

科学研究費補助金研究成果報告書

平成22年 4月15日現在

研究種目：若手研究（B）

研究期間：2007～2009

課題番号：19700159

研究課題名（和文） 気導音声と骨導音声のスペクトル比を用いた新しい話者識別システム

研究課題名（英文） Speaker Identification System Using Spectral Ratio between Air-Transmitted Speech Sounds and Bone-Transmitted Speech Sounds

研究代表者

森 幹男 (MORI MIKIO)

福井大学・大学院工学研究科・助教

研究者番号：70313731

研究成果の概要(和文):本研究は、ユビキタス社会におけるセキュリティー技術の1つとして、新しい話者識別システムを提案し、構築することを目的とする。骨導音声は発声時に骨を介して伝わる音で、周囲騒音の影響を受けにくいという特徴がある。また、気導音声と骨導音声のスペクトル比は、骨導音声から気導音声への変換モデルの伝達特性を表している。そこで、(MFCCを用いた)このスペクトル比を特徴量とすることによって、話者識別システムのさらなる識別率向上と雑音耐性改善を図る。

研究成果の概要(英文): A bone-conductive microphone which transforms bone-transmitted speech sounds through skull into electronic signals has two advantages. One is to operate in hands free mode, and another is to be anti-noise. Therefore, it is expected to match the man-machine interface for wearable computers.

Here, we examined and analyzed the possibility of a speaker identification system with the MFCC of air and bone-transmitted speech sounds to 5 Japanese vowels. This paper shows that the proposed system using spectral ratio between air-transmitted speech sounds and bone-transmitted speech sounds gives better performance than the conventional systems using only air-transmitted speech sounds.

交付決定額

(金額単位：円)

| | 直接経費 | 間接経費 | 合計 |
|--------|-----------|---------|-----------|
| 2007年度 | 1,500,000 | 0 | 1,500,000 |
| 2008年度 | 300,000 | 90,000 | 390,000 |
| 2009年度 | 300,000 | 90,000 | 390,000 |
| 年度 | | | |
| 年度 | | | |
| 総計 | 2,100,000 | 180,000 | 2,280,000 |

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：骨導音声，話者識別，セキュリティー，ユビキタス，雑音耐性

1. 研究開始当初の背景

話者識別・話者照合システムについては従

来から数多くの研究・開発がなされてきているが、識別精度や雑音耐性が不十分のため、今なお実用化が容易でないという問題がある。本研究では、このような話者識別の精度向上に向け、気導音声と骨導音声を用いた新しい話者識別システムを提案する。

また、現状の本人認証システムでは風邪を引いて一時的に声が変わってしまった場合でも会社の中に入れないといったことも考えられる。本研究は、このような話者識別の精度向上に向け、気導音声と骨導音声を用いた新しい話者識別システムを提案する。

ところで、骨導音声は発声時に骨を介して耳に伝わる音で、これを電気信号に変換する骨導マイクロホンにはハンズフリーであるばかりではなく、ヘッドセット型のマイクロホンと比較して周囲騒音の影響を受けにくいという特徴がある。このため、ヘッドマウントディスプレイと組み合わせることによって、ウェアラブルコンピュータの入力装置として使用できることが期待される。

一方、骨導音声は高域が大きく減衰するため明瞭性に劣り、気導音声と比べて少しこもった感じに聞こえることが知られている。このため、音声認識システムに骨導音声を用いると、音声認識率が気導音声を用いた場合に比べ著しく低下する傾向にある。そこで、骨導音声のスペクトルの高域減衰特性を逆に利用し、話者識別のための特徴量として用いる方法を提示し、「気導音声と骨導音声を用いた話者識別」を実験検討した。

2. 研究の目的

本研究は、ユビキタス社会におけるセキュリティ技術の1つとして、気導音声と骨導音声のスペクトル比を用いた新しい話者識別システムを提案し、構築することを目的とする。

気導音声と骨導音声のスペクトル比は、骨導音声から気導音声への変換モデルの伝達特性を表している。したがって、それには頭頸部の形状・組織等に起因する話者固有の特徴が含まれているものと考えられる。そこで、この話者固有の特徴を定量的に表し、話者識別のための特徴量とすることによって、話者識別システムのさらなる識別率向上と雑音耐性改善を図る。

