

深層学習を用いた センサグローブによる指文字認識の改良

土屋智彦^{†1} 白石優旗^{†1}

概要: 我々は、聴覚障害者と聴者の円滑なコミュニケーションを実現するため、深層学習を用いたセンサグローブによる指文字認識システムを開発してきた。しかし、識別対象は平仮名 46 文字と限定されており、かつ識別率も約 83% に留まっているという課題があった。そこで、本論文では、センサからの取得データの安定化、学習データの増加、並びに、複数の学習モデル (Random Forest, Multi-layer Perceptron, Convolutional Neural Network) を用いた比較評価実験により、識別率の向上を目指す。具体的には、開発システムの改良として、サンプリング周波数と識別入力次元数の見直しを行った。識別については、識別率向上のための予備実験として、識別対象に濁音・半濁音・拗音・長音を加えた全 76 文字を識別対象とし、すべての採集データをシャッフルの上で評価実験を行い、その結果について解析を行い、課題を明らかとした。

キーワード: 手話, 指文字, 深層学習, 機械学習, コミュニケーション支援, 聴覚障報保障システム

1. はじめに

近年、音声認識, 及びその関連技術についての研究が進み, 音声による入力機能を備えた情報機器が広く普及してきている。それに伴い, 音声認識による聴覚障害者のための情報保障システムとして, 「こえとら」[1], 「UD トーク」[2], 「Cloud speech-to-text」[3] などのアプリやサービスが公開されている。これにより, 聴者の音声を聴覚障害者が読み取ることが可能になり, 両者のスムーズなコミュニケーションの成長が進んでいる。

一方, 聴覚障害者が発する手話を聴者が読み取ることができるようになるための, 手話認識による情報保障システムについての研究がいくつか報告されている[4-15]ものの, 音声認識による情報保障システムと比較して, 実用レベルに達しているとは言い難い。なお, 聴覚障害者同士の日常会話では, 主要なコミュニケーション手段として手話が使われている。

手話言語は音声言語とは異なる特徴を持つ言語であり, 聴者が手話言語を習得し, 読み取りができるようになるのは容易ではない。したがって, 手話を音声情報, あるいは文字情報に変換し, 聴者に対する情報保障, すなわち手話認識システムが求められている (図 1)。

手話認識システムを実現するためには, 手の位置, 方向, および動きを含む手の形を認識する必要がある。現状の手話を認識する方法は, 大きく分けて非接触式センサであるカメラによる認識[5-6]と, 接触式センサであるセンサグローブによる認識[7-8]とに大別できる。例えば, Luzhnica ら[7]が提案したセンサグローブによる手話認識では識別精度は 98.5% 達成したと報告されているものの, 識別対象クラスは 30 程度であり, 実用化されるまでには至っていない。

一方で, 近年深層学習という技術が注目されている。ここで深層学習とは, 機械学習の一種であるニューラルネットワークの隠れ層を増やしたものであり, 識別率向上に貢献している。例えば, 画像認識によるハンドジェスチャー認識精

度の向上のために, 深層学習を適用した手法[4]が報告されている。

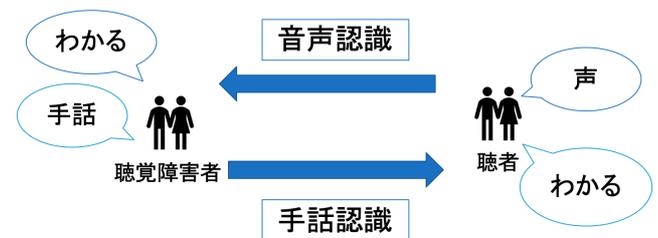


図 1 情報保障システム

本研究では, 聴者と聴覚障害者のコミュニケーションを円滑にするための手話認識への第一歩として, 日本語対応の指文字 (Japanese Finger Spelling, JFS) 入力インターフェースの実現を目指す。ここで JFS とは, 平仮名に 1 対 1 対応した手指の表現のことである。

センサには, 接触式センサであるセンサグローブを採用する。非接触式センサであるカメラを使用する場合は, カメラに手が映るように設置する必要がある, またカメラは環境の影響を受けやすいため, 外出先などでの利用が困難である。一方, センサグローブなど接触式センサによる手形状認識は手に装着するだけでできるため, 利用が容易である。

我々は, センサグローブの重量とコストを抑え, 着用者が手を動かしやすい手法である導電繊維編み込み手法[17]を採用し, 従来の指文字認識手法に深層学習を用いることで識別精度の向上を目指している (図 2) [18]。しかし, 識別対象は平仮名 46 文字と限定されており, かつ識別率も約 83% に留まっているという課題があった。

そこで, 本論文では, 深層学習による JFS 認識手法の改良, 並びに装置としての認識システムの改良を同時に行う。

^{†1} 筑波技術大学
Tsukuba University of Technology

具体的には、センサからの取得データの安定化、学習データの増加、並びに、複数の学習モデルとして Random Forest (RF), Multi-layer Perceptron (MLP), Convolutional Neural Network (CNN) を用いた比較評価実験を行う。比較評価実験については、識別率向上のための予備実験として、識別対象に濁音・半濁音・拗音・長音を加えた全 76 文字を識別対象とし、すべての採集データをシャッフルの上で評価実験を行い、その結果について解析を行う。

