

機関番号：82505

研究種目：若手研究（B）

研究期間：2008～2010

課題番号：20700177

研究課題名（和文） 声道共鳴特性を用いた話者の地域性情報とクラスタリングに関する研究

研究課題名（英文） Residence information of speakers and clustering using vocal tract resonance

研究代表者

鎌田 敏明（KAMADA TOSHIAKI）

警察庁科学警察研究所・法科学第四部・主任研究官

研究者番号：10356173

研究成果の概要（和文）：本研究では、声道共鳴特性に由来する音響特徴量を用いた話者の地域性情報の抽出及び話者の地域分類について検討を行った。ケプストラムによって得られた話者間距離を地域性情報とした特徴量抽出を行い、話者分布によりその有効性及び話者の地域属性による分類の可能性を示した。

研究成果の概要（英文）：This study aimed to extract of residence information of speakers using speech features originating in vocal tract resonance, and local classification of speakers. We extracted the features which made residence information distance between speakers acquired by cepstrum, and speaker distribution showed the possibility of the classification by residence information of speakers.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	1,200,000	0	1,200,000
2009年度	900,000	0	900,000
2010年度	1,100,000	0	1,100,000
総計	3,200,000	0	3,200,000

研究分野：総合領域

科研費の分科・細目：情報学、知覚情報処理・知能ロボティクス

キーワード：音声情報処理・話者認識・法科学

1. 研究開始当初の背景

本研究において話者の地域性情報の観点に着目した背景は、音声認識や話者認識を行う上での話者モデルを、外国語話者と日本語話者を区別して構築することは一般的に行われているにもかかわらず、日本語話者モデルを構築する上では、男女の性別や大人と子供の区別を行うことはあっても、話者の居住地ごとに話者モデル構築するような研究はほとんど行われていないという点である。日本語話者の母音「ア」と米語話者の母音「a」を区別するために異なった話者モデルを構築するというのであれば、大阪居住者の母音「ア」と福島居住者の母音「ア」を区別するための異なる話者モデルを構築すべきで

あると考えられる。ある居住地の母音「イ」は「エ」に似ているが、他の居住地の母音「イ」は「ウ」に似ているといったような、居住地による日本語5母音の母音図のずれというものも存在し、それは方言を構成する要素の1つである。

これまでこのような信号処理技術による居住者ごとの話者モデルや地域性情報について十分な検討が行われてこなかったのは、研究で利用している音声コーパス（データベース）にあると考えられる。一般的に行われている研究で利用されている日本語音声コーパスにおいて、47都道府県を網羅した地域性情報が付加された大規模音声コーパスは市販されていない。一方我々が利用している

音声コーパスは、47 都道府県の居住者それぞれ約 60 名から構成される 3000 人のデータベースであり、地域性情報抽出の実験を行う環境が整備されている。

音声学の分野では方言研究が行われているものの、音声音響工学の分野では方言分離などの研究が行われていることは少ない。そこで本研究では、信号処理技術による方言の一側面であると考えられる話者の地域性情報について検討を行う。

2. 研究の目的

音声による話者認識では、音声中に含まれる個人性情報は声道の伝達特性に由来するスペクトル包絡から得られるケプストラムに基づく研究が行われてきた。また、このような信号処理手法による話者認識において、日本語を母語とする日本人の話者認識を行う研究では、日本人モデルを構築した上での理論及び実験的検討が行われてきた。

しかし本来我々が話者、特に知らない話者の音声を聞いて話者を判断しようとする場合、その話者が自分の知っている誰かに似ていると判断する前に、その話者の居住地を判断する試みをすることがある。例えば、東京居住者が大阪居住者の音声を聞いて大阪弁だと思ふことは頻繁に起こりえることである。

そこで本研究では、話者認識において、ケプストラムの特徴量から話者の居住地として地域性情報を抽出すること、及び得られた地域性情報による話者のクラスタリングを目的とする。

3. 研究の方法

音声コーパスからの特徴量の抽出として、電話を通して録音された約 3000 名の男性話者音声コーパスを利用した実験を行い、地域性情報の抽出に関する実験を行った。利用するケプストラムは LPC ケプストラム係数、メル周波数ケプストラム係数などやそのデルタケプストラムである。得られたケプストラム空間から、2 話者における話者間距離を算出し、これらを実験で用いる特徴量とした。

得られた話者間距離から地域属性ごとの話者分布を求め、地域性情報抽出の有効性及び話者分類の可能性について検討した。

4. 研究成果

(1) 音声データと特徴量抽出

照合実験に使用した音声データは、3000 人程度の音声データベースから抽出された音声である。3000 人の音声データベースは 47 都道府県全ての居住者を網羅しており、各都道府県でそれぞれ約 60 名ずつから構成されている。実験で使用した話者は、複数の発声内容において 1 人当たり 9 回の発声が全て存在するデータを抽出している。実験に使用し