3. 研究の方法

骨導音声は発声時に骨を介して耳に伝わる音で、これを電気信号に変換する骨導マイクロホンには周囲騒音の影響を受けにくいという特徴がある。

一方、気導音声と骨導音声のスペクトル比には、声紋などと同様に、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可能性がある。本論文では、このスペク

トル比を話者識別のための特徴量として用いる方法を提示し、実験検討を行った。

具体的には、日本語音声には必ず母音が含まれていることから母音に着目し、日本語 5 母音発声時の気導音声と骨導音声それぞれから単独で求めたメルケプストラム係数(MFCC)と提案法で求めた特徴量を用いて話者識別実験を行い、識別率を比較することによって提案法の有効性を検証した。

また、骨導音声を用いて裏声を判別し、なりすまし防止や歌声評価等への応用について検討を行う。

(1) 実験用音声 音声の収録は防音室内で行い、気導音声と骨導音声を同時に収録した。気導音声は口唇から 10cm の距離に置いたコンデンサマイクロホン(SONY 製 ECM-999)を用いて、骨導音声は骨伝導イヤホンマイク(JABRA 製 JABRA Ear PHONE)を用いて、サンプリング周波数 48kHz、量子化レベル 16 ビットで録音を行った。

発声者は 18 歳から 26 歳の 70 名(うち女性 12 名)で、日本語 5 母音を 20 回続けて発声し収録を行った。発声者の男女比は等しい方が望ましいが、頭部の伝達特性を特徴量とする本手法では、このことはあまり問題とならない。

(2) 特徴量の抽出 母音定常部(各音声区間の中央を中心とした 2048 ポイントの区間、約 42.7ms) にハミング窓をかけ、求めた気導音声と骨導音声の振幅スペクトルをそれぞれ対数化、最大値による規格化を行い($B(\omega)$, $C(\omega)$), それらの差 $A(\omega)$ を求めた。また、気導音声と骨導音声の振幅スペクトルをそれぞれメルフィルタバンクによる重みづけ、対数化、最大値による規格化を行った後($Bm(\omega)$, $Cm(\omega)$), それらの差の離散コサイン変換(DCT)を求め、 $Amc(\tau)$ とした。図 1 に特徴量抽出処理フローを示す。

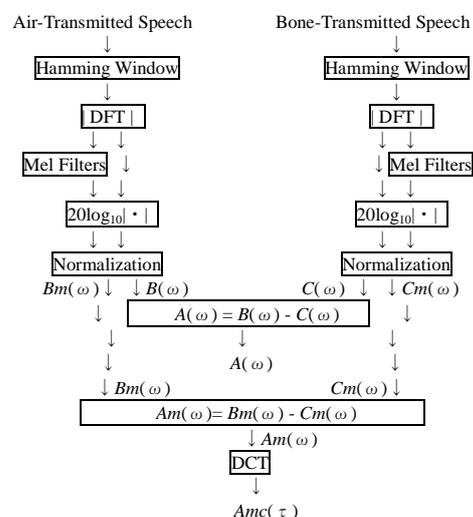


図 1 特徴量抽出処理フロー

(3) テンプレート作成 各話者が日本語 5 母音を 20 回ずつ発声した音声データの各母音それぞれ前半 10 回分をテンプレート作成用データとした。まず、これらのデータそれぞれから特徴量 $A_{mc}(\tau)$ を求めた。これらは、メルフィルタバンクの上限周波数を f_u として、 f_u までの帯域をメルスケールで N 等分割して得られた N 次元ベクトルである。次に、得られた 10 個の N 次元ベクトルの要素ごとに相加平均をとることによって平均ベクトルを求め、各話者の母音毎のテンプレートを作成した。また、 $A(\omega)$ から同様にしてテンプレートを作成した。

(4) 話者識別実験 音声データのうち後半発声 10 回分のデータを用いて話者識別実験を行った。まず、この音声データについてもテンプレート作成時と同様に特徴量を求め、母音種は既知として、テストデータとテンプレートとの間のユークリッド距離を各話者、各母音に対して求めた。次いで、同一母音で距離が最小の話者に 1 ポイントのスコアを与え、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。また、上限周波数 f_u は、骨導音声を用いた本手法では上限周波数を高くしても話者識別率があまり変わらないことから 3kHz とした。0~ $N-1$ 次のメルケプストラム係数を用い、ベクトル次元数 N はすべて 20 とした。

