

**ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN**

**TRẦN TẤN BẢO – 17C 11002
TRẦN THÚY HIỀN – 17C 11026
NGUYỄN HÀ DUY PHƯƠNG – 17C 11032**

**ĐỒ ÁN MÔN HỌC
XỬ LÝ NGÔN NGỮ NÓI**

BÀI TẬP 1: NHẬN DẠNG GIỌNG NÓI

GIÁO VIÊN

PGS. TS. Vũ Hải Quân

TP.HCM - 6/2018

MỤC LỤC

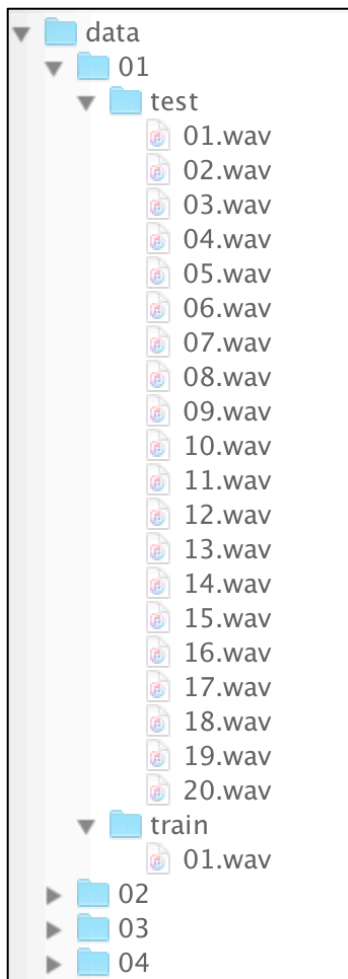
1	Mô tả bài toán	2
2	Phương pháp làm.....	2
3	Kịch bản thử nghiệm.....	4
4	Kết quả, vẽ biểu đồ	5
4.1	Kết quả thử nghiệm.....	5
4.2	Biểu đồ tổng kết.....	8
5	Đánh giá cá nhân	9

1 Mô tả bài toán

- Mục đích của bài tập là nhận dạng người nói. Chương trình sẽ tìm đặc trưng của từng người, sau đó với mỗi file kiểm thử sẽ xác định ai là người nói. Tính tỉ lệ nhận dạng chính xác của từng người nói.
- Tập dữ liệu gồm 65 người nói khác
- Sử dụng công cụ Matlab để viết chương trình huấn luyện và kiểm thử.
- Sử dụng công cụ Audacity để cắt ghép tạo file âm thanh.

2 Phương pháp làm

- Mỗi người nói sẽ được lưu trong folder với số thứ tự tăng dần (“01”, “02”, ...), trong mỗi folder gồm có 2 folder con là “**train**” và “**test**”. Tất cả có 65 người nói.
- Folder “**train**” lưu 1 file “01.wav” có chiều dài 3 phút dùng để huấn luyện. Sử dụng Audacity để nối các file ngẫu nhiên của tập dữ liệu người nói cho trước thành file 3 phút.
- Folder “**test**” chứa 20 file test, mỗi file có chiều dài 3 giây. Lấy ngẫu nhiên 20 file âm thanh trong tập dữ liệu của người nói đó.
- Cấu trúc thư mục như sau:



- Với mỗi người nói, ta sẽ đọc file huấn luyện và rút trích đặc trưng bộ cục khoang miệng MFCC, sau đó dùng mô hình GMM để huấn luyện người nói.
- Với mỗi file test ta cũng xác định đặc trưng MFCC sau đó tính chỉ số mean so với từng GMM của từng người nói và chọn ra giá trị lớn nhất. Nếu giá trị lớn nhất chọn được trùng với người nói mong đợi thì test *pass*, ngược lại là *fail*. Tính phần trăm kết quả đạt và không đạt, lập bảng kết quả và vẽ biểu đồ.
- Tiến hành thử nghiệm với nhiều lần. Mỗi lần thử thay đổi các tham số Gauss, số vòng lặp Init, số vòng lặp huấn luyện, ghi nhận kết quả và vẽ biểu đồ.
- 2 tập tin code để thực hiện là *test.m* và *training.m*.
- Tổng kết bằng 2 biểu đồ:
 - o Biểu đồ thể hiện tổng trung bình phần trăm đúng của mỗi người qua 17 lần thử nghiệm.
 - o Biểu đồ thể hiện 7 người đại diện trong tập dữ liệu với phần trăm kết quả đúng qua 17 lần thử nghiệm.

