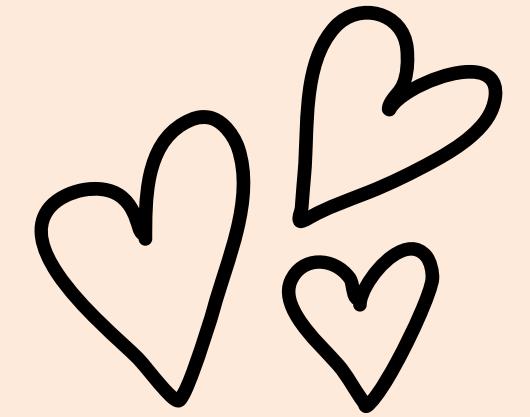




KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO BÀI TẬP XỬ LÝ TÍN HIỆU SỐ

Giáo viên HD : Ninh Khánh Duy



MY TEAM



Phạm Trung Nguyên



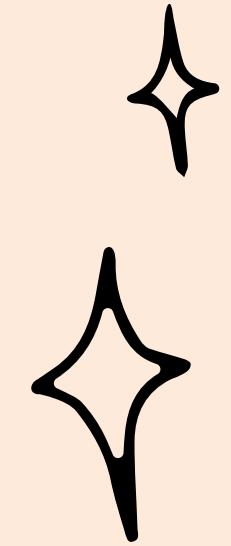
Hoàng Đức Vinh



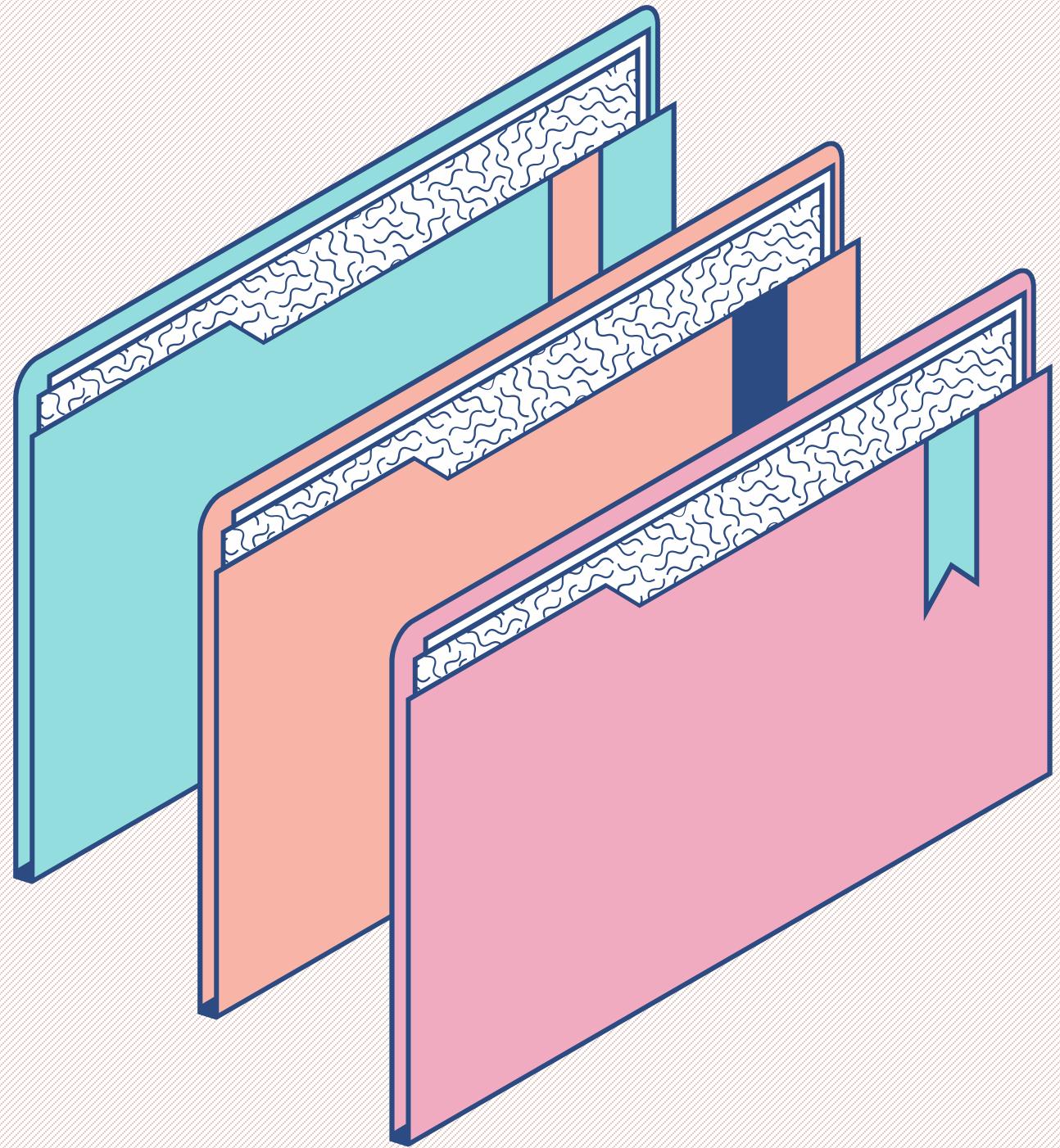
Nguyễn Hoàng Mỹ



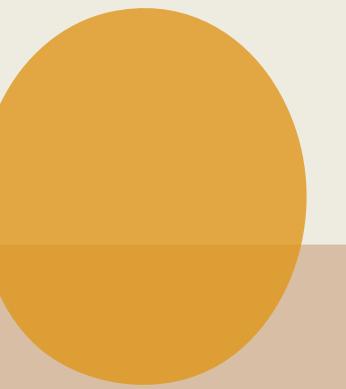
Nguyễn Đình Thúy Vy



Mục tiêu



- Mục Tiêu 1 : Phân tích đặc trưng phổ của các nguyên âm nhiều người nói đưa ra nhận xét.
- Mục Tiêu 2 : Nhận dạng nguyên âm không phụ thuộc người nói dùng đặc trưng phổ FFT.
- Mục Tiêu 3 : Nhận dạng nguyên âm không phụ thuộc người nói dùng đặc trưng phổ MFCC.



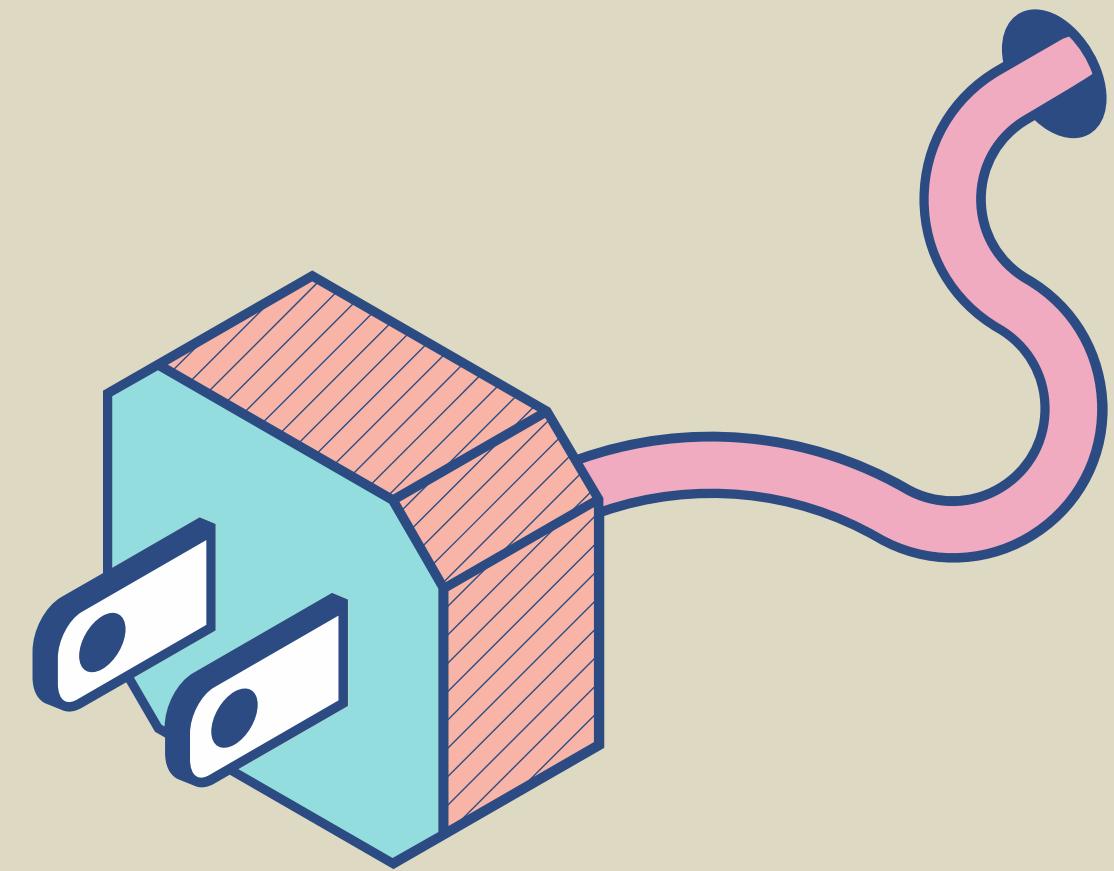
Phân tích đặc trưng phổ của các nguyên âm nhiều người nói



Phân tích đặc trưng phổ của các nguyên âm nhiều người nói



OUTLINE



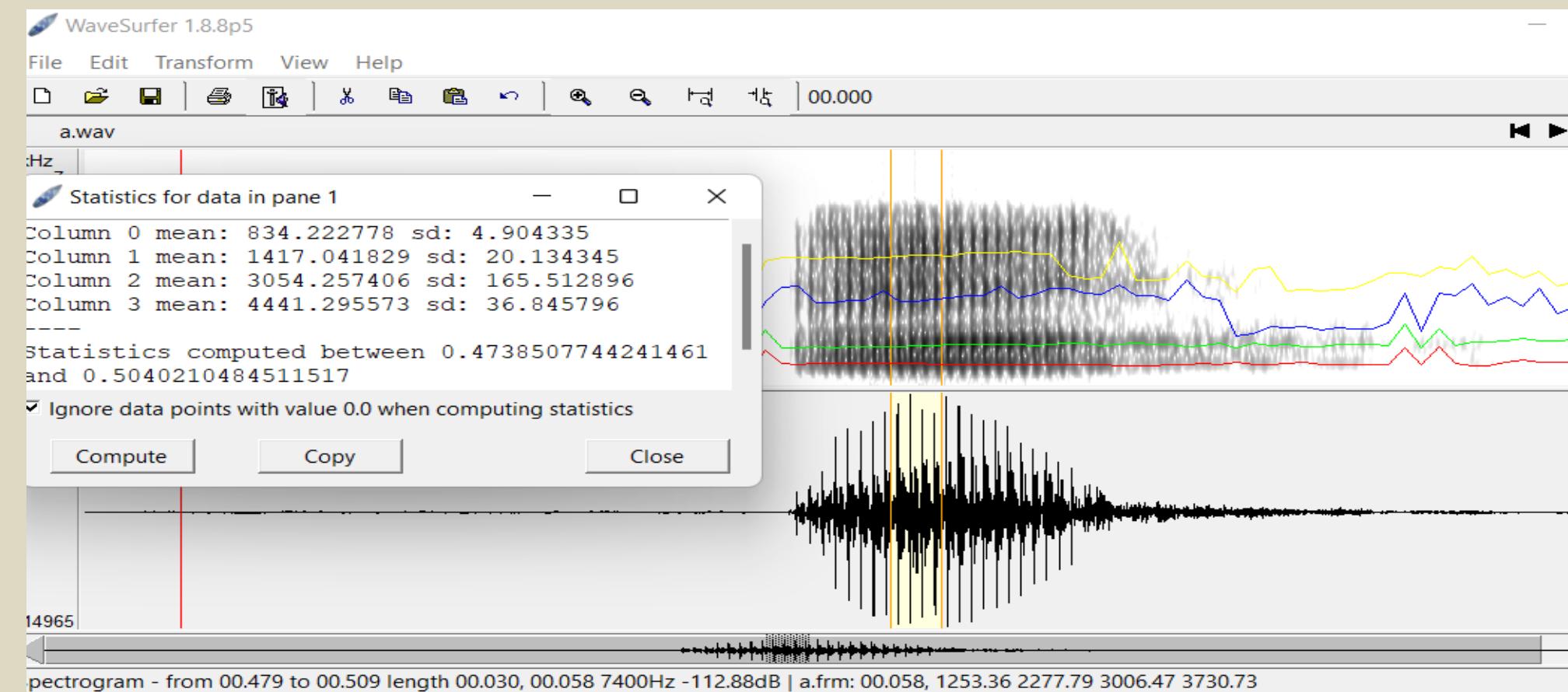
Các bước thực hiện:

