

小型モーションセンサを用いた指文字単語認識 Finger Character Word Recognition Using a Compact Motion Sensor

尾山 匡浩[†]
Tadahiro Oyama

秋田 悠登[†]
Yuto Akita

1. はじめに

聴覚障害者がコミュニケーションをとる方法として、筆談、読話、発話、手話などの方法が挙げられる。その中でも手話は、会話をスムーズに進められるため多くの聴覚障害者が利用している。しかし、手話で会話をするためには、少なくとも 500 種類の手話を覚える必要があるため、手話の知識のない人と手話を用いた会話をするのは困難である。そこで、聴覚障害者とのコミュニケーションの円滑化を図るための手話翻訳システムの需要が高まっている。これまでの手話翻訳に関する研究では、グローブとカメラを用いて行う方法が多く研究されていた[1]。しかし、コストやグローブ装着の手間などを考慮すると実用化までは至っていない。一方で、Microsoft 社の 3 次元モーションセンサデバイスである Kinect は体全体の動きを取得できるため、手だけでなく体全体を使う手話の認識に利用しようとする研究も行われてきている[2,3]。しかし Kinect は装置が比較的大きく、持ち運びが大変なため利便性が良いとはいえない。そこで本研究では、3 次元モーションセンサデバイスの 1 つである Leap Motion を使用することを検討する。Leap Motion は小型、かつ軽量であり、気軽に持ち運ぶことができるため利便性が高い。そのため、これを用いた手話翻訳システムを構築することができれば、実用化の可能性も高まると予想される。

そこで本研究では、Leap Motion を用いた手話認識システムの第 1 段階として、「あ」から「ん」の 46 字の指文字の認識を行う。実際に Leap Motion を用いて指文字を推定する研究も行われており[4]、我々も、これまでに各文字単位で推定を行ってきた[5]。しかしながら、実際に利用することを考えた場合には、各文字単位だけではなく連続して指文字が行われたときの処理に対応する必要がある。そこで本論文では、指文字における単語推定を目的とし、連続的に計測された指文字の情報から各指文字を 1 文字単位で検出する方法と単語の認識手法について検討する。

2. 提案手法

2.1 提案手法の流れ

Leap Motion を用いた指文字単語の認識を実現するために、図 1 に示すような流れで処理を行う。まず、Leap Motion を用いて各指文字単語の計測を行い、認識された各指先や関節などの座標を取得する。取得したそれらの時系列情報から残差平方和を算出し、1 文字の区間を検出する。この区間の中で、平均の特徴量を計算し、それらを用いて線形判別分析(Linear Discriminant Analysis : LDA)により各指文字を推定し、最終的な単語の認識結果とする。

2.2 指文字計測

Leap Motion はモーションセンサデバイスの 1 つであり、手や指を認識し 3 次元空間での座標を取得することが出来る。また、手や指の前後左右や上下の移動だけでなく、手のひねりや各関節の位置も捉えることが可能である。本論文では Leap Motion を用いて 5 本の各指における各関節座標と掌の中心座標を各フレームで取得し、認識に利用する。また、おおよそのフレームレートは 120 fps である。

2.3 各文字区間の検出

指文字計測で得られた情報は、各座標の時系列信号となっている。ある指文字から異なる指文字に変化する際には、この座標値が急激に変化することが予想される。そこで、本論文では(1)式に示す各フレーム間の残差平方和の指標を導入し、この値から指文字区間の検出が可能か検証する。

$$RSS(f) = \sum_{k=1}^5 (F_k(f) - G_k(f) - F_k(f-1) - G_k(f-1))^2 \quad (1)$$

ここで、 $RSS(f)$ は f フレーム目の残差平方和を、 $F_k(f)$ は f フレーム目の指先の 3 次元座標、 $G(f)$ は f フレーム目の掌の 3 次元座標を表している。また、 k は各指を表しており、片手のため 5 が最大である。指の座標 $F_k(f)$ のみの残差平方和では、「の」や「も」などの動きのある指文字の区間を検出することができないため、指の座標から掌の座標を引いて残差平方和を算出することで指の動きのみを検出することを目指している。

この $RSS(f)$ の値があらかじめ設定したしきい値を超えたフレームが文字から次の文字への変化点であると検出し、各変化点の間の複数フレームの情報を推定に利用する。なお、本研究ではしきい値は経験的に決定した。

2.4 利用する特徴量と推定方法

Leap Motion では手を認識した際に各関節の先端と末端の座標を取得できる。この中で手を動かした際に最も動きのある部位である末節骨(指先)の先端、中節骨の先端、掌の中心の三種類を特徴量として利用する。

- 末節骨(指先)の先端座標 $(x_d, y_d, z_d) \times 5$ 本
- 中節骨の先端座標 $(x_i, y_i, z_i) \times 5$ 本
- 掌の中心座標 (x_p, y_p, z_p)

これらの座標を並べたものを 1 フレームあたりの特徴ベクトルとして準備する。

前節で検出された 1 文字の区間において、この特徴ベクトルを算出し、それらの平均をとることでその区間における文字の特徴ベクトルとする。これを利用して LDA によりあらかじめ計測しておいた指文字データと比較することで指文字の推定を行う。

