

上肢の表面筋電位を用いた手形状の認識

岩本万由子^{†1} 石川由羽^{†1} 高田雅美^{†1} 城和貴^{†1}

概要: 本稿では、上肢の sEMG(surface electromyographic)を用いた手形状認識を提案する。sEMG はウェアラブルデバイスである Myo を用いて取得される。対象となる手形状は、グー、チョキ、パーの 3 つである。手形状認識には、FFT とウェーブレット変換から得られた 4 つの特徴量が用いられ、識別器には SVM と改良された k-NN が用いアンサンブル学習を行う。実験の結果、グーの正答率は 86%、パーの正答率は 81.3%、チョキの正答率は 75.6%となる。

1. はじめに

近年、透過型 HMD (Head Mounted Display) が多く発売され、透過型 HMD と AR (Augmented Reality) を用いたアプリケーションが開発されている。透過型 HMD と AR を用いたアプリケーションは、ユーザの直感的な理解を可能にする。例えば、道案内の場合、方向指示や地図が目の前の現実の風景に重なるように表示される。その結果、ユーザはスマートフォンの画面と現実の風景を見比べる必要がなくなり、進むべき方向を直感的に理解することが出来る。有名な透過型 HMD として Google Glass (Google) [1] と Holo Lens (Microsoft) [2] がある。Google Glass は、音声認識によって操作され、Holo Lens は、音声以外にハンドジェスチャで操作される。ハンドジェスチャは Holo Lens の前方に取り付けられたカメラによって認識される。音声認識は、話すことがマナー違反となる場所やプライバシーが保護されていない公共の場所において操作が難しい。カメラを用いたジェスチャ認識には大きな動きが必要であるため、混雑した場所での操作は困難である。以上のことより、音声認識、カメラによるジェスチャ認識以外の透過型 HMD の操作方法が求められる。

本稿では、透過型 HMD に表示される AR の操作方法として筋肉と機械を結ぶ MMI (Muscle-Machine Interface) に注目する。MMI の関連研究[3][4]は、腕の sEMG を用いて車椅子が操縦されるものや顔の sEMG を用いて文字が入力されるものがある。多くの MMI に関する研究は、医療に関するものである。しかし、私たちの提案手法は、MMI を用いて AR を操作することを目的としており、どのような環境でも使用でき、装着脱が容易で、汎用性があることが求められる。EMG を使用する MMI は、身体の電気信号と機械が結ばれるため、外的環境の影響を受けにくい。sEMG は容積伝導によって伝えられる活動電位を皮膚上から計測され、sEMG は筋線維活動全体を評価する。このことより、sEMG は十分に筋肉の状態を取得できる[5]。指の動きの多くは上肢の筋肉の影響を受ける。上肢の筋肉の状態を調べることで、指の動きを間接的に知ることが出来る。本稿において、対象となる手形状は、グー、チョキ、パーの 3 種

類である。これらの手形状は、基本的な動作である手の開閉が含まれる。グーやパーの動作は大きく、チョキの動作は小さい。そのため、これらの手形状を対象とすることで、提案手法を手の基本的な動作によってテストし、さらに大きな動作と小さな動作の比較をする。上肢の sEMG から手形状の認識を行う関連研究[6][7]では、sEMG を取得する機器を装着する際、専門的知識が必要となること、事前にユーザのデータを登録することが必要である。このことはユーザにとって負担である。本稿では EMG に関する専門的知識のないユーザが容易に装着脱できる Myo[8]を用い、且つ、事前にユーザの sEMG を登録せずに手形状判別が可能手法を提案する。

以下本稿では、2 章において、本稿で用いる測定機器と手形状認識の処理手順が紹介され、3 章では手形状認識に用いられる特徴量が記述される。4 章では、単一特徴量での識別とアンサンブル学習の識別実験が行われ、5 章では、まとめが述べられる。