Tín hiệu các nguyên âm

Dùng hàm spectrogram() vẽ ảnh phổ

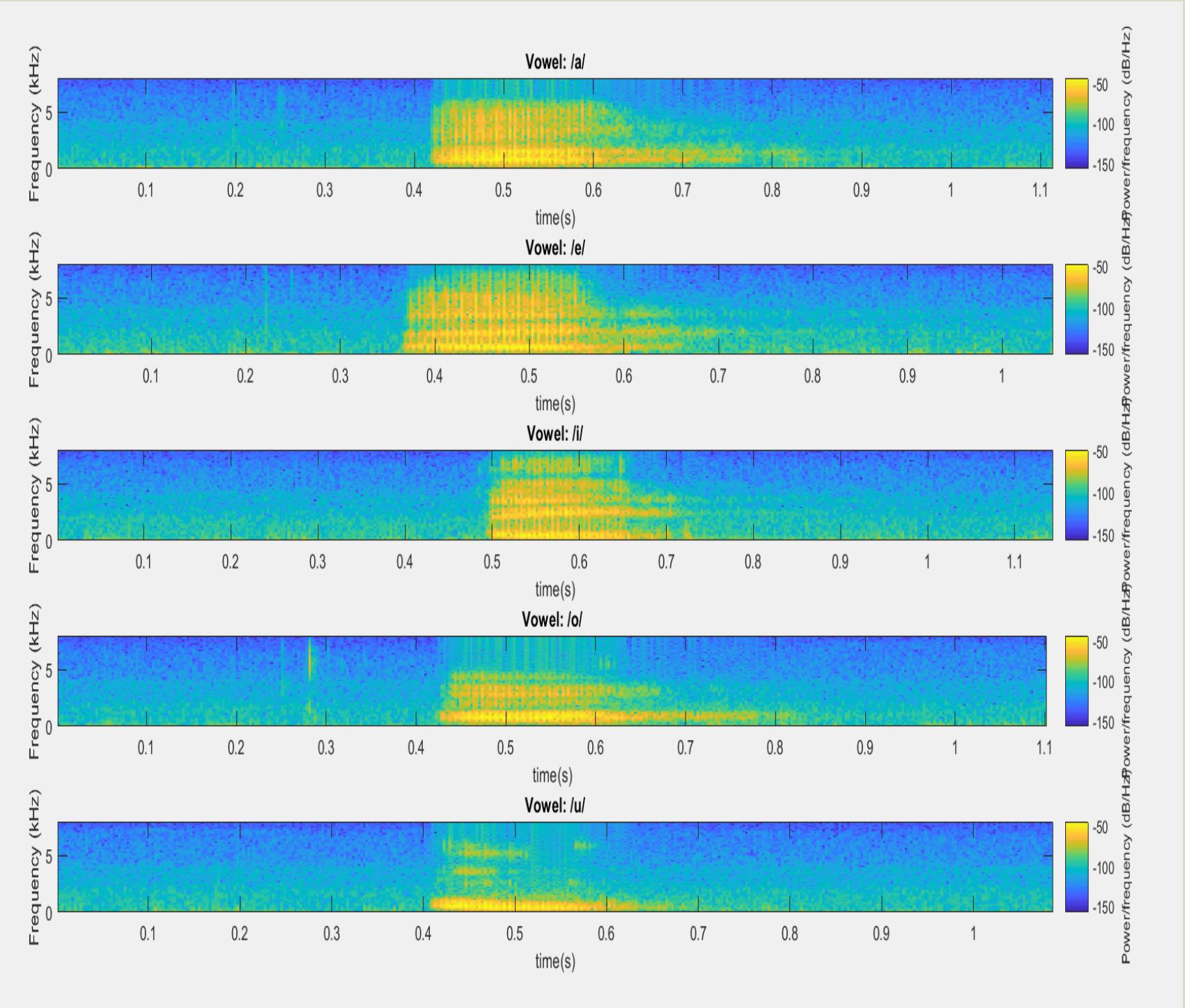
Dùng Wavesurfer đo 3 lần tần số Formant trên 3 khung khác nhau có Khoảng cách là 30ms

Tính trung bình các giá trị F_1 , F_2 , F_3 trên để thu được giá trị trung bình của F_1 , F_2 , F_3

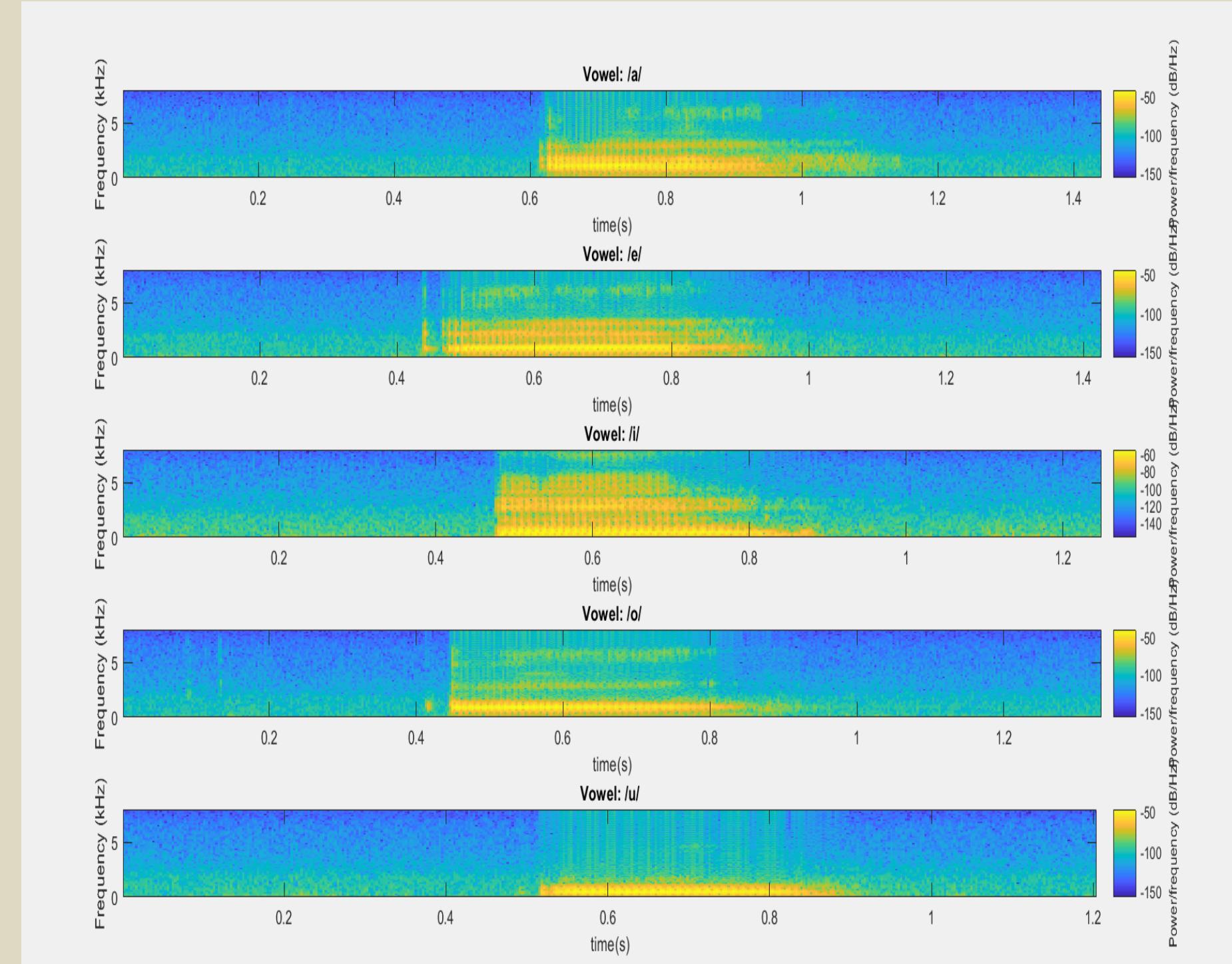


Dựa vào ảnh phổ và bảng số liệu bộ 3 Formant để đưa ra các nhận xét

Ảnh phổ băng rộng:

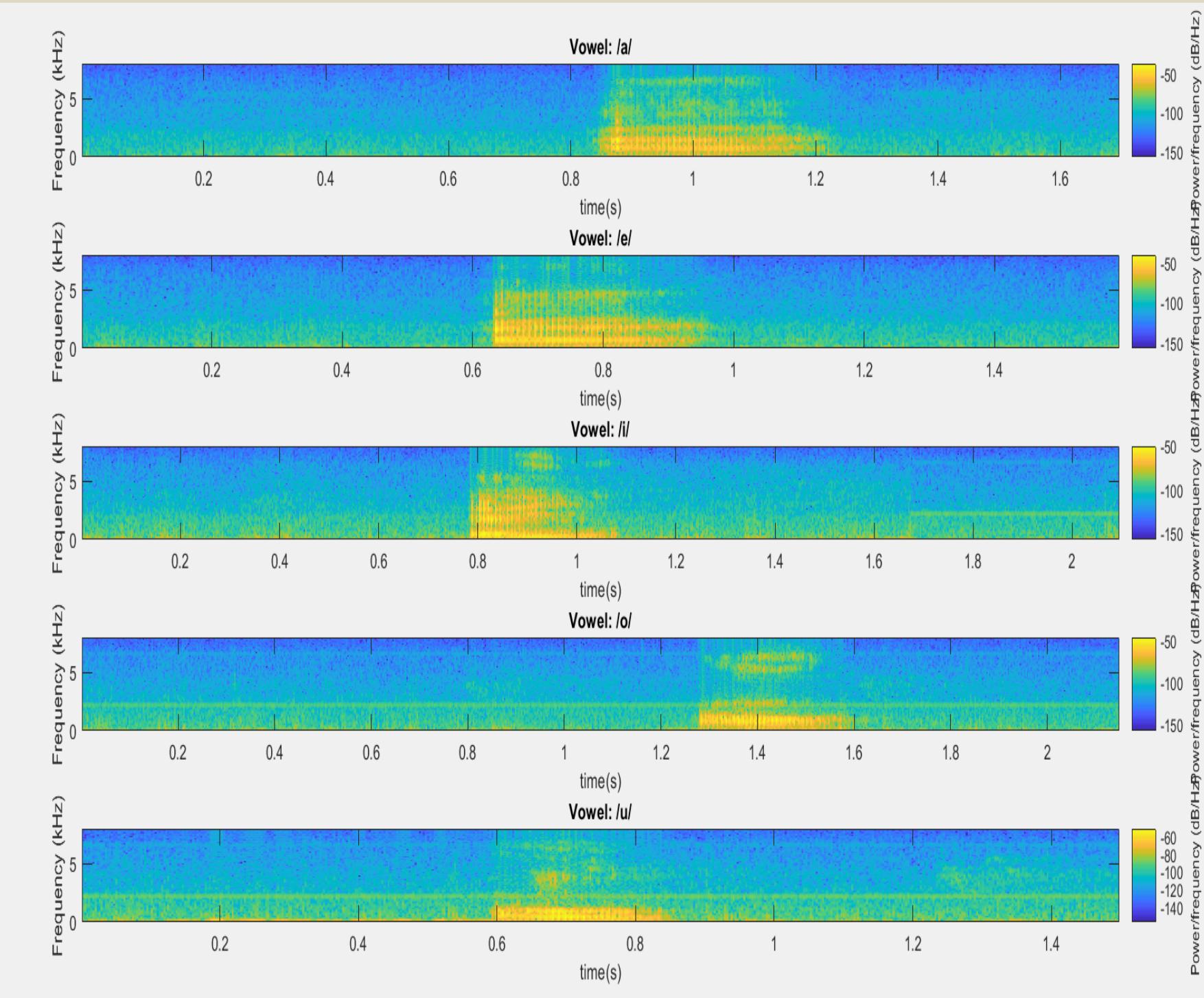


01MDA

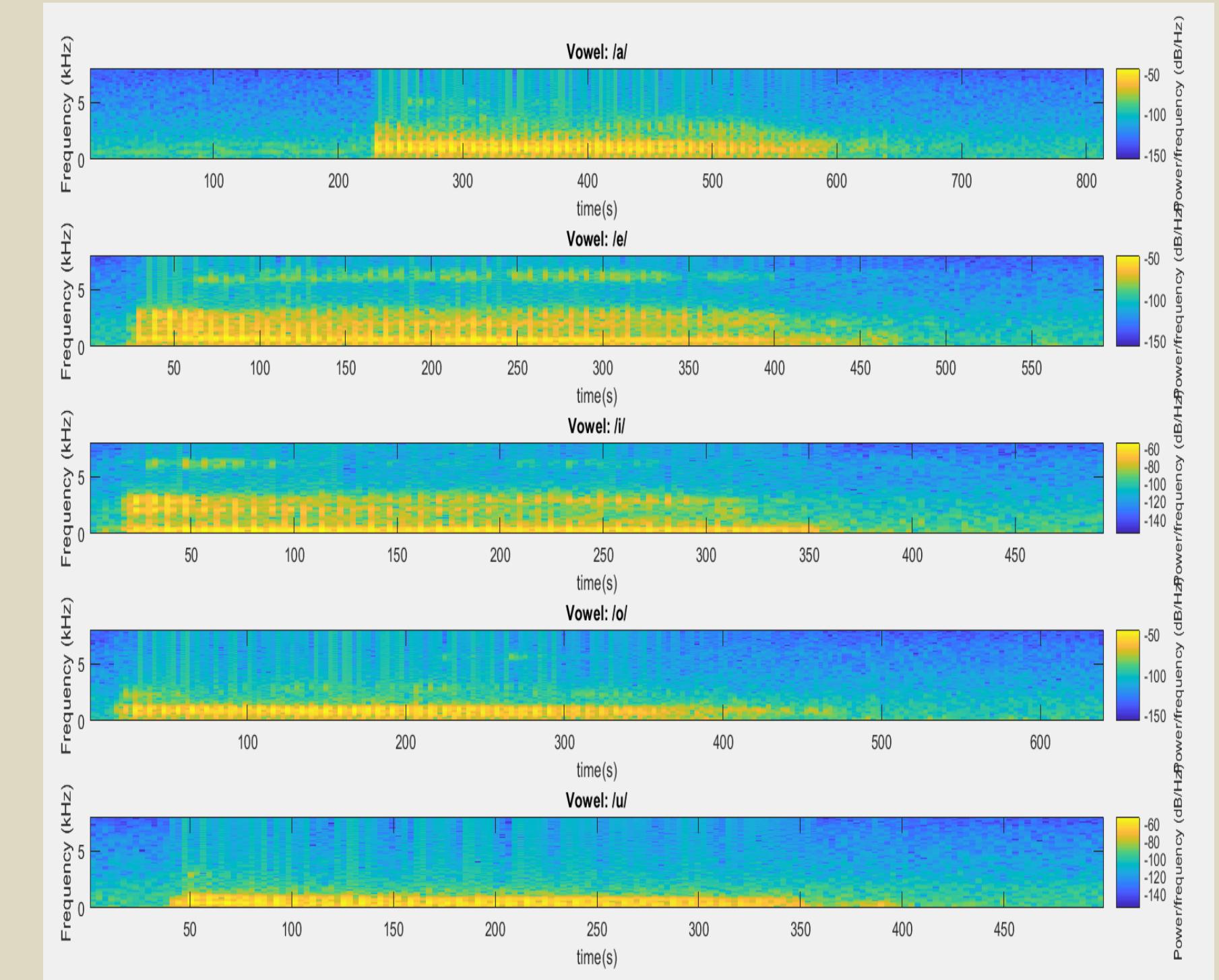


02FVA

Ảnh phổ băng rộng:

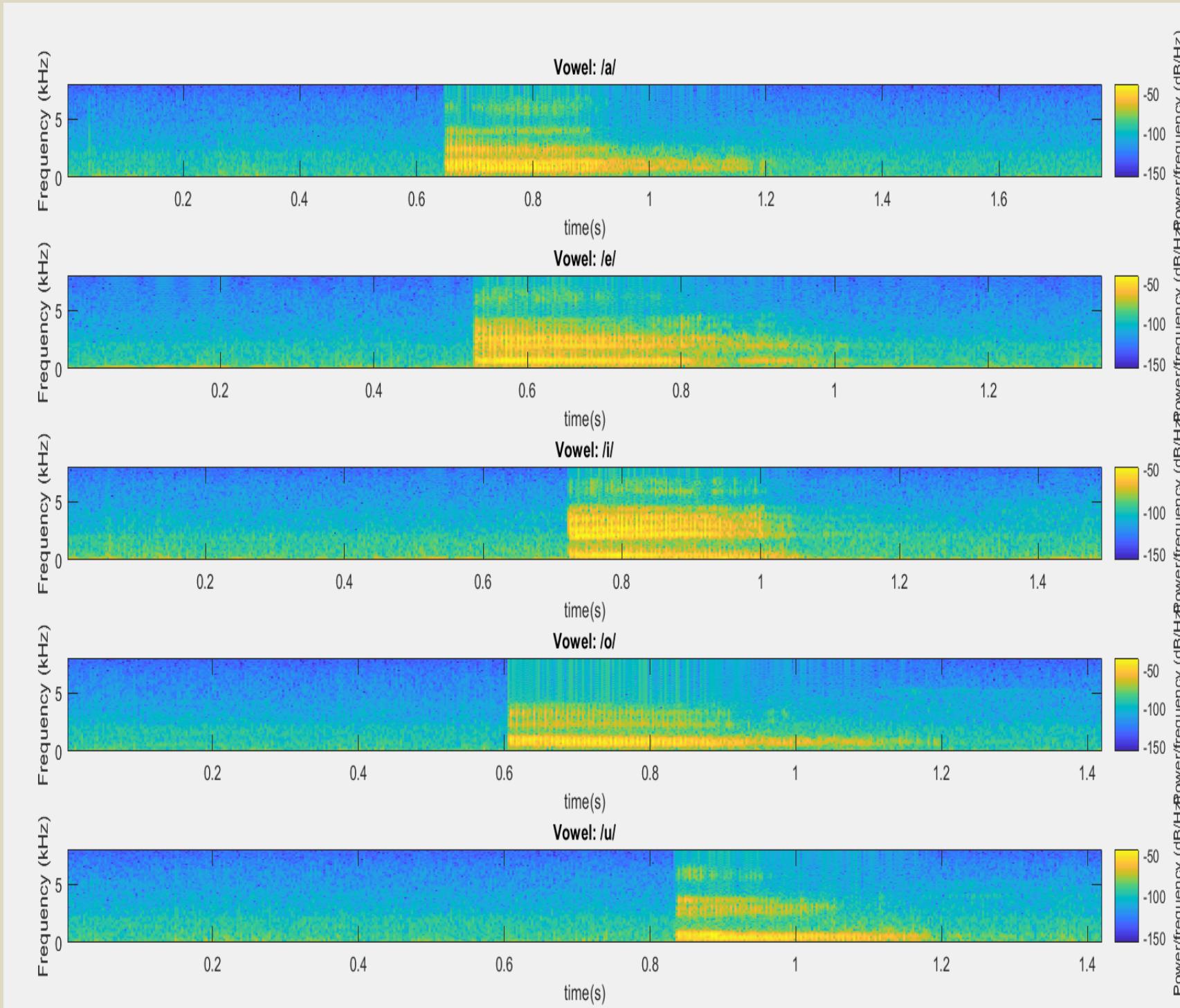


03MAB

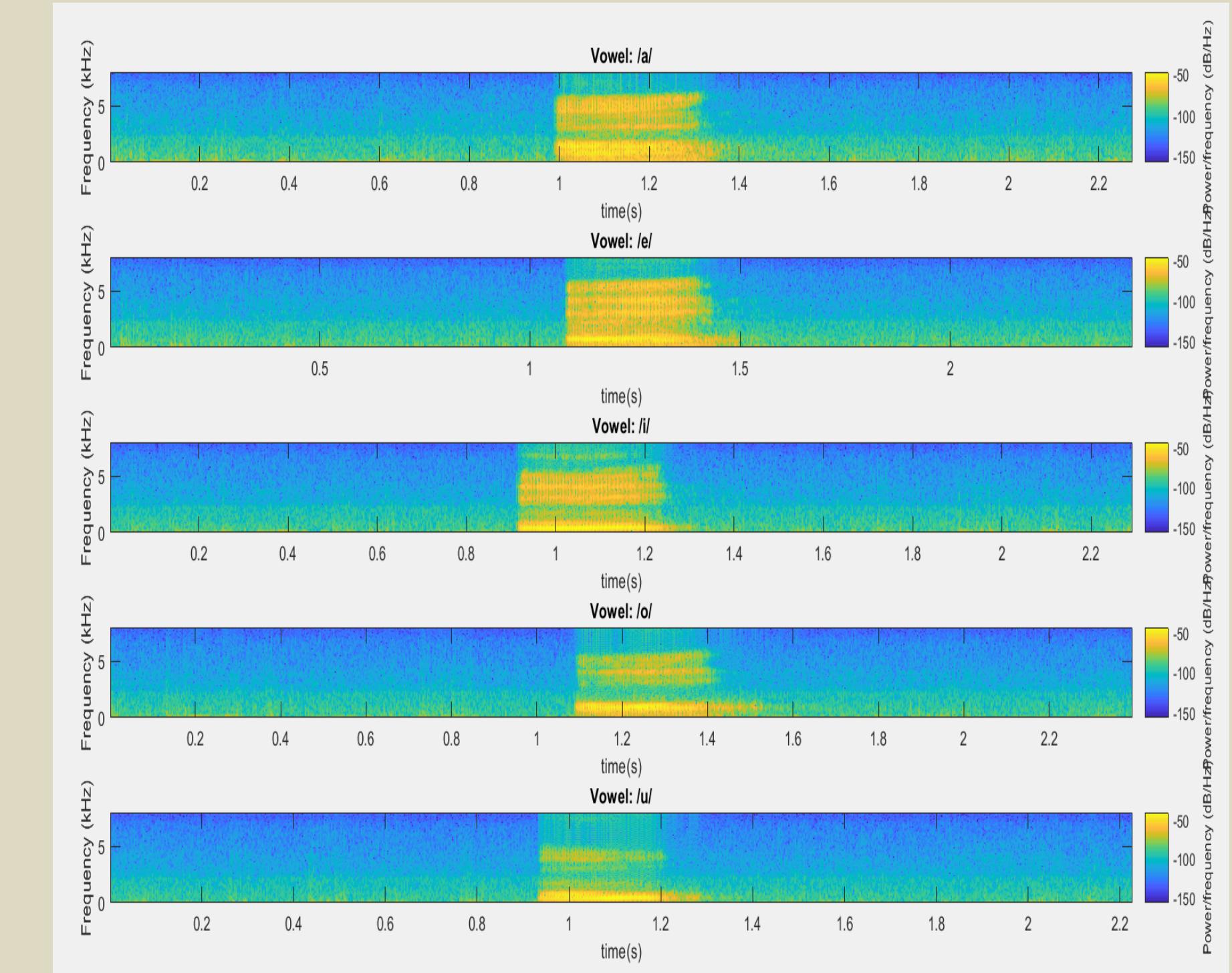


04MHB

Ảnh phổ băng rộng:

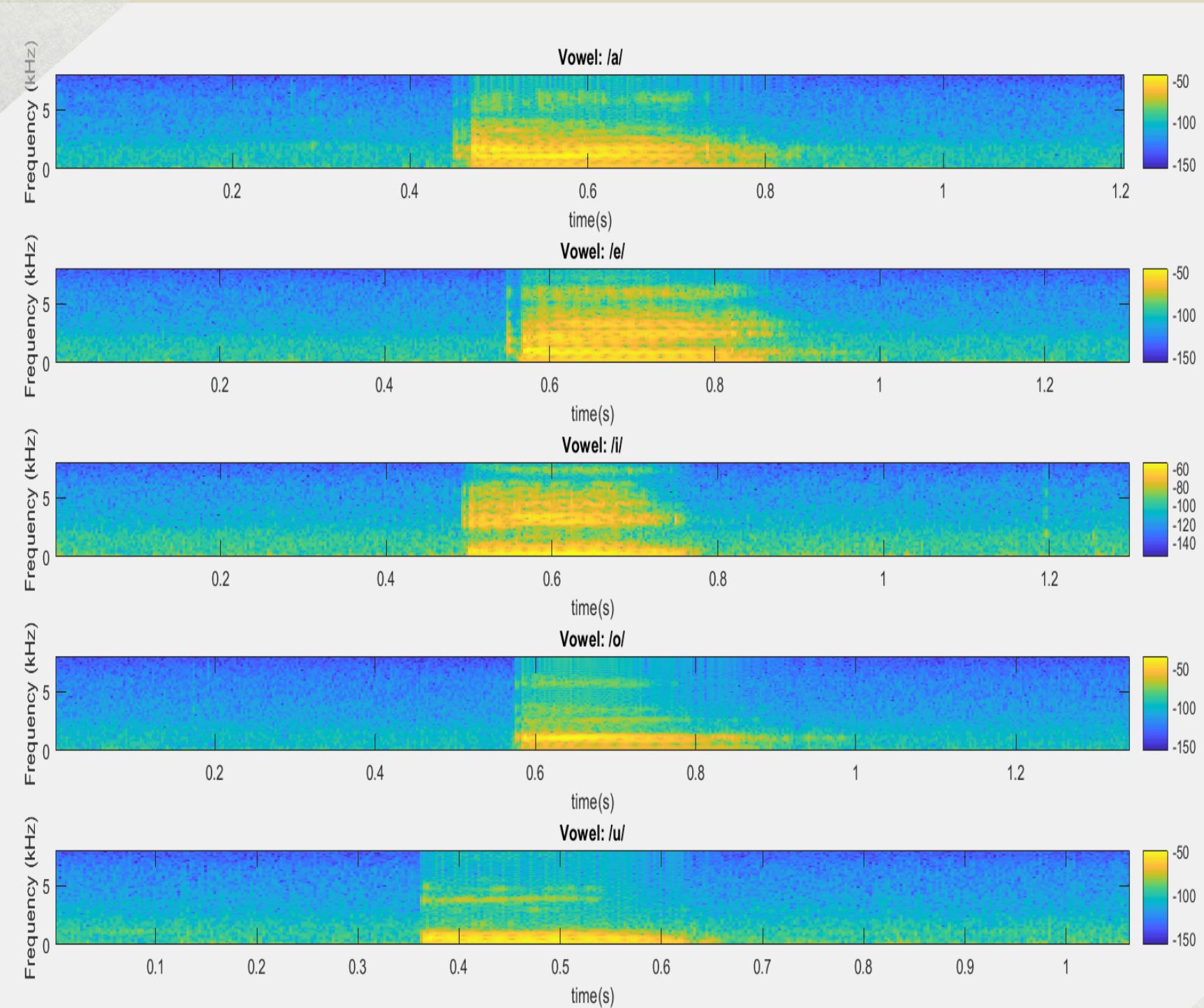


05MVB

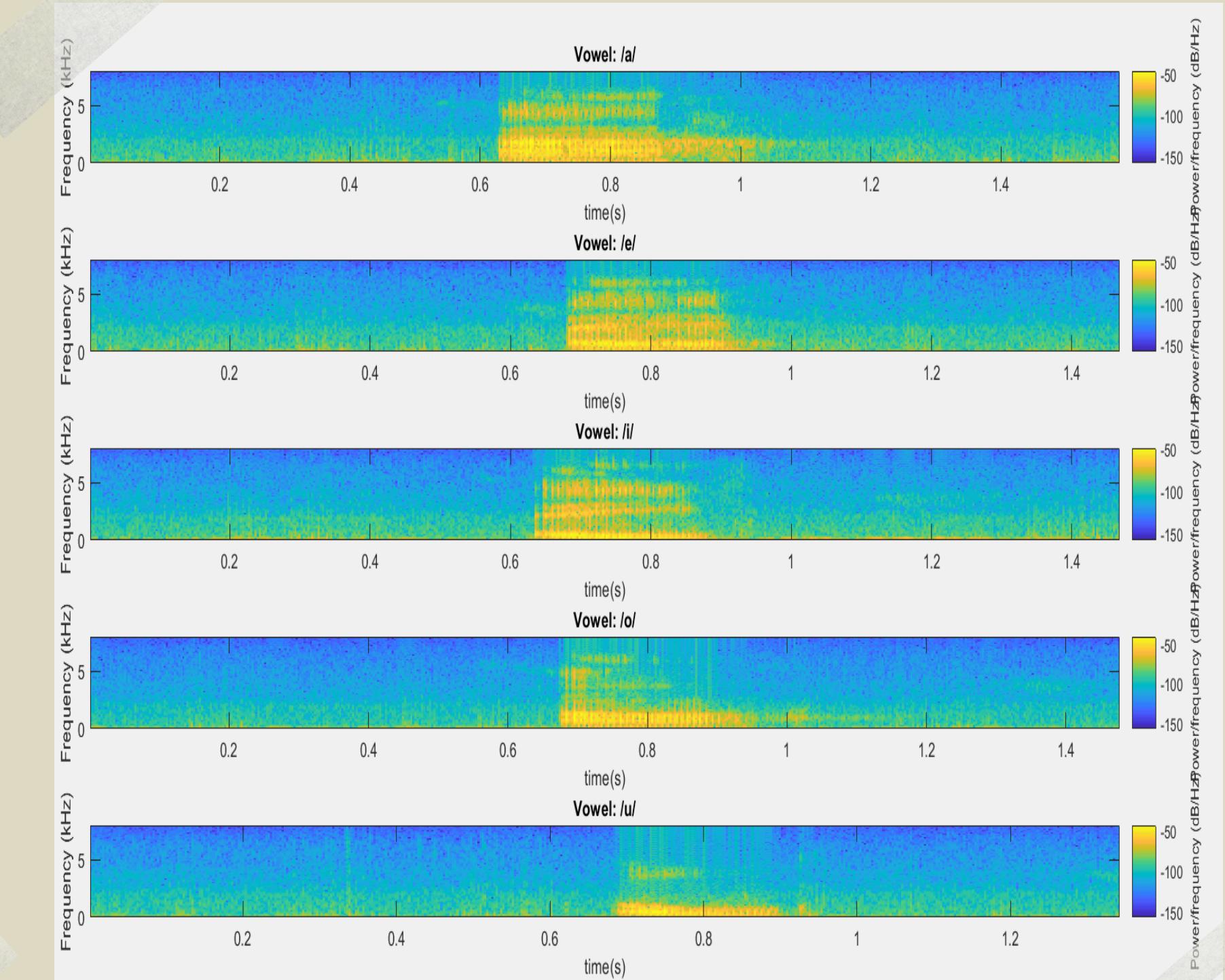


06FTB

Ảnh phổ băng rộng:

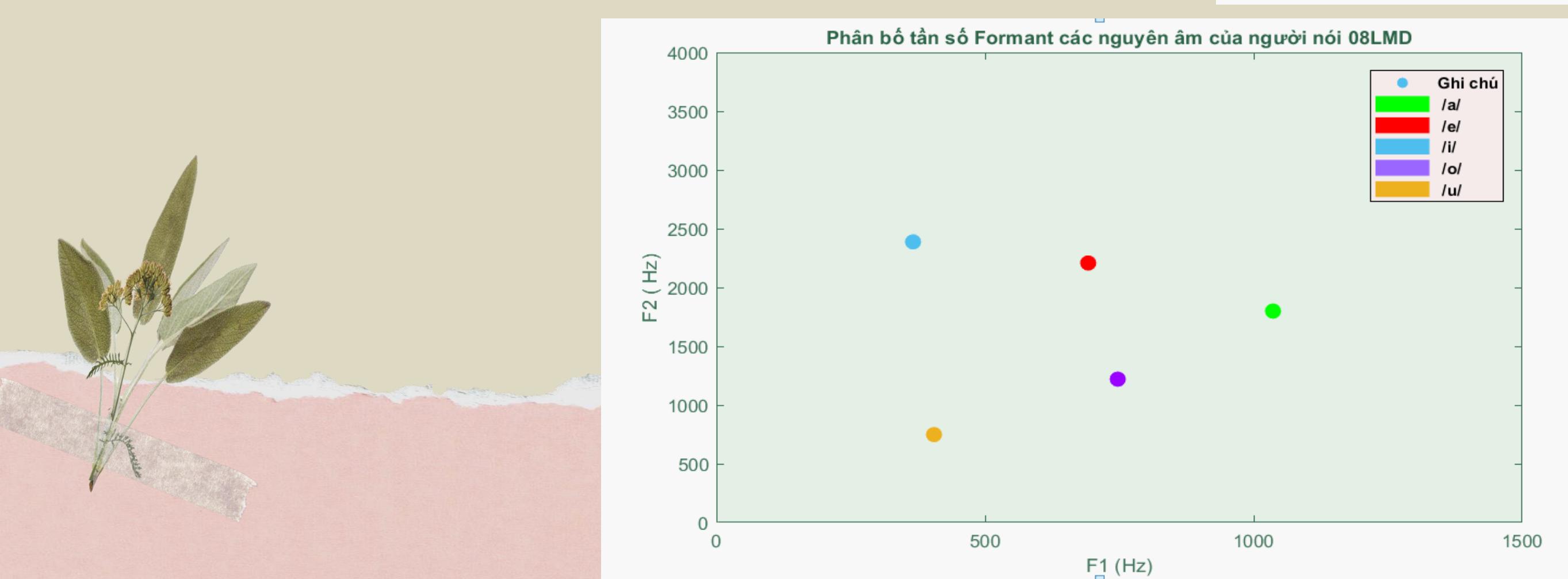
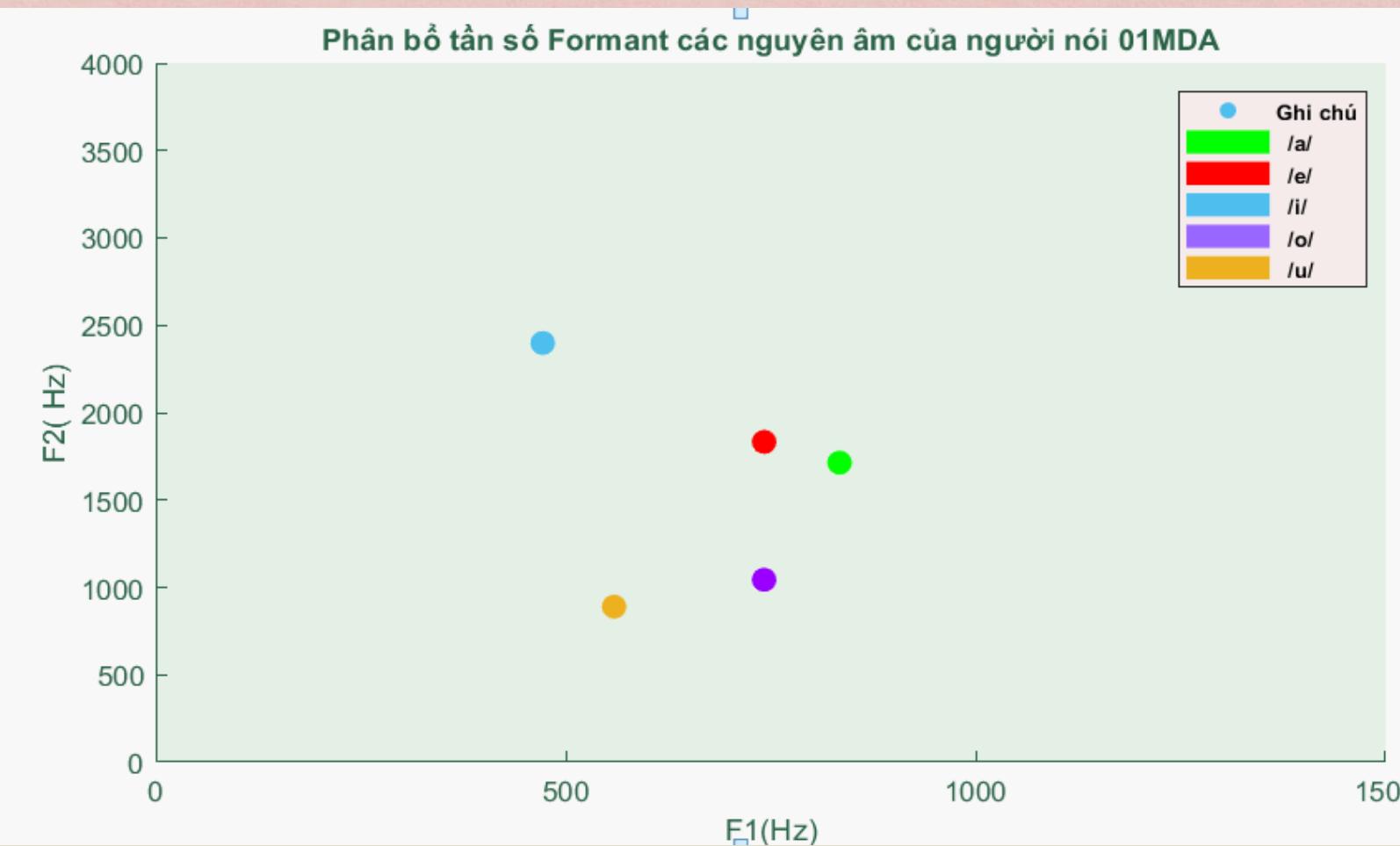


07FTC

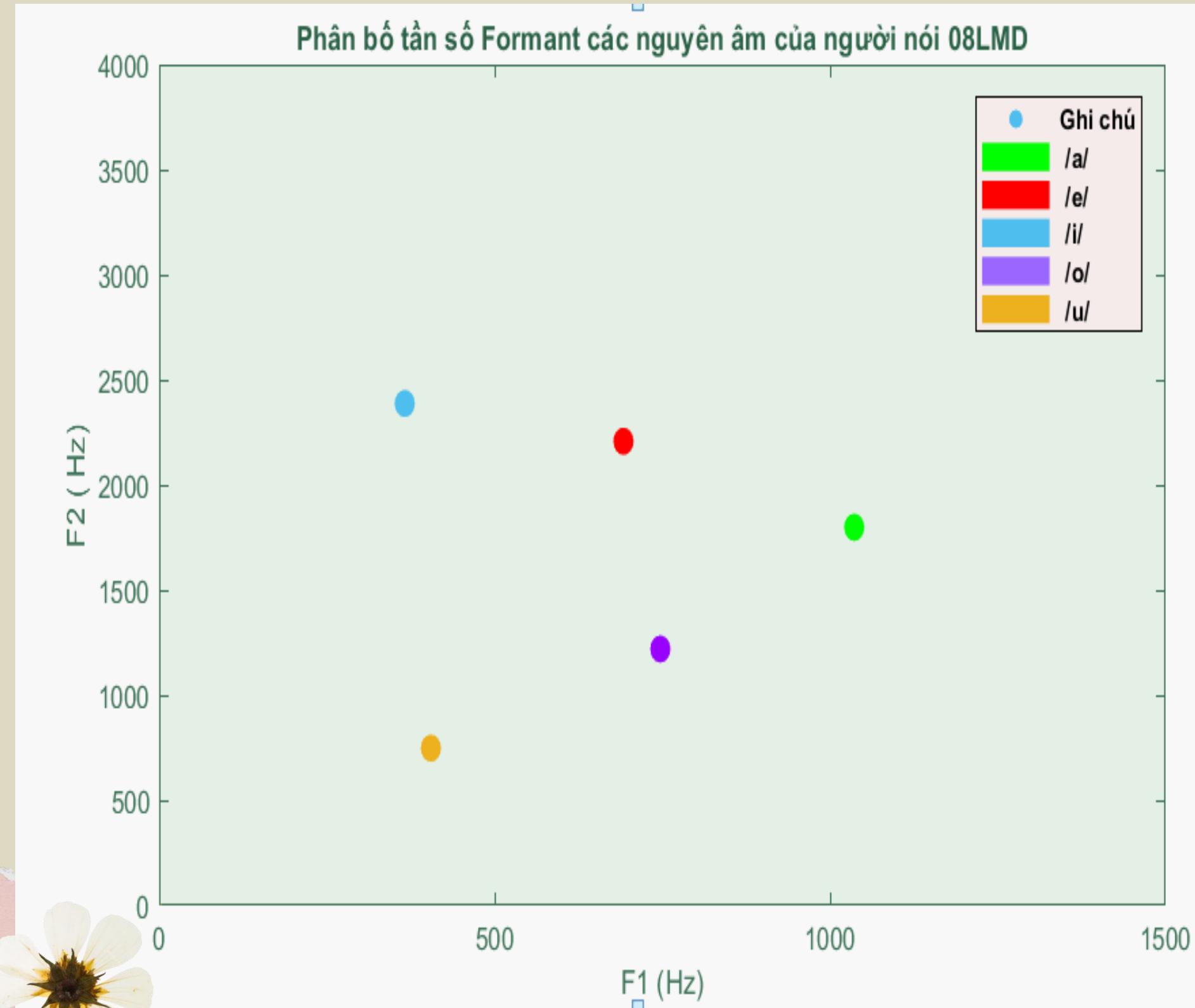


08MLD

Phân bố tần số Formant các nguyên âm của 3 trong 8 người nói



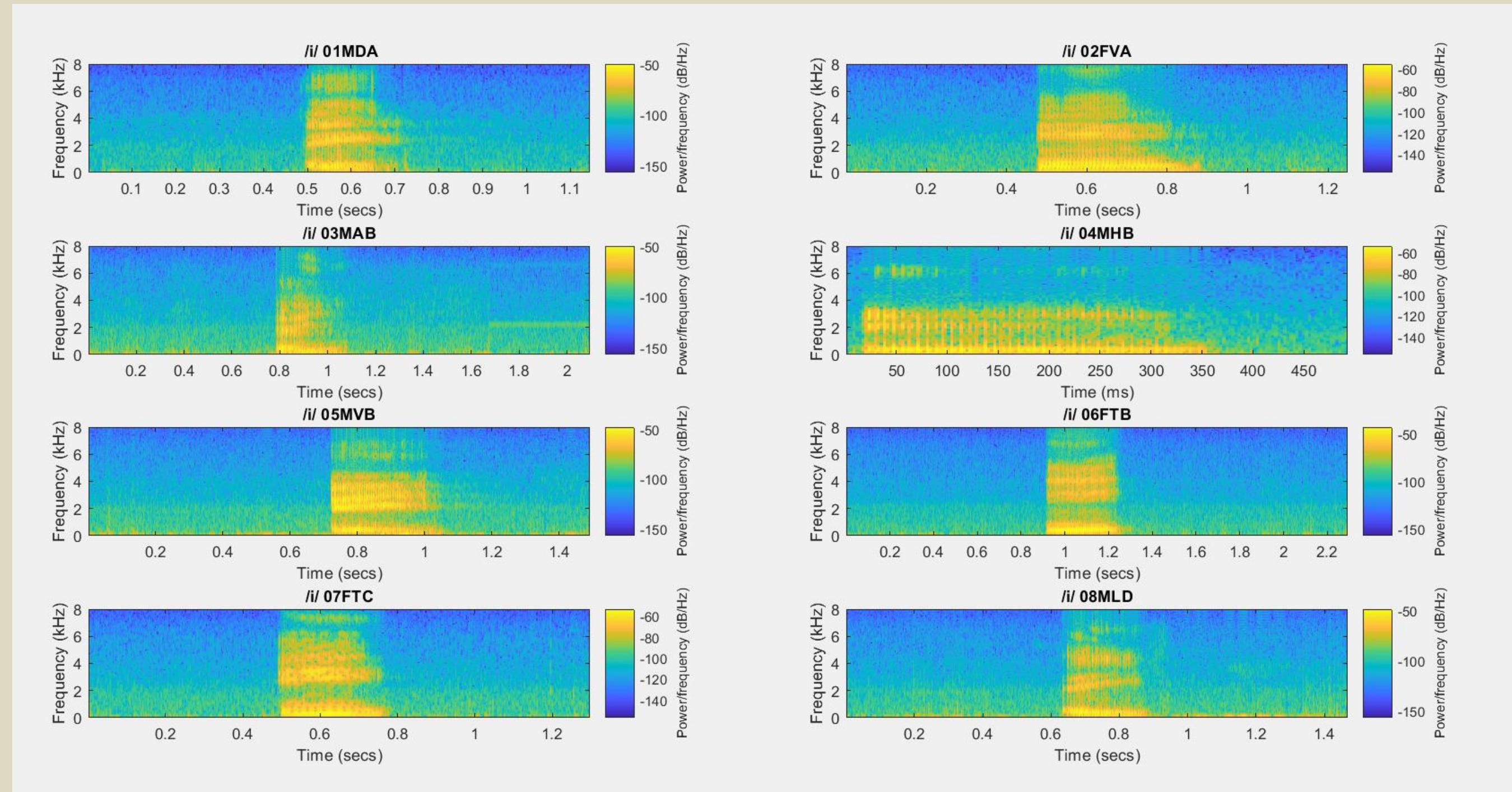
Nhận xét về đặc trưng phổ năm nguyên âm của 1 người nói



