

研究報告書

「脳性麻痺障害者の個人適応型コミュニケーション支援システムの開発」

研究タイプ: 通常型

研究期間: 2015 年 10 月 ～ 2019 年 3 月

研究者: 滝口 哲也

1. 研究のねらい

近年、家庭生活、学校生活、社会生活において様々な機器の情報化が進み、情報機器が身の回りの生活環境にて浸透しつつある。しかし、そのような機器は操作が複雑であり、障害者が使いこなすには困難である場合が多い。福祉の分野では、コンピュータなど情報技術を駆使した、障害者の様々な生活を豊かで便利なものにする支援機器の開発が行われてきている。その過程で、視覚障害、聴覚障害、肢体不自由などの身体障害者に対しては、例えば手話認識、文字の点字化などの研究が行われてきて、これらの支援技術により障害者らの社会的活動への参加、貢献する機会を増大させている。今後ますます情報機器が家庭生活や社会に埋め込まれていくものと考えられる。しかしながらコミュニケーションや発話に障害のある構音障害者が健常者と同じように、発話コミュニケーションをすることは困難であり、情報弱者として取り残されていくものと考えられる。

構音障害者の症状は、その原因により様々存在する。そのような方々の発話コミュニケーションを支援するための音声技術への期待は大きい。本研究課題では、以上のような社会的状況を踏まえ、脳性麻痺(構音)障害者を対象に、ユニバーサルコミュニケーション支援技術の研究を行う。脳性麻痺障害者においては、身体的な障害を伴っている場合もあり、ボタン操作など意図的な動作を行う際には、筋肉の緊張により不随意運動が生じる。そのためボタン操作などによる情報機器の操作も非常に困難な場合がある。また脳性麻痺障害者の発話内容を健常者が聞き取ることが困難な程、障害者の発話スタイルは健常者と異なる。本研究課題では、健常者と障害者が共存した発話コミュニケーション社会の構築を目指し、障害者の身体能力の壁を超越することが出来るコミュニケーションを実現する。

2. 研究成果

(1) 概要

脳性麻痺障害者の発話様式は健常者と異なり、また筋肉の緊張から生じる発話様式の変動が非常に大きくなるため、脳性麻痺障害者の発話内容を健常者が聞き取ることが困難である。さらに脳性麻痺障害者は、健常者と異なり一人一人その障害特性が多様であるため、個人レベルでの適応技術が必要不可欠である。このような問題解決に向けて従来研究では、事前に使用者(障害者)の多量の発話を用いてシステム構築を行っていたが、彼ら/彼女らの身体負担を考慮すると出来るだけ少量の発話のみでシステム構築することが望ましい。少量の障害者発話に対応するため、本研究課題では、使用者(障害者)の発話データのみならず、(I) 健常者、および他の障害者発話データベースとして英語を母国語とする障害者の発話データも利用した新たな枠組みを提案し、脳性麻痺者の発話認識を実現する(研究テーマ 1: 発話認識(図 1))。

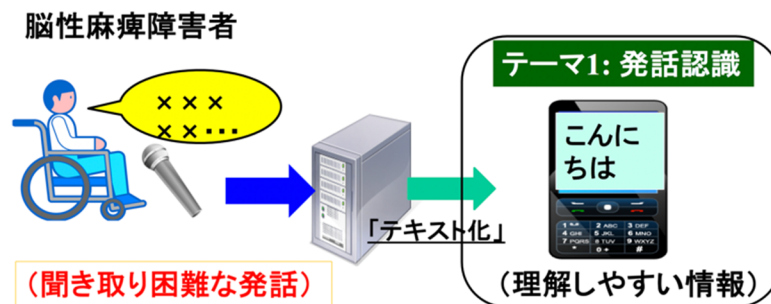


図 1. 研究テーマ1:脳性麻痺者の発話認識

また、聞き取り困難な脳性麻痺障害者の声を健常者へ伝える手段として、声質変換が必要である。障害者からの意見として、出来れば自分自身の声を皆に聞いてもらいたいという要望も存在し、生成される音声は、必ずしも障害者とは全く異なる声質を持った健常者の声にしたいわけではない。よって、(II) 障害者の話者性を保ちながら、聞き取りやすい声に変換する技術確立する(研究テーマ2: 音声変換)。