た音声データを表 1 に示す。電話音声であることから、11.025 kHz から 8 kHz にダウンサンプリングを行った。

表 1 音声資料

話者数	3000 人
発声時期差, 時期	3~4 ヶ月, 3 時期
1 時期の発声回数	3 回
発声内容	6 単語 (はい, 車, 連絡, 爆弾, 銀行, 警察)
サンプリング周波数	8 kHz
量子化精度	16 bit

表 2 分析条件

分析窓	ハミング窓
フレーム長	32 ms
フレーム周期	16 ms
高域強調	1 次の適応型
分析方法	LPC 分析
分析次数	12 次
特徴量	LPC ケプストラム係数 Δ LPC ケプストラム係数

表 2 に示す分析を行い、DP マッチングから得られた DP 距離による話者内、話者間距離をもって、地域性情報が包括されている特徴量とした。

(2) 居住地ごとの距離分布

音声データの話者の居住地を限定した話者間距離の分布についての検討を行った。使用した音声データは 3000 人のデータベースから東京都、大阪府、青森県及び福島県在住の話者それぞれ 50 人を抽出した。なお音声データには話者に対する居住地情報の他に、自己申告ではあるが出身地情報も存在する。今回の実験では話者の出身地については考慮していない。また使用した単語は「はい」である。

DP マッチングによる話者照合過程から、同一人及び別人の距離分布を求めた。同一人の話者間距離 (話者内距離) は、異なる時期の発声全ての組み合わせから得た。別人の話者間距離は居住地の話者ごとに分布を求めた。

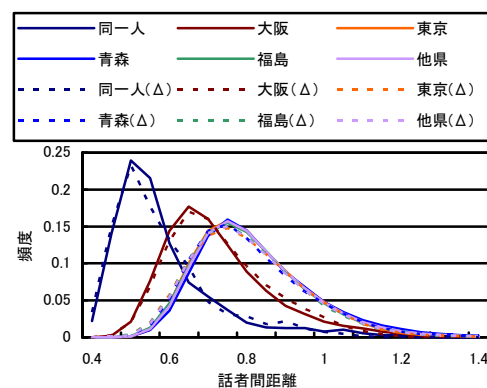


図 1 居住地ごとの話者間距離分布

大阪府在住の話者から見た、LPCC 及び Δ LPCC を特微量とした同一人及び居住地ごとの距離分布を図 1 に示す。明らかにわかるのは、LPCC、 Δ LPCC の場合も、同一地域居住の別人の分布と別地域居住の別人の分布が大きく異なっている点である。LPCC 及び Δ LPCC の分布の差はわずかであるが、その違いは確認できる。

(3) 大規模データベースによる距離分布

前述と実験条件を同一とし、東京及び大阪居住という特定の地域における、居住地を除くすべての都道府県ごとの話者間距離を求め、距離分布による実験を行った。

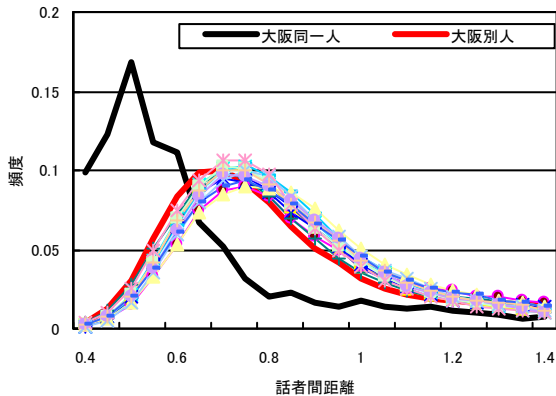


図 2 大阪居住者からみた距離分布

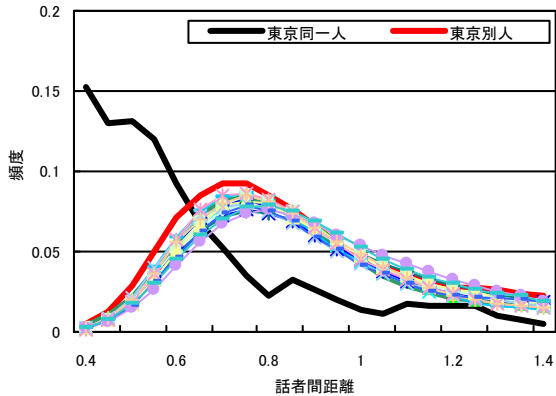


図 3 東京居住者からみた距離分布

大阪在住者からみた距離分布を図 2、東京在住者からみた距離分布を図 3 に示す。同一地域及び異なる地域での別人による話者分布には大きく差があることが示された。既存のデータベースと比較することによって居住地、架電地域の推定、分類ができる可能性があることが示された。

(4) 母音間距離による地域分類

使用した音声データは、3000 人の電話データベースから抽出した東京都、神奈川県、山形県及び福島県在住の話者それぞれ 50 人の音声である。同一時期に発声した 3 回の時間平均 LPC ケプストラム係数を母音ごとに算出し、得られた特微量の母音間距離によって評価を行った。同一人の母音間距離は同一時期の発声データを用い、別人の母音間距離は