2. 手形状予測のための提案手法

2.1 処理手順

上肢の sEMG から手形状認識を行うための処理手順を以下に示す。

手順1. Myo を用いた上肢の sEMG の取得

手順2. 前処理

手順3. 4 つの特徴量の取得

手順4. アンサンブル学習による手形状の認識

手順 1 では、手形状識別のための sEMG を Myo を用いて取得する。Myo は Thalmic Labs 社によって開発されたアームバンド型のジェスチャーコントローラであり、8 チャンネルの筋電センサが搭載されている。図 1 は Myo の各電極に対するチャンネル番号が示される。サンプリング周波数は 200Hz である。取得される sEMG データは EMG の活動量であるため、単位はない。取得データは -127 から 127 の間に収まるように変換される。Myo の装着位置は右前腕の手肘である。Myo 装着の様子を図 2 に示す。図 2 のようにチャンネル 4 が手のひら側の中指の延長線上に装着される。手順 2 では、前処理が行われる。前処理では RMS が計算される。次にノイズ除去が行われる。最後に正規化が行われる。EMG 解析において一般的に使用される RMS を

^{†1} 奈良女子大学
Nara Women's University



図 1 Myo

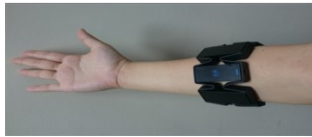


図 2 Myo 装着時の様子

求め、平均振幅に着目した EMG データを取得する。RMS の式を以下に示す。変数 s はある時点の EMG 値である。 T は RMS が適用される時間幅である。

$$RMS_s = \sqrt{\frac{1}{T} \sum_{i=1}^T (s_i)^2} \quad (1)$$

sEMG 取得時に、装着ずれによるノイズ混入が予想されるため、平均 $\pm 2\sigma$ に補正することでノイズ除去を行う。最後に 8 チャンネル全体で 0~1 の範囲で正規化が行われる。正規化によって、sEMG の振幅の個人差を減らす。手順 3 では、FFT と wavelet 変換を用いた特徴量を求める。FFT から求められる特徴量はスペクトル値の総和の順位である。その特徴量はチャンネル間の関係性が示される。wavelet 変換から得られる特徴量は、EMG の時間変化が考慮される。なお、手順 3 の特徴抽出に関しては、3 章で詳細が説明される。手順 4 では、手形状判別のために SVM と改良した k-NN によってアンサンブル学習が行われる。アンサンブル学習によって、独立した複数の識別器によって最終的な答えが導かれ、汎化能力が向上する。SVM で用いられるカーネルは rbf である。なお、k-NN の改良手法は 2.2 節で詳しく説明される。アンサンブル学習では、本稿では 3 手形状の多値判定が行われる必要があるため、AdaBoost.M1 が用いられる。

2.2 k-NN の改良手法

k-NN に改良をすることで、より広域的なクラス分類を行う。k-NN の改良処理手順を以下に示す。

- 手順 a. パラメータ k の計算
- 手順 b. テストデータから近傍 k 個の学習データ選択
- 手順 c. 各選択された学習データから近傍 k 個の学習データ選択
- 手順 d. テストデータクラスの分類

手順 a では、最適な k 値がギャップ統計量によって計算される。手順 b では、テストデータから最近傍 k 個の学習データが選択される。距離関数は、ユークリッド距離が用いられる。手順 c では、手順 b で選択された学習データから、それぞれ最近傍 k 個の学習データが選出される。手順 d では、選択された k^2 個の学習データのクラスを用いた多数

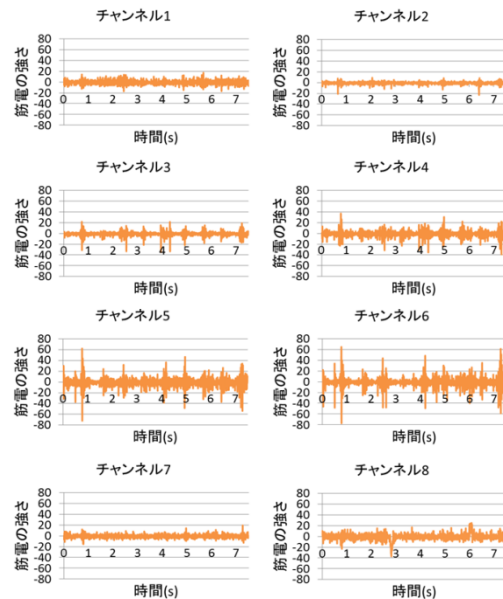


図 3 上肢を上下に振った際の sEMG

決が行われ、テストデータクラスが分類される。一般的な k-NN では、手順 b での k 個の学習データクラスからクラス分類される。選出された学習データに異常値が含まれている場合、識別結果は異常値の影響を受けやすい。本稿で用いる k-NN では、手順 c での k^2 個の学習データクラスによる多数決をし、より広域的なクラス分類を行うことで、異常値の影響を減らす。