3 Kịch bản thử nghiệm

- Tiến hành thử nghiệm 17 lần. Mỗi lần thử nghiệm sẽ thay đổi các tham số Gauss, số vòng lặp gInit, số vòng lặp huấn luyện. Sau đó thống kê kết quả xác suất nhận dạng đúng và sai, và vẽ biểu đồ.
- Bảng chi tiết số lần thử nghiệm:
 - 5 lần thử đầu tiên từ lần 1 đến 5 thay đổi tham số Gauss.
 - 5 lần thử 6 đến lần 10 thay đổi tham số vòng lặp Init.
 - 5 lần thử tiếp theo, từ 11 đến 15 thay đổi tham số vòng lặp huấn luyện.
 - 2 lần thử cuối cùng 16, 17 thay đổi ngẫu nhiên cùng lúc 3 tham số.

Lần	Tham số Gauss	Số vòng lặp Init	Số vòng lặp huấn luyện
1	1	100	100
2	4	100	100
3	8	100	100
4	12	100	100
5	16	100	100
6	16	10	100
7	16	40	100
8	16	60	100
9	16	70	100
10	16	85	100
11	16	100	50
12	16	100	150
13	16	100	300
14	16	100	600
15	16	100	1000
16	12	80	200
17	16	90	500

4 Kết quả, vẽ biểu đồ

4.1 Kết quả thử nghiệm

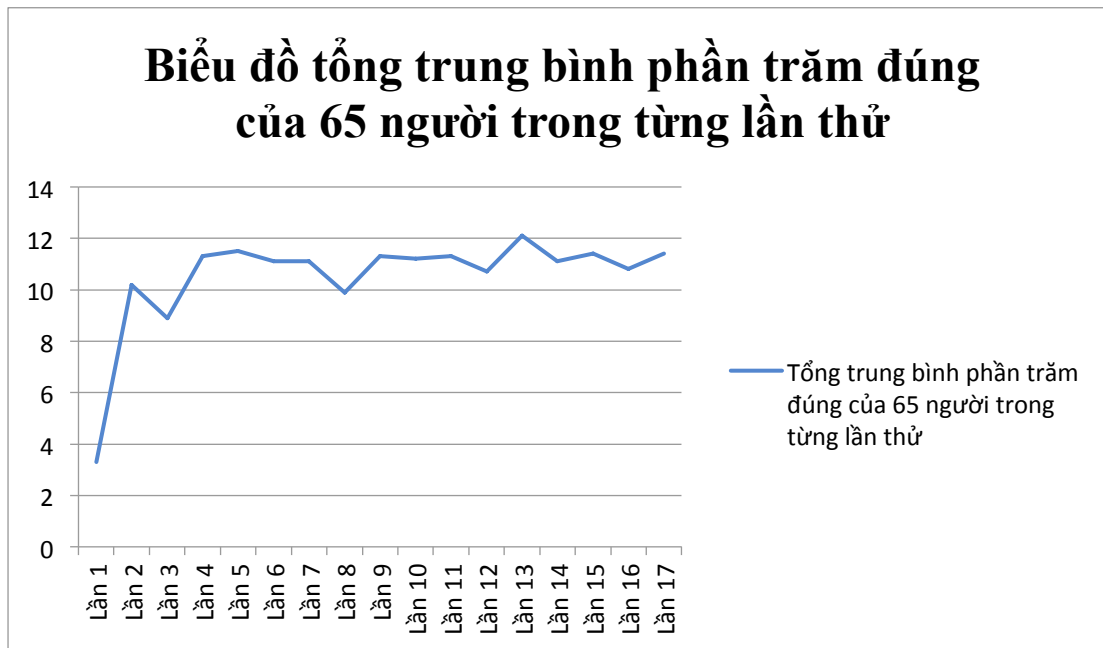
Với mỗi lần thử nghiệm thu được các kết quả khác nhau như sau:

Người nói	Tỉ lệ đúng của mỗi lần thử nghiệm (%)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
01	0	20	30	30	30	50	30	30	40	30	30	35	40	30	30	30	35
02	0	0	0	0	10	10	10	5	10	10	10	5	10	5	10	5	5
03	0	0	0	5	5	5	10	0	5	10	5	10	5	5	10	10	5
04	0	0	5	5	10	5	5	10	10	10	10	5	5	10	10	10	10
05	0	20	35	55	75	55	60	60	75	75	70	60	70	65	75	60	50
06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
07	95	95	80	90	90	80	90	90	80	85	90	85	85	90	90	80	90
08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
09	0	30	25	25	25	30	25	25	30	25	25	25	35	25	25	25	25
10	0	30	25	15	0	10	0	0	0	15	15	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	5	0	5	5	5	5	5	5	5	5	5	5	5	5	5	5
14	0	5	50	25	30	30	30	25	30	25	10	35	10	10	30	15	30
15	0	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0
18	0	20	40	35	40	30	25	20	35	30	35	0	20	10	30	15	35
19	0	15	0	0	0	0	10	0	0	0	0	0	5	5	5	0	0
20	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0

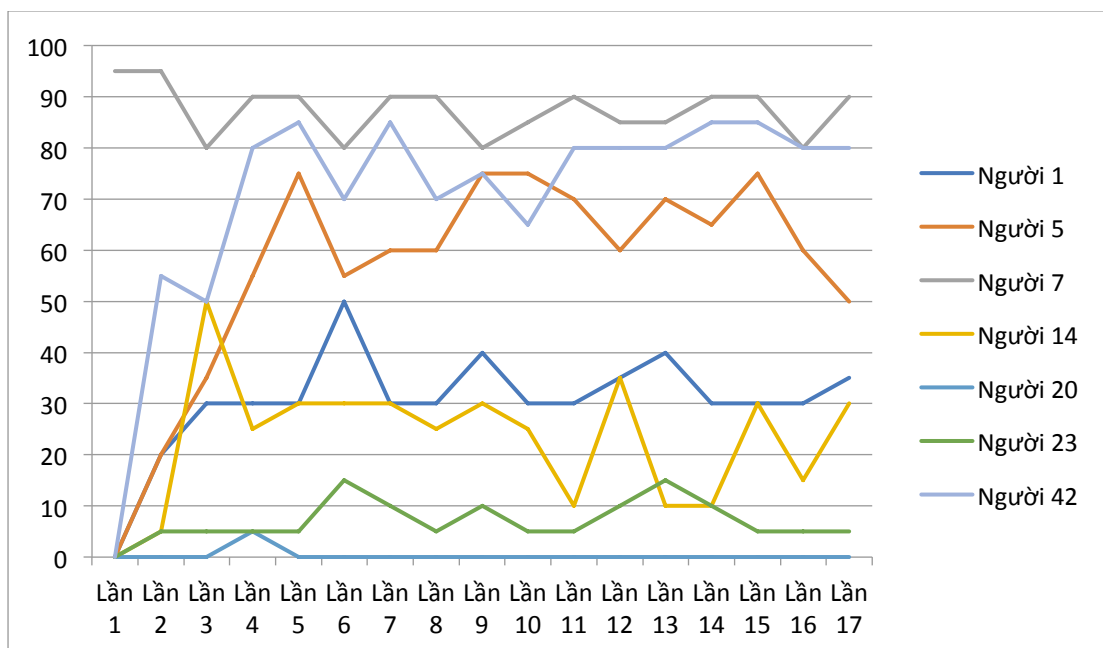
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0
22	0	0	0	5	0	0	0	0	5	5	0	5	5	15	15	0	15
23	0	5	5	5	5	15	10	5	10	5	5	10	15	10	5	5	5
24	0	0	0	0	0	0	0	0	0	0	0	0	5	15	5	0	0
25	0	0	0	0	0	0	0	5	5	0	5	0	0	0	0	5	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	5	0	0	0	0	0	0	0	5	10	0	0	0
28	0	85	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	10	0	10	0	5	5	5	10	0	5	5	5	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	10	10	10	5	5	10	5	15	15	5	15	10	0	0	10	20
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	25	10	10	10	10	10	10	10	15	75	10	0	50	0
38	0	5	0	10	0	0	0	0	0	0	0	0	0	0	0	5	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	20	30	40	30	25	35	25	25	40	20	25	25	35	30	30
41	0	0	10	15	10	20	15	15	15	15	15	15	15	20	15	20	15
42	0	55	50	80	85	70	85	70	75	65	80	80	80	85	85	80	80
43	5	5	5	15	10	5	15	10	20	20	10	10	20	15	10	5	15
44	0	0	5	10	10	15	10	15	5	5	10	5	5	5	10	15	10
45	0	0	5	10	10	10	10	10	10	10	10	10	10	10	10	10	15
46	0	0	10	10	10	10	10	10	10	10	10	15	10	10	10	10	10
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	5	15	5	15	5	5	5	5	5	5	10	5	5	10	15
49	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0

