

研究種目：基盤研究（C）
 研究期間：2006～2008
 課題番号：18500152
 研究課題名（和文）
 音声から発話動作を推定する発話ロボットの構築と対話型発声訓練システムの提案
 研究課題名（英文）
 The development of a talking robot that estimates speech articulation from voice and the proposal of an interactive speech training system
 研究代表者
 澤田 秀之（SAWADA HIDEYUKI）
 香川大学・工学部・准教授
 研究者番号：00308206

研究成果の概要：

人間の発声器官である肺、気道、声帯、声道、鼻腔、舌とそれらを動かす為の筋肉などを全て機械系により再現し、自律的な発話動作によって音声を生成する発話ロボットを構築した。本発話ロボットが、音声の特徴から、その音声を生成するのに必要な口内の発話動作が再現できることを確認した。更に、聴覚障害者や発話障害者が、ロボットの見本となる構音動作や口内の動きを見ながら、対話的に発話訓練をおこなうシステムを構築し、実証実験によりその有効性を確認した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,600,000	0	1,600,000
2007年度	1,300,000	390,000	1,690,000
2008年度	700,000	210,000	910,000
年度			
年度			
総計	3,600,000	600,000	4,200,000

研究分野：ロボティクス

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：発話ロボット、音声、聴覚フィードバック、自己組織化ニューラルネットワーク、適応学習、聴覚障がい者、発話訓練、支援技術

1. 研究開始当初の背景

音声は、特別な道具なしに情報を伝えることができる、最も容易なコミュニケーション手段である。音声には、言語的情報だけでなく、感情的表現や話者の個人性などが含まれており、これらを即時的情報伝達手段として容易に利用できるという利点がある。このため音声を分析し、人間同士のコミュニケーションツールとして役立てようとする研究が、これまでに数多くなされてきた。

音声に関する研究として、音声波から様々

な情報を抽出する音声分析、音声波の情報を圧縮して表現し利用する音声符号化、音声を人工的に生成する音声合成、音声波から内容を文字として抽出する音声認識、音声波から発話者を認識する話者認識などがあり、様々な実用化や応用が提案されている。特に音声の生成については、ソフトウェアを用いた音声合成システムが主流であるが、人間の発話のように、独特な揺らぎや個人性、感情表現を含んだ音声をソフトウェア技術によって実時間で生成することは、未だ困難である。

加えて今後は、人間支援ロボットや無人売店など、音声を用いてより密接に人間を支援するシステムが必要になってくると考えられている。

申請者は、人間の音声を持つ様々な情報を再現することを目的とし、人間と同様の発話機構を持ち、人間のように自律的に発話手法を獲得して音声の生成をおこなう発話ロボットの研究をおこなってきた。

2. 研究の目的

本研究では、人間と同等の発声器官をエアープンプ、人工声帯、声道共鳴管、音響アナライザなどを用いて全て機械的に構成し、計算機による聴覚フィードバック制御によって自律的に音声を獲得、生成することができる発話ロボットを構築することを目的とした。更に、発話ロボットが音声模倣学習によって獲得できる、発話動作および音声を詳細に解析することにより、人間の音声獲得メカニズムの解明につなげていく。また本ロボットは、音声の特徴から、その音声を生成するのに必要な発話動作を再現できることから、聴覚障害者や発話障害者が、見本となる構音動作や口内の動きを見ながら、対話的に発話訓練ができるシステムを構築する。

3. 研究の方法

発話動作をおこなう器官を全て機械的に構成し、計算機による聴覚フィードバック制御によって自律的に音声および発話手法を獲得、生成することができる発話ロボットを構築する。ロボティクス、ヒューマンインタフェース開発の視点から、人間と同等の発話器官を用いて音声対話が可能な、発話ロボットの開発をおこなう。これにより、人間どうしのように発話動作をとらせた対面コミュニケーションが可能となるばかりでなく、発話障害者の発話訓練システムの構築、更にはアミューズメントロボットや発話楽器の提案につなげていく。また発話ロボットが音声模倣学習によって獲得できる発話動作および音声を詳細に解析することにより、人間の音声獲得機構の解明につなげていく。更に本発話ロボットとその自律的発話学習機構を応用し、聴覚障害者や発話障害者が、見本となる構音動作や口内の動きを見ながら、対話的に発話訓練をすることができるシステムを構築する。

