

# Speaker Identification



Yiyin Zhou

# Background

- ❖ **Speech signal conveys message information and speaker identity**
- ❖ **Speech Recognition**
  - Hope to emphasize on speaker-independent information
- ❖ **Speaker Recognition**
  - Emphasize on the difference between speakers
  - A discrimination problem

# Background

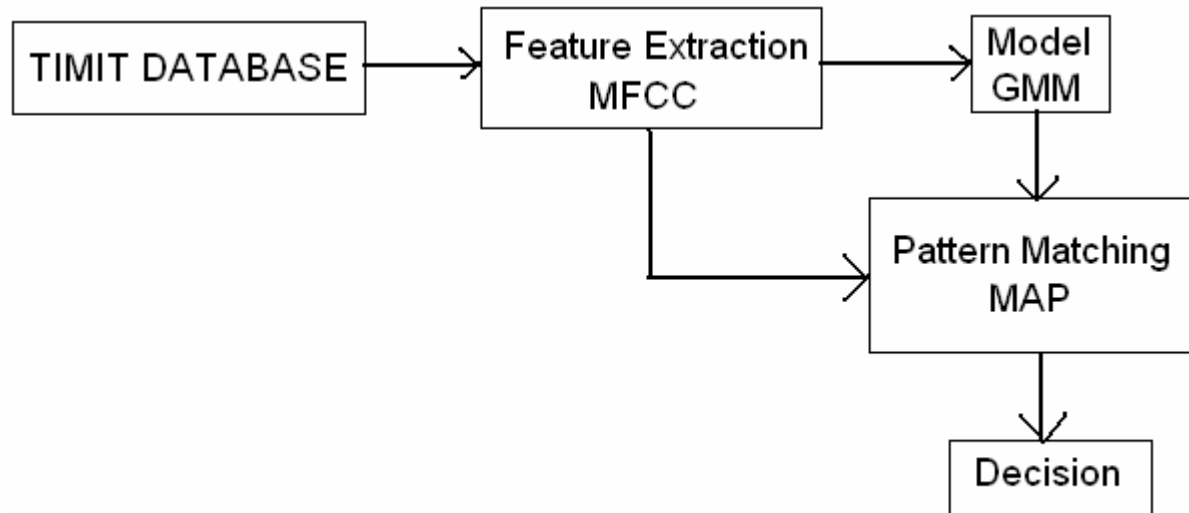
## ❖ Speaker Recognition

- Speaker Verification or **Speaker Identification**
- Text-dependent or **Text-independent**

## ❖ **Apply to daily life**

- Embedded in telephone
- TV or movie subtitle
- Meeting transcript
- When using along with ASR

# System

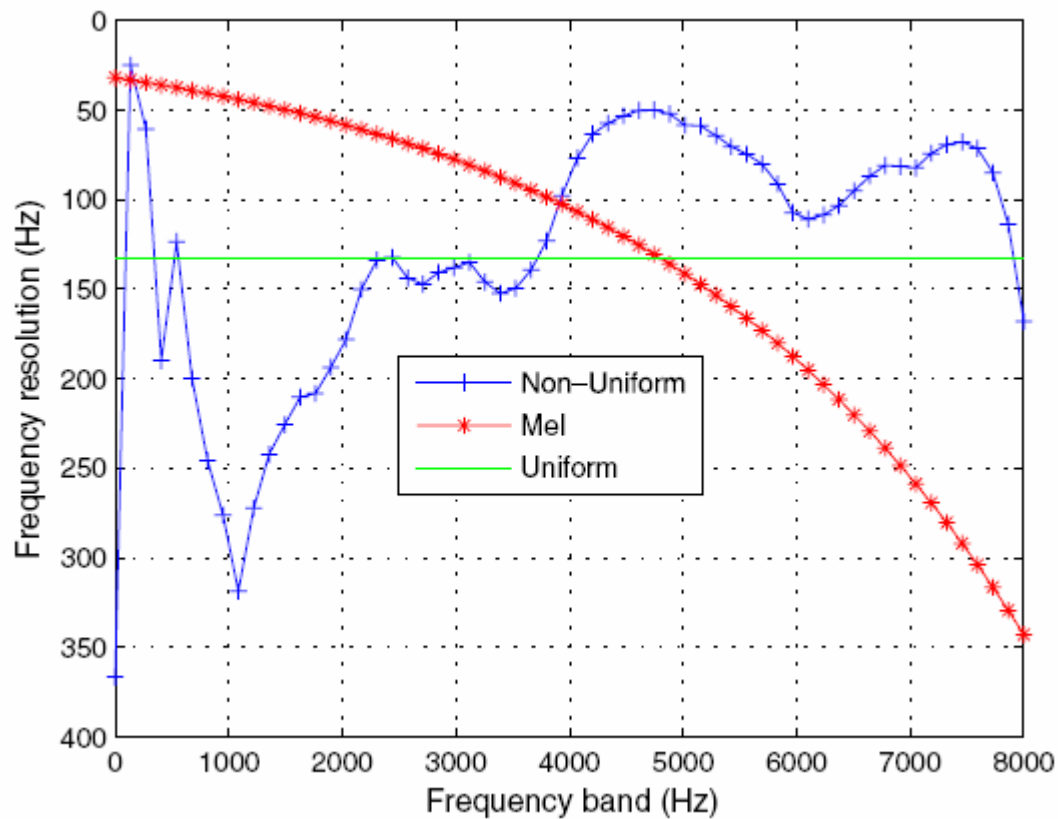


System diagram

# Feature

## ❖ MFCC

- Mel frequency cepstral coefficient



## ❖ Gaussian Mixture Model

- Weighted sum of M Gaussian component density

## ❖ ML estimation using EM algorithm

- ❖ Choose the one with maximum *a posteriori* probability given test data

# System performance

- ❖ **Number of components:  $M$**
- ❖ **Total speakers in the database**
- ❖ **Testing speech length**
- ❖ **Choose a good initial model for EM**

# Goals

## ❖ **Basic Goals**

- Implement a speaker identification with 10 to 20 persons enrollment.

## ❖ **Further Goals**

- Solve several of the following problems
  - Real time recognition
  - Abnormal speaking (laughing, mimic sb. else)
  - Double talking problems

# Reference

- Campbell, J. 1997. Speaker recognition: A tutorial. Proc. IEEE 85(9), 1437-1462
- Lu, X., Dang, J., 2007. An investigation of dependencies between frequency components and speaker characteristics for text-independent speech identification, Speech Comm. 2007
- Reynolds, D.A., 1995. Robust Text-independent speaker identification using Gaussian Mixture Speaker Models, IEEE Trans. Speech Audio Processing, vol.3, no.1, pp72-83
- Gish, H., Schmidt, M., 1994. Text-independent speaker identification. IEEE Signal Processing Mag., vol.11, pp18-32
- Gold, B., Morgan, N., 1999. Speech and audio signal processing.

**Thank You !**

