

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 30 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26750219

研究課題名(和文) 残存身体機能を活かす発話障害者音声支援装置に関する研究

研究課題名(英文) Research on a speech generating device utilizing remaining physical function for speech disorders.

研究代表者

藪 謙一郎 (Yabu, Ken-ichiro)

東京大学・高齢社会総合研究機構・特任研究員

研究者番号：50626215

交付決定額(研究期間全体)：(直接経費) 1,200,000円

研究成果の概要(和文)：本課題は、代表者らが提案する、「舌の動き」を「指の動き」に置き換えて、操作盤面をペンや指でなぞって、リアルタイムに楽器のように連続な音声を生成する方式に関して、試作によりアルゴリズムの改良と提案を段階的に進めた。

まず、ペン入力操作の速度を補うようなアルゴリズムの考案により、初心者でもホルマント軌跡を再現しやすく改良した。また、任意話者の数秒の短母音の録音音声から、その声質を再現した母音音声を自由に生成するアルゴリズムを提案し実装した。さらに、構音のみに障害を持つ者が、残存する不明瞭な声を出しながらペン入力で音韻を加えるシステムを提案し、話者に近い声で元の韻律を保った声を生成できることを示した。

研究成果の概要(英文)：We improved and proposed algorithms step by step with a prototype of a device which substitutes "motion of a finger" for "motion of a tongue" and in which continuous speeches are generated in real time like a musical instrument by the input with a pen or a finger on an operating panel plane.

First, we designed and improved an algorithm of the device with which a velocity of pen-input operation is complemented so that beginner users could also generate formant transitions well. And, we proposed and implemented an algorithm by which any continuous vowel sounds are generated with vocal quality similar to the short vowel sound for several seconds recorded by any speakers. Furthermore, we proposed the system which makes clear an inarticulate voice which remains in a person only with articulation disorder by inputting a motion of a pen, uttering voice. And we showed that the prototype system could generate voices similar to a speaker, keeping original prosody.

研究分野：福祉工学

キーワード：音声生成 ホルマント 構音障害 発話障害 声質 スマートフォン

1. 研究開始当初の背景

発話の障害を支援するための音声支援機器の従来型的方式として、文字を入力したり絵を選択したりして内容を機器に入力し、音声出力する方式がある。このような方式は、意思疎通内容の確実さから、現存の機器の主流となっている。一方で、入力が完了するまでの数秒間は音声出力が開始できず、また、自由にリズムや抑揚をつけたり、文字にならないような中間的な音を表出したりすることができないといった課題があり、日常の円滑なコミュニケーションの障壁となる。そのため文字入力時間の短縮や抑揚の記号入力手段などの改良研究が国内外ともに多く見られる。

代表者らは、従来型とは異なる方式として、「舌の動き」を「指の動き」に置き換えて音声合成のパラメータを制御する音声支援の入力方式を提案してきた^{①,②}。これは、利用者がタッチパネル上を指やペンでなぞると、アナログ楽器のように即時に音が生成され、自由な間(ま)やリズム、抑揚のついた声を生成できる方式である。基本原理は、ヒトの母音の知覚に関わる2つのホルマント周波数、第1ホルマントと第2ホルマントとを図1(a)に示すように、縦軸と横軸に配置し、その操作盤面上をペンやタッチパネルでなぞることで、音声を制御するものである(図2、図3)。

図1(b)に示すように、ヒトが話す際に子音部分に現れる「ホルマント遷移」を、なぞる軌跡として再現させることにより、母音だけでなく子音に近い音も生成可能であることが分かってきており、この原理を発展させた試作器の改良を行ってきた。具体的には、疑似子音を付加する方法、簡易的な抑揚付加の方法^③を提案し、メロディのついた歌声生成を実現^④した。その成果の一部は、スマートフォンで動作するアプリにも実装・製品化され^⑤、一般のユーザが利用可能としてきた。一方で、本方式では、タッチパネル上をなぞる指(またはペン)の速度が遅いと声の明瞭さが低くなるといった課題があった。また、製品化されたアプリの利用者から、声質をもっと明るい声にしてほしいという声があり、改善が求められていた。また、本方式の声帯音源と構音の機能を分離して扱えるという特長を発展させれば、例えば、構音のみに障害を持つ者が、自らの声を利用しながら音韻を付加するといったことも実現可能でありさらなる改善が求められた。

