9. Speaker Recognition

See also:

Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models

Douglas A. Reynolds, Member, IEEE, and Richard C. Rose, Member, IEEE
The Tasks

Speaker verification: is this Mary?

Speaker identification: who is speaking?
Issues

• Text dependent vs. text independent
• Amount of training data?
• Mismatch
  • Channel, noise
  • Cold, mood!
Speaker Verification

Feature Extraction (MFCC)

Speaker Models
(Gaussian Mixture Models)

Yes/no

Classification

\[
P(x \mid \text{speaker}) > \lambda
\]

P(x \mid \text{background})

Adaptation

Yes/no
Speaker Identification

Feature Extraction (MFCC)

Speaker Models (Gaussian Mixture Models)

\[ \hat{I} = \arg \max_{i=1 \ldots M} P(x_1 \ldots x_N | i) \]

\[ P(x_1 | \theta = 1) \]
\[ \vdots \]
\[ P(x_1 | \theta = M) \]
Training the Model

Generic GMM

Adapt generic GMM using speaker specific data (e.g. MLLR)
Effect of Test Data

The longer the test utterance the better

(Number of Gaussians)
Effect of Training Data

The more training data the better

The more training data, the more Gaussians you can afford
NIST - Speaker Recognition Evaluations

General Information

Introduction

NIST has been coordinating Speaker Recognition Evaluations since 1996. Each evaluation begins with the announcement of the official evaluation plan which clearly states the rules and tasks involved with the evaluation. The evaluation culminates with a follow-up workshop, where NIST reports the official results and researchers share in their findings.

Brief History

Since 1996, over 40 research sites have participated in our evaluations. Each year, new researchers in industry and universities are encouraged to participate. Collaboration between universities and industries, is also welcomed.

The overall goals of the evaluations have always been to drive the technology forward, to measure the state-of-the-art, and to find the most promising algorithmic approaches.
Best System Performance 1996-2003