

機関番号：15501
 研究種目：基盤研究(B)
 研究期間：2007～2010
 課題番号：19300078
 研究課題名（和文） 実騒音環境下における発音・発声障害者の音声コミュニケーション手段の確立
 研究課題名（英文） Establishment of Vocal Communication System under Heavy Environmental Noise for Handicapped People with Speech Impediment

研究代表者
 内野 英治 (UCHINO EIJI)
 山口大学・大学院理工学研究科・教授
 研究者番号：30168710

研究成果の概要（和文）：

本研究の目的は、発音・発声障害者の不明瞭な音声を健常者に近い明瞭な音声へ自動変換するシステムを、ソフトコンピューティング技術を用いて開発し、障害者の音声コミュニケーションの手段を確立することである。本課題の基礎となる音声特徴量の抽出、音声特徴量の変換については、従来よりも変換精度の高いモデルが完成し、また、音声の音質変換においても、クリアな変換が実現された。

研究成果の概要（英文）：

The purpose of this research is to establish a vocal communication system for handicapped people with speech impediment. The system automatically converts the unclear speech of a person with a speech impediment to a clear speech by using soft computing technique. In the extraction of features of speech and the conversion of them, the model with higher conversion accuracy than ever has been accomplished, and the clear tone conversion of speech has been obtained.

交付決定額

(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	1,700,000	510,000	2,210,000
2009年度	1,900,000	570,000	2,470,000
2010年度	1,700,000	510,000	2,210,000
年度	0	0	0
総計	7,900,000	2,370,000	10,270,000

研究分野：総合領域

科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング

キーワード：音声変換，ソフトコンピューティング，自己組織化マップ，機械学習，人工知能

1. 研究開始当初の背景

身体に何らかの障害を持つ人を支援するための環境・設備の構築は、社会福祉の分野で解決すべき重要な問題である。特に移動手段に関する発展は目覚しく、運動系に障害を持つ人に対しては、色々な場所に自立的に移

動できるバリアフリー環境が整いつつある。

一方、脳性麻痺などにより発音・発声に障害を持つ人は、発音がはっきりせず言いたいことが相手にうまく伝わらない。特に公共交通機関の利用時や街中では、支援を必要とする機会が多いにも拘わらず、未だ十分な支援

環境が整備されていないのが現状である。このような状況において、発音・発声障害者の音声を聞き取り易い音声に自動変換できれば、これらの障害者の自立に大きく役に立つのは間違いない。

しかし、これまでの音声変換に関する研究は、騒音や雑音がほとんどない静寂な環境を前提とした手法が主流であるのに対し、公共交通機関内や街中では騒音や雑音が溢れ、発話者の音声の特徴を正確に抽出することは極めて困難である。

これらの高騒音下において、目的とする音声のみの情報を取り出す直接的な手段として、骨導マイクを用いることが考えられる。骨導マイクは顎骨の振動を直接音に変換するので、高騒音下でも雑音に汚されずに音声を検出できる。

そこで、発音・発声障害者に骨導マイクを装着してもらい、発音・発声障害者が言いたい内容に自動変換して発話するシステムを構築できれば、実環境下に対応できるコミュニケーション手段となる。

しかし、骨導マイクから収録された骨導音声の音質は、身体や顎骨の振動雑音のため、我々が普段用いる気導音声のものと比較すると非常に悪い。そのため、円滑なコミュニケーションを行うためには、骨導音声をより明瞭な気導音声に変換する必要がある。

これを実現するためには、骨導音声から得られる音声の特徴量と気導音声の音声特徴量との対応関係を事前に学習し、特徴量間の写像(マッピング)を行えばよい。しかしながら、この関係は非線形で非常に複雑であり、新たな変換システムを構築しなければならない。

本研究では、実用化を念頭に置き、実騒音環境下における音声変換アルゴリズムをソフトコンピューティング技術を用いることにより開発し、発音・発声障害者の音声コミュニケーション手段の確立を目指す。

2. 研究の目的

本研究の目的は、発音・発声障害者の不明瞭な音声を健常者に近い明瞭な音声へ自動変換するシステムを、ソフトコンピューティング技術を用いて開発し、障害者の音声コミュニケーションの手段を確立することである。

本研究では、たえず騒音や雑音が存在する

実騒音環境を想定し、周囲環境騒音の中で被験者の音声を、環境騒音に遮られることなく音声信号を撮取する方法として、骨導マイクを採用する。

このとき、骨導音声の音質を通常の気導音並みに改善するために、両者間の対応関係を調べ、骨導音声から対応する気導音声への変換を行う。

そのためには、まず、音声特徴量の適切な選択と、その特徴量間の距離尺度の決定が重要である。これには、メルケプストラム係数など、周波数領域において音声の特徴をよく捉えるものを中心に検討する。