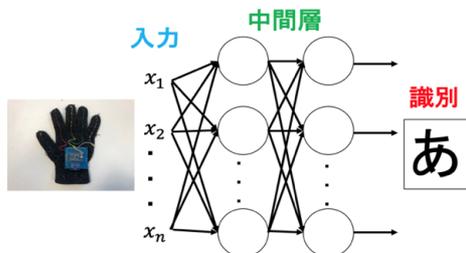


図 2 ニューラルネットワーク

2. 関連研究

これまでの指文字認識に関する研究には、カメラセンサを用い、指文字の一連の動作を画像にし、認識を行う方法と、センサグローブなどを用いた接触式センサを用いる方法がある。以下に、それぞれの方法について述べる。

2.1. 画像認識による認識

カメラを用いて指文字を撮影し画像処理を行い手形認識する手法が複数提案されている。原田ら[9]は指文字を撮影したカラー画像を 2 値化して手の領域を抽出し、手形を認識する手法を提案している。三宅ら[10]が提案する手法は距離値を格納した距離画像に対して画像処理を行い非接触で認識を行う方法である。長嶋ら[12]は撮影された指文字の画像から「手首・拳の方向、孔の有無、指の本数」の情報を取得し、分類アルゴリズムとして決定木を採用し、日本語の静的指文字 (1 つの文字の表現時に手指を動かさない指文字) 41 文字の判別を行っている。このシステムにより、一人の指文字画像データ (41 文字×7 セット) に対して、96.5%の認識率が得られている。Hosoe ら[13]は認識アルゴリズムに深層学習を採用し、日本語の静的指文字の認識を行っており、認識率は 93%と報告されている。

このように、カメラを用いた非接触式センサにより指文字を認識する方法は、手の平にデータグローブなど装着する必要がなく、自然な指文字で比較的高い精度で認識されている。Jalal ら[5]は深層学習を用いて、アメリカ手話 (American Sign Language, ASL) の指文字の画像の識別率は 99%と報告されている。しかし、JFS において実用化のためには認識精度はまだ十分とは言えない。また動的指文字 (1 つの文字の表現時に手指を動かす指文字) についての識別結果はほとんど報告されていない。

2.2. センサグローブによる認識

センサグローブなどの接触式センサを用い、得られた計

測データを元に手形状認識する手法が複数提案されている。例えば、センサにより 5 本指の屈折、手の位置・方向の計測を行い、そのデータを PC やマイコンに計測データを送り、分類アルゴリズムにより手形状を認識する[14]というものである。Cabrera ら[14]は、5DT Data Glove 5 Ultra[16]と加速度センサを組み合わせ、各指の屈曲度と手首の向きに関する情報を取得している。得られた測定値は、オフラインかつニューラルネットワークにより、「J」と「Z」を除いた 24 個の American Sign Language (ASL) の静的指文字の分類のテストを行なった。ニューラルネットワークに 5,300 パターンで訓練を行い、1,200 のテストパターンで 94.07%の精度が得られたとしている。Mumjadi ら[15]は、手袋に複数の小さな慣性センサーユニット (IMUs) を埋め込んだセンサグローブプロトタイプを提案している。57 人のフランス手話 (LSF) の指文字のデータを集め、テストを行なった結果、提案されたシステムにより平均認識率 92%、F 値 91%を得られたとしている。JFS 認識を行った手法の中で導電繊維編み込み手法[17]では曲げセンサの代わりに導電繊維が編み込まれている手袋自体をセンサとして利用し、更に方向ジャイロセンサを組み込む事で手形状と手の動きを認識する手法を提案している。ユーリッド距離による指文字 (「あ」～「お」) の認識率は 60%と報告されている。

3. 開発システムの概要

本研究では、いつでもどこでも円滑なコミュニケーションを実現するため、軽量で着用しやすいセンサグローブを用い、指文字の動作を行うと同時にリアルタイムで高い精度により識別し情報伝達するシステムを設計する。

認識システムは、センサ値計測部と識別部で構成されている。本研究で開発した JFS 認識システムの外観を図 3 に示し、ソフトウェアの全体構成を図 4 に示す。



図 3 改良したプロトタイプ

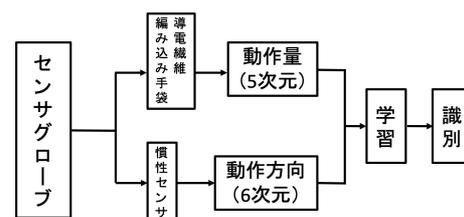


図 4 ソフトウェア全体構成

3.1. センサグローブ

手指と手首のそれぞれの動きのある指文字を効率良く認識するために、センサグローブを用いることで動作量と動作方向を検出して認識を行う。

本研究では、長時間身につけていて負担が少ない、かつ安価であり、比較的軽量である導線繊維編み込み手法 [17] を採用する。動作方向はジャイロセンサを用い、動作量はグローブに編み込まれた導線繊維の抵抗の変化を利用し、指の動作量と手の動作方向を検出する。