2. 研究の目的

前節のような背景から、本課題は、製品化に至ったスマートフォン向けアプリのアルゴリズムの改良を行うと同時に、利用者本人に残された声や身体機能を活用できる新しい方式の開発を目標とした。この過程で、音声の知覚メカニズムや音声合成・認識の研究に貢献を目的とした。

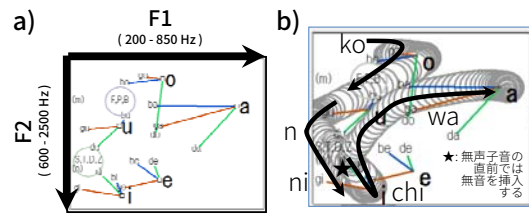


図1 (a)入力操作盤面と (b)「こんにちは」発声時の入力軌跡

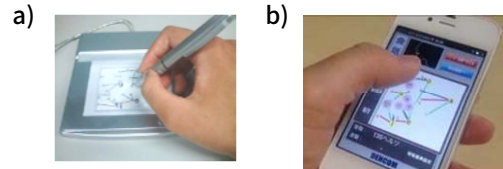


図2 (a)PCのペンタブレットによる入力と (b)スマートフォンアプリの入力画面

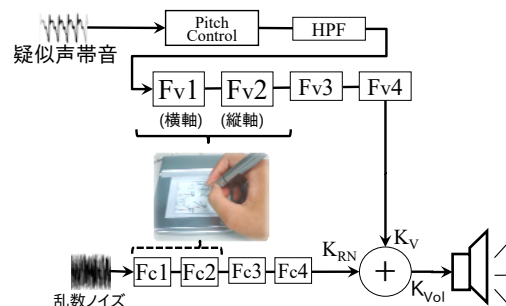


図3 音声生成器のブロック図

3. 研究の方法

本研究では、試作機の改良と生成音の評価を繰り返すことによって、アルゴリズム改良と新しい方式の提案を段階的に進めていった。

(1) 平成26年度には、ヒトの入力操作の速度の限界に基づいて、それを補うようなアルゴリズムの考案により、初心者でも導入しやすいように改良することを目指した。また、実験機において、所望の声質を再現したまま音韻を制御できる方法を考案することを目指した。前者の入力速度を補うアルゴリズムについては、音声パラメータの変化が激しい子音らしい音の再現部分においてのみ速度の補完が特に必要で、母音部分では速度の補完が必要ではないことを利用し、アルゴリズムを改良・実装して音声生成実験を行った。特に先行研究の改良として、入力値の連続な変化軌跡を維持することと、速度補完タイミング検出のためのパラメータ閾値の変更機能を加えた。

(2) 平成27年度には、以上のことを踏まえて、利用者の声質をできる限り維持したままの声を生成することに焦点を絞った。まず、共鳴を再現するホルマントフィルタ群の同定を行う方法を検討したが、同定には音声の