50	0	35	20	25	20	20	25	20	15	20	20	15	20	25	25	25	20
51	0	10	0	0	5	0	0	0	0	0	0	0	5	15	0	0	10
52	0	35	25	35	55	35	35	40	45	40	30	30	45	50	50	35	50
53	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0
54	0	0	0	15	15	0	0	0	0	0	5	15	0	0	10	5	5
55	100	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
56	0	0	0	0	0	0	5	5	0	0	0	0	5	0	5	0	0
57	0	30	15	10	10	15	10	10	15	5	15	5	15	5	5	20	5
58	0	0	5	5	5	5	5	10	5	5	0	10	0	5	5	5	5
59	0	0	0	0	0	0	5	5	5	5	0	5	5	5	0	0	0
60	0	0	15	25	25	35	20	20	20	35	25	45	25	25	25	20	30
61	0	10	10	10	5	20	20	10	20	20	10	20	20	15	15	10	20
62	0	0	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
63	0	10	0	0	5	5	5	0	5	5	0	0	0	5	5	0	5
64	0	15	30	35	30	25	35	25	35	35	30	35	30	35	30	25	35
65	15	100	35	30	25	25	25	25	25	30	25	25	25	25	25	25	25
Tổng				11	11	11	11	9.	11	11	11	10	12	11	11	10	11
cộng	3.3	10.2	8.9	.3	.5	.1	.1	9	.3	.2	.3	.7	.1	.1	.4	.8	.4

4.2 Biểu đồ tổng kết



Biểu đồ tổng trung bình phần trăm đúng của 65 người trong từng lần thử



Biểu đồ phần trăm đúng của 7 người đại diện qua từng lần thử

5 Đánh giá cá nhân

- Việc tăng số lượng Gauss của mỗi GMM, tăng số lượng vòng lặp Init không mang lại nhiều thay đổi trong kết quả nhận dạng.
- Việc tăng số vòng lặp huấn luyện mang lại ảnh hưởng lớn nhất.
- Tăng số lượng vòng lặp huấn luyện thì độ chính xác sẽ tăng lên tuy nhiên đến một giá trị cố định thì độ chính xác không được cải thiện nữa.
 - o Giá trị trung bình tổng phần trăm của 65 người cao nhất là 12.1 ở lần thử thứ 13 (số gauss: 16, số vòng lặp init: 100, số vòng lặp huấn luyện: 300)
 - o Giá trị trung bình tổng ở số vòng lặp huấn luyện là 1000, đạt được là 11.4%.
 - o Giá trị trung bình không dao động nhiều khi đạt đến vòng lặp huấn luyện là 1000, chỉ dao động trong khoảng 11%.