具体的には、動作検出ボードには Arduino を採用し、PC と I2C 接続を行うことで検出されたセンサグローブの計測値を Arduino から PC に転送し、CSV 形式で保存する。機械学習並びに動作認識は PC 上で Python 言語を用いて行う。

JFS の動作のデータグローブのセンサ取値は、JFS を行う人によって異なるスケールを有する。したがって、個人の手の動きの違いを考慮した線形正規化を行う。その際、本システムの活性化関数、尤度関数は確率を扱うため、ネットワーク入力の前処理として、0~1 にスケール変換を行う。

動作量は指の曲げ伸ばしによる導電繊維の抵抗の変化を検出している。その際、分圧式を用いて、入力電圧を算出している。

$$V_{in} = \frac{R_1}{R_1 + R_2} * V_{out} \quad (式1)$$

ここで、 R_1 は導電繊維による可変抵抗、 R_2 は固定抵抗である。指を伸ばした時は導電繊維の抵抗の大きさが大きくなる。逆に、指が屈めた時は導電繊維の抵抗の大きさが小さくなる。

3.2. 識別アルゴリズム

本研究では、先行研究 [18] で識別結果が良かった RF, MLP, CNN の 3 つの学習モデルを採用する。RF にはオープンソースのライブラリ scikit-learn [19] を用い、MLP, CNN にはオープンソースライブラリ TensorFlow [20] を用いる。MLP, CNN の活性化関数として ReLU 関数

$$f(u) = \max(u, 0) \quad (式2)$$

を、誤差関数には交差エントロピー関数

$$E = - \sum_k t_k \log y_k \quad (式3)$$

を、学習アルゴリズムには RMSprop [21] を用いる。ここで t_k は正解ラベル (one-hot 表現) を、 y_k はネットワークの出力を表す。

CNN は画像認識で識別に使われることが多く、一般に高い認識率を得ることができている。畳み込み層とプーリング層があることが特徴であり、学習過程で自ら特徴量抽出

を行いながら学習することができる。CNN を採用するため、我々はデータグローブから得られた計測データを 2 次元データに変形し、CNN で学習・評価することを試みた。CNN のシステム概要を図 5 に示す。

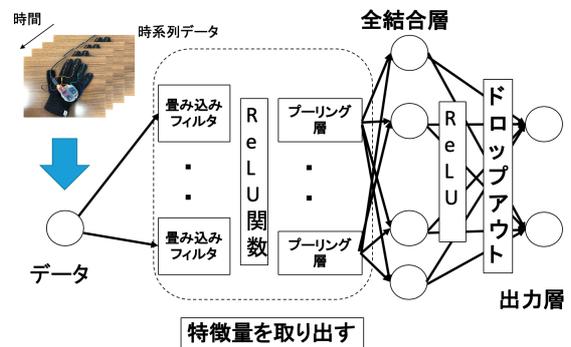


図 5 畳み込みニューラルネットワーク

4. 開発システムの改良

先行研究 [20] における認識精度の向上のため、まずはセンサグローブの調査を行なった。具体的には、動作量と動作方向のそれぞれについて調査と改良を行った。なお、先行研究では、動作方向 (図 4 参照) として、角度情報のみを識別に用いていた。

最初に、動作量と動作方向の両方に共通の検討事項として、それぞれ手を動かした時の変化の様子を時系列データとしてグラフ表示し確認した。その際、サンプリング周波数が小さくなると、各情報が漏れてしまうため、サンプリング周波数の見直しを行った。具体的には、サンプリング周波数の大きさを、20Hz から 200Hz へと大きくした。例として、手を上げて静止した状態と、手を横にふり「ばいばい」をした状態の x 軸方向の加速度を 200Hz で取得した時系列データを図 6 に示す。明らかに情報を漏れなく取得できることがわかる。以上により、動作量と動作方向の各情報を漏れなく取得できることが確認できる。

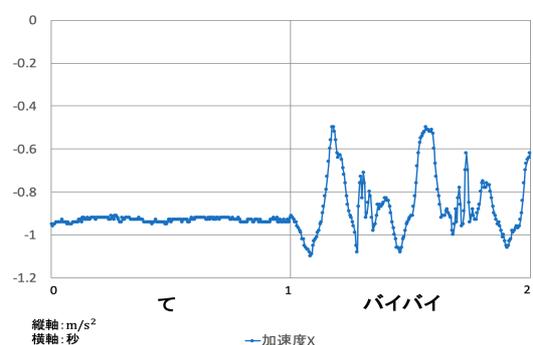


図 6 加速度 (x 軸) の時系列データ (200Hz)

次に、動作量の調査として、指同士が接触した場合や指が少し曲げた場合について、センサにより得られる読取値について調査した (図 7)。結果、いずれの場合でも、入力電圧の大きさが、手を大きく開いた時よりも小さくなることを確認した。

