

## 気導音声と骨導音声のスペクトル比を用いた機械学習による話者識別の検討

福永 琢人・森 幹男 (福井大学)

## 1. はじめに

本研究室では、気導音声と骨導音声のスペクトル比に、頭頸部の形状・組織等に起因する話者固有の特徴が含まれていると考え、骨導音声と気導音声のスペクトル比を特徴量とした話者識別システムを提案している[1]。話者識別実験の結果、風邪をひいた話者の識別が可能なが示唆される結果が得られた一方で、訓練に用いたデータの不足から高い識別率は得られなかった[2]。

そこで本研究では音声データを追加し、機械学習を用いた場合との話者識別精度の比較を行う。

## 2. 実験用音声

音声収録は21~25歳の男性10名に対して研究室で行った。このとき、PC用マイクと咽喉上部に装着した咽喉マイクを用いてステレオ録音することによって、気導音と骨導音を同時収録した。ここで、サンプリング周波数は22.05 kHzとした。そして日本語5母音発声時の音声収録を、各被験者に対して日を変えて13回行うこととした。

## 3. 特徴量の抽出

同時収録した気導音声と骨導音声の振幅スペクトルに対してそれぞれメルフィルタによる重みづけ、対数化、最大値による規格化を行い  $B_m(\omega)$ ,  $C_m(\omega)$  を求める。そして、それらの差に対して離散コサイン変換(DCT)を行うことで特徴量  $A_m(\tau)$  を得る。図1に特徴量抽出フローを示す。なお、メルフィルタバンクの上限周波数は5200 Hz, 下限周波数は100 Hzとして、メルスケールで50等分割している。

## 4. 話者識別実験

各話者が日本語5母音を13回ずつ発声した音声データのうち、9回分を訓練用データ、4回分をテストデータとした。このとき、訓練データは13回分のデータから9回分を無作為に選び、残り4回分をテストデータとした。この操作を5回行い、従来法と提案法(機械学習)による話者識別実験を行う。

## 4.1 テンプレートマッチング(従来法)

まず、特徴量  $A_m(\tau)$  を求め、得られた9個の50次元ベクトルの要素ごとの平均ベクトルを求め、各話者の母音ごとのテンプレートを作成した。そして作成したテンプレートとテストデータとのユークリッド距離を算出し、同一母音で最小の話者に得点を付与する。それを5母音で合計したスコアが最も高い話者を識別結果とした。

## 4.2 機械学習(提案法)

まず、特徴量  $A_m(\tau)$  を求め、各母音ごとに Python ライブラリ `sklearn` のサポートベクターマシン(SVM)で学習を行った。この時、SVMのカーネル関数は放射基底関数(RBF)とした。その後、テストデータで各母音ごとに識別を行い、予測された話者に点数を付与する。

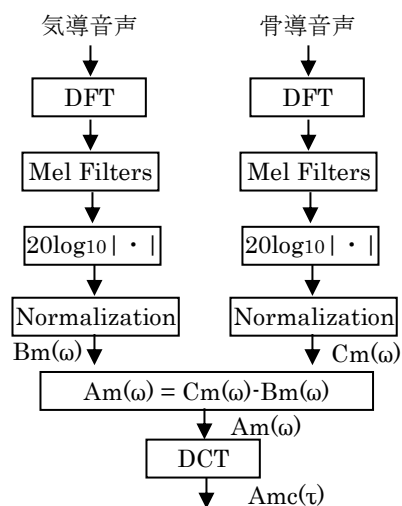


図1 特徴量抽出フロー

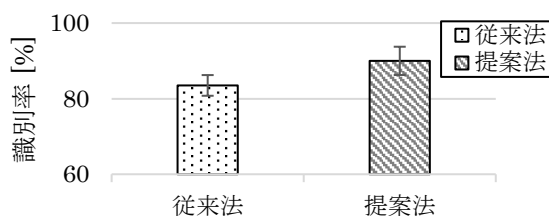


図2 従来法と提案法との平均識別率の比較

それを5母音で合計したスコアが最も高い話者を最終的な識別結果とした。

## 5. 実験結果

5回の話者識別結果の平均識別率を従来法と提案法とで比較した結果を図2に示す。従来法では識別率が83%となったが、機械学習を用いた提案法の場合、識別率は90%となった。

## 6. まとめと今後の予定

音声データを新たに追加し、機械学習を用いた場合との話者識別精度の比較を行った結果、機械学習を用いることでより精度の高い識別結果が得られることがわかった。

今後は、機械学習における細かなパラメータの調整や特徴量の再検討を行い話者識別精度の向上を目指す。また、風邪をひいた話者に対する話者識別実験を行い、音声変動耐性についても調べる予定である。

## 参考文献

- [1] 森, 竹田, 萩原, 谷口, 高橋: 「気導音声と骨導音声のスペクトル比を用いた話者識別の検討」 電学論 C, 127 巻 3 号 pp.456-457 (2007)  
 [2] 福永, 森: 「咽喉マイクを用いた個人認証の検討」 電気・情報関係学会北陸支部連合大会講論集 (2020)