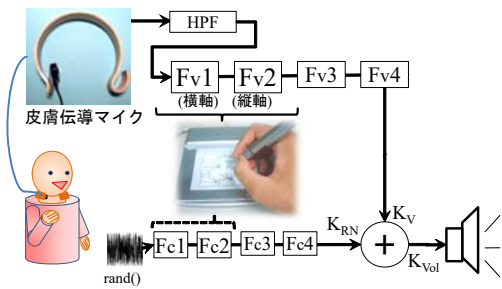


図4 首に装着した皮膚伝導マイクで構音障害者の自身の声を利用するシステム

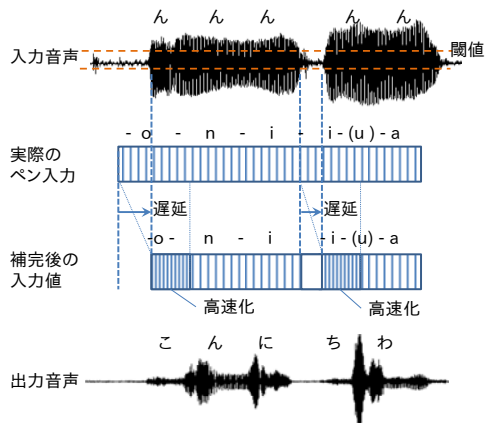


図5 入力音声とペン入力値との同期方法

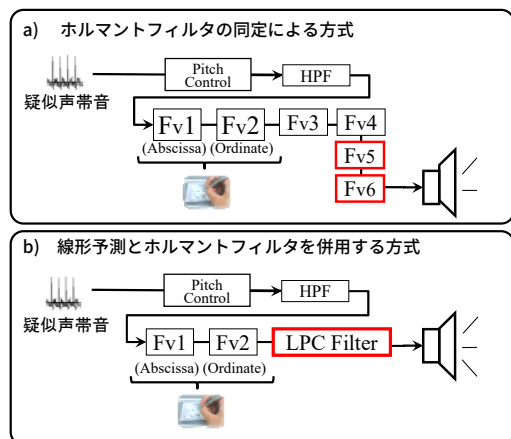


図6 声の個人特性の再現手法
(a) 高次ホルマントの同定による方法と
(b) 線形予測フィルタによる方法

専門家による知見が必要であるとともに、一意に定まりにくいことが分かった。そこで、第1第2ホルマントを除去した音声に対して線形予測フィルタの同定を行うことよって高次の周波数スペクトル包絡を再現する手法を検討した。実験により、一定周波数が続く数秒間のサンプル音声があれば、本手法で個人の声をある程度再現したまま、第1第2ホルマントを自在に移動させることが分かった。

(3) 28年度には、平成26年に行った速度補完による喉頭マイクからの音声との同期アルゴリズムと、平成27年度の個人の声質の再現手法を統合させて、指の動きと本人に残された発声機能とを同時に使うことで、音韻が補完された音声を生成させるように、音声生成器を改良した。すなわち、構音機能に障害を持つけれども発声機能が残されているような利用者が、残存している発声機能を活かして、抑揚や声質を残したまま音韻を付加することを、汎用パソコン上で可能にした。また、平成27年度の声質維持のアルゴリズムを改良し、専門知識のない一般のユーザが利用可能な、簡単に自分自身の声を設定して、音声を生成できるように改良した。結果、数秒間の「あー」の録音音声のみから、声質を再現する母音を生成可能であることを示した。子音を含む音声の明瞭度等に関しては引き続き検討が必要である。

4. 研究成果

(1) 速度補完のアルゴリズムとして、特に子音部分において、実際のペンの速度よりも、高速なF1-F2の変化速度を補う方式を考案・提案した。すなわち、入力ペン速度が閾値を超えたタイミングで、入力値を固定し、ペン速度が閾値以下になるまで次の入力を遅延させたのち、高速なホルマント遷移で現在の入力値に戻すことで、子音部のホルマント遷移を高速化する手法を提案した。これにより、子音部がより聞きやすい音声を再現できたが、入力速度の閾値設定により、再現可能な話速が制限されてしまう課題も見られた。このため、本機能は、初心者練習時の補助機能とするのが妥当と考えている。

また、所望の声質を再現する、初期段階として、発話障害者の「原音生成」の機能が正常に残っている場合を想定し、本人の声を喉頭部分に取り付けた皮膚伝導マイクで取得して原音として利用するシステムを構築した(図4)。これにより、抑揚やアクセントをそのまま反映できることを確認できた。さらにこのシステムで、話者の声の振幅に合わせた疑似子音が付加されるように改良した上で、声入力とペン入力のずれを補正するアルゴリズムを考案した。このアルゴリズムには、速度補正機能と同様の遅延方式を採用し、声よりも先にペン操作が行われた場合には、ペン入力の動きを遅延させて、タイミングをそろえる動作と、ホルマント遷移の高速化を同時に行うことで、子音開始部分を明瞭化できることを示した。(図5)

簡易の音声生成実験により、話速は通常よりもゆっくりした速度であったが、音声とペン操作との同期と自然な抑揚とアクセントで、音声が出力され、有効性を示すことができた。以上から、残された正常な声帯機能を利用するシステムの可能性を示した。