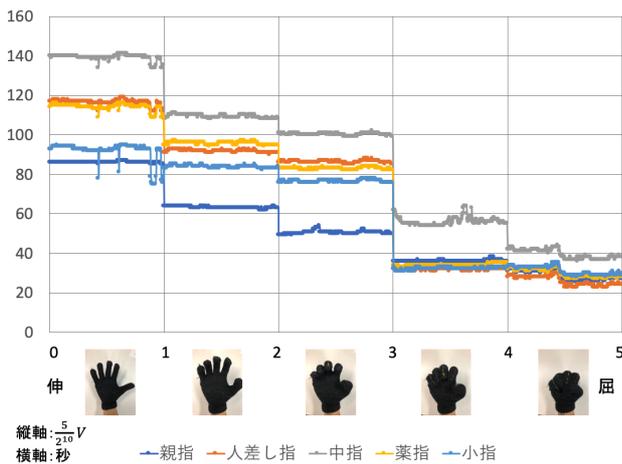


図 7 動作量の変化

最後に、動作方向に関する調査として、加速度と角速度のそれぞれについて調査した。その際、加速度には、重力加速度が含まれていることから、加速度のみを用いることで静止状態の手の向きのある程度推定できる。また、もし手が動いた場合、重力加速度にその加速度が加わることになる。

角速度を用いることで、手を回転させた時の情報を取得できる。しかし、角速度から角度を算出する際は、時間が経っていくと、ジャイロドリフトが入ってしまい、誤差が大きくなってしまいう問題を確認した(図8)。ここで、図8に示す各指文字は、「う」から「と」において手首が回転したものである。なお、先行研究ではカルマンフィルを用いてジャイロドリフトを取り除いていたが、完全に排除することは困難であった。

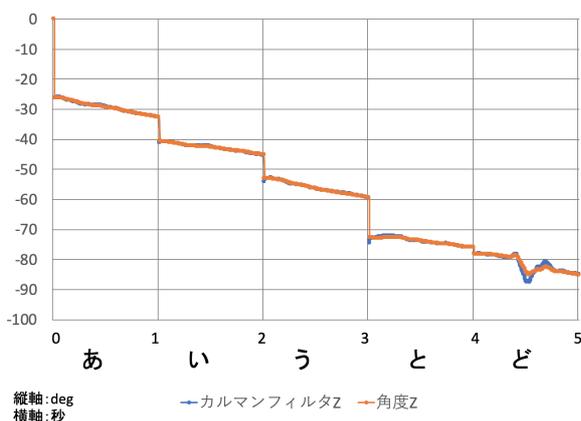


図 8 Z 軸の角度と生データ (200Hz) と
 それにカルマンフィルタを適用した時系列データ

以上の検出方法を見直した後、予備実験を行った。濁音・半濁音・拗音・長音を含めた JFS76 文字を用いて、手指の動作量と動作方向を調査した。特に、手指と手首が動く指文字については、動作量と動作方向の大きさが変化していることを確認した。その結果の一部を図9に示す。

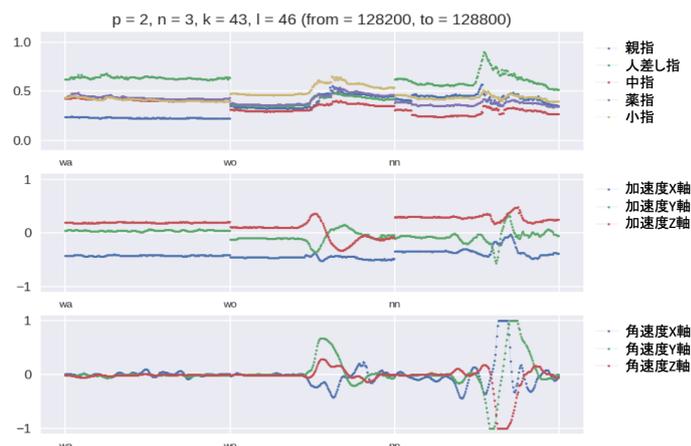


図 9 動作量(5次元)・加速度(3次元)・角速度(3次元)の時系列データ (-1~1 にスケール変換後)

「wa : わ」は、人差し指と中指と薬指を伸ばして、手を表にして静止した状態で表現している。「wo : を」は全ての指を丸くして、胸の方面に動いて表現している。「nn : ん」は人差し指を伸ばして、片仮名の「ん」を書くイメージで表現する。3文字の結果を見ると、静止状態と動作状態の違いが現れている。加速度には重力加速度が含まれていることから、位置情報が明確に分かる。角速度について、手首が静止している時は変化が見られないものの、動作している時は変化がある。このことから、サンプリング周波数を 200Hz にすることで、加速度と角速度の各情報を漏れなく取得できることを確認した。

5. 実験

本実験では、開発したプロトタイプを用いて、センサグローブから得られた動作量・動作方向データを用いて指文字が正しく識別できるか検証する。その際、識別率向上のための予備実験として、すべての採集データをシャッフルの上で評価実験を行い、解析する。

5.1. 指文字データ収集

JFS76 文字を対象とし、実験協力者は 20 名 (20 歳 : 7 名, 21 歳 : 9 名, 22 歳 3 名, 27 歳 : 1 名) とした。実験では、実験協力者がセンサグローブを着用し、指文字の動作を一文字ずつ 1 秒間行った。同時に、図 10 のように、実験協力者の手首と手指の動きを捉えるためにビデオ撮影も行った。JFS76 文字の 1 秒間 200 個のセンサグローブの動作量 (5 次元), 動作方向 (加速度 3 次元, 角速度 3 次元合計 6 次元) の計測データ (11 次元) を取得した。データの収集と同時に手動でデータのラベル付けを行った。この一連の動作を 5 回繰り返した。結果、一人当たり 1 文字 5 個, JFS76 文字, 1 秒 200 個の合計 76,000 個の動作計測データを収集した。更に、実験協力者 20 名, 計測データの 11 次元も含めて合計 16,720,000 個のデータサンプル数を集めることができた。