次に決定された特徴量を用いて、骨導音声と気導音声間の特徴量の非線形なマッピングを行う必要がある。これには、従来のLBG(Linde-Buzo-Gray)法などのベクトル量子化では、十分な精度が得られない。そこで、本研究では、ソフトコンピューティング技術の一部である自己組織化マップ、ニューラルガスネットワークなどの手法を用いて、この非線形マッピングを表現することを試みる。

最後に変換された特徴量を用いて音声合成を行い、時間領域の音声を出力する。

これら一連のシステムの構築が、本研究の目的である。

3. 研究の方法

発音・発声障害者の音声を健常者の音声へ自動変換するには、まずそれぞれの音声の特徴量の抽出、次に抽出された音声特徴量間の関係を学習し、その変換を実現、最後に変換された音声特徴量から音声の合成が必要である。

この過程は、以下のように3つのブロックに分割できる。

(1) 骨導マイクの使用

実騒音環境下での応用を考え、発音・発声障害者の音声を通常のマイク(気導マイク)ではなく、周辺雑音の混入がほとんどない骨導マイクの使用を考える。

ただし、骨導マイクから収録された骨導音声は音質が悪いので、この骨導音声を通常の気導音声の音質へ変換する必要がある。ここでは、ソフトコンピューティングを用いた各種変換法により音質の改善を試みる。

(2) 音声特徴量の抽出

従来のメルケプストラム係数だけでなく、デルタケプストラムや、データ空間の広がりも考慮したマハラノビス距離を用いて新たな音声特徴量の抽出を行い、骨導音声から気導音声の写像（マッピング）を念頭に、より適切なパラメータで音声を表現することを目指す。

(3) 音声特徴量の変換

同時収録した2音声の各時間フレームの音声特徴量間の関係を、新たに開発した双子型自己組織化マップにより記述し、その変換モデルを構築する。従来のスペクトルマッピングコードブックを用いた方法と比較し、提案手法の有効性を調べる。

また、音声特徴量変換の計算量を低減させるため、音声特徴量間の変換マッピングコードブックをニューラルガスネットワークにより量子化し、計算量の低減を図る。その際、変換精度の低下を抑えるため、クラスごとに構築した局所変換モデルを用いる。変換された音声に対して動的計画法（DP）マッチングによる単語認識を行い、従前のモデルと単語認識精度を比較する。

4. 研究成果

本課題の基礎となる音声特徴量の抽出および音声特徴量の変換については、従来法よりも変換精度の高いモデルが完成し、また、骨導音声の音質変換においても、クリアな変換が実現された。

本研究では、最初に、骨導音声の音質を気導音声並みの音質に改善する新しいタイプの自己組織化マップ（双子型自己組織化マップ）を考案し、それをを用いた音質改善の有効性を確認した。双子型自己組織化マップの概念図を図1に示す。

この双子型自己組織化マップでは、学習により骨導音声と気導音声間の非線形写像関係を自己組織化マップの双子型ユニットに保持しておき、未知の骨導音声が入力されると、それに対応する気導音声が出力される。音声特徴量としてはメルケプストラム係数が用いられた。

変換前と変換後の骨導音声と、それに対応する気導音声間の単語ごとの量子化歪みの平均値評価では、変換後の量子化歪みは変換

前のそれと比べて、20単語と100単語の時それぞれ約70%と50%の改善が見られた。

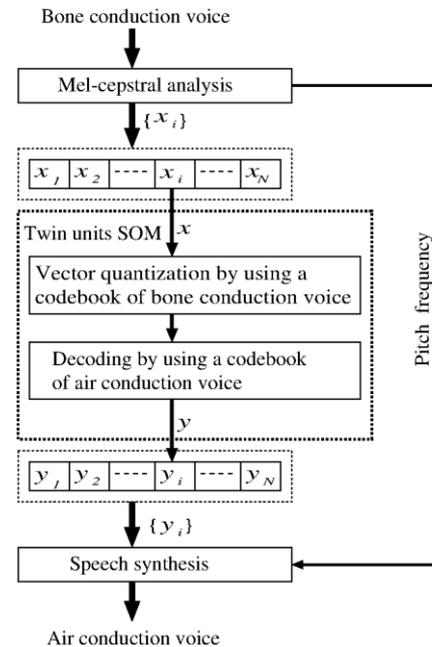


図1. 双子型自己組織化マップを用いた骨導音声から気導音声への音声変換

また、被験者14名による三刺激二重盲験法を用いた主観評価では、変換前の骨導音声の品質と比べて、変換後の骨導音声の品質の優位性が心理実験的にも確認された。

さらに上記の手法を改良し、双子型自己組織化マップの入力側すべてのユニットに線形部分空間を保持させた、双子型部分空間自己組織化マップを新たに提案し、従前の方法よりも変換速度を1/3に短縮することが出来た。