4. 研究成果

(1) 発話ロボットの構成

発話ロボットの構成を図1に示す。本ロボットは、エアープンプ、人工声帯、鼻腔、共鳴管、マイクロフォン、音響アナライザから構成され、これらはそれぞれ人間の肺、声帯、鼻腔、共鳴管、聴覚に対応している。エア-

ポンプから送られた空気流が人工声帯を振動させる事で、音源波が生成される。シリコンゴムで作成した共鳴管をステンレス棒の上下運動で変形させることにより、音響の共鳴特性を変化させ、任意の音声を生成する。生成された音響は、マイクロフォンを通してデジタル信号として計算機に入力され、音声特徴の解析と聴覚フィードバック学習に使われる。またマイクロフォンより人間の音声を与えることにより、ロボットの発話動作の強化学習や声真似も可能である。

(2) 発話動作の獲得

本ロボットは、自己組織化ニューラルネットワーク(SONN)を用いて、発話動作を自律的に獲得する。学習は、上向きに自己組織化学習(SOM)、下向きにバックプロパゲーション法による3層パーセプトロン学習を使用した。学習フェーズでは、上向き学習でロボットによりランダムに発声された音声の特徴マップにマッピングされる。その後、下向き学習により入力音声から抽出された音響パラメータとモータ制御パラメータが関連付けられる。学習後にターゲットとなる音声を入力すると、競合層のパターンにより出力層から声道形状が逆推定され、その結果ロボットから音声生成される。

また本システムでは上向き学習にSOMを用いることにより、比較的少ない学習回数でマップ上に、音響パターンの類似性をノルム空間として持つ位相構造を作ることができる。つまり、音の特徴が似ていれば近くに、似ていなければ遠くのセルに配置されたマップを得ることができ、パターンの違いに応じて、

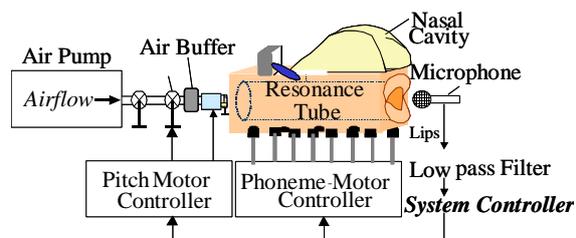


図1 発話ロボットの構成

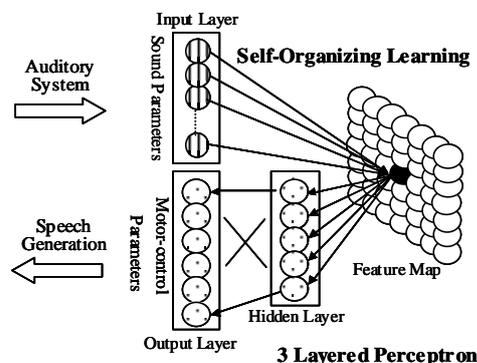


図2 SONNの構成

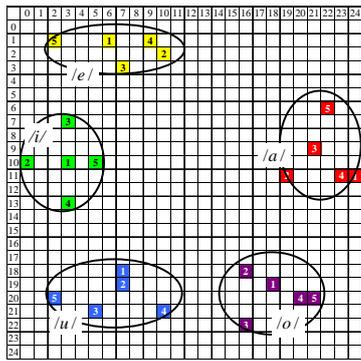
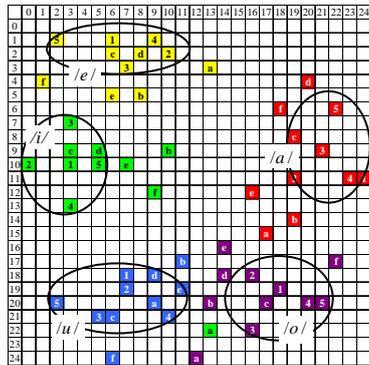


図3 健聴者の日本語5母音マッピング

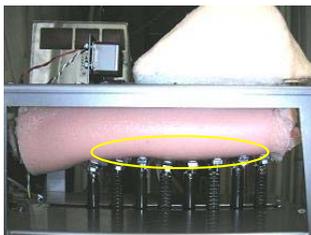
未学習な音もマップ上のセルとして推定することができる。SONN 学習後、健聴者5名にそれぞれ日本語5母音を発声させ、その特徴を特徴マップ上のセルへ配置した。マッピング結果の例を図3に示す。マッピング結果から、各母音の特徴点がカテゴライズされており、音の特徴が似ている母音は近くに、似ていないものは遠くにマッピングされており、良好に位相構造が獲得できていることがわかった。