3. 特徴量の抽出方法

3.1 FFT から求められる特徴量

FFT から求められる特徴量の処理手順を以下に示す。

- 手順 A. 各チャンネルのスペクトル総和の計算
- 手順 B. 各チャンネルの順位付け
- 手順 C. チャンネル 3, 7, 8 の順位取得

手順 A では、FFT を求め、チャンネルごとのスペクトルの総和を求める。これにより、各チャンネルが取得する筋肉の活動量に着目する。じゃんけんの都度、スペクトルの総和は変化するが、チャンネル間のスペクトル総和の関係性はほとんど変化しない。そのため、手順 B ではスペクトルの総和を降順に順位付けを行う。ただし、順位付けを行う際、チャンネル 5, 6 は除く。sEMG が計測される際、上肢を上下に振る可能性があるが、この動作はじゃんけんの手形状の判別には関係ない。そのため、この特徴量では、そのような sEMG を取得するチャンネルはノイズになりうる。上肢を 7 秒間上下に振る際の各チャンネルの sEMG が図 3 に示す。図 3 より、この特徴量においてチャンネル 5, 6 はノイズになりうると考えられる。手順 C では、チャンネル 3, 7, 8 の順位が 3 次元の特徴量となる。特徴量として用いられるチャンネルは手形状ごとの特性を特に表す必要がある。各チャンネルが取得する sEMG を発生させる筋肉の推定を行うことで、特徴量として用いられるチャンネルが適切かどうか判断する。チャンネル 3 が主に取得す

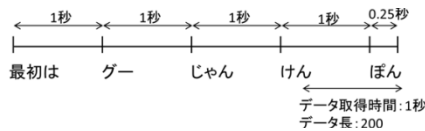


図 4 データ計測

る sEMG は、浅指屈筋[9]によるものと推測される。この筋肉は、手関節の屈曲に関する筋肉であり、手が握られる動作を行ったときに sEMG が取得される。チャンネル 7 が主に取得する sEMG は、総指伸筋[9]によるものであると推測される。この筋肉は、指の伸展に関するものであり、特にチョコキパーで指を伸ばす動きをするときに sEMG が取得される。チャンネル 8 が主に取得する sEMG は、小指伸筋[9]によるものであると推測される。特にパーで小指を伸ばす際に sEMG が取得される。よって、FFT から得られる特徴量にはチャンネル 3, 7, 8 を用いる。

3.2 wavelet 変換から求められる特徴量

wavelet 変換から求められる 3 つの特徴量の処理手順が以下に示される。

- 手順 i. 時間経過を考慮したスペクトル値の算出
- 手順 ii. スペクトル値の平均をエリアごとに計算
- 手順 iii. 各エリアと各チャンネルの組み合わせ

手順 i では、EMG の時間的変化を考慮するために wavelet 変換が用いられる。連続ウェーブレット変換を用い、マザーウェーブレットは Morlet である。σ の値は 4 である。手順 ii では、スペクトルの時間変化に注目し、wavelet 変換後、周波数と時間について 6 つのエリアに分け、各エリアのスペクトルの平均値を求める。時間軸においては、0 秒～0.5 秒と 0.5 秒～0.75 秒、0.75 秒～1 秒の 3 つのエリアに分ける。周波数軸においては、0Hz～50Hz と 50Hz～100Hz の 2 つのエリアに分ける。手順 iii では、手順 ii で求められる各エリアと各チャンネルを組み合わせる。その組み合わせは手形状の特徴をより表すものである。チャンネルとエリアの組み合わせについては 4.2 節で詳細が述べられる。

4. 実験

4.1 データ計測

被験者は右利きの健常な 20 代女性 30 人である sEMG の計測は、じゃんけんの掛け声に合わせて行われる。計測時の掛け声、データ取得時間を図 4 に示す。計測時、被験者に対し、できるだけ指先のみを動かすように指示が出される。被験者は、ゲー、チョコキ、パーそれぞれ 10 回、合計 30 回計測を行う。

4.2 予備実験

本節では、3.2 節で述べた wavelet 変換から得られる特徴量の組み合わせについての実験を行う。組み合わせられる特徴量の次元数は 2～4 とし、その組み合わせは総当たりで調べられ、各手形状のクラスタが生成されるか k-means 法[16]によって確かめる。クラスタ数は 3～6 に変化さる。(n, a) はチャンネル n のエリア a を表す。クラス分類の結果、