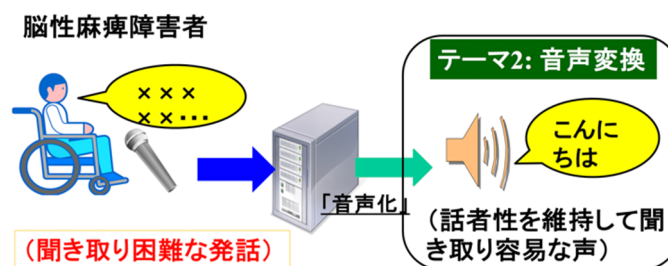


図 2. 研究テーマ2: 脳性麻痺者の話者性を維持して聞き取り容易な声へ変換

(2) 詳細

研究テーマ1:脳性麻痺者の発話認識

- マルチモーダル情報を用いた教師無し特徴量抽出

構音障害者はテキストと同じ発話ができない場合があり、発話データとテキストにミスマッチが生じる。そのような問題に対処するため、教師無し機械学習による特徴量抽出法を提案した。また健常者を対象にしたマルチモーダル発話認識に関する報告はあるが、構音障害者を対象にした研究は、これまで見当たらない。本研究では構音障害者を対象にして、マルチモーダル情報を用いた Deep Canonical Correlation Analysis に基づく教師無し特徴量抽出を行ない、有効性を示すことができた(学会発表 3)。またコミュニケーション支援において重要な要素ともなる顔表情認識についても、新たに回転不変特徴量を提案し有効性を示した(論文発表 2, 4)。

- 複数データベースを使用した発話データ拡張

使用者(脳性麻痺者)の発話データが少ないことに対処するため、日本人構音障害者の音声だけでなく、外国人構音障害者と日本人健常者の音声を用いることによる、新しい発話データ拡張を提案した。図 3 に提案手法の概要を示す。本手法の特徴は、英語障害

者音声と日本語障害者音声を用いて共通の音響モデル(障害者音響モデル)を学習し、日本語障害者音声と日本語健常者音声を用いて共通の言語モデル(日本語言語モデル)を学習するところにある。

図4に脳性麻痺者の発話認識結果を示す。日本語障害者音声だけでは誤り率50.2%であるが、日本語健常者と英語障害者の音声も用いることにより誤り率を23.2%まで改善することができた(学会発表1)。おおよそ目的を達成することができたと言える。