(3) 脳性麻痺患者音声の不明瞭要因解析

発話ロボットの自律的音声獲得能力を利用し、脳性麻痺患者の不明瞭な音声を入力音声として用い、どのような声道形状が再現されるかを検証した。実験には脳性麻痺患者6名分の5母音の音声データを用いた。例えば脳性麻痺患者 a の母音/i/音声を見ると、健常者母音領域からかなり離れた/u/と/o/領域



(a) 実験後の特徴マップ



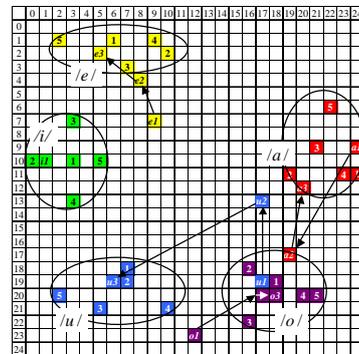
(b) 再現された声道形状
図4 特徴マップと声道形状

域の間にプロットされた。音声を聞いたところ、こもった/i/のような音声で/u/に近いように聞こえた。発話ロボットにより再現された声道形状は、健常者音声/u/と/o/の間のような形状だった。つまりこの患者は、構音動作を良好におこなえていないため、音声不明瞭になることがわかった。また今回の実験で、不明瞭な音声でも声道形状を推定して再現できることがわかった。

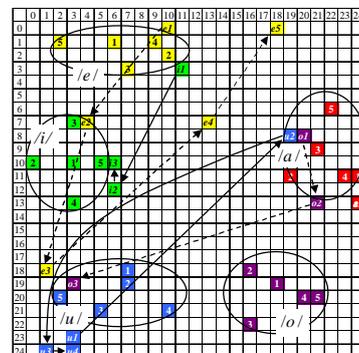
(4) 発話ロボットとの対話型発話訓練実験

発話ロボットを対話型発話訓練装置として応用し、実際に聴覚障害者に対して対話型発話訓練実験をおこなった。被験者は香川県立聾学校の高等部(12名)と中等部(7名)の生徒計19名である。発話訓練では、日本語5母音の発話と構音動作の獲得訓練を目的とした。

訓練の流れは、まず健聴者の音声から得られた理想的な母音の声道形状を目標形状として、被験者に提示する。次に、被験者が同じ音声を発話し、自身の発声音声からロボットが推定した声道形状を目標形状と比較する。更に、違う部分を自ら意識しながら発声し直すことで、発話訓練おこなう。また、ロボットによる声道形状比較に加え、特徴マップを用いる。特徴マップ上には、発声毎にその声の音響的特徴がセル上に特徴点となつて表示される。聴覚障害者が発声した際に示される特徴点と、目標音声領域の距離を参考



(a) 少ない回数で成功の例



(b) 訓練失敗の例

図5 発話訓練の結果例

に、目標領域に近づくよう意識しながら学習を進める。訓練者は、発話ロボットの動作と特徴マップ上の距離を参照しながら、共に目的の形状と領域が得られるまで、発声と観察を繰り返していく。

図5に少ない回数で訓練が良好におこなわれた例と訓練が成功しなかった例を示す。発話訓練結果から、訓練を良好におこなえた被験者は高等部の学生が多かった。聾学校の中でも発話訓練がおこなわれており、中等部の学生に比べ高等部の学生の方が発話訓練の経験が多いため良好に訓練をおこなうことができたと考えられる。このことから、発話訓練の達成度は訓練を繰り返すことにより上昇し、実験でおこなった音声特徴点のディスプレイとロボットの声道形状による視覚的提示が有用であることがわかった。

(5) まとめと今後の課題

本研究では、発話障害者が自分自身で声道形状と声の音響特徴を確認しながら、対話的に発話学習訓練ができるロボットを構築した。発話ロボットが自律的に日本語音声の発話動作を獲得するための学習として、SONNを用いた。上向き学習にSOMを用いることによって、少ない学習でも未学習な音を推定することが可能であることがわかった。SONN学習後、脳性麻痺患者の不明瞭音声を入力として用い、如何に声道形状が再現されるかを検証した結果、不明瞭な音声に対応した声道形状を得ることができると解った。