Nhận xét

- Dựa vào bộ 3 tần số formant từ ảnh phổ ta thấy được rằng:
 - ❖ 5 nguyên âm có sự khác biệt rõ ràng đáng kể giữa mỗi người nói.
- ➔ **Phân biệt được 5 nguyên âm của cùng một người nói.**

Nhận xét về đặc trưng phổ 1 nguyên âm nhiều người nói



Đặc trưng phổ âm /i/ của 8 người nói

Nhận xét về đặc trưng phổ 1 nguyên âm nhiều người nói

/a/	F1	F2	F3
mean(Hz)	926	1522	2622
std	108.06	224.28	464.46
CV(%)	11.67%	14.73%	17.71%

/e/	F1	F2	F3
mean(Hz)	747	1857	2792
std	92.07	403.03	488.04
CV(%)	12.32%	21.70%	17.48%

/i/	F1	F2	F3
mean(Hz)	429	2239	3050
std	74.84	434.29	303.04
CV(%)	17.47%	19.40%	9.94%

/ɔ/	F1	F2	F3
mean(Hz)	793	1079	2683
std	82.77	79.57	326.80
CV(%)	10.44%	7.37%	12.18%

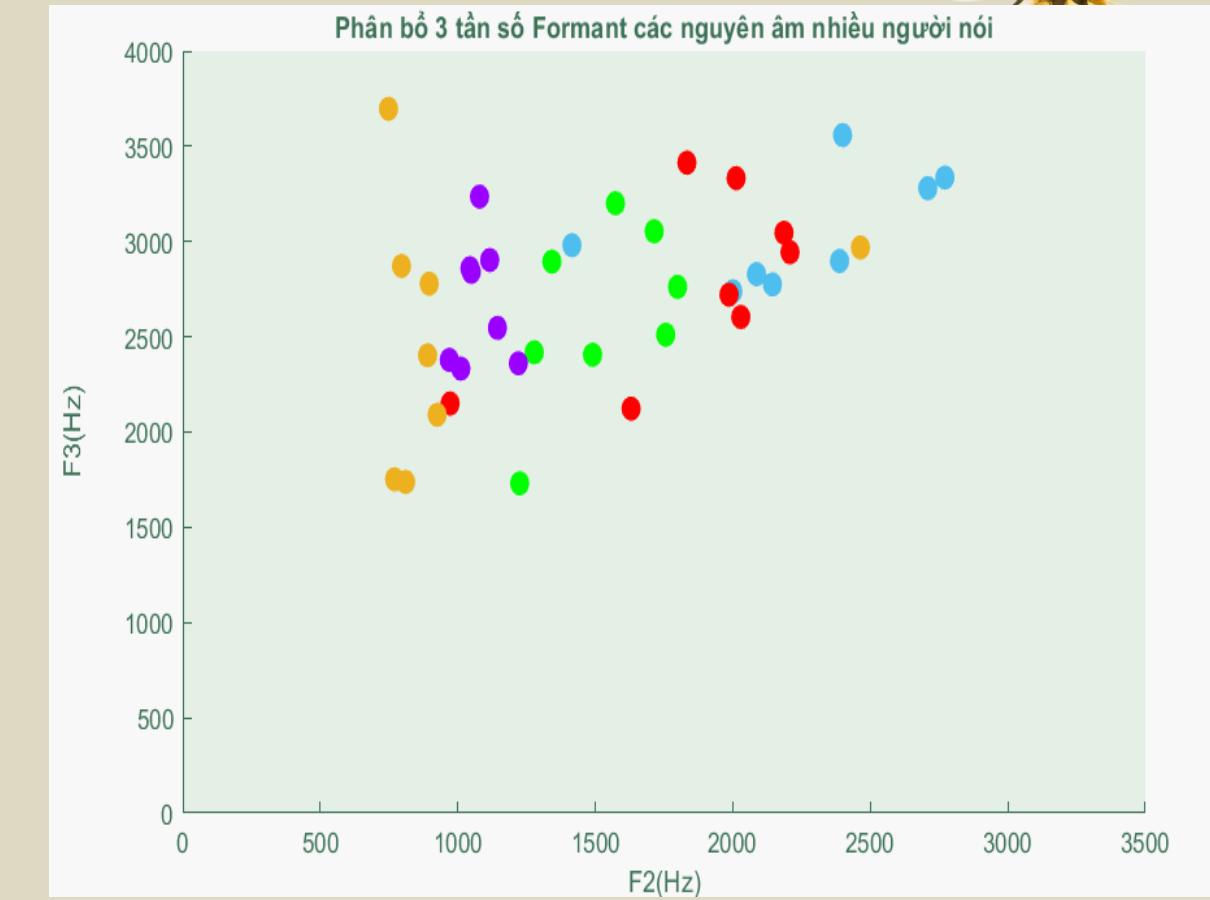
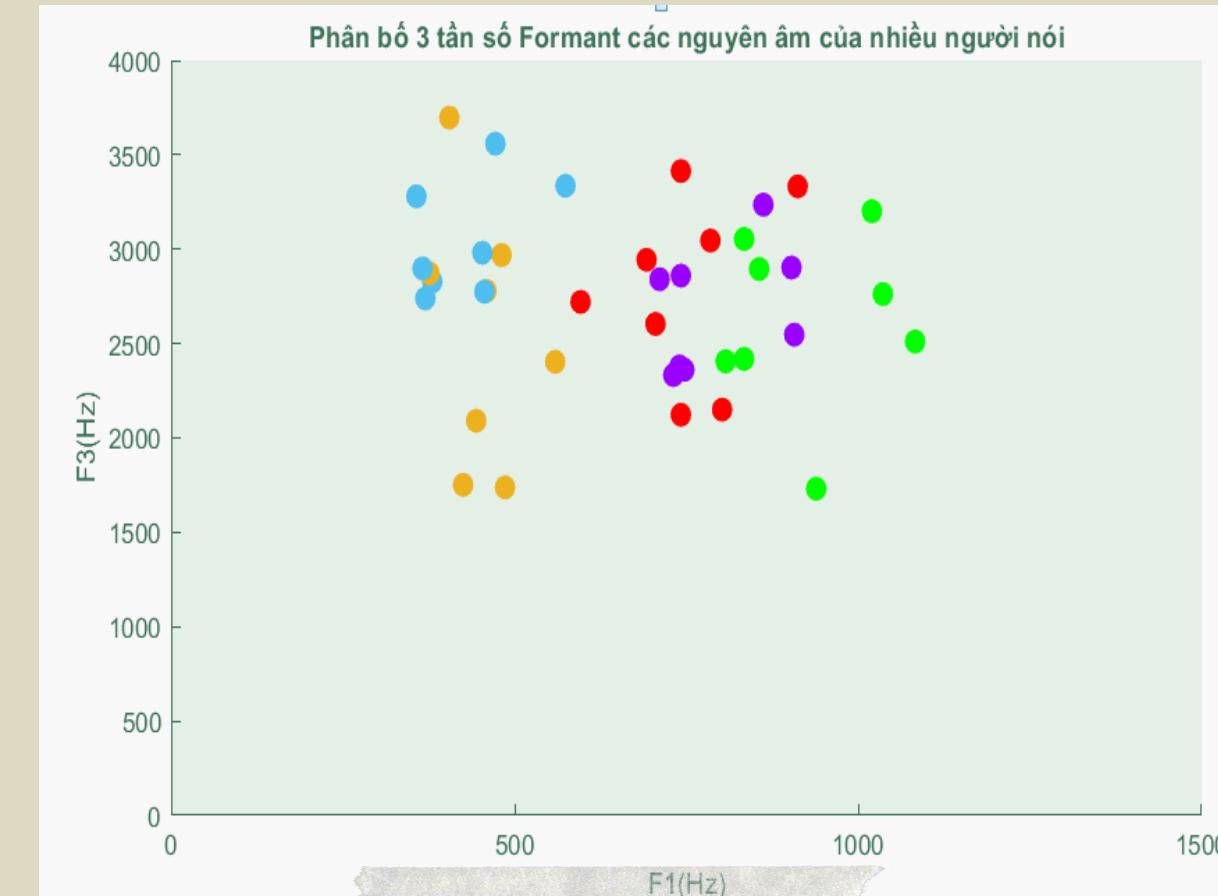
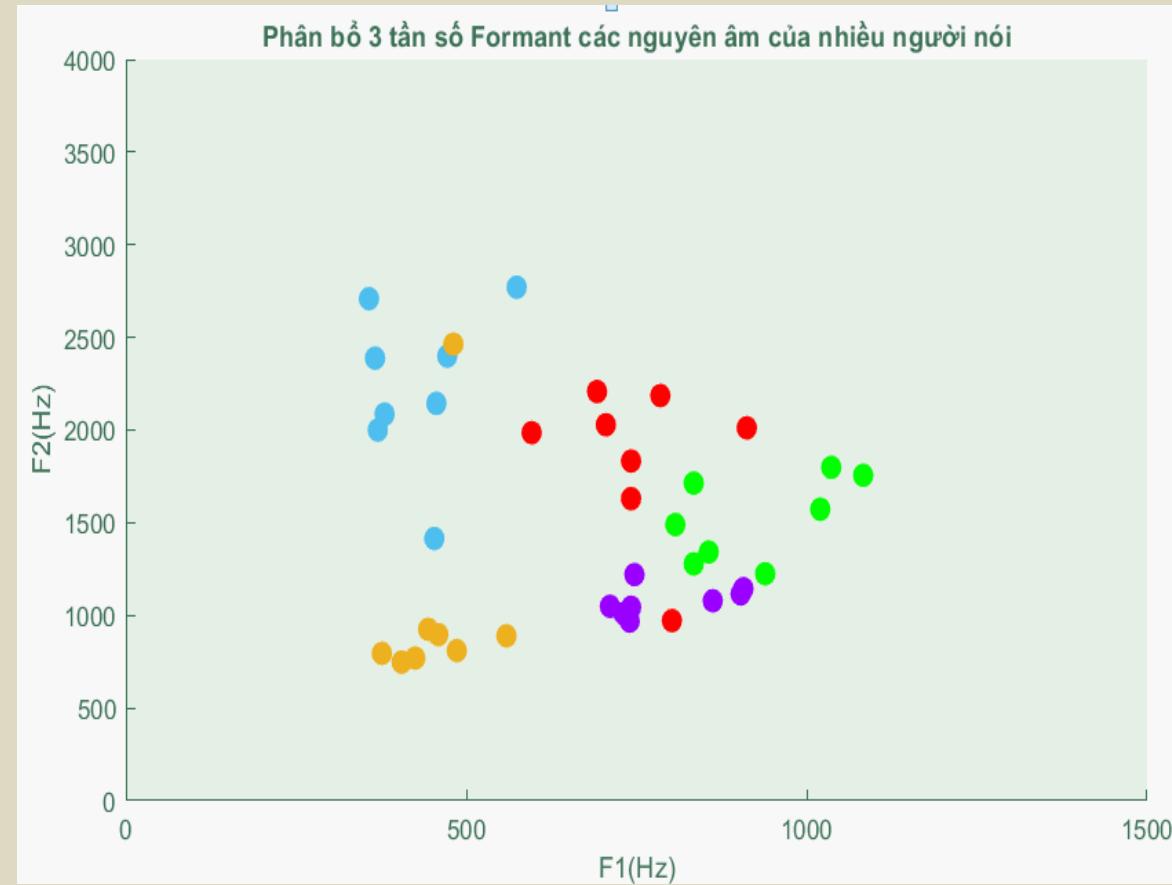
/u/	F1	F2	F3
mean(Hz)	454	1037	2538
std	56.35	580.06	673.42
CV(%)	12.40%	55.92%	26.54%