次に、音声特徴量として新たにデルタケプストラムを採用した。デルタケプストラムとは、前後 k 時刻分のメルケプストラム係数に対する次元ごとの線形回帰係数であり、最も簡単な動的特徴量である。これにより、時間軸上の各時刻で写像が1対1と仮定しているメルケプストラム係数の欠点が補われる。さらに、距離尺度としてマハラノビス距離を採用し骨導音声の音質改善を行った。

具体的には、変換すべき音声特徴量間の対応関係をデルタケプストラムにより1対1に修正した後、ニューラルガスネットワーク

により音声特徴量の分類を行った。本提案手法の概念図を図2に示す。

ニューラスガスネットワークは、クラスタリング手法の1つである。ここでは、データ空間におけるデータ間の類似性のみに基づいた順位学習を行い、特徴ベクトルを学習後に得られた参照ベクトルをもとに各クラスに分類した。その後、分類された各クラス毎に複数の局所変換モデルを作成し、これらのモデルを適応的に選択することにより音声変換を行った。

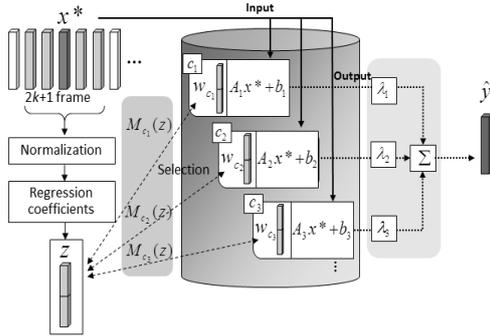


図2. デルタケプストラム及びニューラルガスネットワークを用いた骨導音声から気導音声への音声変換

実験では、一人の話者が発話した17単語5組を用いて音声変換精度の評価を行った。実験では、本手法を用いることにより、一例として、量子化歪14.8dBを持つ骨導音声は、変換により量子化歪5dB以下にまでその音質が改善され、本手法の有効性が確認された。

今後はさらに、スパースコーディング(Sparse Coding)や独立成分分析(Independent Component Analysis)、適応部分空間自己組織化マップ(Adaptive Subspace Self-Organizing Map)などの手法を用い、発音・発声障害者の不明瞭な音声から、音声の特徴付けにロバストな新たな音声特徴量の抽出を試みる。

また、発音・発声障害者の音声特徴量から健常者の音声特徴量への変換においては、自己組織化マップやその他のソフトコンピューティング技術を駆使し、データ数の少ない空間(疎な空間)でも正確な変換(写像)の出来る更なる音声特徴量の高精度な変換法の開発を行う。

さらには、これら一連のシステムを実際に使用した感想をフィードバックさせシステムの完成度を高める。

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4件)

- ① 矢野和昭, 畔津忠博, 内野英治:
"片耳装着型骨導マイク音声の自己組織化マップによる音質改善と対話型船用機関運転支援システムへの応用可能性の検討," 電子情報通信学会論文誌D, Vol. J92-D, No. 7, pp. 1066-1070, 2009, 査読有
- ② R. Kubota, E. Uchino and N. Suetake:
"Hierarchical k-Nearest Neighbor Classification Using Feature and Observation Space Information," IEICE Electronics Express, Vol. 5, No. 3, pp. 114-119, 2008, 査読有
- ③ E. Uchino, K. Yano and T. Azetsu:
"A Self-Organizing Map with Twin Units Capable of Describing a Nonlinear Input-Output Relation Applied to Speech Code Vector Mapping," Information Sciences, Vol. 177, pp. 4634-4644, 2007, 査読有
- ④ T. Azetsu, E. Uchino and N. Suetake:
"Speech Extraction of a Target Speaker from One Mixed Speech Signal," 電気学会論文誌C, IEEJ Trans. EIS, Vol. 127, No. 6, pp. 970-971, 2007, 査読有

[学会発表] (計9件)