なお、本実験は、筑波技術大学研究倫理委員会の承認を受けて実施した(承認番号:H30-17)。

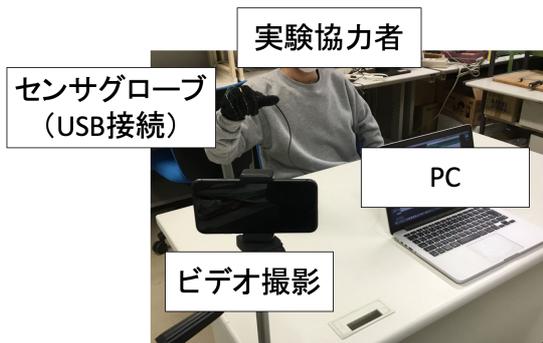


図 10 実験の様子

5.2. JFS46 文字に対する識別予備実験

本論文では、ファーストステップとして、平仮名 46 文字に絞って解析を行う方針を採用した。すなわち、今回の実験で対象とした JFS76 文字から 46 文字を絞り、学習モデルに学習させた。46 文字は濁音・半濁音・拗音・長音を除いた指文字である。

具体的には、1 文字 5 個、46 文字合計 46,000 個の動作計測データ (11 次元) に対して学習・評価を行った。その際、識別器の基本性能を評価し、課題を明らかとするため、すべてのデータをシャッフルして実施した。以上より、20 人のサンプルデータは合計 920,000 個になる。

RF, MLP, CNN に対して 5-fold CV 法でモデルの識別評価を行ったを表 1 と図 11 に示す。ここで、Google の Colaboratory [22]を用いて識別評価を行った。なお、識別時間は評価データ 1 つあたりの識別時間を示している。

表 1 JFS46 文字の識別結果

学習モデル	識別率	識別時間 (ms)
RF	0.9974	0.4701
MLP	0.9529	0.06796
CNN	0.9938	0.04365

表 1 より、RF, MLP, CNN の識別率は、それぞれ 99.7%, 95.3%, 99.4%であり、RF と CNN は同程度の識別率であることがわかる。一方、識別時間は CNN の方が最速であり、特に RF と比較して 10 倍程度高速であることがわかる。更に、図 11 より、一文字ずつに対する識別の正確性が高いことがわかる。

図 11 より、「o : お」と「wo : を」、「no : の」と「ri : り」、「su : す」と「hu : ふ」について、一部識別ミスが生じていることがわかる。

「o : お」と「wo : を」は、手形が同じだが、動作のみが異なっている。「o : お」は静止状態で、「wo : を」は手前に向けて動く。表現は異なっているが、動作スピードには個人差があり、これが誤認識の原因の一つとして考えられる。

「no : の」と「ri : り」は手の動作は同じだが、手形は異若干異なり、人差し指のみを突き出すか人差し指と中指を同時に出すかの違いがある。その際、人差し指と中指の密着がしているかどうかについては個人差があり、これが誤認識の原因の一つとして考えられる。

「su : す」と「hu : ふ」も同様に、手の方向は同じだが、人差し指と中指の密着がしているかどうかについては個人差があり、これが誤認識の原因の一つとして考えられる。

