

不特定多数の使用者を対象とした sEMG による多クラスのハンドジェスチャ識別手法

伊東 和輝† 田村 仁†

日本工業大学 工学研究科†

1. はじめに

ハンドジェスチャは、腕の動作によるコミュニケーション手段であり、発話や顔の表情とともにボディランゲージの一部として用いられている。また、手話として体系化され、使用されることもある。近年、ハンドジェスチャはハンズフリーで直感的な操作ができるという点からロボットや xR (VR, AR, MR) のインターフェースとして着目されることが多い。

デバイスがハンドジェスチャを認識する方法としては、カメラや深度センサを用いることで手の形状や動作を認識するコンピュータビジョンによる手法と IMU (慣性計測ユニット) センサーやフレキシブルセンサーを搭載したグローブ型や指輪型デバイスを手手に装着して手指の形状や 3 次元加速度などを取得し、それを用いてハンドジェスチャを認識する手法に大別される。前者ではカメラや深度センサが必要となるため、撮影の範囲や距離など、対象の位置と角度に制限が課せられることが多い。また、後者においては手指や上腕などにセンサを装着するため、手指形状の個人差で装着時に違和感が生じる場合がある。いずれの方式にせよ、実際にデバイスを使用するにあたり、撮影環境や個人に合わせた調整が必要となる。

本論文で着目するのは、着脱が容易な市販の表面筋電信号 (surface Electromyogram ; sEMG) センサを使用し、計測した sEMG データを深層学習手法により分類し、ハンドジェスチャを識別する手法である。使用したデバイスは 8 個の sEMG センサが腕輪状に連結されている Talmic Labs (現 North) 製の Myo Gesture Control Armband (以下 Myo) (図 1) である。このデバイスは簡便に前腕への着脱が可能であり、計測した 8 箇所 sEMG は Bluetooth 接続によりデータを PC へ送信できる。また、3 軸加速度センサや 3 軸ジャイロスコープを搭載しており、重さは約 93g、センサの厚みは 11mm であることから VR や AR など、直感的な操作を行い、デバイスの装着感が重要視される分野においても適しているデバイスだと考えられる。



図 1 Myo Gesture Control Armband

Method for Classifying Multi-Class Hand Gestures with sEMG for Unspecified Users

† Kazuki Itoh, Hitoshi Tamura, Nippon Institute of Technology Graduate School of Engineering.

ハンドジェスチャの解析において sEMG を従来の機械学習手法で識別する研究は多く見られる。例えば、手話認識に対して決定木と HMM を使用する手法[1]や義手制御に HMM とサポートベクターマシン (Support Vector Machine; SVM) を用いた手法[2]、ハンドジェスチャ解析に決定木と k 最近傍法 (k-Nearest Neighbor; k-NN) を使用する手法[3]などがある。また、sEMG を高速フーリエ変換し、10 個の周波数領域をサポートベクターマシンで識別する手法としては[4]がある。

ハンドジェスチャから計測した時系列データの sEMG に対して機械学習による識別を行う場合、一般的に時系列データに適しているといわれる再帰型ニューラルネットワーク (recurrent neural network; RNN) を用いることが考えられる。特に時系列データ用に拡張された LSTM (long short term memory) モデルを用いられることが多い。

計測したハンドジェスチャの sEMG データに対して LSTM を用いて分類した例としては[5]がある。この研究には LSTM を用いた分類により正解率が向上していることが示されている。

問題点としては、sEMG データを用いてハンドジェスチャの判別を行う場合、Myo などの簡便な sEMG センサで計測を行った際にわずかな装着時のズレで sEMG が大きく影響を受けることである。体表面から筋電位を測定する方法であるため、3 次元形状の前腕部分の複雑な筋肉の動きを検知するにはそもそも不安定である。従来手法の対策としては、それを大量のデータを用いて機械学習手法により解決を図っている。しかしその反面、大量のデータを必要とすることから誰でもすぐ可以使用の状態とすることは困難である。

機械学習手法により sEMG からハンドジェスチャ分類を行う提案は多いが、学習データについて使用者自身の sEMG データのみを使用して学習を行うものが多く、学習済み分類器を不特定多数の使用者に適応可能とする観点の研究はわずかしかない。

本研究では、計測した sEMG データに対して効果的なデータ拡張処理を施し、不特定多数に適応可能なハンドジェスチャ分類器を作成することを目的とする。

2. 提案手法

機械学習手法で不特定多数の使用に対応するためには、おそらく数百人単位の不特定多数からデータを計測する必要があると考えられる。しかし、その量のデータを収集するのは現実的ではない。不特定多数の使用への対応を実現するには、現実的に 10 人程度の被験者から計測したデータを拡張し、数百人分相当のデータセットにすることを提案する。

データ拡張手法としては、一般的に音声信号の特微量抽出で使用されるケプストラムを用いる。ケプストラムは低次部分がスペクトル包絡、高次部分が微細構造となっていることから、sEMG の特徴を維持しつつ、微細構造部分にのみノイズを加え、データを拡張させることが可能となる (図 2)。ケプストラムにより拡張したデータは

最終的に対数振幅スペクトルから振幅スペクトルに変換されるため、学習・評価にはスペクトルデータを使用する。

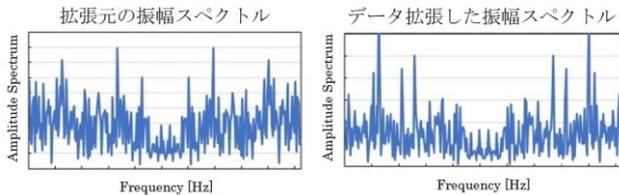


図2 データ拡張(左:元データ, 右:拡張後データ)