4. 研究成果

(1) 骨導音声データの作成 日本語 5 母音、及び単語音声（特に数字）発声時の気導音声と骨導音声の収録を、防音室内で行った。また、防音室内で電子協騒音データベース DAT の駅構内騒音を再生し、同様の収録を行うことで、雑音環境下音声データを作成した。

(2) MFCC を用いた話者識別実験結果

図 2 に話者識別実験結果を示す。対数パワースペクトルをそのまま用いた場合（特徴量 $A(\omega)$, baseline），話者数が 12 名では 99.2% の話者識別率を得ることが出来たが、70 名では識別対象話者数が増えたために 52.3% に低下した。一方、MFCC を用いた場合（特徴量 $A_{mc}(\tau)$ ）は、話者数が 12 名で 100%、70 名で 88.4% の話者識別率が得られた。このことから、MFCC を用いた場合の方が、話者識別率が高く、話者人数増加に対する識別率低下も小さいことが分かる。

(3) 上限周波数の検討 上限周波数を 3kHz から 7kHz まで 0.5kHz 刻みで変化させて、話者識別実験を行った結果、6.5kHz としたとき、話者識別率が最も高くなることが分かった。このとき、MFCC を特徴量として用い

ると、話者数が 70 名で 97.1% の話者識別率が得られた。

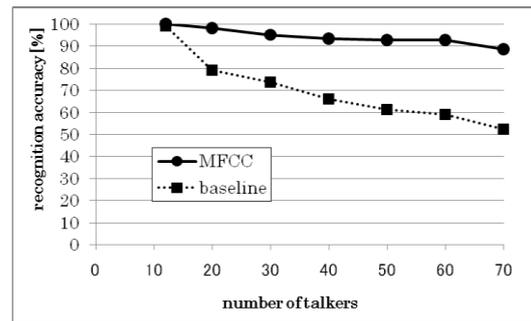


図 2 話者数を増加させたときの識別率

(4) 風邪などによる声質変化への対応やなりすまし防止等への応用についての検討

40 名に対して、日本語 5 母音発声時の気導音声と骨導音声を鼻声母音（鼻をつまんで発声した母音）について話者識別実験を行った結果、鼻声母音を入力した場合も、MFCC を用いた場合の識別率は対数パワースペクトルをそのまま用いた場合より高くなることが分かった。しかし、鼻声母音の場合、通常母音のときに比べて識別率が著しく低下した。

(5) 骨導口笛音の分析とその応用

音声と口笛音それぞれについて、気導音に対する骨導音の音質劣化の程度を聴取実験によって比較した。その結果、骨導口笛音では音質がほとんど劣化しないことを定量的に示すことができた。また、骨導口笛音を分析した結果、音質劣化の原因となる音がこもる現象が起こらない理由が明らかとなった。しかし、骨導口笛音の周波数特性を実測した結果、人が異なれば共振周波数が異なることが分かった。この特徴を利用した骨導口笛音による話者識別の可能性が示唆される。

5. まとめ

気導音声と骨導音声のスペクトル比を用いた話者識別システムを提案し、70 名の話者に対する識別実験を行った。その結果、MFCC を用いた場合の方が、対数パワースペクトルをそのまま用いた場合よりも話者識別率が高く、話者人数増加に対する識別率低下も小さいことから、両者のスペクトル比を用いる本方式の有効性は MFCC を用いた場合においてより強力に表れることが分かった。また、上限周波数を 6.5kHz としたとき、話者識別率が最も高くなることが分かった。このとき、MFCC を特徴量として用いると、話者数が 70 名で 97.1% の話者識別率が得られた。

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)

[雑誌論文] (計6件)