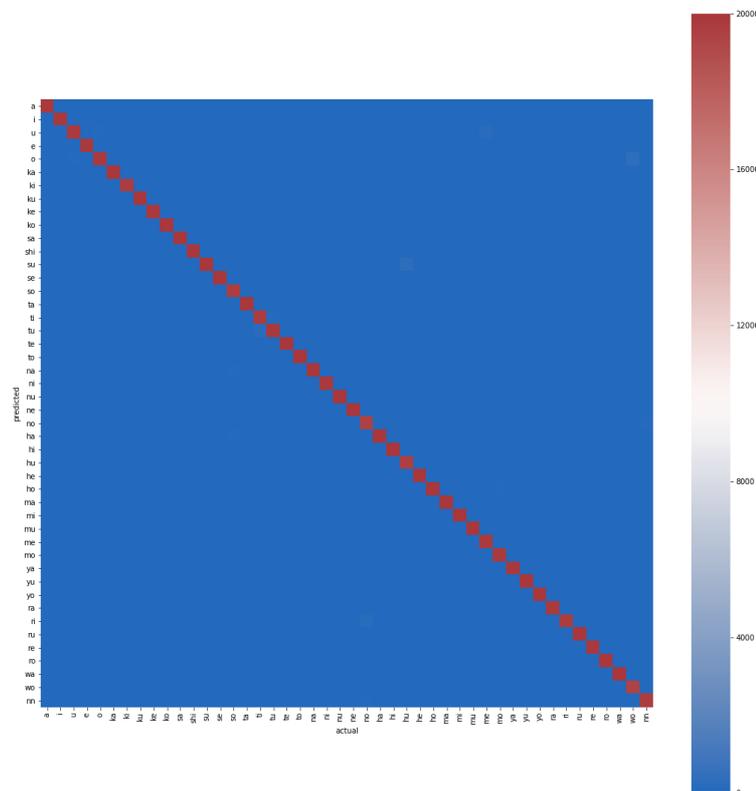


図 11 CNN における JFS46 文字の混合行列

5.3. JFS76 文字に対する識別

次に、濁音・半濁音・拗音・長音を含むすべての JFS76 文字に対する識別評価実験を行った。具体的には、1 文字 5 個、76 文字合計 76,000 個の動作計測データ (11 次元) に対して学習・評価を行った。その際、識別器の基本性能を評価し、課題を明らかとするため、すべてのデータをシャッフルして実施した。以上より、20 人のサンプルデータは合計 1,520,000 個になる。

全ての JFS 76 文字合計 76,000 個の動作計測データ (11 次元) に対し、5.2.において識別率と識別時間が共に良好であった CNN に対して、ホールドアウト検証法を用いて識別評価を行った。その際、学習データは全体の 8 割をランダムサンプリングにより選出した。ここで、Google の Colaboratory [22]を用いて識別評価を行った。

結果を表 2、混合行列を図 12 に示す。なお、識別時間は評価データ 1 つあたりの識別時間を示している。

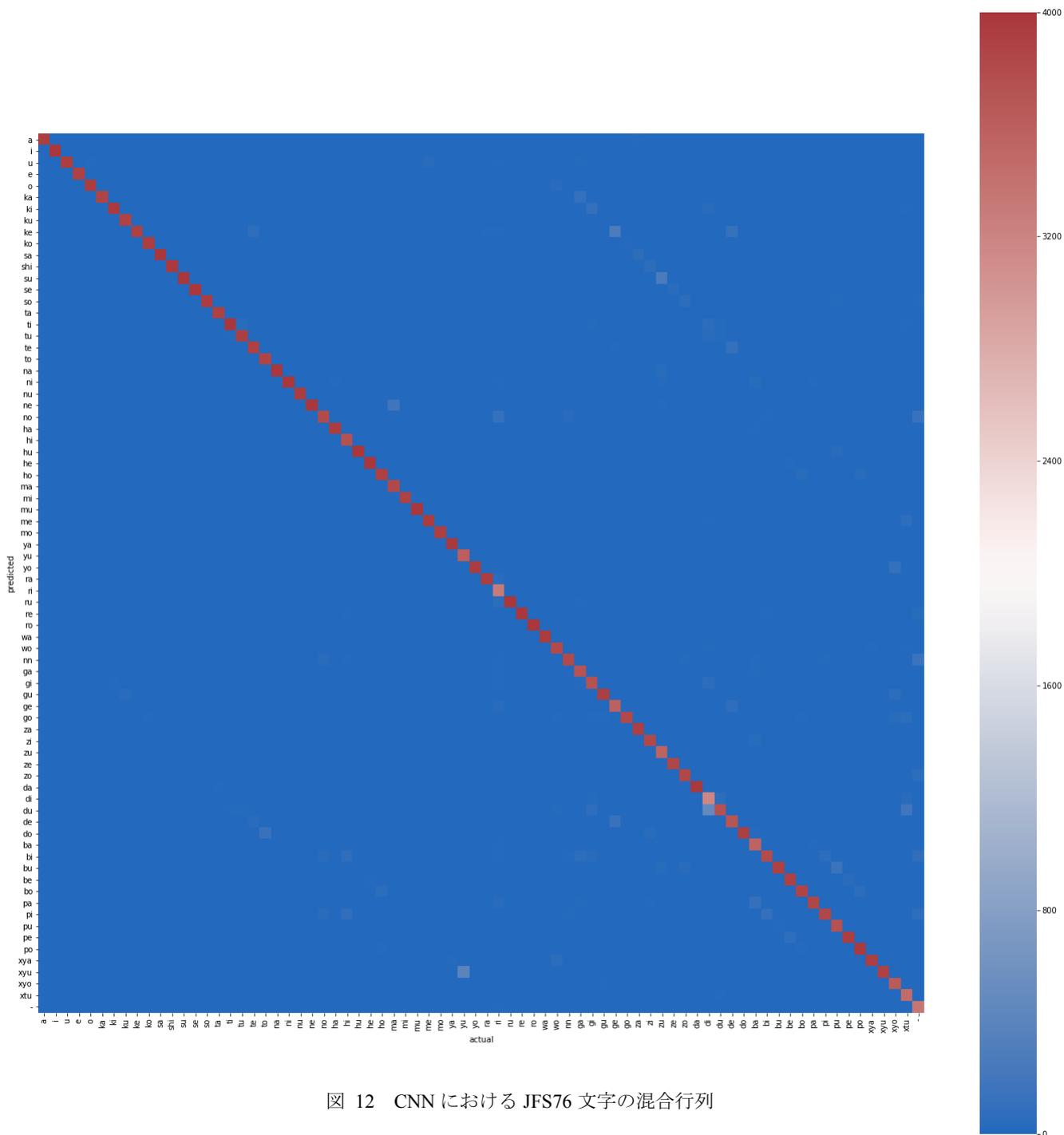


図 12 CNN における JFS76 文字の混合行列

表 2 JFS76 文字の識別結果

学習モデル	識別率	識別時間(ms)
CNN	0.97149	0.07059

表 2 から分かる通り、CNN の識別率約 97.1%を確認した。しかし、JFS46 文字の識別結果と比較すると、約 3% 低下した。JFS76 文字は JFS46 文字に濁音・半濁音・長音・拗音を追加したものであり、これらの追加した JFS はすべて手指または手首が動作するものである。これが識別率低下の原因と考え、図 12 により誤識別が生じた指文字のセットを捉え、ビデオ分析を行った。

