

# 平成 18 年度 情報工学コース卒業研究報告要旨

大西 研究室	氏 名	木 下 裕 司
卒業研究題目	テキストに依存しない話者照合システムの研究	

## 背景と目的

近年、携帯電話やコンピュータによる音声通信が盛んに行われている。そこでは金融を扱った取引などもあり、他人の不正利用に対してのセキュリティが必要となってくる。こうした他人の不正利用を防ぐために使用者が本人であることの確保することが重要である。そこであらかじめある音声インターフェースを使った話者認識システムが最も効率が良いと考えられる。本研究ではテキストに依存しない話者照合システムを検討することで、発話内容に依存しない逐次的な照合を行うことを目標とする。

## システムの概要

システムの処理の手順を図 1 に示す。入力音声から特徴量を抽出してあらかじめ学習された話者モデルとの尤度を計算し、閾値によって判定を行う。特徴量は MFCC(メルケプストラム)12次元とそのケプストラム 12次元と対数パワーをあわせた 25次元を用いる。MFCC は人間の聴覚の特性を考慮した特徴量で音声認識などでもよく利用される頑強な特徴量である。話者モデルでは従来から話者識別システムで用いられる GMM(混合ガウス分布話者モデル)を用いる。GMM は話者の音響的特徴をうまく捉えることができるため、頑強な話者モデルであることが知られている。GMM の学習では EM アルゴリズムを用いた。パラメータを繰り返し更新することにより、話者の特徴量の分布を混合ガウス分布で表した。判定において話者の尤度のばらつきに対応するために判定時間の尤度に基づいた判定率を使うことで、安定したシステムをめざした。

## 実験と結果

実験では学習サンプルとして約 90 秒の音声、テストサンプルとして本人約 50 秒の音声と他人 9 名のそれぞれ約 20 秒の音声を用いた。学習サンプルを用いて GMM 話者モデル  $\lambda$  を作成し、その話者モデルに対しての各テストサンプルの尤度を求めた。その中で尤度の低かった話者のテストサンプルを用いて GMM 話者モデル  $\lambda_0$  を作成した。判定では対数尤度比  $\log L(x) = \log p(\vec{x}|\lambda) - \log p(\vec{x}|\lambda_0)$  の分布を用いた。実験結果は図 2 のようになり、横軸は判定のために使用した時間、縦軸はその率 (%) である。テストサンプルから判定時間ごとにデータ 100 個、本人のサンプルデータ 50 個を用いて判定率を閾値に対して照合を行った。およそ 1 秒間のデータをまとめた判定率を用いた場合には本人認識率はほぼ 100%、他人受率率は 0% という結果となった。

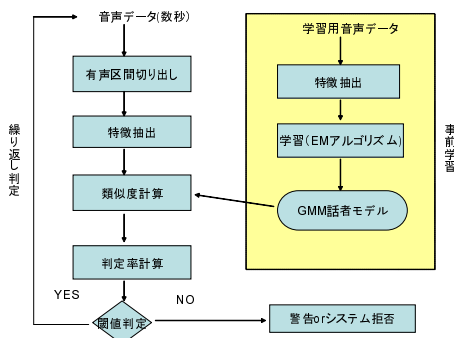


図 1. システムの概要

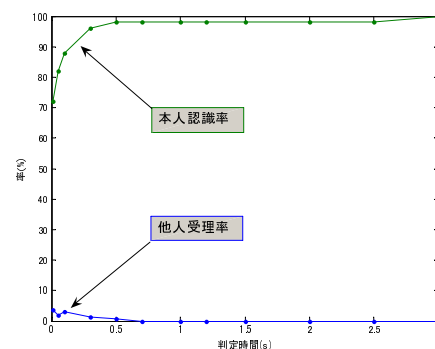


図 2. 判定時間に対する率