Thống kê bộ 3 formant của mỗi nguyên âm từ 8 người nói

Tên file	/a/			/e/			/i/			/o/			/u/		
	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
01MDA	834	1714	3054	742	1833	3414	472	2399	3559	742	1044	2860	559	890	2403
02FVA	939	1225	1731	802	972	2150	574	2771	3336	903	1116	2903	486	810	1738
03MAB	807	1490	2406	742	1630	2123	370	2000	2739	731	1011	2333	444	925	2091
04MHB	856	1342	2895	596	1986	2721	380	2086	2829	711	1049	2841	459	896	2779
05MVB	834	1278	2419	705	2029	2604	456	2144	2775	740	969	2379	481	2464	2969
06FTB	1020	1573	3201	785	2186	3046	453	1415	2981	862	1079	3236	425	770	1752
07FTC	1083	1756	2511	912	2012	3333	357	2709	3280	907	1144	2547	376	795	2872
08MLD	1036	1799	2762	692	2208	2944	366	2388	2898	747	1220	2361	405	748	3697

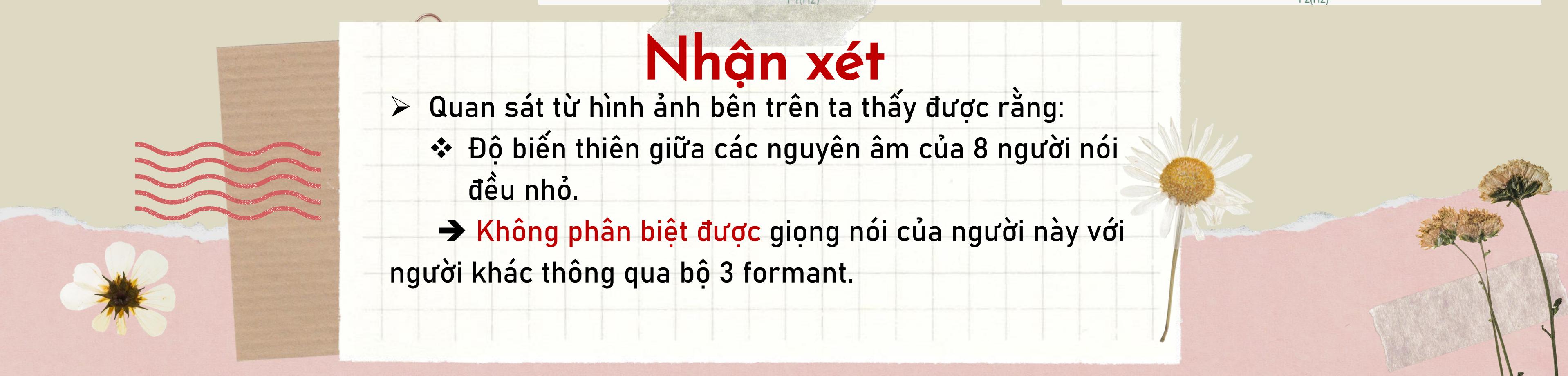
Thống kê bộ 3 formant 5 nguyên âm của 8 file tín hiệu

Nhận xét về đặc trưng phổ 1 nguyên âm nhiều người nói



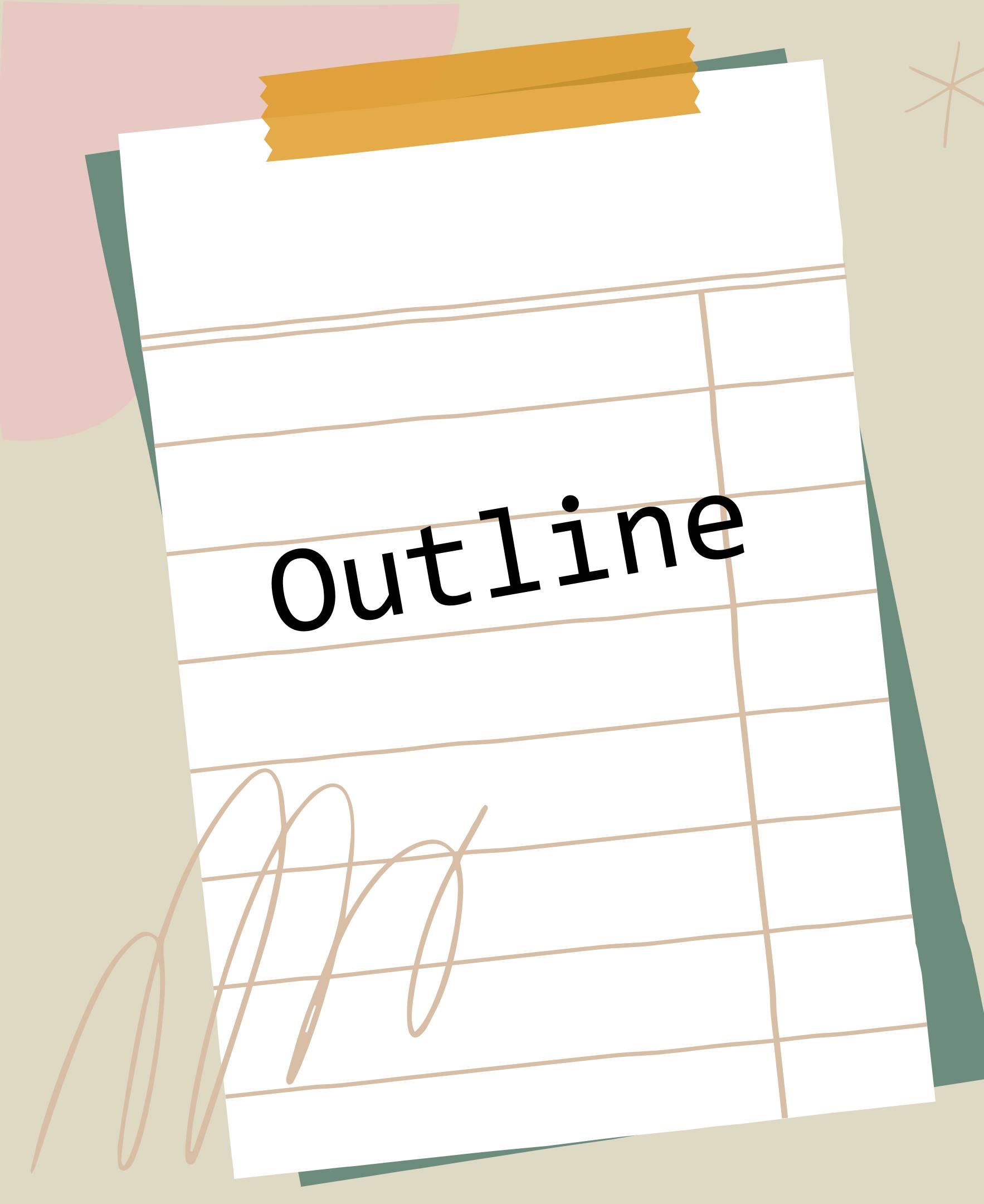
Nhận xét

- Quan sát từ hình ảnh bên trên ta thấy được rằng:
 - ❖ Độ biến thiên giữa các nguyên âm của 8 người nói đều nhỏ.
 - ➔ Không phân biệt được giọng nói của người này với người khác thông qua bộ 3 formant.