結果、手の動作スピードの個人差が大きく、手首の方向と位置情報も人によって異なっていた。特に、手の動作スピードが遅い場合は、静的指文字と誤って認識してしまう可能性があると考えられる。静的指文字というのは、濁音・半濁音・拗音・長音・「の」・「も」・「り」・「を」・「ん」を除いた指文字である。

「di : ぢ」と「du : づ」において、誤認識することがあった。薬指と中指の密着の有無である。更に、濁音が付いていることから、手が動く時に指関節も動いていることによって、導電繊維の短絡による電圧変化が生じたと考えられる(図 13, 図 14)。



図 13 実験協力者 A の「ぢ」と「づ」



図 14 実験協力者 B の「ぢ」と「づ」

「一：長音」について、人差し指を、上から下に向けて動作表現した実験協力者や、左から右に向けて動作表現した実験協力者がいた。また、手の握り方が異なり、図 15 は手指が密着した状態、図 16 は手指があまり密着していない状態になっている。そのため、個人差が生じてしまい、認識精度が低下してしまったと考えられる。



図 15 実験協力者 C の「一：長音」の動き

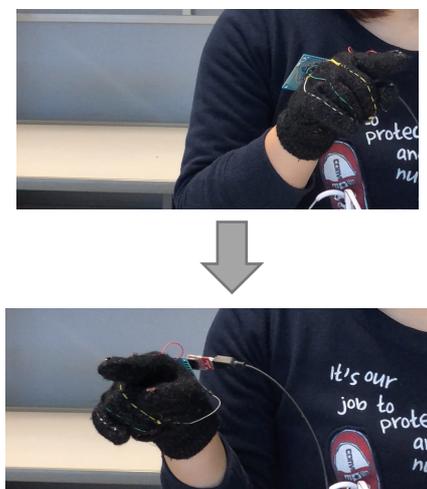


図 16 実験協力者 D の「一：長音」の動き

結果として、JFS76 文字は、手指と手首のそれぞれの動きが多く、人による手の動きのスピードによって静的指文字と誤認識してしまう時があることがわかった。また、手の握り方も異なっていることを確認した。そのため、指関節の変化が生じることから、導電繊維の短絡による電圧変化に影響してしまう可能性を確認した。

もう一つの誤認識の要因として、今回は全サンプリングデータを個別に識別していることから、動的指文字として採取したデータの一部に、実際には手指が静止してしまっている、すなわち静的指文字との区別が困難なデータが混在してしまっていることが考えられる。したがって、データ解析により、手指が動いているデータ区間の特定と切り出しが必要と考えられる。

6. まとめと今後の課題

本研究では、聴覚障害者と聴者が円滑なコミュニケーションを実現すべく、軽量のセンサグローブを採用し、先行研究[18]における課題を洗い出し、識別率の向上を目指して評価実験を行った。

開発システムの改良として、サンプリング周波数と動作方向の次元数を見直し、手指動作時の動作量と動作方向の各情報を漏れなく取得できていることを確認した。また、JFS データ採取実験として、実験協力者 20 名に対し、JFS76 文字毎に、11 次元入力で 1 秒間 (200Hz) のデータ採取を 5 回繰り返し実施した。その後、全データをシャッフルして識別評価を行った結果、JFS46 文字の認識に関しては、RF、MLP、CNN のそれぞれの 5-fold CV による識別率が 99.7%、95.3%、99.4% となった。一方、JFS76 文字の認識に関しては、CNN のホールドアウト法 (学習データは 80% ランダムサンプリング) のよる識別率が約 97.1% となった。

今回、サンプリング周波数を 200Hz にすることで情報を漏れなく取得できているものの、逆にノイズが入ってしまう欠点が存在する。ノイズを抑えるためには、移動平均化、LPE、あるいはカルマンフィルタを用いる方法などが考えられる。また、ビデオ分析を行った結果からも、多くの場合で個人差が生じていることを確認したため、その吸収方法についても調査する。

更に、動的指文字として採取したデータの一部に、実際には手指が静止してしまっている、すなわち静的指文字との区別が困難なデータが混在してしまっている課題に対処するため、手指が動いているデータ区間の特定と切り出しアルゴリズムについて検討する。

センサグローブについては、データ採取実験を終えた後に実験協力者に記入してもらったアンケートを分析しつつ、改善方法を考案する。特に多かったのは、手の負担を軽減するために Bluetooth があつた方が良いという意見だった。また、人によって手のサイズが異なっていることも確認した。以上の点を踏まえ、センサグローブをより良くするよ