全ての時期の発声データを用いた。

母音 [a] 及び母音 [i] の母音間距離を同一人内で算出し、居住地ごとの距離分布結果を図 4 に示す。

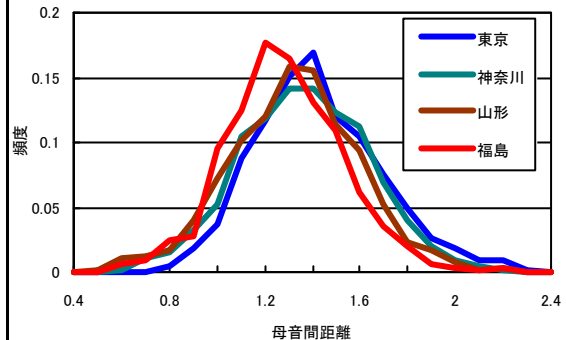


図 4 [i]-[e] による母音間距離分布

福島の居住者分布が、他の 3 地域における距離分布よりも小さいところにあることが示された。これは福島県居住者の [i] 及び [e] の発声が他の地域よりも似ている母音であるということを示しており、特微量からの居住地あるいは出身地推定の可能性が示された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 12 件)

- ① 蒔苗久則、鎌田敏明、長内隆、「音声明瞭化ソフトウェアの開発」、日本法科学技術学会、2010年11月12日、ホテルフロラシオン青山 (東京都)
- ② 長内隆、鎌田敏明、蒔苗久則、網野加苗、「テキスト独立型話者照合における発話内容の共通性の評価」、日本法科学技術学会、2010年11月12日、ホテルフロラシオン青山 (東京都)
- ③ 鎌田敏明、長内隆、蒔苗久則、網野加苗、「母音間距離を利用した話者の地域性情報に関する検討」、日本法科学技術学会、2010年11月12日、ホテルフロラシオン青山 (東京都)
- ④ 長内隆、鎌田敏明、蒔苗久則、網野加苗、「テキスト独立型話者照合における発話内容の共通性に関する検討」、日本音響学会秋季研究発表会、2010年9月15日、関西大学 (大阪府)
- ⑤ 蒔苗久則、鎌田敏明、長内隆、「周波数帯域強調 MFCC による話者認識実験」、日本法科学技術学会、2009年11月12日、ホテルフロラシオン青山 (東京都)
- ⑥ 四宮康治、蒔苗久則、鎌田敏明、長内隆、「尤度比を用いた話者照合法」、日本法科学技術学会、2009年11月12日、ホテルフロラシオン青山 (東京都)
- ⑦ 鎌田敏明、長内隆、蒔苗久則、「話者認識における話者の地域性情報の抽出に関する

る検討」、日本法科学技術学会、2009年11月12日、ホテルフロラシオン青山（東京都）

⑧ 長内隆、鎌田敏明、蒔苗久則、「テキスト独立型話者照合における話者モデルとその学習に用いるデータ量に関する考察」、日本法科学技術学会、2009年11月12日、ホテルフロラシオン青山（東京都）

⑨ 四宮康治、蒔苗久則、鎌田敏明、長内隆、「テキスト依存型話者照合を用いた法科学的検査法の検討」、日本音響学会秋季研究発表会、2009年9月17日、日本大学(福島県)

⑩ 長内隆、鎌田敏明、蒔苗久則、「骨導音によるテキスト依存型話者照合におけるスペクトルのローカルピークの利用」、日本法科学技術学会、2008年11月6日、ホテルフロラシオン青山（東京都）

⑪ 蒔苗久則、鎌田敏明、長内隆、「電話音声によるテキスト依存型話者照合におけるMFCCの利用」、日本法科学技術学会、2008年11月6日、ホテルフロラシオン青山（東京都）

⑫ 鎌田敏明、長内隆、蒔苗久則、「雑音環境下におけるデジタル圧縮技術が話者照合に及ぼす影響に関する検討」、日本法科学技術学会、2008年11月6日、ホテルフロラシオン青山（東京都）

〔図書〕（計2件）

① Kanae Amino, Takashi Osanai, Toshiaki Kamada, Hisanori Makinae, and Takayuki Arai, "Historical and procedural overview of forensic speaker recognition." , in Forensic Speaker Recognition: Law Enforcement and Counter-Terrorism, pp.3-20, Springer-Verlag, 2011.

② Kanae Amino, Takashi Osanai, Toshiaki Kamada, Hisanori Makinae, and Takayuki Arai, " Effects of the phonological contents and transmission channels on forensic speaker recognition." , in Forensic Speaker Recognition: Law Enforcement and Counter-Terrorism, pp.275-308, Springer-Verlag, 2011.

6. 研究組織

(1) 研究代表者

鎌田 敏明 (KAMADA TOSHIAKI)

警察庁科学警察研究所・法科学第四部・主任研究官

研究者番号：10356173

(2) 研究分担者

なし

(3) 連携研究者

なし