

脳性麻痺構音障害者の音声認識による情報家電操作の検討*

松政宏典, 滝口哲也, 有木康雄 (神戸大・工), 李義昭 (追手門大), 中林稔堯 (神戸大・発達)

1 はじめに

情報技術が向上し, 近年, 福祉分野での情報技術の重要性が増している.

重度障害者の自立支援機器として, 環境制御装置は重要な役割を担っている. これまでにも, 幾つかの環境制御装置が開発, 実用化されている. 従来の環境制御装置では操作方法として, 押しボタンスイッチを用いたものや, 呼気を用いたもの, 音声認識を用いて操作を行うものもある. 手足が不自由な重度障害者において音声認識は非常に重要である.

我々は, これまでに構音障害者の音声認識の実現に取り組んできた [1]. また, 文献 [2] では, 重度障害者の自立移動を支援するために構音障害者の音声認識を用いた電動車イスの開発に取り組まれている. しかし, 言語障害者などの障害者を対象としたものは非常に少ない. 現在, 日本だけでも構音障害者も含めた言語障害者が 3 万 4000 人もいることから十分なニーズがあり, さらなる研究の必要性がある.

従来の音声認識を用いた環境制御装置では, 健常者の音声を用いて音響モデルを作成している. 構音障害者の場合には, 発話スタイルが健常者と大きく異なるため, 認識が困難である. そこで, 本稿では構音障害者の音声認識が可能な環境制御装置の実現に向けて, 構音障害者の音響モデルを作成した.

本稿ではアテトーゼ型の脳性麻痺による構音障害者を対象とした. アテトーゼ型とは脳性麻痺患者の約 10~15% に発生する症状であり, 知能障害を合併していないケースや比較的知能障害の程度が軽いケースも多いのが特徴である [3].

環境制御装置「みてら」[4] を用いて音声認識評価を行い, 日々の発話変動に対する改善として音響モデルの適応を行った.

2 音声認識システム

従来の健常者の音声を用いて作成した音響モデルでは, 発話障害者と健常者との発話スタイルが大きく異なるため, 発話障害者の音声認識は非常に困難である. しかし, 構音障害者の音声を用いて作成した音響モデルを用いることによって, 認識精度の大幅な改善が得られる [1].

そこで, 本稿では構音障害者の音響モデルを事前に作成し, 音声認識に用いた. 特徴量抽出, および音声認識には julian [5] を用いた. 図 1 にシステム全体の流れを示し, 図 2 に使用した評価装置を示す. 実際の操作方法は, 利用者 (重度障害者) は音声認識

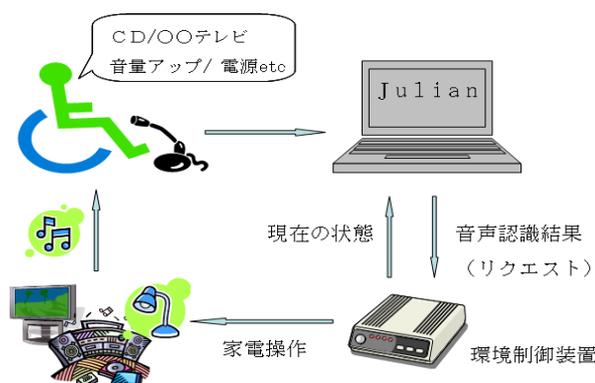


Fig. 1 システム全体の流れ



Fig. 2 実際の評価装置

部 (PC) の支持画面に従って, 操作コマンドを発話 (入力) する. 操作可能なコマンドは環境制御装置の操作画面の状態に従う. 認識結果によっては, 状態遷移が生じる. 図 3 に状態遷移例を示す. 認識に関しては, 各状態ごとに認識文法を変更し, シーン内のコマンドのみを認識結果として出力する.

3 認識実験

3.1 実験条件

今回の認識実験では表 1 に示した 7 つのシーン (機器選択, テレビ操作, CD 操作, 照明操作, チャンネル操作, 番組表操作, メニュー操作), 41 個のコマンドからなるタスクに対して, 459 発話 (各コマンド 8~15 回) を構音障害者 1 名から収録した. 認識実験は, 各シーン毎に行った. 機器選択シーンの「戻る」と TV 操作シーンの「戻る」のように, 各シーンで共通のコマンドに関しては, 一つの発話を重複して使用した.

収録データのサンプリング周波数は 16 kHz である. 音響モデルは monophone (3 状態) を用い, フレーム周期は 10 ms, ハミング窓長は 25 ms とした. 特徴量は 25 次元 (12 次 MFCC+ 12 次 MFCC+ power) を用いた.

*A Study on Operation for Home Information Appliances Using Speech Recognition of a Person with Articulation Disorders. by Hironori MATSUMASA, Tetsuya TAKIGUCHI, Yasuo ARIKI (Kobe Univ.) and Ichao LI (Otemon Univ.) and Toshitaka NAKABAYASHI (Kobe Univ.)