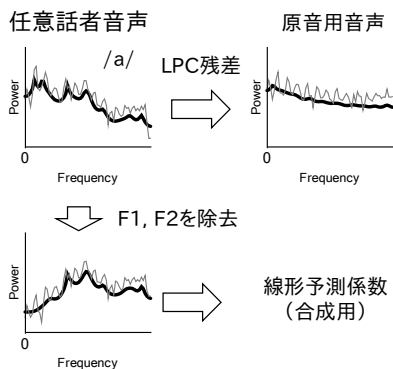


図7 個人性再現のための原音生成方法と線形予測パラメータの推定方法

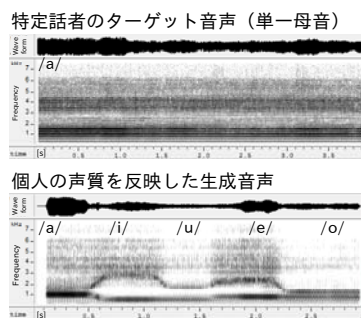


図8 短母音「あ」から個人特性を再現した母音音声

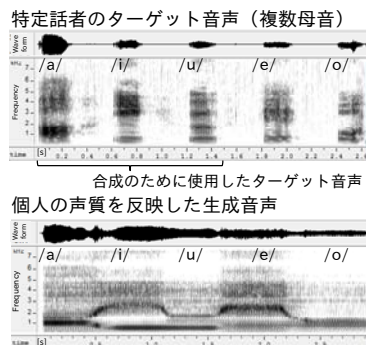


図9 /a/, /i/, /u/の3母音から個人特性を再現した母音音声

(2) 以上の結果を踏まえ、特定話者の声質をできる限り維持したままの声を生成する手法として、ホルマントフィルタ群の同定を行う方法と、線形予測 (LPC) フィルタにより高次のスペクトル包絡全体を再現する手法を検討した (図6)。その結果、前者では、声質の再現が困難であることが分かった。後者では、ペン入力により操作される F1-F2 成分と、声質に関わる包絡成分との分離が課題であったが、LPC 係数算出の前に F1-F2 を除去しておく手法 (図7) により、聴感上で個人特性を維持した声が生成するアルゴリズムを提案し、本人に近い声の生成に成功した。音声生成実験により、一定周波数が続く数秒間の単一母音のサンプル音声があれば、本手法で個人の声をある程度再現したまま、第1第2ホルマントを自在に移動させることができることが分かった (図8)。

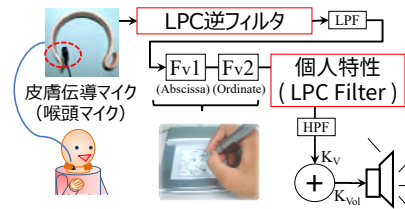


図10 リアルタイムに話者自身の声に対して処理を行うシステム

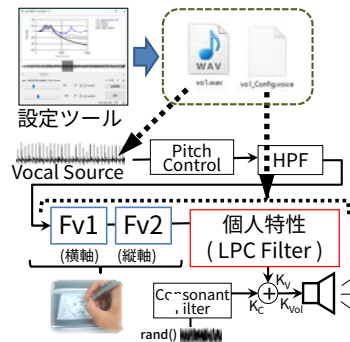


図11 個人特性のパラメータ生成ツール

さらに、線形予測係数の線形補完により、複数母音をターゲットとして、F1-F2 平面上の任意の母音に対しても連続的に音声を生成することが可能となった (図9)。

ただし、子音音声の生成に対する調整方法などについては、今後、検証が必要である。

(3) 最終年度には、(1) で見通しを付けた、本人に残された発声機能に対してペン入力で構音機能を補う方式について、(2) で得られた声の個人特性の再現アルゴリズムと統合させた。すなわち、喉頭部分に取り付けた皮膚伝導マイクからの入力信号に対して、リアルタイムに疑似声帯音を生成し、ペン入力に従った F1-F2 を加えたのち、個人性の特徴付けを行うという一連のアルゴリズムを提案し実現させた (図10)。これにより、構音だけに障害がある話者では、自身が声を出しながらペン操作を行い、その声に対して構音機能を補えるため、抑揚やリズム等の他、本人の声質に近い音声生成でき、表現の幅が大きく広がるものと考えている。

また、(2) で得られた声の個人特性の適用機能について、スマートホンのために誰もが簡単に設定することができるような自動設定アルゴリズムを提案し、汎用 PC 上で検証のための声質設定ツールを実装した (図11)。結果、数秒間の「あー」の録音音声ファイルのみを入力すれば、声質を再現するパラメータを半自動的に設定でき、録音音声に近い母音音声を生成できることを実証できた