そこで、聴覚障害者に対して、対話的に発話訓練をおこなうシステムを構築し、聾学校の生徒19名に対して発話訓練実験をおこなった。実験結果から、発話訓練の達成度は訓練を繰り返すことにより上昇し、実験でおこなった音声特徴マップ表示とロボットの声道形状による視覚的提示が有用であることがわかった。

これらの成果を国内外の学会で発表をおこなった結果、ヒューマンコミュニケーション賞、ベストプレゼンテーション賞、New Technology Foundation Award Finalist、Best Presentation Awardなど、6つの賞を受賞した。その他にも国内ばかりでなく海外の各種学会や大学、研究所などから、多くの招待講演や招待講義を依頼されており、本研究の成果は大きなインパクトがあるといえる。

今後は、発話障害者の発話訓練装置の実用化に向けた研究を進めていく。特に被験者から、ロボットの動作によって口内の動きが見られるのは面白いという意見があった一方で、ロボットの声道と口内形状の対応関係が解りにくいという指摘もあった。口内形状をより明確にかつ直感的に理解できるように提示するために、発話ロボットの声道と人間の口内の対応について詳細に解析をおこなっ

ていく必要がある。更に、発話ロボットの音声模倣学習課程を解析することにより、人間の音声獲得機構の解明につなげていく。

5. 主な発表論文等

[雑誌論文] (計14件)

- ① Hideyuki Sawada, Mitsuki Kitani and Yasumori Hayashi: "A Robotic Voice Simulator and the Interactive Training for Hearing-Impaired People", Journal of Biomedicine and Biotechnology, Volume 2008, Article ID 768232, 2008, 査読有
- ② Mitsuki Kitani, Yasumori Hayashi and Hideyuki Sawada: "Interactive training of speech articulation for hearing impaired using a talking robot", International Conference on Disability, Virtual Reality and Associated Technologies ICDVRAT2008, pp.293-301, 2008, 査読有
- ③ Mitsuki Kitani, Yasumori Hayashi, and Hideyuki Sawada: "A Robotic Voice Simulator and the Articulatory Reproduction of Impaired Voices", Mecatronics2008, Paper #203, 2008, 査読有
- ④ Hideyuki Sawada and Minoru Ohkado, "Sensing of Particular Speakers for the Construction of Voice Interface Utilized in Noisy Environment", Journal of Electrical Engineering in Japan, Vol.162, No.3, pp.78-86, 2008, 査読有
- ⑤ **ヒューマンコミュニケーション賞 受賞** (電子情報通信学会 ヒューマンコミュニケーショングループ)、澤田秀之、中村光宏、林恭守、木谷光来:「発話ロボットを用いた聴覚障害者のための発話訓練」, 電子情報通信学会 技術研究報告 HCS2007-24, pp. 125-130, 2007, 査読無
- ⑥ **New Technology Foundation Award Finalist**, Hideyuki Sawada and Mitsuhiro Nakamura: "Mechanical Voice System and its Singing Performance", International Conference on Intelligent Robots and Systems (IROS2007), Proc. of IROS2004, pp. 1920-1925, November 2nd, 2007, 査読有
- ⑦ Hideyuki Sawada and Toshiya Takechi, "A Robotic Auditory System that Interacts with Musical Sounds and Human Voices", Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.11,

- No.10, pp.1177-1183, 2007, 査読有
- ⑧ Hideyuki Sawada, Mitsuki Kitani and Yasumori Hayashi, "Articulatory Reproduction of Voices of Hearing-impaired by a Talking Robot", *Frontiers in the Convergence of Bioscience and Information Technologies (FBIT2007)*, pp.603-608, 2007, 査読有
- ⑨ Hideyuki Sawada, Yasumori Hayashi and Mitsuki Kitani, "A Robotic Mechanical Voice Simulator to Train Auditory Impaired People", *Malaysia Japan International Symposium on Advanced Technology (MJISAT2007)*, MJISAT-319_SS, 2007, 査読有
- ⑩ Mitsuki Kitani, Yasumori Hayashi and Hideyuki Sawada, "A Talking Robot and the Reproduction of Human Voice", 14th Tri-University International Joint Seminar and Symposium, pp. 347-352, 2007, 査読有
- ⑪ Atsushi Todo and Hideyuki Sawada, "Estimation of musical pitch by using comb filters for the identification of musical instruments", *SICE Annual Conference 2007*, pp. 660-663, 2007, 査読有
- ⑫ 澤田秀之、大加戸稔:「雑音環境下における音声インタフェース構築のための特定話者のセンシング」, *電気学会論文誌D*, Vol.126, No.11, pp.1446-1453, 2006, 査読有
- ⑬ Mitsuhiko Nakamura and Hideyuki Sawada, "Talking Robot and the Analysis of Autonomous Voice Acquisition", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2006)*, pp. 4684-4689, 2006, 査読有
- ⑭ Hideyuki Sawada, Atsushi Todo and Toshiya Takechi, "Realtime Speaker Tracking for Robotic Auditory System", *Annual Conference of the IEEE Industrial Electronics Society (IECON2006)*, pp.5474-5479, 2006, 査読有