**Nhận dạng nguyên âm không phụ
thuộc vào người nói**



outline

01

Các bước tiến hành

02

Trình bày thuật toán

03

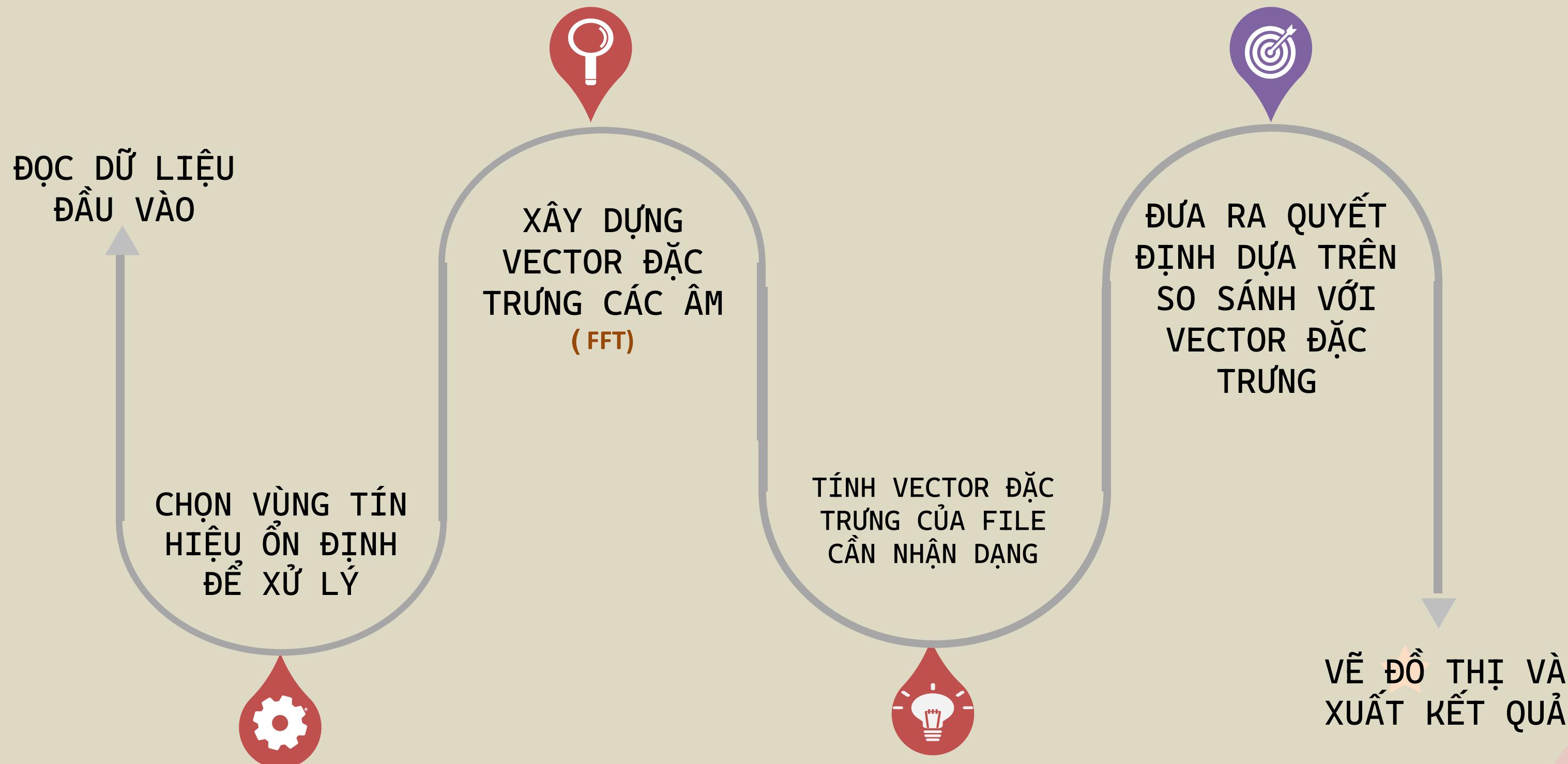
Kết quả thực hiện

04

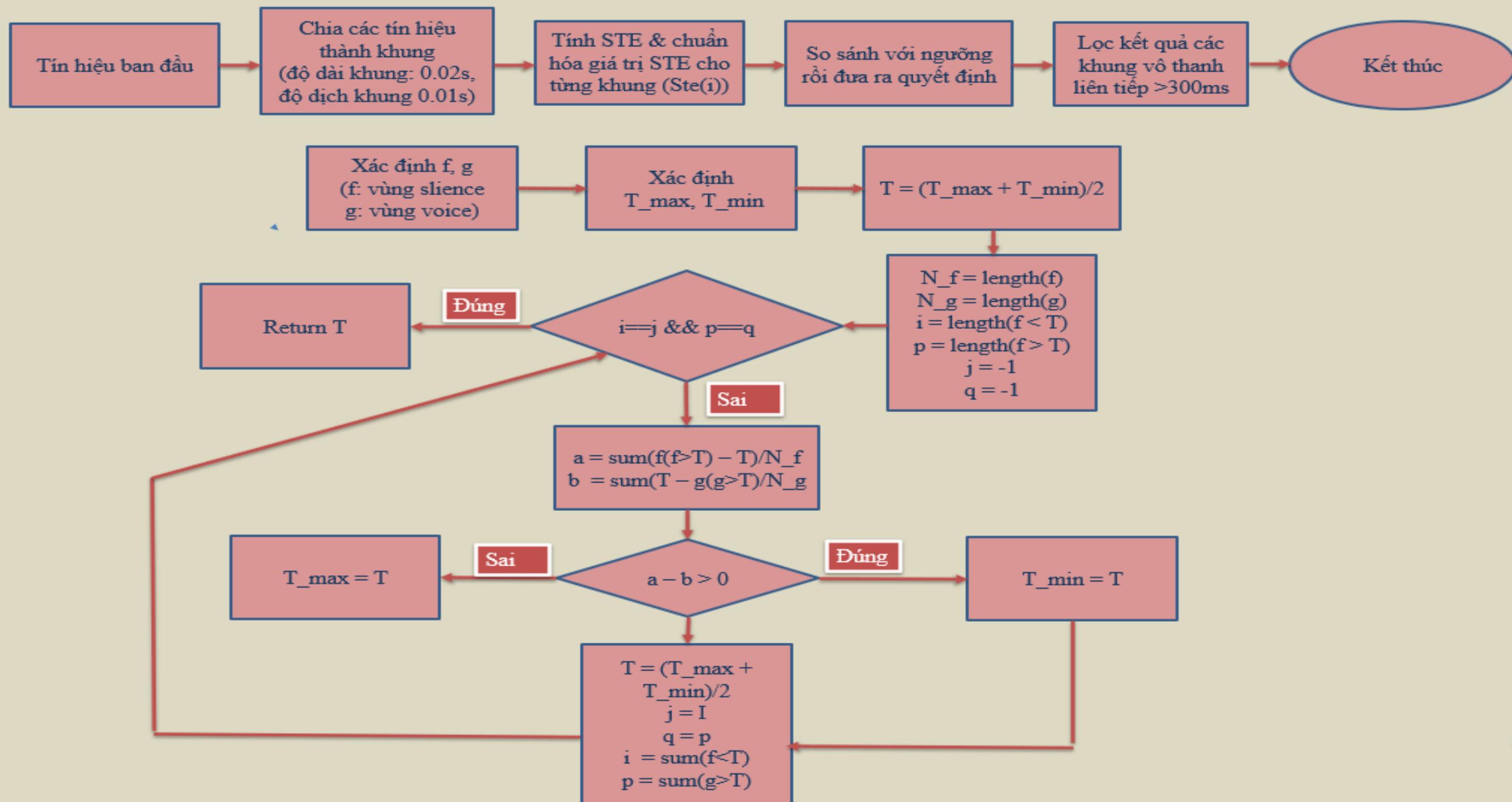
Nhận xét kết quả

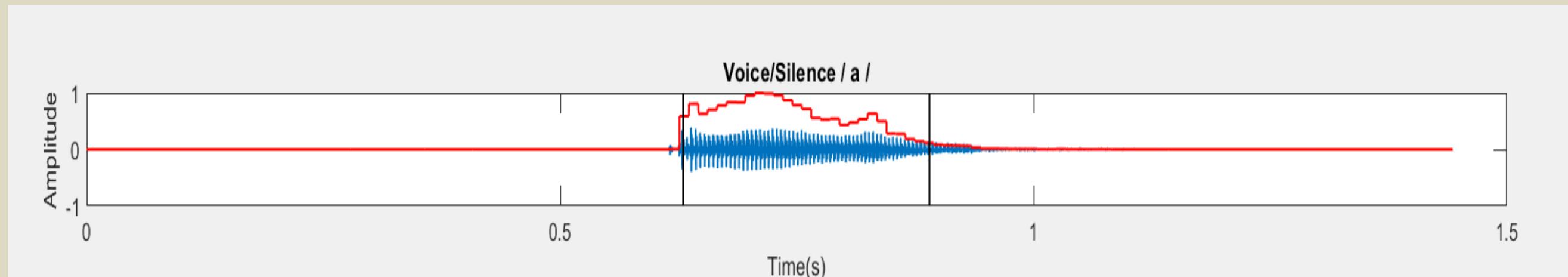
01

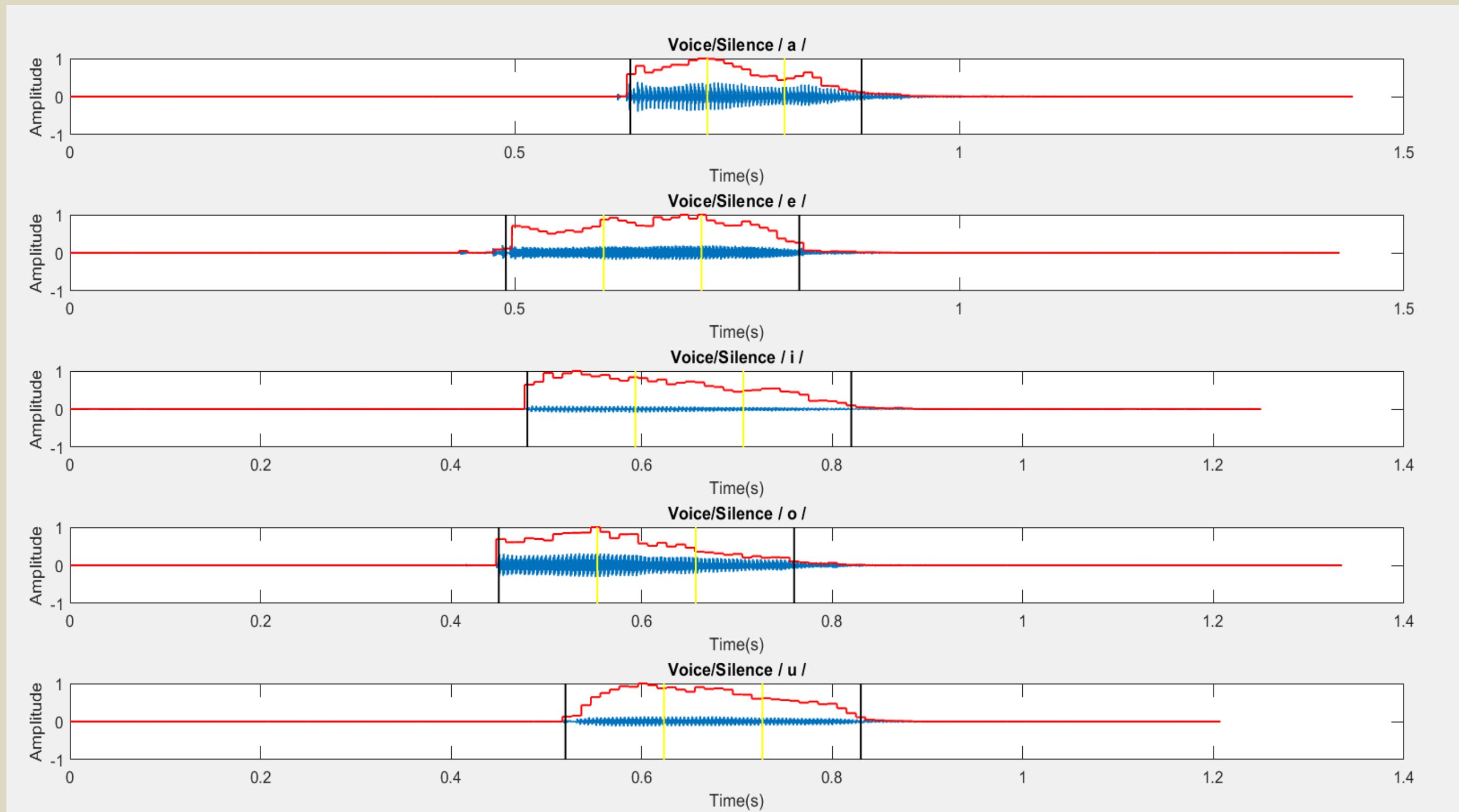
Các bước tiến hành



Thuật toán phân đoạn Voice/Silence



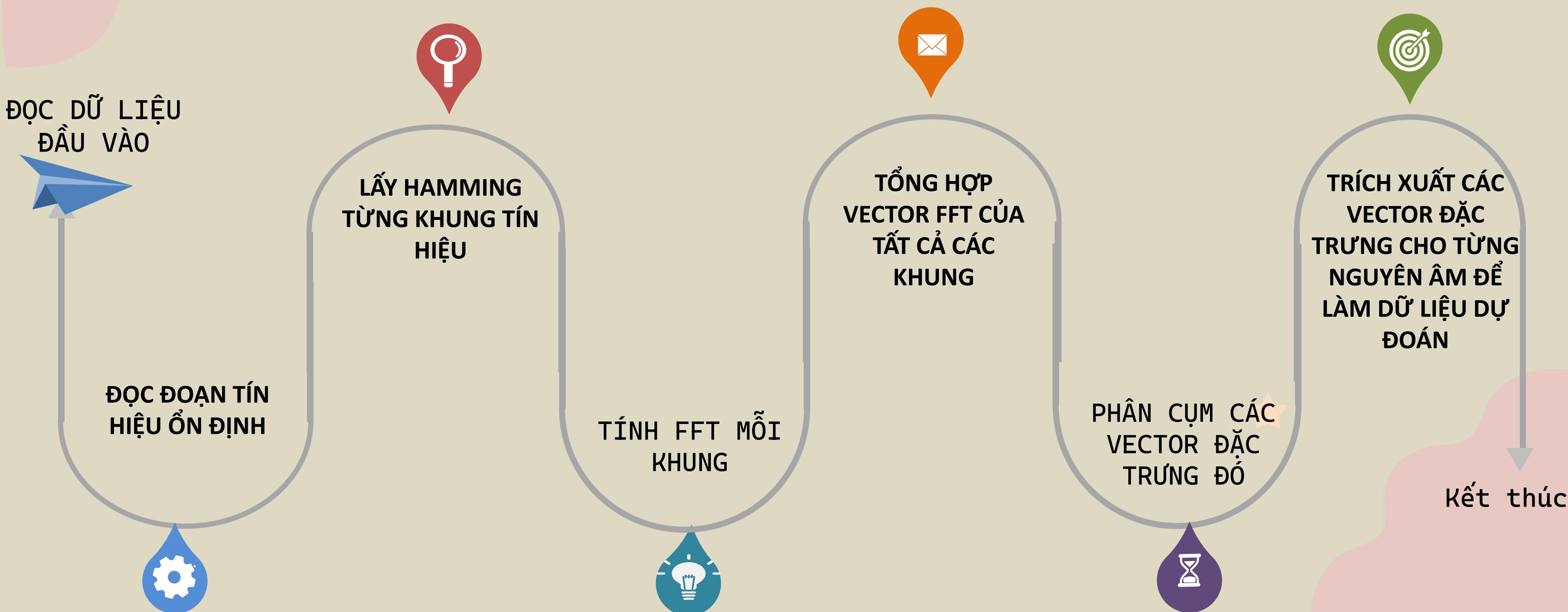
Hình ảnh minh họa cho các bước của thuật toán**Bước 1 : Phân đoạn nguyên âm của file tín hiệu đầu vào****Bước 2: Chia phần nguyên âm thành 3 phần bằng nhau và đoạn chính giữa**

Bước 3: Xuất kết quả cho cả 5 nguyên âm

02

Trình bày thuật toán

Thuật toán nhận dạng nguyên âm không phụ thuộc vào người nói (dựa trên vector đặc trưng FFT)



Thuật toán phân cụm

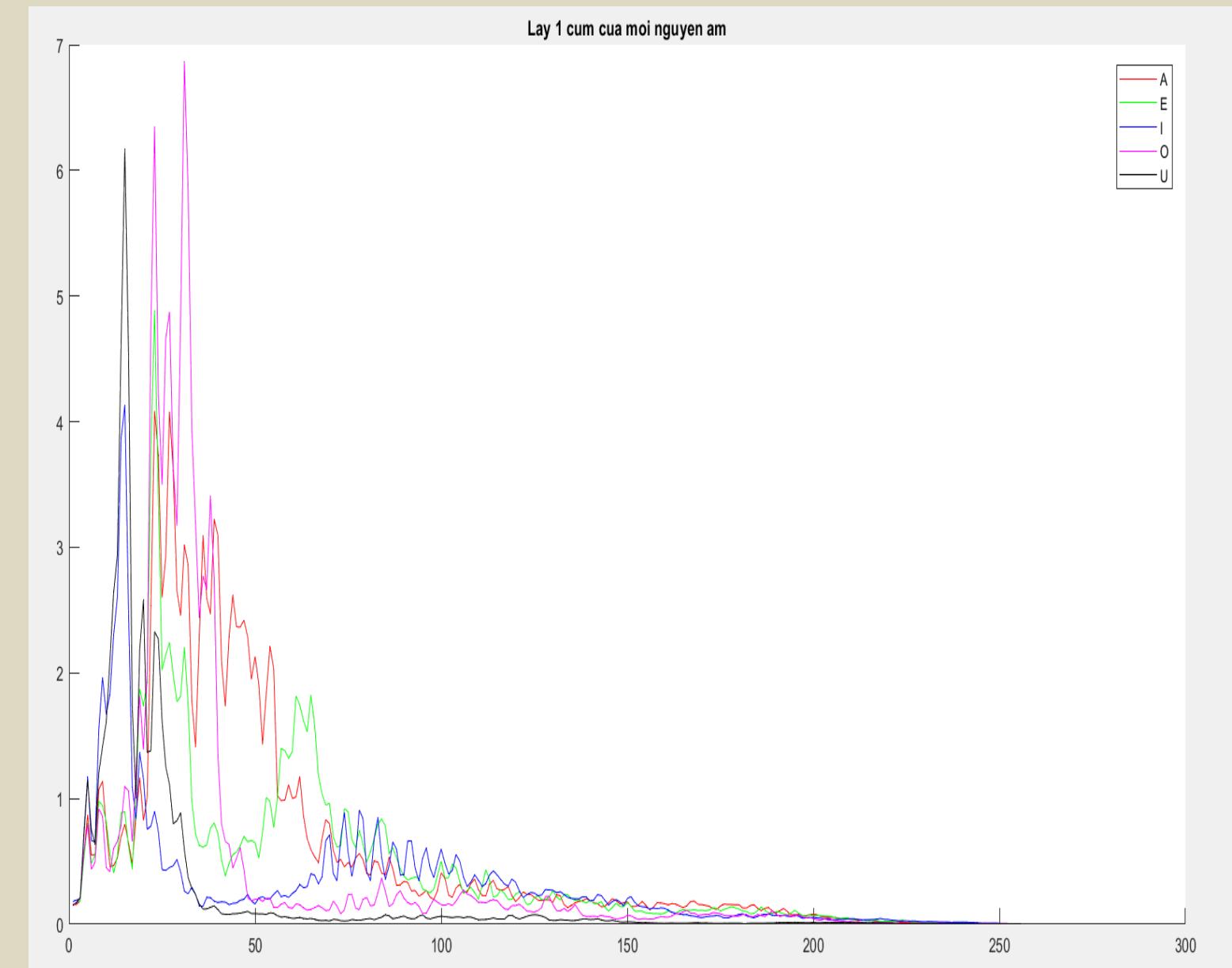
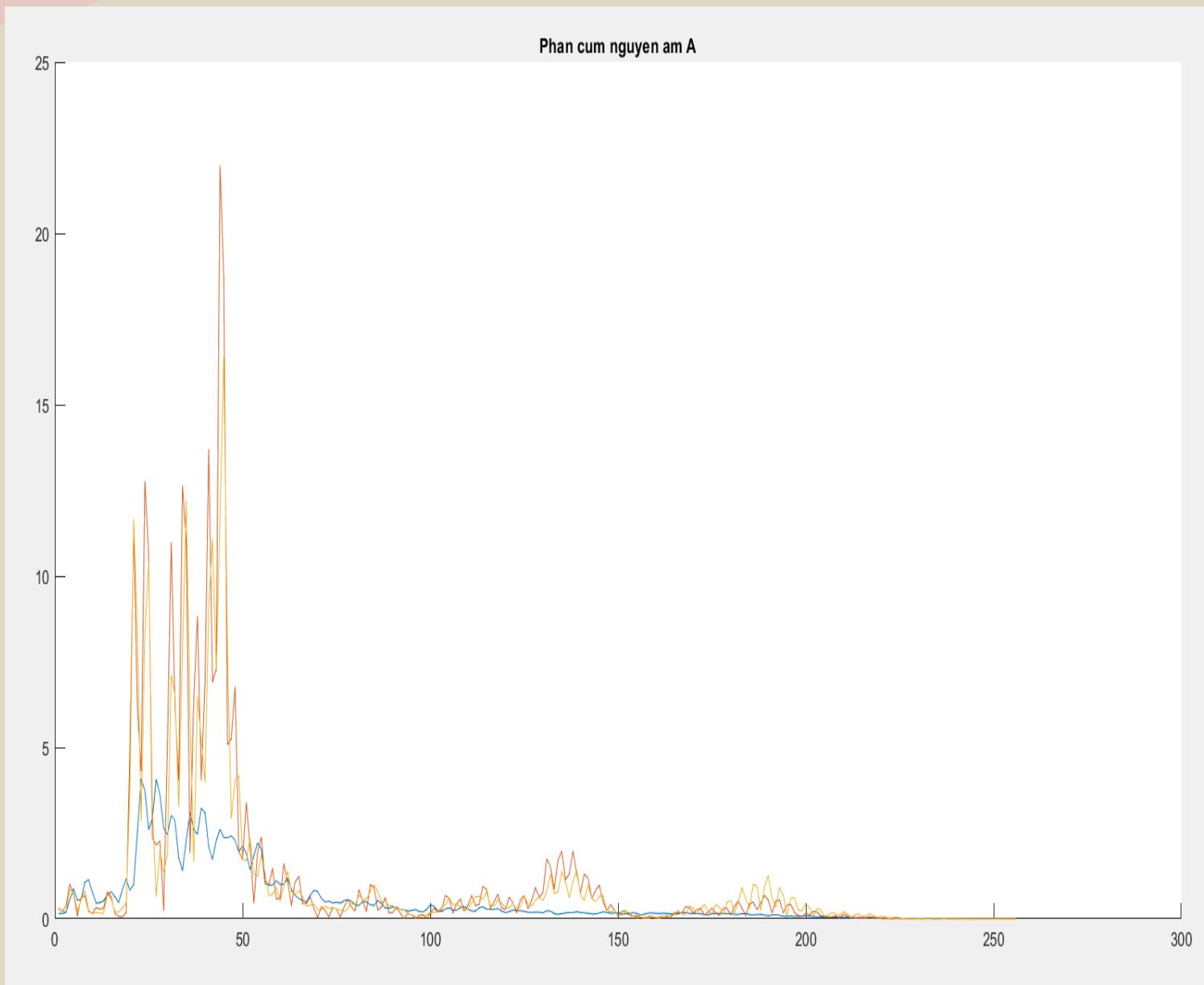
```
[idx,C] = kmeans(vecto,k_mean,'Start',aa);
```