今後、上記で得られた知見から、スマートフォンアプリの改良を進め、一般のユーザにも声質再現機能等が利用可能となるようにしていく予定である。

<引用文献>

- ① 橋場 参生, 藪 謙一郎, 井野 秀一, 泉隆, 伊福部 達, “F1-F2 平面で操作する会話補助装置の提案と試作”, ヒューマンインタフェースシンポジウム 2005 論文集, pp. 585-588, 2005.
- ② 藪, 謙一郎 伊福部, 達 青村, 茂, “発話障害者支援のための音声合成器の基礎的設計”, 信学技報, .SP, 音声, 105, pp:59-64, 2006.
- ③ Klatt, D. H., “Software for a cascade/parallel formant synthesizer, Journal of the Acoustical Society of America, 67(3), pp. 971-995, 1980.
- ④ 藪 謙一郎, 伊福部 達, “構音機能障害者のための音声生成器の抑揚制御方式に関する基礎的検討”, 電子情報通信学会 技術研究報告 . SP, 111(225), pp. 43-48, 2011.
- ⑤ 藪 謙一郎, 伊福部 達, “タッチ面上のなぞりで操作する音声生成器—メロディデータに基づく歌声生成機能—”, 電子情報通信学会総合大会講演論文集 2013 年_基礎・境界, pp. 272, 2013.
- ⑥ 藪 謙一郎, 橋場 参生, 高瀬 雅由, 須貝 保徳, 伊福部 達, “指の位置と動きで操作するスマートフォン向け音声生成アプリの応用例の考察”, 日本音響学会研究発表会講演論文集, pp. 1475-1478, 2014

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 1 件)

- ① 藪 謙一郎, 伊福部 達, 拡声器能を備えたウェアラブル電気式人工喉頭の設計, 日本バーチャルリアリティ学会論文誌 21(2), 295-301, 2016/6/30.

[学会発表] (計 5 件)

- ① 藪 謙一郎, 伊福部 達, ペンの位置と動きで操作する音声生成支援器における声質パラメータ設定方法の検討, 日本音響学会 2017 年春季研究発表会, 平成 29 年 3 月 15 日~17 日, 明治大学生田キャンパス (神奈川県川崎市)
- ② 藪 謙一郎, 伊福部 達, 構音障害者自身の声を用いたペンの位置と動きで操作する音声生成支援機器, 日本音響学会 2016 年秋季研究発表会, 平成 28 年 9 月 14 日~16 日, 富山大学 五福キャンパス (富山県富山市)
- ③ 藪 謙一郎, 伊福部 達, 発話障害者のた

めの指の位置と動きで操作する音声生成器における声質再現手法, 日本音響学会 2015 年秋季研究発表会, 平成 27 年 9 月 16 日~18 日, 会津大学 (福島県会津若松市)

- ④ 藪 謙一郎, 伊福部 達, ペンの位置と動きで操作する音声生成器による本人の声を活かした発話補助方法の考察, 音響学会 2015 年春季研究発表会, 2015 年 03 月 16 日~18 日, 中央大学 (東京都)
- ⑤ 藪 謙一郎, 橋場 参生, 高瀬 雅由, 須貝 保徳, 伊福部 達, 指の位置と動きで操作するスマートフォン向け音声生成アプリの応用例の考察, 日本音響学会 2014 年秋季研究発表会, 2014 年 09 月 03 日~05 日, 北海学園大学 (北海道).

[図書] (計 1 件)

- ① 日本音響学会 (編), コロナ社, 音響キーワードブック, 2016 年 03 月, 総ページ数: 494 頁, 分担ページ: pp. 362 - 363 「発話障害者のための支援機器」

[その他]

- ① ゆびで話そうサポートページ(電制)
<http://www.dencom.co.jp/yubihana/yubihana.htm>
- ② ダウンロードページ(iTunes store)
<https://itunes.apple.com/jp/app/id626142603>

6. 研究組織

(1) 研究代表者

藪 謙一郎 (Ken-ichiro Yabu)
東京大学 高齢社会総合研究機構
特任研究員
研究者番号: 50626215