

研究種目：若手研究 (B)

研究期間：2007～2010

課題番号：19700172

研究課題名 (和文)

長・短期間の音声変動の解明および音声変動に頑健な音声認識手法に関する研究

研究課題名 (英文)

A study of robust speech recognition method for intra-speakers' speech variability

研究代表者 徳島大学・大学院ソシオテクノサイエンス研究部・講師

柘植 覚 (Satoru Tsuge)

研究者番号：00325250

研究分野：総合領域

科研費の分科・細目：情報学・知識情報処理・知能ロボティクス

キーワード：音声認識、特定話者、話者内音声変動

1. 研究計画の概要

本研究は、同一発声者が特定の音声認識（話者依存音声認識）システムを使用した場合においても必ず毎回同一の音声認識結果とならないという問題を取り扱うものである。この個々人が同一内容を発声したとしても、気分や発声環境により変動（音声変動）する原因を解明し、音声認識精度の変動を抑制することが本研究の目的である。

2. 研究の進捗状況

本研究は過去三年間（平成 19 年から 21 年）において以下の事を行った。

●特定話者の音声変動の解明および音声変動と音声認識精度の関連

発声時期や発声時間帯が異なることによる発声変動を明確することを目的に以下の事を調査検討した。

- 女性話者 1 名、男性話者 1 名の 1 年分の音声データを用いた連続音声認識を行い、発声時間、発声時間帯による音声認識精度の変動を調べた。
- 認識結果とアンケート情報との相関関係を調べ、特定話者の音声変動が起りやすい要因を調査した。
- 特定話者の連続音声認識ではつかみきれない音声変動を孤立単語音声認識を行うことにより詳細な音声変動の分析を行った。

●発声変動に頑健な音声認識手法の確立

音声認識手法の確立のため、以下の手法を提案、検討した。

- 音響モデル選択による発話変動に頑健な音声認識手法の検討
- 発話時期、時間帯に適応した音響モデル学習による頑健性の向上。

●長期間特定話者音声データベースの構築

データベース処理に関しては、音声データ収集の際には発声者の負担軽減を目的に一発話ずつ区切って録音せず、全発話セットを区切らず録音し 1 ファイルとして保存している。そのため、各発話の切り出し作業が必要となるが、その作業は人手で行った場合には多大の労力と時間が必要となる。そこで、自動的にこれらの発話切り出し作業を行う手法を提案し、ツールとして作成した。このツールを使用し、録音データのデータベース化を 3 年間で行った。

また、録音時に行ったアンケート結果もデジタル化を行った。

3. 現在までの達成度

現在までに長短時間音声データベースに関しては録音が大量に行え、これは当初の予定通りに達成ができています。

また、音声変動の分析を行う上では連続音声認識結果との関連性のみならず、孤立単語音声認識結果を行った点においては計画より達成ができていていると思われるが、因子分解などの統計的手法による検討が

今年度の課題である。

頑健な認識手法の確立に関しては、音声変動に追従する発声時間帯の音響モデルの選択などで行えることが解っている点より、達成を行えていると思われる。

4. 今後の研究の推進方策

科学研究費補助金 若手研究(B)の補助を頂き、研究する最終年となる平成22年度は、変動に頑健音声認識手法の確立を目指し検討を行う予定である。現在までは、おもに音響情報に対し適応などを行っていたが、辞書適応などの言語面の制約を行う予定である。

さらに、まだファイリングなどの音声データベース化が行えていないファイルに対しデータベース化を行う予定である。

5. 代表的な研究成果

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計0件)

[学会発表] (計13件)

