

前腕の表面筋電位を用いた RNN による指書き文字認識

Finger-writing Character Recognition with Recurrent Neural Network
Using Surface Electromyogram of the Forearm

坂井 浩志郎† 中井 満†

Koshiro Sakai Mitsuru Nakai

1. はじめに

空中での文字入力に関する研究では、カメラを用いたものや加速度を用いたものがある [1][2]。しかし、カメラを用いた手法は場所を制限され、加速度を用いたものでは、大きな動きが必要になるという問題がある。どこでも、小さな動きで入力できる方法として、指運動時の前腕の表面筋電位を用いた入力の研究が行われている [3]。そこで指先で文字を書くときに生じる前腕の表面筋電位を用い、RNN (Recurrent Neural Network) で文字認識する手法を提案する。また、筋電位のような生体信号は個人差が大きいので、個人用途の RNN を学習することにする。

2. 指書き文字認識の構成

2.1 表面筋電位信号

データの収集には North 社製のアームバンド Myo を用いる。Myo は 8 個の筋電センサを内蔵しており、前腕周りの 8 箇所表面筋電位を同時に測定できる。これを装着し図 1 のような姿勢で指先で文字を書く。本研究では、8 個の筋電位センサから取得できる 8 次元の表面筋電位を入力として用いる。表面筋電位はサンプリング周波数 50Hz で取得する。観測した表面筋電位の時系列を

$$O = o_1 o_2 o_3 \cdots o_T$$

とおく。ここで、 T はサンプル O の時系列の長さである。時刻 t で観測される筋電位は

$$o_t = (o_{(1,t)}, o_{(2,t)}, \dots, o_{(8,t)})$$

であり、 $o_{(i,t)}$ はセンサ番号 i の短時間二乗平均平方根である。ある時刻での表面筋電位を図 2 に示す。ここで手の向きは紙面奥向きである。

2.2 識別器

識別器に RNN を用いる。入力層は 8 個のノードで構成した。これは入力データである 8 次元の表面筋電位に対応しており、ノード i ($1 \leq i \leq 8$) には各センサから取得した表面筋電位信号 $o_{(i,t)}$ が入力される。中間層は 1 層あたり 128 個の LSTM セルで構成した。出力層はひらがな 46 文字に対して 46 個のノードで構築した。これを Python のライブラリである PyTorch の torch.nn.LSTM クラスを用いて構築した。



図 1: 収集時の姿勢

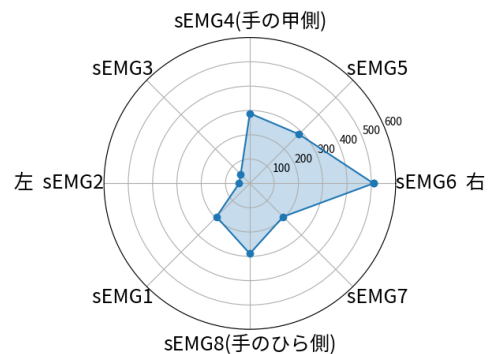


図 2: 指書き中のある時刻の表面筋電位

3. 文字認識実験

3.1 データ収集

ひらがな 46 文字を 1 セットとして 1 人の被験者から、110 セット収集した。なお、2 セット収集するごとに Myo の着脱を行い、1 日 10 セットで 11 日間に分けて収集した。

3.2 RNN による評価実験

ひらがな 46 文字の認識実験を行った。評価に 1 日分の 10 セットを用い、学習に残りの 10 日分の 100 セットを用いて交差検証を行った。中間層の数を 1 層、3 層、5 層でそれぞれ学習を行った。

結果を表 1 に示す。1 層での認識率は 30.1% と低いが、3 層では 86.8%、5 層では 87.0% となり、中間層の数を増やすことで、認識率が上昇することが分かった。また、3 層の場合と 5 層の場合では認識率に大きな差がないため、中間層の数は 3 層で良いと考えられる。

†富山県立大学, Toyama Prefectural University

表1: 指書き文字の認識率 [%]

中間層の数	1層	3層	5層
認識率	30.1	86.8	87.0

3.3 サンプルの正規化の効果

個々のサンプルを値が -0.5 から 0.5 の範囲で値を取るよう原点移動とスケールの正規化をした。「あ」を筆記したサンプルの正規化前を図3に、正規化後を図4に示す。これらのサンプルを用いて3.2節と同様に交差検証を行った。

結果を表2に示す。正規化した場合と正規化しない場合の結果を比較すると、中間層が1層の場合は正規化することで認識率が高くなることが分かった。一方で、3層以上では認識率に大きな差はなかった。このことから、中間層が1層の場合は正規化したほうが良いが、多層の場合は正規化をする必要がないと考えられる。