- ‘vecto’ : danh sách các vector cần phân cụm
- ‘k_mean’: số cụm
- ‘Start’: phương pháp chọn vị trí bắt đầu cụm trung tâm
- ‘C’: kết quả vector trả về

02

Trình bày thuật toán

Hình ảnh phân cụm

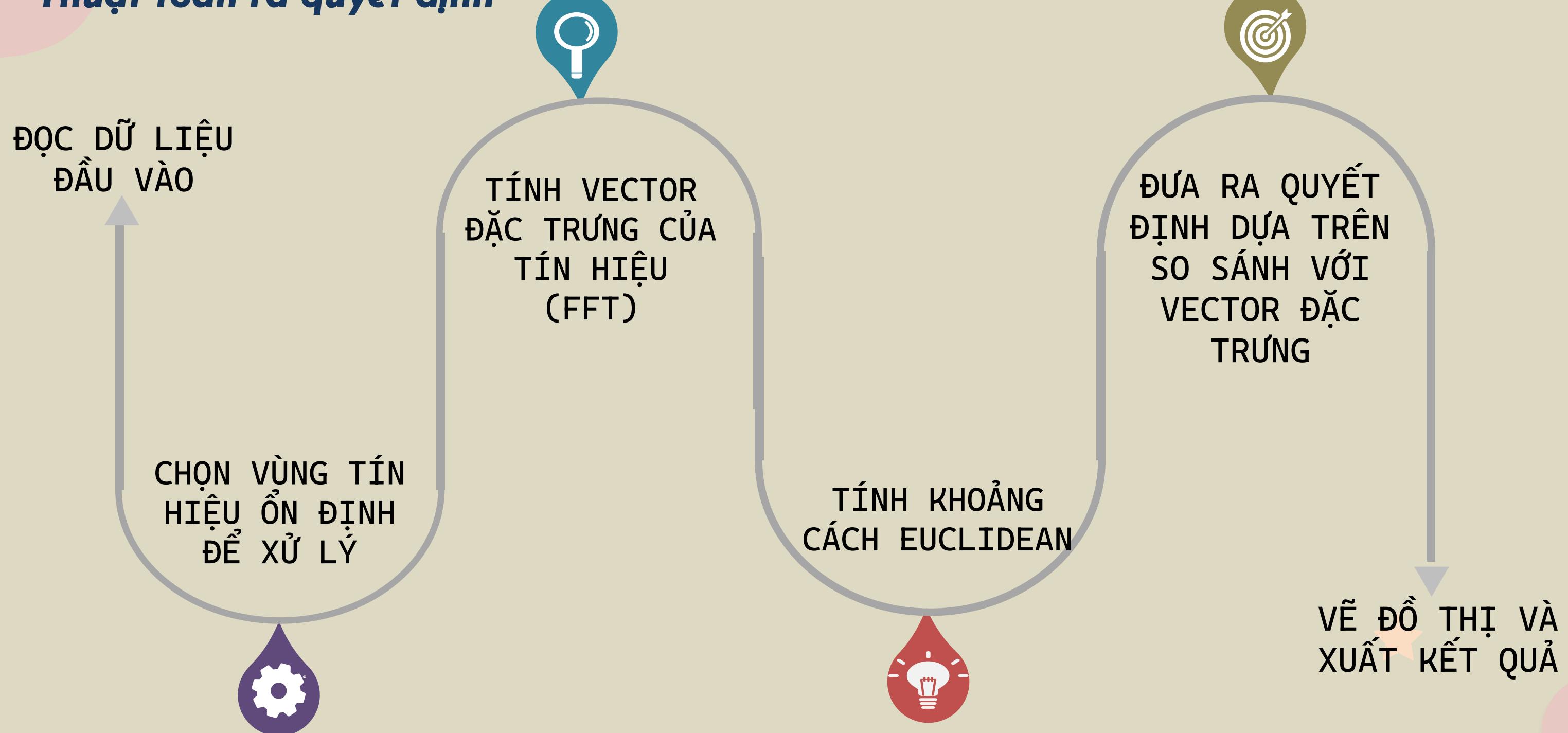


K-mean = 3 ; ffffpoint = 512

02

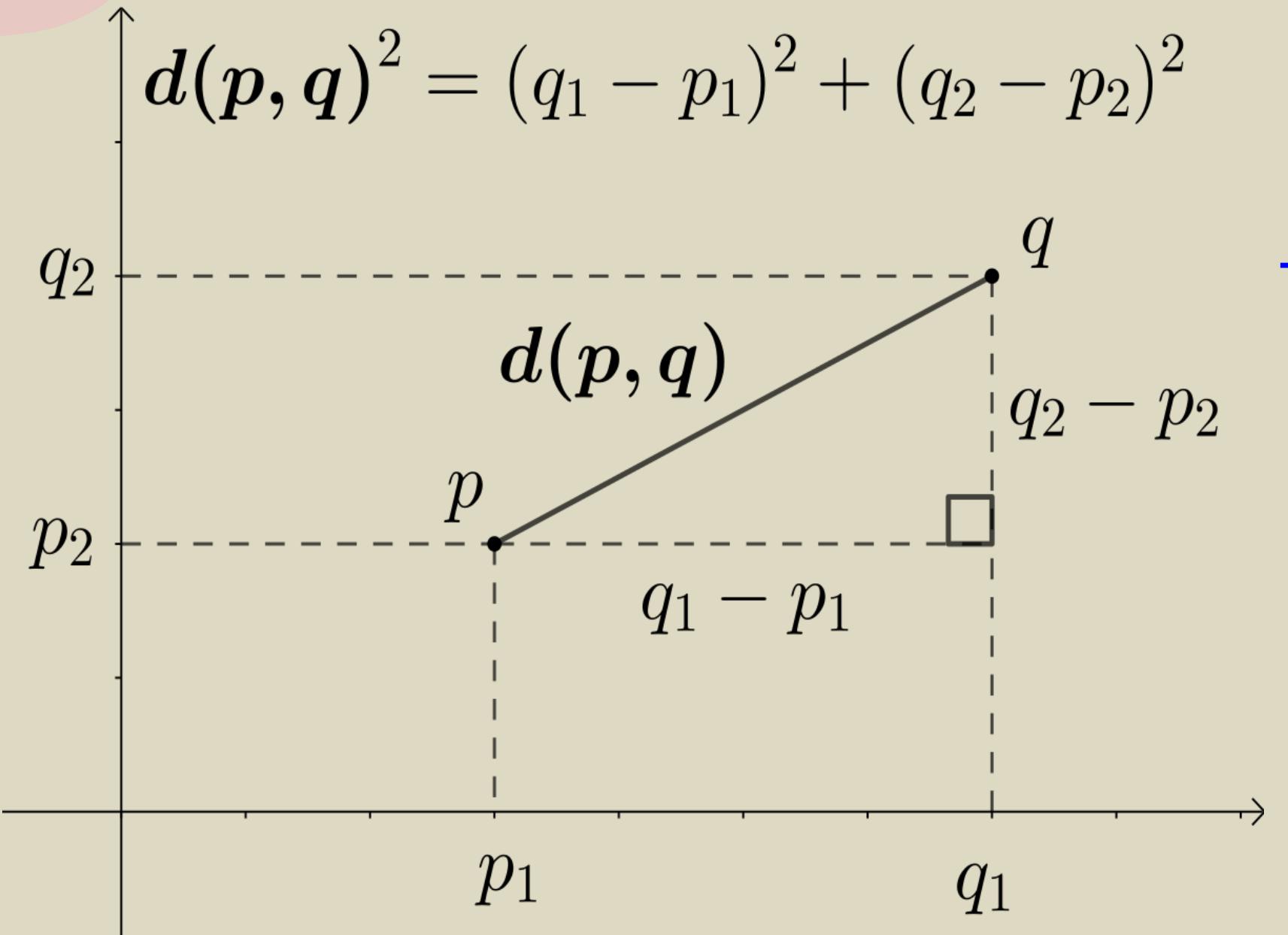
Trình bày thuật toán

Thuật toán ra quyết định



02

Trình bày thuật toán



Hàm tính khoảng cách
function $d = \text{distance_v}(\text{vectoA}, \text{vectoB})$

TÍNH KHOẢNG CÁCH EUCLIDEAN

03

Kết quả thực hiện

Ngưỡng
tính từ
file lab
GK

Thuật toán		Số cụm							
FFT	Số điểm FFT	2		3		4		5	
		Độ chính xác (%)	Thời gian(s)						
		512	75,23	2,18	75,23	2,20	75,23	2,30	74,28
FFT	1024	75,23	2,21	75,23	2,21	75,23	2,23	74,28	2,25
FFT	2048	75,23	2,32	75,23	2,34	75,23	2,39	74,28	2,53

Ngưỡng tự
cho = 0,01

Thuật toán		Số cụm							
FFT	Số điểm FFT	2		3		4		5	
		Độ chính xác (%)	Thời gian(s)						
		512	78,09	1,07	80	1,18	80	1,22	80
FFT	1024	78,09	1,20	80	1,22	80	1,24	80	1,26
FFT	2048	78,09	1,30	80	1,32	80	1,33	80	1,55

03

Kết quả thực hiện

MA TRAN NHAM LAN:

	a	e	i	o	u
a	17	1	2	1	0
e	0	17	3	1	0
i	0	0	19	0	2
o	2	4	2	13	0
u	0	0	8	0	13

MA TRAN NHAM LAN:

	a	e	i	o	u
a	17	1	2	1	0
e	1	16	3	1	0
i	0	0	19	0	2
o	2	4	2	13	0
u	0	0	8	0	13

K_mean = 3 ; ffffpoint = 512
Độ chính xác 75,23%

K_mean = 5 ; ffffpoint = 512
Độ chính xác 74,28%

03

Kết quả thực hiện

MA TRAN NHAM LAN:

	a	e	i	o	u
a	16	0	3	2	0
e	0	16	4	1	0
i	0	0	19	0	2
o	2	4	2	13	0
u	0	0	3	0	18

MA TRAN NHAM LAN:

	a	e	i	o	u
a	17	0	3	1	0
e	0	16	4	1	0
i	0	0	19	0	2
o	2	3	2	14	0
u	0	0	3	0	18

K_mean = 1 ; ffffpoint = 512
Độ chính xác 78,09%

K_mean = 5 ; ffffpoint = 512
Độ chính xác 80%



- ❸ **Kết quả cho được với độ chính xác tương đối cao lắm**
- ❹ **Việc thay đổi các điểm FFT từ 512 , 1024 , 2048 không làm thay đổi kết quả của độ chính xác**
- ❺ **Việc phân cụm k_mean cho vector đặc trưng cho kết quả tốt hơn nhưng việc tăng số cụm một cách gượng ép không làm tăng độ chính xác.**
- ❻ **Nguyên âm /i/ là nguyên âm có độ chính xác cao nhất trong 5 nguyên âm với kết quả chính xác là 19/21.**
- ❼ **Việc chọn hay tính ngữ nghĩa đúng cũng sẽ ảnh hưởng đến độ chính xác của kết quả.**

Cảm ơn thầy
và các bạn đã
lắng nghe!



@reallygreatsite