

EMG インターフェースによる XR での入力システム開発

伊東 和輝[†] 田村 仁[‡]日本工業大学 工学部 創造システム工学科[‡]

1. はじめに

近年、様々な操作手法の VR コントローラーが登場し、VR 空間内で高い臨場感を体感することが可能になっている。例えば、Leap Motion 社の Leap Motion の様に赤外線カメラで手のモーションを認識し、入力操作を行うデバイスや CyberGlove Systems 社の CyberGlove2 の様にグローブ型コントローラーなどが挙げられる。しかし、赤外線カメラで手のモーションを取得する場合、高い精度で指の座標をとらえることが可能であるが、手が隠れた時に認識できない。また、グローブ型コントローラーの中では各指の細かな動作を読み取ることが可能であるが、ボタンによる操作ができない為、操作のコマンド数が少ないなどのデメリットがある。

その為、本研究では握力の様な視覚的に判別が難しいジェスチャーに対応でき、ハンズフリーな状態での入力操作が可能である腕の表面筋電位を用いた操作方法の開発を試みた。また、操作コマンドの数を考慮し、ハンドジェスチャーによる操作や表面筋電位の個人差を調整する為、深層学習で表面筋電位の判別を行う。本研究では深層学習を用いて、表面筋電位による XR での入力操作システムの開発を目的とする。なお、今回は VR に着目するものとする。

2. EMG の測定方法と学習環境

本研究では 8 つの筋電センサからデータを取得する株式会社 FOX の筋電センサアームバンド Myo を用いた。Myo では 200Hz のストリーミング速度の筋電図 (EMG: Electromyogram) データを測定できる。筋電によるジェスチャー操作システムで必要となるジェスチャーの判別は次の様に考える。

- 1) 画像認識では難しいジェスチャーの判別
 - 2) 他のジェスチャーの筋電データが混合する
- ジェスチャーの判別

以上の事を踏まえて、本研究ではパー、グー、強く握る、人差し指を立てる、チョキ、脱力の 6 つのジェスチャーで学習用データを作成した。(図 1) ジェスチャーのデータ取得には、ジェスチャーを行い、筋電データが安定してから測定する方法とジェスチャーを行い始めた瞬間も含め、データを測定する方法がある。本研究では実用性を考慮し、行い始めた瞬間も含めてジェスチャーのデータを測定した。



図 1 6 種類のジェスチャー

また、深層学習の統合開発環境は SONY の NeuralNetworkConsole を用いて学習と評価を行った。

3. 実装

本研究では、Unity を使用して Myo から取得した筋電データを CSV ファイルに保存するプロジェクトを作成した。(図 2) CSV ファイルに保存する際、正規化処理をすることで CSV ファイルを NeuralNetworkConsole にそのまま学習データとして使用できるようにした。CSV ファイルのデータは、8 つの筋電センサそれぞれに 4 秒間、0.01 秒ずつ記録する。その為、8 列 400 行となっている。



図 2 実験画像

A motion control system for XR using EMG interface

[†] Kazuki Itoh, Hitoshi Tamura, Nippon Institute of Technology faculty of innovative system engineering.

筋電データの学習には、畳み込みニューラルネットワークと再帰型ニューラルネットワークのLSTM(Long short-term memory)のネットワークを使用した。畳み込みニューラルネットワークは、畳み込み層とプーリング層の2層で構成される順伝播型のネットワークで画像のパターン認識に高い性能を発揮する。LSTMは、ネットワークの出力を別のネットワークに入力する様な再帰的な構造をしており、中間層のLSTM Block 入力ゲート、出力ゲート、忘却ゲートとメモリを持つ事で時系列データに対し、高いパターン判別性能を発揮する。(図3)

どちらのニューラルネットワークでも学習用データ300個、評価用データ60個で学習させ、畳み込みニューラルネットワークの正解率(Accuracy)は33.3%、LSTMの正解率は83.33%であった。

