**XỬ LÝ NGÔN NGỮ NÓI**

Bài tập 1: Nhận dạng giọng nói

GVHD:

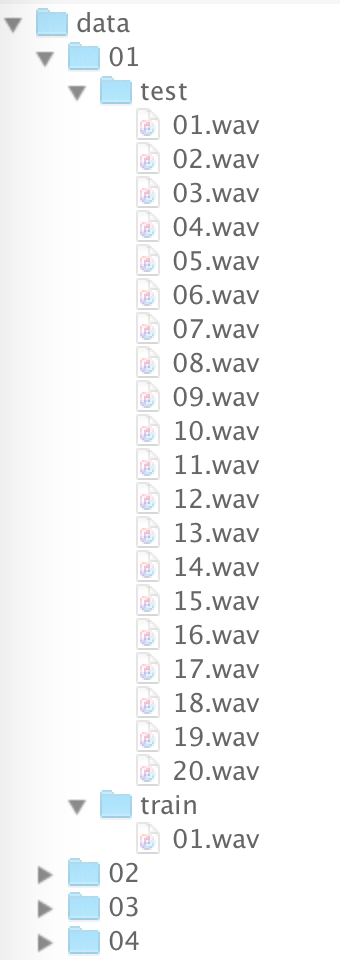
SV thực hiện: Trần Thuý Hiền – Trần Tấn Bảo

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.

1. **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”.
* Folder “train” lưu 1 file “01.wav” có chiều dài 3 phút dùng để huấn luyện.
* Folder “test” chứa 20 file test, mỗi file có chiều dài 3 giây.
* Cấu trúc thư mục như sau:

****

* Với mỗi người nói, ta sẽ đọc file huấn luyện và rút trích đặc trưng bố 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 đọi thì test *pass*, ngược lại là *fail*. Tính phần trăm kết quả đạt và không đạt, lặp bảng kết quả và vẽ biểu đồ.

1. **Kịch bản thử nghiệm**
2. **Kết quả, vẽ biểu đồ**

**Lần thử nghiệm 1:**

* Với số lượng Gauss: 16
* Số lượng vòng lặp Init: 100
* Số lượng vòng lặp huấn luyện: 100

Thu được kết quả như sau:

|  |  |
| --- | --- |
| Người nói | Tỉ lệ đúng từ 20 file test |
| 01 | 35% |
| 02 | 5% |
| 03 | 25% |
| 04 | 15% |
| 05 | 80% |
| 06 | 0% |
| 07 | 90% |
| 08 | 0% |
| 09 | 30% |
| 10 | 0% |
| 11 | 0% |
| 12 | 0% |
| 13 | 10% |
| 14 | 30% |
| 15 | 30% |
| 16 | 15% |
| 17 | 10% |
| 18 | 40% |
| 19 | 35% |
| 20 | 0% |

1. **Đánh giá cá nhân**

* Việc tăng số lượng Gauss của mỗi GMM, tăng số lượng vòng lặp Init, tăng số lượng vòng lặp huấn luyện không mang lại nhiều thay đổi trong kết quả nhận dạng.