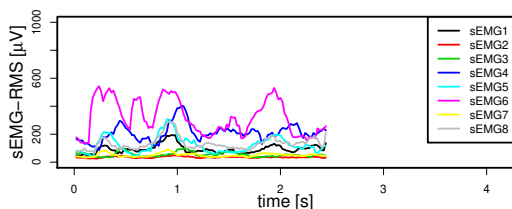


図3: 正規化前のサンプル

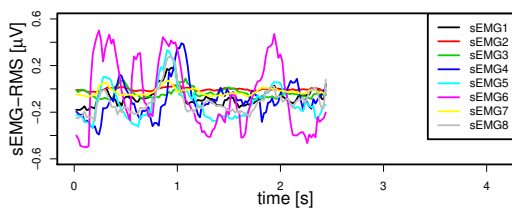


図4: 正規化後のサンプル

表2: 正規化による効果 [%] (は表1の再掲)

中間層の数	1層	3層	5層
正規化なし	30.1	86.8	87.0
正規化あり	82.0	89.0	89.4

3.4 日数の経過による影響

3.2節でモデルを学習した後、半年後に新たにデータを収集して評価した。結果を表3に示す。学習用データと同時期に収集したデータを評価に用いた場合と、約半年後に収集したデータを評価に用いた場合では認識率に有意な差は見られなかった。このことから、収集した時期による影響は小さく、1度学習すればしばらくは再学習の必要はないと考えられる。

表3: 半年後の文字認識率 [%] (は表1の再掲)

中間層の数	1層	3層	5層
学習直後	30.1	86.8	87.0
学習から半年後	34.0	91.6	91.8

3.5 位置ずれのデータ拡張の効果

学習用データよりセンサ番号をずらして仮想的に位置ずれしたデータを作成し、データ拡張をする。ここで、原データを

[0], 原データをセンサ1個分だけ右回転させたデータ拡張を[+1], 原データをセンサ1個分だけ左回転させたデータ拡張を[-1]と表記する。原データ[0]に[+1][-1]のデータを組み合わせることで学習し、3.2節の実験と同様に交差検証を行った。中間層は3層で学習した。

それぞれの認識率を表4に示す。原データのみで学習した場合に対して、データ拡張したデータを組み合わせることで学習した場合は認識率が約5ポイント上昇した。このことから、データ拡張の方法としてセンサを仮想的に回転させる方法は有効であると考えられる。

表4: データ拡張による学習効果 [%]

データ拡張	[0]	[+1][0][-1]
[0]の認識率	86.8	91.5

3.6 未学習の位置ずれに対する頑健性

前節と同様にデータ拡張を行った。原データのみの場合と、原データに[-1]のデータを組み合わせることでそれぞれ学習し、3.2節の実験と同様に交差検証を行った。ここで評価に用いたデータは、学習に用いなかった位置ずれデータ[+1]である。中間層は3層で学習を行った。

結果を表5に示す。原データのみで学習した場合の認識率が7.8%であったのに対して、原データに[-1]のデータを組み合わせることで認識率は51.6%であった。このことから、未学習の位置ずれに対しては、他の位置ずれのデータを学習に加えることで、認識率の向上を図ることが可能であると考えられる。一方で、認識率は位置ずれがない場合と比べると低いため、今後は中間層を増やしたり、学習に用いるデータを増やしたりしてさらなる認識率の向上を図る必要があると考えられる。

表5: 未知の位置ずれに対する認識率 [%]

データ拡張	[0]	[0][-1]
[+1]の認識率	7.8	51.6

4. まとめ

前腕の表面筋電位を用いてRNNによる指書き文字の認識の実験を行った。本手法では、データを収集した時期による影響は小さく、1度学習すればしばらくは再学習の必要はないと考えられる。さらに、データ拡張の方法としてセンサを仮想的に回転させる方法は有効であると考えられる。

謝辞 本研究はJSPS 科研費17K00275の助成を受けて行った。

参考文献

- [1] 園田智也, 村岡洋一, “空中での手書き文字入力システム,” 信学論, D-, Vol.J86-D2, No.7, July 2003.
- [2] 中井満, 米澤久光, “加速度センサを用いた空中手書き文字認識,” 情報科学技術フォーラム (FIT), Sept. 2009.
- [3] 石川圭佑, 他, “表面筋電信号を用いた実時間指運動認識インタフェースとその応用,” 情報処理学会インタラクショナル2011, 4DEX-7, March 2011.