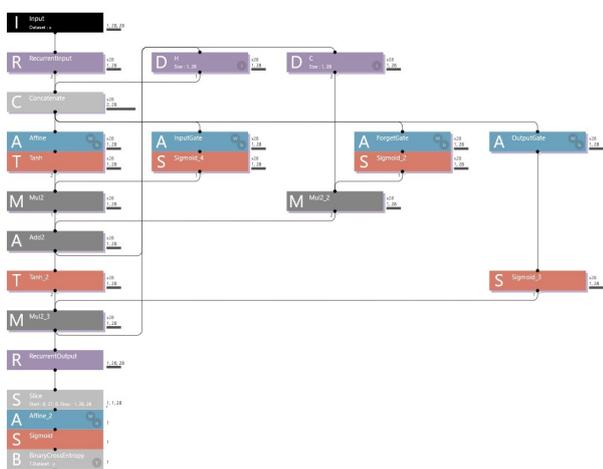


図3 LSTMの構造

4. 実験結果

LSTMが筋電データの学習に適していると考え、学習用データ2880個、評価用データ360個で学習を行った。正解率は96.38%、F値(F-Measure)が96.35%でコントローラーとして使用できる高い判別精度を出すことが出来た。(表1)

また、Myo装着時のズレによる影響を調べるため、可能な限り前回と同じ向き、位置でMyoを装着し、測定を行った。そのデータを評価用データ60個として設定し、学習用データは上記の2880個のデータを使用して推論を行った。結果は正解率が、35.00%となり、装着時のズレによる影響で正解率が大幅に低下する事が分かった。次に表面筋電位の個人差における正解率の変化を調べるため、2人分の筋電データを計測した。そのデータをそれぞれ評価用データ60個として設定し、学習用データは上記に記した2880個を使用して推論を行った。結果、正解率は16.67%と23.33%となり、個人差による影響が大きい為、個人で学習させた筋電データを他人に用いることが難しいと分かった。

表1 実験結果

Accuracy(正解率)	96.38%
Precision(適合率)	96.60%
Recall(検出率)	96.38%
F-Measures(F値)	96.35%

5. 考察

以上の結果からジェスチャーの筋電データで深層学習を行った場合、高い精度を得られることが分かった。その為、手首の角度や腕のひねりなどの再現も可能であると考えられる。しかし、Myoを一度外してしまうと正解率が大幅に低下する為、装着時に学習したデータから前回の装着状態へ誘導するシステムを作成する必要がある。また、学習用データを用意する時間や筋電データを学習させる時間に4時間近くかかってしまうと分かり、実用性を考えると判別精度だけでなく、準備時間の短縮を図る必要がある。

デバイスのズレに関しては様々な向きでデバイスを装着し、計測を行う。そのデータで学習・推論を行う事でズレを修正する必要がなくなり、時間の短縮にも繋がる可能性がある。

そして、筋電によるジェスチャー認識はハンズフリーで画像認識による判別を行わない事からARやMRの分野でも応用できると考えられる。

6. おわりに

表面筋電位をVRにおけるモーションコントローラーとして使用するにあたり、深層学習による判別性能の向上を本研究で行った。デバイスを外さず学習用・評価用の筋電データを計測し、学習を行った場合、高い正解率を得られることが分かった。しかし、個人の学習したデータを他人に用いることが難しい事や推論をリアルタイムで処理することを考えた場合、最小限の学習用データで学習を行うなど準備時間の短縮を図る必要があると分かった。

参考文献

[1] 廣田 衛, 佐藤 淳紀, 坪井 歩武, 横山 正幸, 柳澤 政生”フィンガーレスグローブ型筋電インターフェースのための5指タッピング動作のCNNによる識別”, 情報処理学会, 研究報告コンピュータビジョンとイメージメディア (CVIM), 2018-CVIM-210, 18号, pp1-6
 [2] 山田知之, 棟方渚, 小野哲雄, ”アクティブ音響センシングを用いた手のジェスチャー認識手法の検討”, 情報処理学会, 研究報告ゲーム情報学 (GI), 2016-GI-36, 23号, pp1-8
 [3] 御手洗彰, 棟方渚, 小野哲雄, ”物を把持した状態における筋電センサを用いたハンドジェスチャー入力の問題抽出と新手法の提案”日本バーチャルリアリティ学会論文誌, vol. 22, no. 1, pp. 41-50, 2017.