[学会発表] (計12件)

- ① **Invited Talk**, Hideyuki Sawada, "A Talking Robot and the Interactive Speech Training for Hearing-Impaired", *General Meeting of the Acoustical Society of America*, May 18th, 2009, Portland, USA
- ② **IEEE Best Presentation Award 受賞**,

- Mitsuki Kitani and Hideyuki Sawada: "Proposal of 3D SOM for Autonomous Voice Acquisition of a Talking Robot", *電気関係学会 四国支部大会 講演論文集*, p.370, 2008年9月27日、徳島大学
- ③ Atsushi Todo and Hideyuki Sawada: "Active sensing of sound location using a robotic arm": *電気関係学会 四国支部大会 講演論文集*, p.399, 2008年9月27日、徳島大学
- ④ Hideyuki Sawada: "Active Tracking of Particular Person Using Visual and Auditory Information", *International Symposium for the Establishment of Sustainable and Recycling-based Society*: p.21, 2008年10月16日、香川大学
- ⑤ 東藤篤史、澤田秀之:「ロボットアームを用いた音場の能動センシング」, *計測自動制御学会 第9回システムインテグレーション部門講演会 (SI2008)*: pp.657-658, 2008年12月6日、岐阜 長良川国際会議場
- ⑥ **IEEE Best Presentation Award 受賞**, Yasumori Hayashi and Hideyuki Sawada, "Analysis of Acquired Vocal Sounds by a Talking Robot": *電気関係学会 四国支部大会 講演論文集*, p.335, 2007年9月29日、徳島大学
- ⑦ **招待講演**, 澤田秀之:「発話ロボットによる発話動作の再現と肢体不自由児の発話訓練への応用」, *第52回西日本肢体不自由児施設運営研究会*, 2007年9月14日、高松リーガホテルベスト
- ⑧ **Invited Lecture**, Hideyuki Sawada, "Evolution of Talking Robot", *Universiti Teknologi Malaysia*, 2007年11月19日、マレーシア工科大学 (マレーシア)
- ⑨ Takuma Ise and Hideyuki Sawada, "Extraction of voice characteristics for speech training of auditory impaired people", *電気関係学会 四国支部大会: 講演論文集*, p.343, 2007年9月29日、徳島大学
- ⑩ **ベストプレゼンテーション賞 受賞**, 中村光宏、林恭守、澤田秀之:「発話ロボットの発話動作獲得と歌声の生成」, *情報処理学会 音楽情報科学研究会 夏のシンポジウム: 情報処理学会研究報告 2006-MUS-66*, pp. 135-140, 2006年8月8日、香川 ヴィラ塩江
- ⑪ **IEEE Presentation Award 受賞**, Mitsuhiko Nakamura and Hideyuki Sawada, "Voice Acquisition of a Talking Robot Using a Self-organizing Map", *電気関係学会 四国支部大会: 講*

演論文集、p.343, 2006年9月26日、愛媛大学

- ⑫ Takuma Ise and Hideyuki Sawada, "Acoustic analysis of esophageal speech", 電気関係学会 四国支部大会: 講演論文集、p.344, 2006年9月26日、愛媛大学

[図書] (計1件)

- ① Hideyuki Sawada, "Talking Robot and the Autonomous Acquisition of Vocalization and Singing Skill", Chapter 22 in Robust Speech Recognition and Understanding, Edited by Grimm and Kroschel, I-Tech Education and Publishing, pp.385-404, June 2007, 460 pages, ISBN: 978-3-902613-08-0

[その他]

澤田研究室ホームページ

<http://www.eng.kagawa-u.ac.jp/~sawada/>

6. 研究組織

(1) 研究代表者

澤田 秀之 (SAWADA HIDEYUKI)

香川大学・工学部・准教授

研究者番号: 00308206