- ① R. Kubota, Y. Kawamura, N. Suetake and E. Uchino:
"Tone Quality Improvement of Bone Conduction Voice by Cepstrum-based Local Conversion Models," Proceedings of the 11th International Conference on Multimedia Systems and

Signal Processing (MUSP 2011),
pp. 74-78, 2011-3-8 Venice Italy

- ② 河村庸平, 久保田良輔, 内野英治, 末竹規哲:

“デルタケプストラムとマハラノビス距離に基づく骨導音声の音質改善法,” 第19回計測自動制御学会中国支部学術講演会論文集, 120, pp. 58-59, 2010-11-28 島根, 島根大学

- ③ 久保田良輔, 河村庸平, 内野英治, 末竹規哲:

“ニューラルガスネットワークに基づく骨導音声の音質改善法とその認識性能評価,” 第14回日本知能情報ファジィ学会中国・四国支部大会講演論文集, pp. 11-12, 2009-12-12 広島, 広島大学

- ④ 矢野和昭, 内野英治:

“適応部分空間ユニットを備えた双子型自己組織化マップによる骨導音声変換の速度向上実験,” 平成21年度電気関係学会九州支部連合大会論文集, 03-2P-07, 2009-9-28 福岡, 九州工業大学

- ⑤ 久保田良輔, 内野英治, 末竹規哲:

“ニューラルガスネットワークと局所変換モデルを用いた骨導音声の音質改善法,” 第25回ファジィシステムシンポジウム講演論文集, 2B3-04, pp. 2B3-04.1-2B3-04.2, 2009-7-15 茨城, 筑波大学

- ⑥ R. Kubota, E. Uchino, N. Suetake and M. Kunihiro:

“A Code Vector Conversion Model for Bone and Air Conduction Voices by Using Self-Organizing Map with Additional Learning,” Proceedings of 2009 International Workshop on Nonlinear Circuits and Signal Processing (NCSF 2009), pp. 109-112, 2009-3-1 Hawaii USA

- ⑦ 国広真実, 久保田良輔, 古賀崇了, 末竹規哲, 内野英治:

“追加学習機能を有する自己組織化マッ

プを用いた骨導一気導音声変換モデル,” 第17回計測自動制御学会中国支部学術講演会論文集, 109, pp. 34-35, 2008-11-15 広島, 広島大学

- ⑧ 国広真実, 久保田良輔, 末竹規哲, 内野英治:

“自己組織化マップを用いた音声変換モデルによる骨導音声の音質改善,” 第24回ファジィシステムシンポジウム講演論文集, TF2-3, pp. 721-722, 2008-9-4 大阪, 阪南大学

- ⑨ T. Azetsu, E. Uchino, R. Kubota and N. Suetake:

“Feature Extraction and Signal Reconstruction of Air and Bone Conduction Voices by Independent Component Analysis,” Proceedings of the 2008 IAENG International Conference on Artificial Intelligence and Applications, pp. 55-59, 2008-3-20 Hong Kong China

〔図書〕(計1件)

- ① 大北正昭, 徳高平蔵, 藤村喜久郎, 権田英功 編:
“自己組織化マップとそのツール,” 第10章担当, 内野英治, 畔津忠博, シュプリンガー・ジャパン, pp. 179-195, 226頁, 2008

〔産業財産権〕

○出願状況 (計0件)

○取得状況 (計1件)

名称: クラス分類方法及び装置

発明者: 内野英治, 末竹規哲, 久保田良輔, 松崎益徳, 廣高史

権利者: 国立大学法人山口大学

種類: 特許権

番号: 第4280834号

取得年月日: 平成21年3月27日

国内外の別: 国内

[その他]

ホームページ

<http://www.ic.sci.yamaguchi-u.ac.jp/>

6. 研究組織

(1) 研究代表者

内野 英治 (UCHINO EIJI)

山口大学・大学院理工学研究科・教授

研究者番号：30168710

(2) 研究分担者

末竹 規哲 (SUETAKE NORIAKI)

山口大学・大学院理工学研究科・准教授

研究者番号：80334051

畔津 忠博 (AZETSU TADAHIRO)

山口県立大学・情報化推進室・准教授

研究者番号：70285451

森田 博彦 (MORITA HIROHIKO)

(H19年度)

(財) ファジィシステム研究所・研究部・
主席研究員

研究者番号：60270584

久保田 良輔 (KUBOTA RYOSUKE)

(H20年度～H22年度)

宇部工業高等専門学校・制御情報工学科・
助教

研究者番号：50432745

(3) 連携研究者

なし