

平成21年3月31日現在

研究種目：基盤研究（C）

研究期間：2007～2008

課題番号：19500481

研究課題名（和文） 体内伝導音認識を用いた発声機能障害者支援装置の開発

研究課題名（英文） Development of the speech disorder support system using body-conducted speech recognition

研究代表者

石光 俊介（ISHIMITSU SHUNSUKE）

広島市立大学・大学院情報科学研究科・准教授

研究者番号：70300621

研究成果の概要：本研究では発声機能障害者のコミュニケーション障害を解決すべく、骨導などの体内伝導音を利用した発声支援システムを開発した。初めに健常者における場合のシステムを構築し、そのシステムとアルゴリズムの有効性を検証した上で機能障害者への適用を行った。同時に発声障害者音声データベースの構築も行った。発声機能障害者はその手術形式により多様であり、本手法による性能向上が認められるものの、実際の受聴効果を高めるためにさらなる工夫が必要であることが明らかになった。

交付額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	2,600,000	780,000	3,380,000
2008年度	900,000	270,000	1,170,000
年度			
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域

科研費の分科・細目：人間医工学・リハビリテーション科学・福祉工学

キーワード：生活支援技術

1. 研究開始当初の背景

音声は人間にとって重要なコミュニケーション手段であり、それを失うことは不便であるばかりでなく、多大な精神的苦痛を伴う。現在、咽頭癌患者の数は年々増加しており、手術により発声機能を失うことに伴う機能障害者も増えている。

これら機能障害者の内、約80%が食道発声法を行っている。これは口から空気を吸って、食道の途中でいったん止めて逆流させ、食道の入り口部や下咽頭粘膜などの新声門を振動させて声を出す方法である。しかし、この

発声法による音声は必ずしも明瞭ではなく、さらに発声音の基本周波数の低下、音量の減少により、屋外などでは騒音に埋もれてしまいコミュニケーションに障害が起きているのが現状である。

そこで本研究では機能障害者のコミュニケーション障害を解決すべく、骨導などの体内伝導音を利用した発声支援システムを開発することにした。応募者はこれまで体内伝導音を利用した高騒音下での音声通信システム開発を行ってきた。騒音が音声より大きな環境下では音声は通信に使えないため、騒音の影

響を受けにくい骨導をはじめとする体内伝導音によるシステムを構築した。この検討の中で体内伝導音により明瞭度のある音声を作り出すことができるのであれば、咽頭を摘出した機能障害者の音声も体内伝導音から高品質に変換できる可能性への着想に至った。つまり、食道発声法の新声門付近体内伝導音から明瞭な音声信号が生成できれば、コミュニケーション障害も緩和できる。

この体内伝導音はそのまま取り出すだけでは明瞭な音声にはならないため、補正する必要がある。これまで体内伝導音から音声への伝達関数を求め、明瞭度の高い音声を得る研究も行った。図1に声帯信号と音声との関係

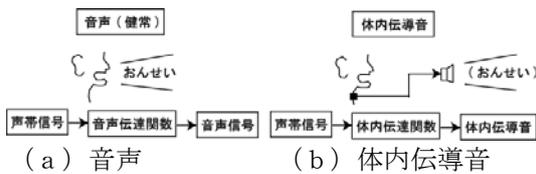


図1 声帯信号と声の関係

を示す。図のように音声、体内伝導音共に、ある伝達関数でそれぞれに声帯信号を関連づけることができるので、体内伝導音から明瞭度の高い音声への変換も可能となる。

食道発声をする場合も新声門付近の体内伝導から食道発声による音声への変換が可能であると考えられる。このことは体内伝導音と食道発声音との間に伝達関数が存在することを意味している。同様に、健常者の音声においてもその伝達関数関係は存在するわけであるが、機能障害になる前の音声ビデオなどの記録に残っている場合、環境差の影響は受けるものの食道発声時の体内伝導音からその音声への伝達関数を求めることができる。この手法を用いて、食道発声音から自然な音声への変換を行うことを実現する。

これに似た発想として、個人音声での規則合成 (CHATR, ATR) がある。しかしこれはあくまでもテキストを入力した上での音声再生であり、今回のように機能障害者が発声した音声が入座に変換されるというものではない。しかし、このCHATRと連続音声認識の手法を体内伝導音に適用することで瞬時性は得られるものと考えられる。ただし、認識と規則合成を組み合わせた場合、誤認識時にはその誤認識した言葉がそのままはっきりと発話される。一方、今回の提案のような伝達関数をたまたみ込む場合、誤認識であっても発話再生に近い伝達関数特性が認識結果として選択されるため、発声の主体を失わず、より実際的であると判断できる。また、非可聴つぶやき