3. 評価実験

3.1 データ拡張

データ拡張を施した際、ケプストラムによる処理が効果的であったのか、ただ学習用データの量が増えたことが正解率向上に影響したのかを検証する為、単純に計測データにランダムノイズを±12.7[mV]の範囲で加え、11倍に拡張した場合の学習も行った。学習に使用したジェスチャは、脱力、パー、軽く握る、強く握る、人差し指を立てる、チョキの6種類であり、深層学習にはLSTMを用いて行った。また、ジェスチャ1回分に対して2.56秒間の計測を100Hzとして行った。

被験者1人から計測した3,240回分のデータで学習を行った結果、正解率は75.03%となり、データ拡張を施した場合の正解率は78.48%となった。総合的な正解率は向上したが、ジェスチャによっては正解率が低下したものも複数あり、このことから単純にランダムノイズを加算するだけではデータ量が増えても効果的でないことが分かった。

ケプストラムを活用したデータ拡張手法の有効性を検証した実験では、被験者1人から3度の計測(1度に1,260回分の計測)を行い、その中から1度分の計測データを用いて11倍のケプストラムによるデータ拡張を行った。3度の計測データで学習を行った場合、98.21%の正解率となり、1度分の計測データのみで学習を行った場合、91.19%となった。そして、データ拡張を施した場合では99.28%の正解率となった。

3.2 不特定多数への対応

次に実際、不特定多数に対応可能であるかを評価実験によって検証する。検証方法としては被験者7人のデータで交差検証(cross validation)を行う。交差検証とは学習と検証のデータを分割し、評価することで少ないデータでも汎化性能を測定することができる検証手法である。特に今回使用したK分割交差検証は、K個にデータセットを分割、そのうち1個分を検証用データとして学習・評価を行う。これを全てのデータセットが検証用データとなるようにK回繰り返して性能の平均をとる検証手法である。これによりモデルの平均的な汎化性能を検証することが可能である。今回は被験者7人のうち、1人のデータ全てを検証用データセットとして学習・評価することで不特定多数への汎化性能を検証した。

また、深層学習に関してはsEMGをスペクトルデータとして学習・評価を行う場合、8(センサ数)×128(サンプリング数)の画像データとしても扱うことができることから、LSTMだけではなく、畳み込みニューラルネットワーク(convolutional neural network; CNN)のVGGnetも使用

し、学習を行った。

実験には図3に示す5種類のジェスチャ(wave in, wave out, パー, グー, 親指と中指でダブルタップ)を使用した。また、ジェスチャ1回の計測時間を1.28秒間とし、被験者7人からそれぞれ315回分のデータ(1ジェスチャあたり63回)を計測、その中から学習に使用するデータをケプストラムにより、11倍に拡張した。



図3 対象とするハンドジェスチャの一覧

結果、LSTMで学習を行い、K分割交差検証で評価した場合、正解率は84.12%となり、VGGnetで学習した場合では正解率が86.02%となった。

4. 考察

実験結果からケプストラムでデータ拡張を行い、学習を行った場合では3倍のデータ量で学習した分類器より高い正解率を得られることが分かった。このことからケプストラムによる拡張手法はsEMGデータの学習に有効であると考えられる。また、VGGnetで学習した場合でも高い正解率を得られたことから、sEMGのスペクトルデータはCNNによる学習に対して有効であると考えられる。交差検証に関しては、どの被験者が検証用データになった場合でも、安定して正解率が80%を超えていた為、このことからsEMGをデータ拡張する手法により、学習済み分類器を不特定多数の使用者に適用させることが可能であると考えられる。

5. おわりに

本研究では着脱が容易なデバイスのMyoを使用し、ハンドジェスチャのsEMGを計測した。そして、データ拡張を施したデータで学習を行い、不特定多数の被験者を前提としたハンドジェスチャ識別の実験を行った。実験結果からケプストラムの活用により、学習に効果的な人工的sEMGデータを作成することが可能だと示すことができた。そして、これは不特定多数の使用者に適用できる有効なデータ拡張手法であると考えられる。

参考文献

- [1] X.Zhang, X.Chen, Y.Li, V.Lantz, K.Wang, and J. Yang, A framework for hand gesture recognition based on accelerometer and EMG sensors. IEEE Trans. Syst., Man, Cybern. A, Syst. and Humans, 2011, Vol. 41, No. 6, p.1064-1076.
- [2] M. Rossi, S. Benatti, E. Farella, and L. Benini, Hybrid EMG classifier based on HMM and SVM for hand gesture recognition in prosthetics, Proc. Int. Conf. Ind. Technol., 2015, pp. 1700-1705.
- [3] Kuang-Yow Lian, Chun-CHieh Chiu, Yong-Jie Hong and Wen-Tsai Sung. Wearable Armband for Real Time Hand Gesture Recognition. IEEE International Conf. Syst., Man, Cybern.. 2017, p. 2992-2995.
- [4] Frederic Kerber, Michael Puhl, Antonio Krüger, User-Independent Real-Time Hand Gesture Recognition Based on Surface Electromyography. MobileHCI '17. Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services September 2017 Article No.36 Pages 1-7.
- [5] Yuheng Wu, Bin Zheng and Yongting Zhao, Dynamic Gesture Recognition Based on LSTM-CNN, Chinese Automation Congress, 2018, pp.2446-2450.