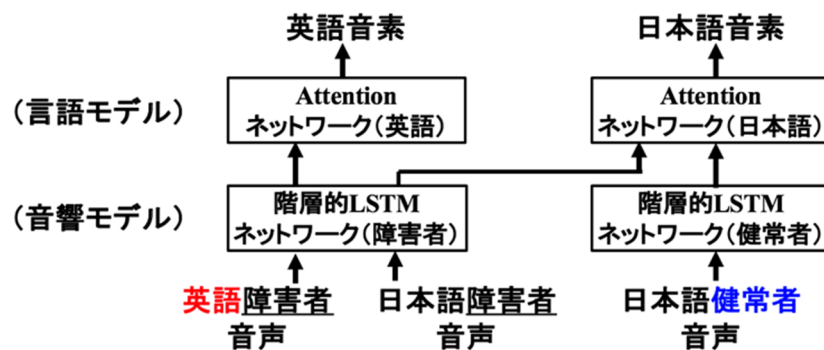


図3. 複数データベースを使用した発話データ拡張

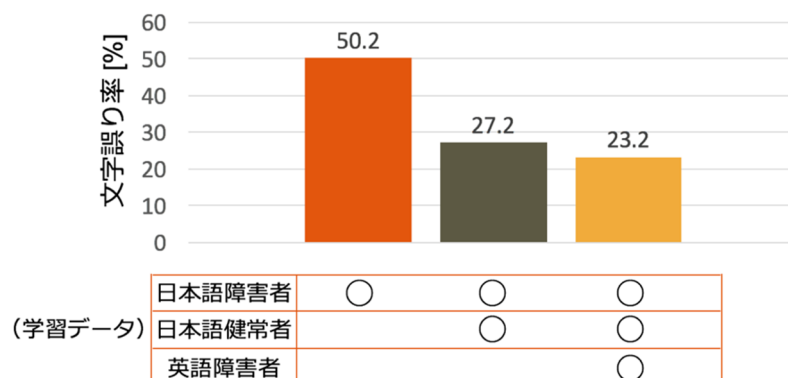


図4. 脳性麻痺者の発話認識結果

研究テーマ2: 脳性麻痺者の話者性を維持した音声変換

- 脳性麻痺者と健常者音声を用いたハイブリッド音声変換

従来の音声変換では、多量の発話データが必要となる。しかしながら脳性麻痺者においては、筋肉の緊張により不随意運動が生じるため、多量の発話収録を行うと健常者と比べて体力的負担がかなり大きくなる。本研究では、脳性麻痺者の少量学習データを用いた新しい音声変換の枠組みとして、(多量の発話データが存在する)健常者の音声合成と、脳性麻痺者の少量発話データによる声質変換システムの組み合わせ手法を提案した(図5)。

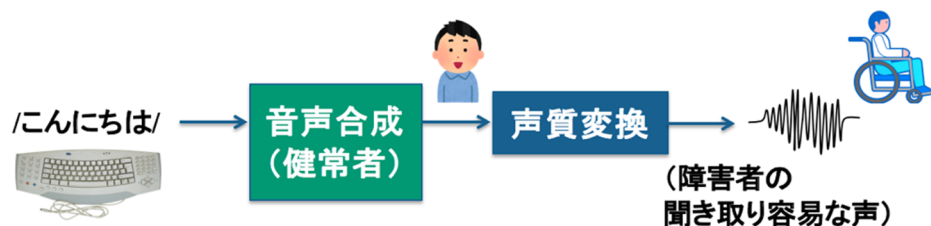


図 5. 音声合成と声質変換を組み合わせた脳性麻痺者少量発話データへの対応

発話認識の箇所で記述したように、構音障害者はテキストと同じ発話ができない場合があり、発話データと教師のテキストにミスマッチが生じる。そのような問題に対処するため、教師無し機械学習による声質変換として、Cycle-Consistent Adversarial Networks に基づく手法を検討した(論文発表 1, 学会発表 2)。脳性麻痺者の発話と比べて、提案手法により生成された音声は、聞き取りやすくなっており、また話者性も若干の劣化は存在するが、(話者性が)保存されていることが示された。従来手法と比較して、学習データ量を 1/5 に削減することもでき、おおよそ目的を達成することができたと言える。

3. 今後の展開

- 脳性麻痺者の発話認識

表 1 に脳性麻痺者の発話認識結果の例を示す。この結果より子音が欠落しやすい傾向が確認できる。脳性麻痺者の音声は、筋肉の不随意運動により特に子音を発音するのが難しく、(脳性まひ者の発話) 音声聞き取りづらくなる。母音は比較的正しく認識できており、モデルがこのような障害者音声の特徴を捉えていると考えられる。得られた誤りの傾向を分析することで、その話者が発話しづらい音の発見に繋がると期待される。さらに、誤り訂正技術を応用することで、より認識精度を向上させられると考えられる。

表 1. 脳性麻痺者の発話認識結果の例

正解	u r a y a- m a sh ii
認識結果	u m a a a a s ii
正解	d a ky ou
認識結果	d a k ou

現在の発話認識の正解率は 100%ではないが、人が聞いても理解するのが困難な状況を考えれば、正解率が 70%から 80%程度でもコミュニケーション支援システムとして活用できる。まずは少ない語彙数のタスクにおいて実用化を目指していく。

- 脳性麻痺者の話者性維持した音声変換

脳性麻痺者の発話と比べて、提案手法により生成された音声は聞き取りやすくなっている。あらかじめテキストを入力することができれば、コミュニケーション支援システムとして活用できる。今後は、Vocoder による音質劣化を改善するために、End-to-End、Vocoder-free な音声合成手法、また Cycle-Consistent Adversarial Networks における損失関数を検討する。さらに、発話認識も組み合わせた End-to-End なシステムへ展開していく。