音を用いるシステムの研究も行われているが、本提案のような体内伝導音連続認識システムを使う方法とは異なっている。

この体内伝導音を利用したシステムは研究代表者の2001年における講演以降様々な機関で研究が推進されるようになってきた。これまで主として体内伝導音を利用した認識システムの構築を研究してきたが、この認識システムが今回の開発では重要な役割を果たした。それは体内伝導音から音声の変換に用いられる伝達関数を全音声に対して同じものを用いるわけにはいかないからである。音声は非定常信号として、時々刻々とその特性が変化する。よって、音声のできるだけ小さな単位である個々のサブワードレベル (音節・半音節レベル) での伝達関数のペアを逐次探し、つなぎ合わせて、その言葉 (音声) の伝達関数を得る必要がある。これはUnited States Patent 5,903,865 (石光俊介, 1999年) より着想に至った。この特許はカーナビゲーション音声認識システムにおいて、新しいモデルを音声により追加する手段として考えたものである。

2. 研究の目的

研究においてはまず健常者における場合のシステムを構築し、そのシステムとアルゴリズムの有効性を検証した上で、機能障害者への適用を行う。以下のことを研究期間中に明らかにすることを目的とした。

(1) 健常者でのシステム構築

- ①体内伝導音によるサブワードレベルでの連続認識可能性の検討
- ②各サブワードレベルでの伝達関数の導出
- ③誤認識時の音声明瞭度に及ぼす影響調査
- ④明瞭度評価手法の検討

(2) 機能障害者へのシステム適用と構築

- ①評価用機能障害者食道発声データベース作成
- ②食道発声における体内伝導音連続音節認識システムの構築
- ③食道発声時の各サブワードモデルでの伝達関数導出
- ④誤認識時の影響調査
- ⑤実システム構築時における問題点

3. 研究の方法

まず、健常者でのシステム構築による有効性の検証と評価用食道発声データベース作成を行った。図2にシステムの概念を示す。まず、体内伝導音を連続サブワード認識し、それに対応する伝達関数を呼び出す。体内伝導音は順次その連続した伝達関数にたまたみ込

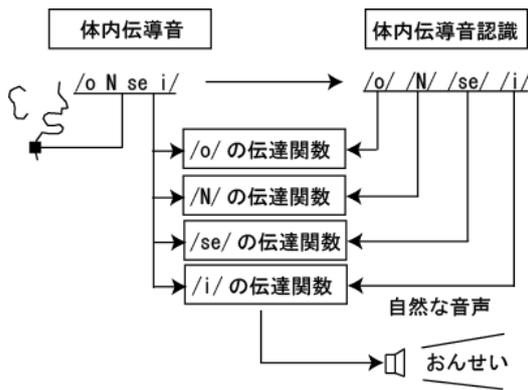


図2 提案システム

れ、自然な音声を生成するというものである。

(1) 健常者の体内伝導音によるサブワードレベルでの連続認識可能性の検討

まずサブワードレベルの認識確認実験を行った。

(2) 健常者の各サブワードレベルでの伝達関数の導出

音声作成のための各サブワード間の伝達関数を作成した。導出にはクロススペクトル法や適応フィルタを用いた。

(3) 誤認識時の音声明瞭度に及ぼす影響調査

以上の検討により、システム構築がほぼなされるので、ここで体内伝導音連続認識の結果に基づいた骨導音のたたみ込みを行い、音声明瞭度の確認と誤認識の影響についても聴感的に調査した。

(4) 明瞭度評価手法の検討

音声明瞭度については、距離尺度による音声との比較の他、音声認識率でも評価を行った。

(5) 評価用機能障害者食道発声データベースの作成

次に、機能障害者用に特化したシステムを実現する際に評価用音声として機能障害者による音声発声データベースを作成した。体内伝導音取り出し位置はこれまでの研究に基づき、食道発声時の最適位置を再検討した。そのための前実験として、明瞭度のチェックや簡易認識実験による体内伝導音最適取り出し位置の確認も行った。その後、そのデータベースを用いて、システム構築および評価を行い、実システム構築に向けた問題点の検討を行った。