- ① Satoru Tsuge, Keiji Seida, Masami Shishibori, Kenji Kita, Fuji Ren, Minoru Fukumi, and Shingo Kuroiwa, Analysis of Variation on Intra-Speakers Speech Recognition Performances, Proc. of IEEE NLP-KE2007, pp. 387-392, Oct. 2007, Beijing China.
- ② Mitsuhiro Ozawa, Satoru Tsuge, Masami Shishibori, Kenji Kita, Fuji Ren, Minoru Fukumi, and Shingo Kuroiwa, Automatic Utterance Segmentation Tool for Speech Corpus, Proc. of IEEE NLP-KE2007, pp. 401-406, 2007, Beijing China.
- ③ Katsuyuki Oobayashi, Satoru Tsuge, Minoru Fukumi, Seiji Tsuchiya, Ryosuke Sumitomo, Fuji Ren, and Shingo Kuroiwa, A Study of speaker identification using phoneme-information, Proc. of NCSP' 2008, pp. 164-167, March 2008, Gold Coast, Australia.
- ④ 柘植 覚, 黒岩 眞吾, 喜多 雅彦, 蒔苗 久則, 長内 隆, 鎌田 敏明, 谷本 益巳, 福見 稔, 任 福継, 日本語大規模話者認識用データベースを用いた話者識別実験, 日本音響学会 春季研究発表会, pp. 97-98, 2007年9月, 山梨.
- ⑤ 喜多 雅彦, 柘植 覚, 黒岩 眞吾, 任 福継, 大規模話者骨導音声データベースを用いたテキスト独立型話者照合実験, 情報処理学会音声言語情報処理, pp. 7-12, 2007, 京都.
- ⑥ 柘植 覚, 喜多 雅彦, 長内 隆, 蒔苗 久則, 鎌田 敏明, 谷本 益巳, 福見 稔, 任

福継, 黒岩 眞吾, 骨導音と気導音を併用した話者識別手法, 日本音響学会 秋季研究発表会, pp. 205-206, 2008, 千葉

- ⑦ Satoru Tsuge, Osanai Takashi, Makinae Hisanori, Kamada Toshiaki, Minoru Fukumi and Shingo Kuroiwa, Combination method of Bone-conduction Speech and Air-conduction Speech for Speaker Recognition, Proc. of Interspeech 2008, pp. 1929-1932, 2008, Brisbane Australia.
 - ⑧ Satoru Tsuge, Minoru Fukumi and Shingo Kuroiwa, Specific speakers' speech corpus over long and short time period, Proc. of oriental COCOSA, pp. 45-48, 2008, Kyoto Japan.
 - ⑨ Yasunori Kashiwara, Satoru Tsuge, Kipsang H. Choge, Tadahiro Oyama, Minoru Fukumi and Shingo Kuroiwa, Non-Stationary Noise Robust Speech Recognition Method using Repetitive Phrase, Proc. of NCSP' 2009, pp. 221-224, 2009, Hawaii U.S.A.
 - ⑩ 柏原康徳, 柘植 覚, 尾山匡浩, Choge Hillary Kipsang, 福見 稔, 黒岩眞吾, 繰り返し発声を用いた突発性雑音に頑健な音声認識に関する研究, 平成21年度電気学会電子情報システム部門大会, pp. 710-712, 2009, 徳島.
 - ⑪ 中尾 昌史, 柘植 覚, 福見 稔, 気導音と骨導音を併用した話者ベクトルを用いた話者識別, 平成21年度電気学会電子情報システム部門大会, pp. 720-723, 2009, 徳島.
 - ⑫ Satoru Tsuge, Daichi Koizumi, Minoru Fukumi, and Shingo Kuroiwa, Speaker verification method using bone-conduction and air-conduction speech, Proc. of ISPACS 2009, pp. 449-452, 2009, Kanazawa Japan.
 - ⑬ Haruka Okamoto, Amira Abdelwahab, Masahumi Nishida, Satoru Tsuge, Yasuo Horiuchi and Shingo Kuroiwa, Text-Independent Speaker Verification Using Rank Threshold in Large Number of Speaker Models, Proc. of InterSpeech 2009, pp. 2367-2370, 2009, Brighton U.K..
- [図書] (計0件)
[産業財産権]
○出願状況 (計0件)
○取得状況 (計0件)
[その他]