[†] 神戸市立工業高等専門学校, Kobe City College of Technology

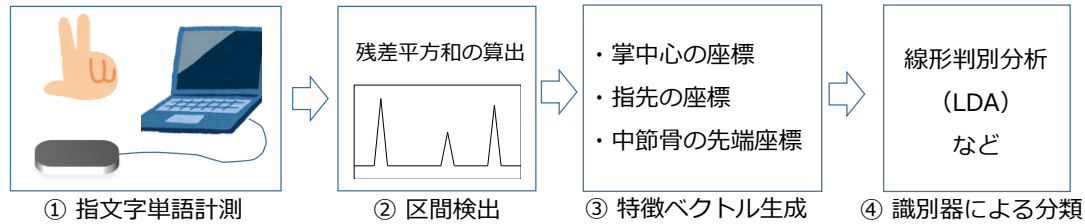


図 1 提案手法の流れ

3. 検証実験

3.1 計測条件

本研究では、PC モニタの上に Leap Motion のセンサ面が自分側に向くように 3D プリンタを用いて制作した固定器具で固定し、指文字の計測を行う。

識別の際に利用する学習データは、Leap Motion の前で各指文字の動作を行い、その状態で 1 秒間静止する。これを 1 文字につき 10 回実施する。計測後、取得したデータを確認し、右手を左手と認識している等の明らかな誤認識が見られた場合は、データの再計測を行う。

検証用の単語の指文字データは手を開いた状態から測定を開始し、約 3 秒ごとに指文字を変えて単語として表現する。計測する指文字単語として、指の動きが大きい「さけ」と指の動きが小さい「あさ」について検証を行う。なお、それぞれ 5 回ずつ計測を実施した。

3.2 文字区間検出実験

上述した計測条件に従って計測した、「さけ」と「あさ」の単語データを利用して各指文字区間の推定が可能か検証した。それぞれ 5 回ずつ測定したときの区間検出の成功回数を表 1 に示す。「さけ」の区間検出はすべて成功したが、「あさ」の場合では 2 回区間検出に失敗した。「あさ」の区間検出に失敗したときの RSS の変化の様子を図 2 に示す。

表 1 文字区間検出実験結果

単語	成功回数
さ→け	5/5 回
あ→さ	3/5 回

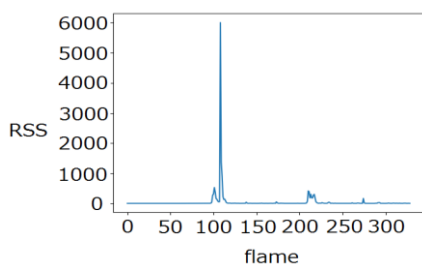


図 2 「あさ」の RSS の変動(区間検出失敗時)

図 2 から、「あ」から「さ」の区間の検出に失敗していることが分かる。区間検出に失敗した要因として、「あ」から「さ」へ指の形を変えるときに動く指が親指だけであることが挙げられる。動かす指が少ない場合、残差平方和の値が小さくなってしまい、区間の検出に失敗してしまうと考えられる。

3.3 単語推定実験

計測条件は、3.2 節と同様である。この単語データと事前に計測した各指文字の学習データを用い、区間検出後の特徴ベクトルを比較することで、区間ごとに推定を試みた。

「あさ」、「さけ」について単語推定を行った結果を表 2 に示す。「さけ」の単語は 100%、「あさ」の単語は 60%の精度で推定することができた。

単語の推定に失敗している「あさ」の単語データは、3.2 節の区間検出実験で区間検出に失敗しているデータである。また、表 2 より、失敗しているデータの推定結果を見ると、どちらも正解の単語のうちの 1 文字の推定には成功している。これより、単語推定が失敗している要因として区間検出の精度の低さが考えられる。

表 2 単語推定結果「さけ」、「あさ」

入力単語	区間推定	推定結果
さ→け	成功	さ→け
	成功	さ→け
	成功	さ→け
	成功	さ→け
	成功	さ→け
あ→さ	失敗	さ
	失敗	あ
	成功	あ→さ
	成功	あ→さ
	成功	あ→さ

4. おわりに

本論文では、手話未経験者と聴覚障害者のコミュニケーションの円滑化を図るための手話認識・翻訳システムの前段階として、指文字区間の検出と単語の推定を試みた。本論文で検討した残差平方和を用いた区間検出手法では、指の動きが小さい区間の検出での精度の低さや、緩やかな指の動きに対応できない等の問題があることが分かった。今後は、これらの問題を解決することができる新たな区間推定手法を検討する必要があると考えられる。

参考文献

- [1] 大里宗之他, "カラーグローブを用いた指文字認識における特徴量の統合性", PRMU, Vol.105(375), pp.73-78, (2005)
- [2] 西村洋介他, "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", PRMU, Vol.111(430), pp.161-166, (2012)
- [3] 井上快他, "Kinect を利用した指文字認識に関する検討", MBE, Vol. 112(417), pp. 45-50, (2013)
- [4] 河原圭佑他, "装着型機器を用いた音声翻訳による対人コミュニケーション支援", IPSJ 第 77 回全国大会, pp. 619-620, (2015)
- [5] 尾山匡浩他, "3 次元モーションセンサデバイスを用いた指文字の認識", 電気学会 C 部門大会, pp.1140-1143, (2017).