うに改善に取り組む予定である。

将来は、以上の考察を踏まえ、本研究を応用して連続する指文字認識に関する研究に取り組む予定である。

謝辞 本研究の一部は、筑波技術大学平成 30 年度学長のリーダーシップによる教育研究等高度化推進事業による助成、並びに JSPS 科研費 JP16K16460 の成果であり、ここに記して謝意を表すものとする。

7. 参考文献

- [1] 国立研究開発法人情報通信研究機構(NICT): こえとら, FEAT(online),available-from<<http://www.koetra.jp/>>(accessed, January,2018)
- [2] Shamrock Records, Inc. :UD トーク ,Shamrock Records, Inc.(online),available-from<<http://udtalk.jp/>>(accessed, January,2018)
- [3] Google Cloud: Cloud speech-to-text, available-from<<https://cloud.google.com/speech-to-text/?hl=ja>>(accessed February 2019)
- [4] Srujana Gattupal :” Evaluation of Deep Learning based pose Estimation for Sign Language Recognition”, Proeedings of the 9th ACM International Conference on Pervasive Technologies Related to AssistiveEnviroments Article No.12, 2016
- [5] Md Asif Jalal, Ruilong Chen, Roger K Moore, Lyudmila Mihaylova:”American Sign Language Posture Understanding with Deep Neural Networks”,10-13 Jul 2018, Cambridge, UK. IEEE , UK , pp. 573-579. ISBN 978-0-9964527-6-2
- [6] 西村洋介, 今村大輔, 堀内靖男, 篠崎隆宏, 黒岩眞 吾:”Kinect とパーティクルフィルタを用いた HMM 手話 認識手法の検討”, 電子情報通信学会, pp.161-166 (2012)
- [7] Granit Luzhnica, Elizabeth Lex, Viktoria Pammer: “A Sliding Window Approach to Natural Hand Gesture Recognition using a Custom Data Glove”, 3D User Interfaces (3DUI); 2016 IEEE Symposium on ; 2016 Mar 19 ; New York : IEEE; 2016 ; p.81-90.
- [8] K. Murakami and H. Taguchi: “Gesture recognition using recurrent neural networks”, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '91, pages 237–242, New York, NY, USA, 1991. ACM.
- [9] 原田直人, 張英夏, 向井信彦: “カラー画像による指文字認識に関する基礎検討”, テレビジョン学会技術報告,Vol.17, No.14, pp.19-24,1993.
- [10] 三宅太一, 若月大輔, 内藤一郎: “距離画像を用いた動きのある指文字の非接触認識手法の検討”, 電子情報通信学会 2012,pp.270-275,2012.
- [11] 三浦航平, 張英夏, 向井信彦: ”爪と手首の位置検出に基づく日本語手話の指文字認識”, 映像情報メディア学会技術報告, Vol.37, No.17
- [12] 長嶋裕二, 藤井昌紀, 長嶋秀世: “決定木を用いた日本語手話における指文字の認識”, 映像情報メディア学会年次会,ROMBUNNO.7-3,2014.
- [13] Hana Hosoe, Shinji Sako and Bogdan Kwolek, “Recognition of JSL Finger Spelling Using Convolutional Neural Networks”, 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA) Nagoya University, Nagoya, Japan, May 8–12, 2017.
- [14] Maria E Cabrera, Juan Manuel Bogado, Leonardo Fermin, Raul Acuna, and Dimitar Ralev, “Glove-based gesture recognition system”, In *Proc. of Intl. Conf. on climbing and walking robots and the support technologies for mobile machines*.747–753, 2012.
- [15] Chaithanya Kumar Mummadi,, Frederic Philips Peter
Leo,Keshav Deep Verma, Shivaji Kasireddy, Philipp Marcel Scholl,Kristof Van Laerhoven ,“Real-time Embedded Recognition of Sign Language Alphabet Fingerspelling in an IMU-Based Glove”, the 4th international Workshop on Sensor-based Activity Recognition and Interaction, Rostock Germany, ISBN:978-1-4503-5223-9, 2017
- [16] 5DT : Virtual Reality Hardware. Available-from<<http://www.5dt.com/vr-hardware/>>(accessed, April, 2018))
- [17] 高田峻介, 志築文太郎, 高橋伸: “導線織維編み込み手袋を用いた指の曲げ計測手法”, 情報処理学会 2017, Vol.2017-HCI-171, No.25, 2017
- [18] 小松周生, 白石優旗:”深層学習を用いたセンサグローブによる指文字認識の検討”,情報処理学会アクセシビリティ研究会(IPJS SIG AAC)第 6 回研究会; 2018-3
- [19] David Courmapeau; scikit-learn, INLIA(online), available-from<<https://www.scikit-learn.org/stable/>>(accessed November 2017)
- [20] Google Brain: TensorFlow, Google(online), available-from<<https://www.tensorflow.org/>>(accessed November 2017)
- [21] Geoffrey Hinton, Nitish Srivastava, Kevin Swersky. 2014. “Lecture 6e: Rmsprop: Divide the gradient by a running average of its recent magnitude”, (CSC321 Winter 2014)
- [22] Google Colaboratory, Google(online), available-from<<https://colab.research.google.com/notebooks/welcome.ipynb?hl=ja>>(accessed January 2019)