(6) 食道発声における体内伝導音連続音節認識システムの構築

まず、食道発声時における体内伝導音認識を可能とするため信号適応処理を行い、連続サブワード認識を行い、認識状況を調査する。

認識率が信号適応処理では全くシステム構築ができないほどのレベルである場合は食道発声における特定話者体内伝導音認識モデルを新たに作成する。そしてこのシステムにおいて、健常者で動作を確認したシステムへの導入を図る。この際には体内伝導音のみではなく、発声障害者の音声特徴解析も行った。

(7) 食道発声時の各サブワードモデルでの伝達関数導出

健常者の際と同様に音声作成のための各サブワード間（音声－体内伝導音）の伝達関数を作成した。この際に健常者と比較した解析も行った。さらに健常者音声への伝達関数も同時に求め、この利用についても検討した。

また、ビデオの発声機能障害前の音声との対応付けも試みた。この場合、環境差やサブワードの不均一性などの問題が生じる可能性もあるが、それらの再生音に及ぼす影響も確認した。

(8) 誤認識時の影響調査

ここでは主として音声意味理解度を焦点とした評価になるが、前年度で実施した健常者における音声品質評価手法を用いて同様な評価を行うことにより比較検討した。

以上のようにこれまでの体内伝導音（骨導音）認識システムの技術を活かして、実時間応答性のある高品質な発声機能障害者支援装置の開発を目指した。

4. 研究成果

本研究では発声機能障害者のコミュニケーション障害を解決すべく、骨導などの体内伝導音を利用した発声支援システムを開発した。まず、健常者における場合のシステムを構築し、そのシステムとアルゴリズムの有効性を検証した上で機能障害者への適用を行うため、健常者でのシステム構築を行った。そこで、以下の点を検討した。

- (1) 体内伝導音によるサブワードレベルでの連続認識可能性の検討
- (2) 各サブワードレベルでの伝達関数の導出
- (3) 誤認識時の音声明瞭度に及ぼす影響調査
- (4) 明瞭度評価手法の検討。

以上の検討から、サブワードレベルでの連続認識が可能であり、さらに、誤認識時も音声の明瞭度に及ぼす影響が少ないことを明らかにした。サブワードレベルでの伝達関数作成については、サブワード長による誤差解析も行った。

図3及び4に/asahi/の音声及び体内伝導音

を示し、図5及び6に対応した伝達関数を用いた場合の変換音声と認識誤りが生じたとして異なる伝達関数を用いた変換音声の音声波形及びスペクトログラムを示す。図5の変換音声において、損失していた高域成分の周波数特性を回復していることが確認できる。また、図6の変換音声でも同等の回復が見られた。

また、明瞭度をさらに上げるために、加速度差分法を開発した。本手法は非常に簡単なアルゴリズムで音声成分の明瞭度化と抽出が可能である。これにより、信号そのものの明瞭度を上げ音声認識率を向上させることに成功した。本提案手法は音声認識により伝達関数を次々と呼び出すので、その有効性をさらに向上させることができた。また、音声の明瞭度自体を向上させるのに適応フィルタが有効であることも示した。これらの評価には音声認識率はもちろんであるが、時間周波数解析による比較や対数断面積比(Log-Area Ratio: LAR)等の評価指標によりその有効性を明らかにした。図7にLAR距離の評価結果を示す。LAR距離は音声と体内伝導音: BCS (Acc), 体内伝導加速度差分信号: Pre(Acc_{diff}), 音節単位適応フィルタ適用: Pre-adapted (Syllable), そして、単語単位適応フィルタ適用: Pre-adapted (Word unit)について示す。体内伝導音と比較して、各提案法によって処理された信号間の距離は近くなっている。さらに、処理後加速度差分信号に対して各適応フィルタ法を併用することにより、より音声に近い信号を推定できていることが確認できる。

次に、日本大学病院の協力を得て発声機能障害者の音声並びに体内伝導音を収録し、発声障害者音声データベースの構築を行った。発声機能障害者はその手術形式により、多様であった。つまり、食道発声のみではなく、音声再建法など特殊な音声獲得方法もあり、そのような発声方式についても提案手法が対応可能であるかを調査した。まず、障害者音声における体内伝導音連続音節認識システムの構築については、連続サブワード認識を行い、認識状況を調査した。