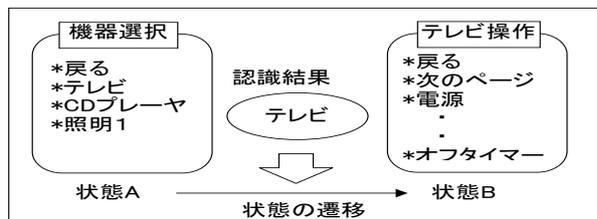


Fig. 3 状態の遷移例

Table 1 タスク内容

状態 1	コマンド名	状態 2	コマンド名	状態 3	コマンド名
機器選択	戻る TV CD 照明 1	TV 操作	戻る 次のページ 電源 CH 操作 CH アップ CH ダウン 音量アップ 音量ダウン 番組表操作 メニュー操作 消音 音声切替 オフタイマー	CH 操作	戻る 次のページ NHK サンテレビ 毎日放送 テレビ大阪 ABC テレビ 関西テレビ 読売テレビ NHK 教育
		CD 操作	戻る 電源オフ 再生する 一時停止 停止 次の曲 前の曲	番組表操作	戻る 番組表 上へ移動 下へ移動 右へ移動 左へ移動 決定
		照明操作	戻る 照明をつける 照明を消す 豆球にする 明るさを調整する	メニュー操作	戻る メニューの表示 上へ移動 下へ移動 右へ移動 左へ移動 決定

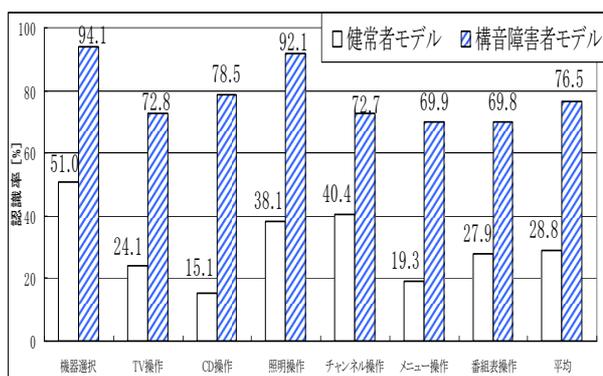


Fig. 4 実験結果

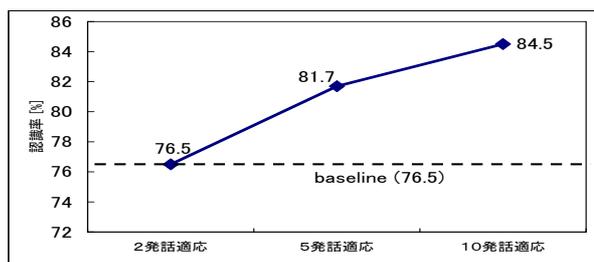


Fig. 5 適応による認識率の変化

3.2 健常者モデルの認識実験

初めに、健常者の音声を用いて作成した不特定話者モデルでの認識実験を行った。認識実験には「Julian 記述文法音声認識キット v3.1」[5]に含まれている音響モデル(43音素16混合)を使用した。実験結果を図4に示す。平均認識率は28.8%と低く、従来の環境制御装置の操作は困難であることがわかる。

3.3 構音障害者音響モデルでの認識実験

健常者の音声を用いた音響モデルでは、認識が困難なため、構音障害者の音響モデル(54音素6混合)の作成を行った。モデル作成には事前に収録した1050単語の音声データを用いた。実験結果を図4に示す。構音障害者の音声を用いた音響モデルを用いることで、平均認識率が28.8%から76.5%まで改善した。評価データと学習データの収録には時期差があり、かつ構音障害者の場合には、日々の体調や環境によって発話スタイルの変動が大きいと考えられる。次に、MLLR法とMAP法[6]を用いて音響モデルをテストデータに適応させた結果を図5に示す。適応データ数は2, 5, 10発話と変化させて実験を行った。10発話を用いて適応を行う事により84.5%まで認識率が改善された。

4 おわりに

構音障害者の音声認識が可能な環境制御装置の実現に向けて、構音障害者の音響モデルを作成し、音声認識評価を行った。従来の健常者音声を用いた音響モデルの場合では、平均で28.8%しか認識が出来ず、操作が困難であった。しかし、構音障害者の音声を用いて、音響モデルを作成したところ、84.5%まで認識精

度が改善した。

今回は、音声のみの評価であったが、音声インターフェイスにおける非コマンド発話への対応も必要である。非コマンド発話への対処方法に関しては、タスクベリフィケーションを行い、コマンドと非コマンドの区別を行うことが考えられる。今後は、そのような課題への取り組みなども行っていきたい。

参考文献

- [1] 松政宏典, 滝口哲也, 有木康雄, 李義昭, 中林稔堯, “ 構音障害者の音声認識の検討, ” 日本音響学会講演論文集(春), 1-Q-30, pp. 323-324, 2007.
- [2] 児島宏明, 佐宗晃, 石洙永, 李時旭, 佐土原健, “ 重度障害者のための音声認識電動車いすの開発, ” 日本音響学会講演論文集(春), 2-8-18, pp. 371-372, 2007.
- [3] S.Terry Canale, 落合直之, 藤井克之, “ キャンベル整形外科手術書 第4巻 小児の神経障害/小児の骨折・脱臼, ” エルゼビア・ジャパン, 2004.
- [4] 三菱電機コントロールソフトウェア株式会社 “ 環境制御装置 みてら, ” <http://www.mcr.co.jp/product/fukushi/mitera/mitera.htm>
- [5] “ 大語彙連続音声認識システム Julius, ” <http://julius.sourceforge.jp/index.htm>
- [6] S. Young et. al., “ The HTK Book, ” Entropic Labs and Cambridge University, 1995-2002.