- ① 森 幹男, 荻原慎洋, 南本智哉, 谷口秀次, 加藤省三, 荒木睦大, 口笛音楽検定試験システムのための周波数変調による口笛音合成, 電気学会論文誌C, 査読有, Vol.130, No.1, 2010, pp.156-157
- ② 森 幹男, 荻原慎洋, 清金誉行, 谷口秀次, 荒木睦大, 気導音声と骨導音声のスペクトル比を用いた話者識別システムの識別精度に関する評価, 電気学会論文誌C, 査読有, Vol.129, No.9, 2009, pp. 1770-1771
- ③ 森 幹男, 荻原慎洋, 邱 添, 谷口秀次, 荒木睦大, 骨導口笛音と骨導音声の音質比較評価, 電気学会論文誌C, 査読有, Vol.129, No.8, 2009, pp. 1613-1614
- ④ 森 幹男, 荻原慎洋, 口笛音楽検定試験システムの検討—気導音を用いた吹音・吸音判定と骨導音を用いたピッチ判定—, 電気学会論文誌C, 査読有, Vol.129, No.1, 2009, pp.151-152
- ⑤ 森 幹男, 吉田千夏, 荻原慎洋, 谷口秀次, 高橋謙三, 自己聴取音に占める気導音と骨導音の割合の推定, 電気学会論文誌C, 査読有, Vol.127, No.8, 2007, pp 1268-1269
- ⑥ 森 幹男, 竹田 勉, 荻原慎洋, 谷口秀次, 高橋謙三, 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, 電気学会論文誌C, 査読有, Vol.127, No.3, 2007, pp.456-457

[学会発表] (計12件)

- ① 寺井祐樹, 荻原慎洋, 森 幹男, 谷口秀次, 気導音声と骨導音声を用いた話者識別の再検討, 平成21年度電気関係学会北陸地区学生による発表会, 2010.03, 野々市
- ② 内倉裕人, 荻原慎洋, 森 幹男, 谷口秀次, 気導音・骨導音同時受聴における音量ならびに音質の評価, 平成21年度電気関係学会北陸地区学生による発表会, 2010.03, 野々市
- ③ 森 幹男, 口笛音楽検定試験システムの開発, 日本音響学会北陸支部第2回音響・超音波ジョイント研究会, 2010.03, 7, 野々市
- ④ 成瀬 悟, 荻原慎洋, 森 幹男, 谷口秀次, 荒木睦大, 口笛音に対する瞬時ピッチ抽出と考察, 平成21年度電気関係学会北陸支部連合大会, 2009.09, 能美市
- ⑤ 金井重樹, 荻原慎洋, 森 幹男, 谷口秀次, 荒木睦大, 気導音声と骨導音声を用いた話者識別—男女間の誤識別に対する考察—, 平成21年度電気関係学会北陸支部連合大会, 2009.09, 能美市

- ⑥ Mikio Mori, Mitsuhiro Ogihara, Shuji Taniguchi, and Chikahiro Araki, A Musical Whistling Test System Using Air-Transmitted Sounds and Bone-Transmitted Sounds, the Second International Symposium on Universal Communication (ISUC2008), 2008.12, Osaka
- ⑦ 金井重樹, 荻原慎洋, 森 幹男, 谷口秀次, 気導音声と骨導音声を用いた話者識別—話者数を変えたときの識別率—, 平成20年度電気関係学会北陸支部連合大会, 2008.09, 福井市
- ⑧ 森 幹男, 荻原 慎洋, 谷口 秀次, 高橋 謙三, 骨導音を用いた口笛音楽検定試験システムの開発, 第7回情報科学技術フォーラムFIT2008, 2008.09,
- ⑨ 森 幹男, 荻原慎洋, 気導音と骨導音を用いた口笛音楽検定試験システムの開発, 情報処理学会音楽情報科学研究会, 2008.08, 名古屋市
- ⑩ 金井重樹, 荻原慎洋, 森 幹男, 谷口秀次, 高橋謙三, 気導音声と骨導音声を用いた話者識別システム, 平成19年度北陸地区学生による研究発表会, 2008.03, 鯖江市
- ⑪ 森 幹男, 黒谷直人, 上野有輝, 荻原慎洋, 谷口秀次, 高橋謙三, 口笛音楽検定試験システムの開発, 電子情報通信学会2008年総合大会, 2008.03, 北九州市
- ⑫ 竹田 勉, 荻原慎洋, 森 幹男, 谷口秀次, 高橋謙三, 気導音声と骨導音声を用いた話者識別システム, 平成19年度電気関係学会北陸支部連合大会, 2007.09, 福井市

6. 研究組織

(1) 研究代表者

森 幹男 (MORI MIKIO)

福井大学・大学院工学研究科・助教

研究者番号: 70313731

(2) 研究分担者

なし

(3) 連携研究者

なし