以下に示す被験者3名の音声と体内伝導音を収録し、音声認識を行った。

- ・ 22才男性健常者
- ・ 28才声帯ポリープ男性患者、男性
- ・ 61才声帯全摘出男性患者、音声再建法

各収録点における文章認識の結果を図8に示す。全被験者において音声と比べて体内伝導音の認識率は非常に低いことが分かる。また、音声では健常者と機ポリープ患者に違いは見られなかったが、体内伝導音では20%以上の認識率の低下を確認した。また、全摘出患者と比べると認識率はさらに低下し、体内

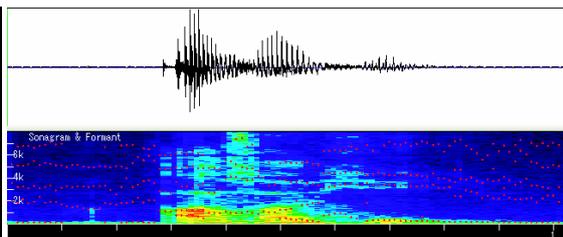


図3 音声

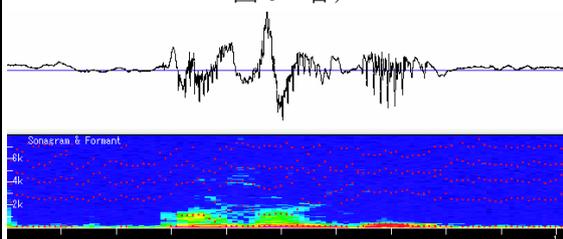


図4 体内伝導音

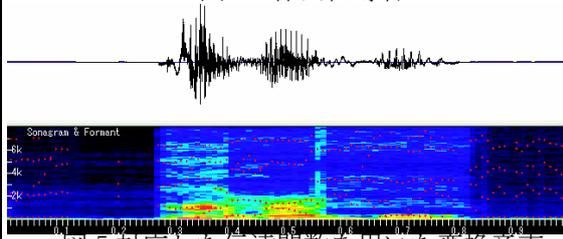


図5 対応した伝達関数を用いた変換音声

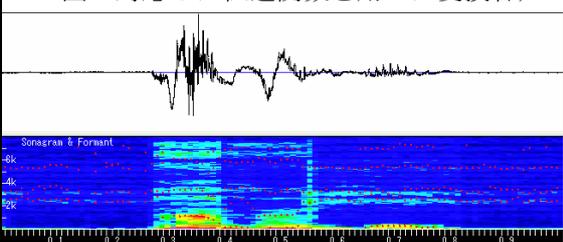


図6 異なる伝達関数を用いた変換音声

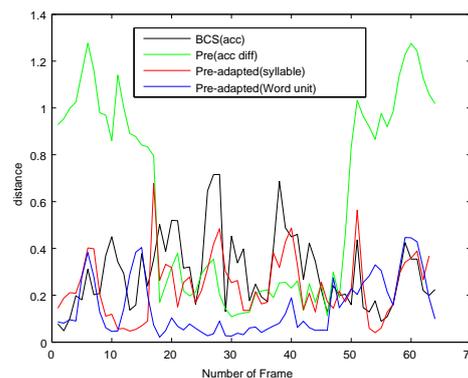


図7 LAR 距離による評価

伝導認識率は健常者の半分となり、適応処理が必要であることを確認した。適応処理により認識率は向上したため、特定話者体内伝導音認識モデルは新たに作成する必要はないこともわかった。発声機能障害者の音声特徴量を解析したところ、症状が深刻になるにつれてパワーが弱くなり、音声認識特徴量であ