4. 自己評価

研究目的の達成状況

“発話認識”については、日本語障害者音声だけでは誤り率 50.2%であるが、日本語健常者と英語障害者の音声も用いることにより誤り率を 23.2%まで改善することができた。また、“聞き取り容易な音声への変換”については、脳性麻痺者の発話と比べて、提案手法により生成された音声は、聞き取りやすくなっており、また話者性も若干の劣化は存在するが、(話者性が)保存されていることが示された。学習データ量を 1/5 に削減することもでき、おおよそ目的を達成することができたと言える。

研究の進め方(研究実施体制及び研究費執行状況)

毎年、数名の大学院生を雇用することにより、順調に研究データの収集、解析を進めることができた。また研究費に関しては、物品費(計算機サーバーなど)、旅費(研究成果の学会発表)について適切に執行して、効果的に研究を遂行することができた。

研究成果の科学技術及び社会・経済への波及効果(今後の見込みを含む)

本研究では、多量の学習データを集めるのが困難な課題に対して、対象者の少量学習データにより対応する枠組みを提案した。本研究で対象にした脳性麻痺者だけでなく、別の疾患のコミュニケーション支援においても、同様の少量学習データの課題が存在する。世界中にて高齢社会が問題となっているが、高齢者の明瞭度の劣化した発話や、脳梗塞などの麻痺者発話に関しても、本研究成果の枠組みの適用(波及効果)について大いに期待できる。

5. 主な研究成果リスト

(1)論文(原著論文)発表

1. Jinhui Chen, Zhaojie Luo, Zhihong Zhang, Faliang Huang, Zhiling Ye, Tetsuya Takiguchi, Edwin R. Hancock,
“Polar Transformation on Image Features for Orientation-Invariant Representations,”
IEEE Transactions on Multimedia,
doi: 10.1109/TMM.2018.2856121, 2018. (Jul. 2018)
2. Zhaojie Luo, Jinhui Chen, Tetsuya Takiguchi and Yasuo Ariki,
“Emotional Voice Conversion Using Neural Networks with Arbitrary-Scales F0 based on Wavelet Transform,” EURASIP Journal on Audio, Speech, and Music Processing, 2017,
13 pages. doi:10.1186/s13636-017-0116-2 (Aug. 2017)
3. Jinhui Chen, Zhaojie Luo, Tetsuya Takiguchi, Yasuo Ariki,
“Multithreading Cascade of SURF for Facial Expression Recognition,” EURASIP
Journal on Image and Video Processing, Vol. 2016 (1), 2016:37,
doi: 10.1186/s13640-016-0140-7, 13 pages. (Oct. 2016)
4. Ryo Aihara, Tetsuya Takiguchi, Yasuo Ariki,
“Multiple Non-negative Matrix Factorization for Many-to-many Voice Conversion,”
IEEE Transactions on Audio, Speech and Language Processing, 2016, Vol. 24, No. 7,
pp. 1175–1184. doi: 10.1109/TASLP.2016.2522643 (Jan. 2016)

(2)特許出願

研究期間累積件数:0 件

(3)その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

学会発表

1. 高島悠樹, 滝口哲也, 有木康雄, “複数データベースを使用した end-to-end 構音障害者音声認識” 日本音響学会 2019 年春季研究発表会講演論文集, 2-9-4.
(2019 年 3 月)
2. 南坂 竜翔, 滝口 哲也, 有木 康雄, “CycleGAN に基づくノンパラレル声質変換を用いた構音障害者音声合成” 日本音響学会 2018 年秋季研究発表会講演論文集, 1-R-41, pp. 1185-1188, 2018. (2018 年 9 月)
3. Yuki Takashima, Tetsuya Takiguchi, Yasuo Ariki, Kiyohiro Omori,
“Audio-Visual Speech Recognition for a Person with Severe Hearing Loss Using Deep Canonical Correlation Analysis,” International Workshop on Challenges in Hearing Assistive Technology (CHAT), 2017, pp. 77-81. Stockholm, Sweden (Aug. 2018)
4. 滝口 哲也, “発話障害者の音声コミュニケーション支援”
招待講演, 電子情報通信学会音声研究会 (2018 年 10 月)

プレスリリース

1. 神戸新聞 5 月 15 日朝刊 特集記事「AI 駆使し発話を支援」(2017 年 5 月)

以上