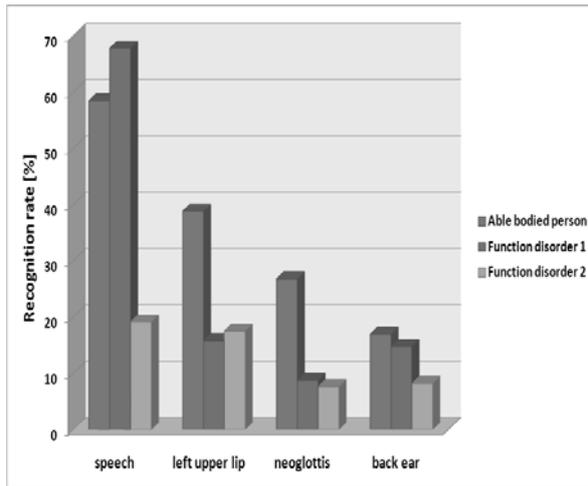


図8 音声認識率

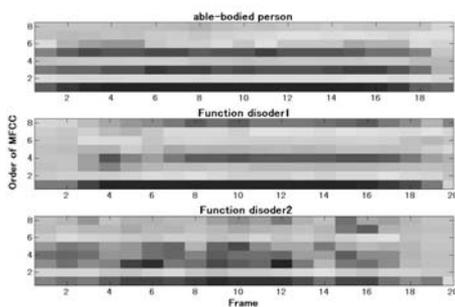


図9 音声認識パラメータによる比較

るメルケプストラム値も大きくずれていくことを確認した。その様子を図9に示す。次に、各サブワードモデルでの伝達関数の推定を行った。これには発声機能障害者の喉頭摘出前のビデオ音声を利用した。サブワード認識による区間推定誤差回避のため、今回は手作業により区間推定を行った。このため、この作業にかなりの時間を要した。ある程度、サブワード伝達関数データベースが構築できたところでシステムを稼働させ、その有効性を確認した。時間周波数平面上で評価をすると高域成分が回復し、障害音声に比べ、品質が向上したように観測できたが、実際に受聴すると、障害前の音声にまでは達していない。このため、音声認識特徴量的には性能向上が認められるものの、実際の受聴効果を高めるためにさらなる工夫が必要であることが明らかになった。しかし、問題点も明らかになったため、今後は音声品質をさらに向上させた実用システムを構築することを検討する。

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 S. Ishimitsu, M. Nakayama, T. Yoshimi and H. Yanagawa, NOISE-ROBUST RECOGNITION SYSTEM MAKING USE OF BODY-CONDUCTED SPEECH MICROPHONE, AES 122nd Convention No.7105, Proceedings CDROM, 8 pages, 2007, 査読有
- ② S. Ishimitsu, M. Nakayama and T. Yoshimi, Construction of the Noise-Robust Body-Conducted Speech Recognition System, IEEE Second International Conference on Innovative Computing Proceedings CDROM, 4 pages, 2007, 査読有
- ③ M. Nakayama, S. Ishimitsu and S. Nakagawa, Speech recognition with body-conducted speech using differential acceleration, Acoustics'08 Paris Proceedings, pp.1407-1412, 2008, 査読有
- ④ S. Ishimitsu and M. Nakayama, Construction of Speech Support System Using Body-Conducted Speech Recognition for Disorders, IEEE Third International Conference on Innovative Computing, Information and Control Proceedings, 4 pages, 2008, 査読有
- ⑤ M. Nakayama, S. Ishimitsu and S. Nakagawa, Speech Estimation from Body-conducted Speech with Differential Acceleration, IEEE International Conference on Networking, Sensing and Control Proceedings, pp. 858-863, 2009, 査読有

〔学会発表〕(計7件)

- ① 中山 仁史, 石光 俊介, 中川 誠司, 加速度差分を用いた体内伝導音明瞭化のための基礎検討, 電子情報通信学会第22回信号処理シンポジウム, 2007年11月7日~9日, 東北大学
- ② 小田康平, 中山仁史, 石光俊介, 発声機能障害者の音声特徴分析, 日本機械学会中国四国支部 第47期総会・講演会, 2009年3月6日, 山口大学

〔図書〕(計2件)

- ① H. Abut, J.H.L. Hansen, K. Takeda and S. Ishimitsu, Advances for In-Vehicle and Mobile Systems : Challenges for International Standards, Springer Science, 2007, 284
- ② Shunsuke Ishimitsu, Speech Recognition Chapter 14 Construction of a Noise-Robust Body-Conducted Speech Recognition System, IN-TECH, 2008, 550

〔産業財産権〕

○出願状況（計3件）

- ① 発声支援方法，石光俊介，中山仁史，広島市立大学，特願 2007-340381，2007年12月28日，国内
- ② 信号再生装置，中山仁史，石光俊介，中川誠司，広島市立大学，産総研，特願 2007-271204，2007年10月18日，国内
- ③ 音声認識装置および音声変換装置，石光俊介，中山仁史，中川誠司，広島市立大学，産総研，特願 2009-068545，2009年3月19日，国内

○取得状況（計0件）

〔その他〕

○報道関連

日刊工業新聞，2009年1月16日，26面

中国新聞，2009年1月16日，8面

朝日新聞，2009年1月27日，31面（地域）

6. 研究組織

(1) 研究代表者

石光 俊介 (ISHIMITSU SHUNSUKE)

広島市立大学・大学院情報科学研究科・准教授

研究者番号：70300621

(2) 研究分担者

なし

(3) 連携研究者

なし