表 1 クラス分類結果

Feature a			
Number of cluster	Rock	Scissors	Paper
1	87	7	13
2	174	54	15
3	8	88	124
4	28	135	81
5	3	16	67

Feature b			
Number of cluster	Rock	Scissors	Paper
1	65	4	4
2	35	135	84
3	191	57	21
4	2	13	67
5	7	91	124

Feature c			
Number of cluster	Rock	Scissors	Paper
1	6	85	129
2	193	53	16
3	2	15	61
4	29	131	85
5	70	16	9

特に各手形状の特徴を表す組み合わせは 3 つある。それは (4,4), (4,5), (7,4), (8,5) の 4 次元の特徴量 a と、(4,5), (7,4), (8,5) の 3 次元の特徴量 b と、(4,4), (7,4), (8,5) の 3 次元の特徴量 c である。表 1 は特徴量 a～c のクラス分類の結果を示す。表 1 にある全ての組み合わせについて、クラスタ数は 5 である。特徴量 a では、ゲーがクラス 2 に、チョコキがクラス 4 に、パーがクラス 3 に多く分類されている。特徴量 b では、ゲーがクラス 3 に、チョコキがクラス 2 に、パーがクラス 5 に多く分類されている。特徴量 c では、ゲーがクラス 2 に、チョコキがクラス 4 に、パーがクラス 1 に多く分類されている。それらの組み合わせは手形状ごとのクラスタを生成するため、特徴量 a～c を wavelet 変換から得られる特徴量として用いる。

4.3 実験

本稿では、汎用性の高い分類を目標とする。単一特徴量では十分な識別が困難であると考えられるため、アンサンブル学習を行う。

実験 1. 単一特徴量の識別

実験 2. AdaBoost.M1

全ての実験において、学習データ数は 720(被験者: 24 人分)、テストデータ数は 180(被験者: 6 人分)であり、5 回交叉検定を行う。識別器である k-NN の k 値はギャップ統計量により決定する。FFT から求められる特徴量は k=5、特徴量 a, c は k=11、特徴量 b は k=9 である。実験 1 では、FFT と wavelet 変換から得られる 4 つの特徴量と 2 つの識別器によって識別が行われ、8 つの結果を得る。実験 2 では、正答率向上のため、弱識別器の重みを改良した AdaBoost.M1 を行う。ブースティング回数は 20 回で打ち切る。

4.4 実験結果と考察

本節では、4.3 節で説明した 2 つの実験の結果と考察が述べられる。

(1) 単一特徴量の識別結果と考察

実験 1 の手形状の識別結果を表 2 に示す。以下、FFT から得られる特徴量を特徴量 f とする。特徴量 f を用いた SVM

表 2 単一特徴量の識別結果

	SVM			
	Feature <i>f</i>	Feature <i>a</i>	Feature <i>b</i>	Feature <i>c</i>
Rock	84.3%	86%	86%	85%
Scissors	75%	70%	69.3%	62.9%
Paper	77%	74%	72.6%	75.3%
average	78.7%	76.6%	75.9%	74.4%

	An improved k-NN			
	Feature <i>f</i>	Feature <i>a</i>	Feature <i>b</i>	Feature <i>c</i>
Rock	76.3%	89.3%	89%	88%
Scissors	72.3%	57.6%	62.3%	55.6%
Paper	69%	72.3%	71.6%	73.6%
average	72.5%	73%	74.3%	72.5%

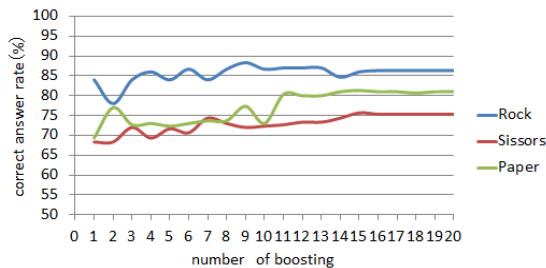


図 5 adaboost.M1 の正答率の変化

は、平均正答率が 78.7%となり、最も正答率が高い。特徴量 *f* を除いて、改良された k-NN はグーの正答率が SVM より高い且つ、手形状ごとの正答率は、グーが最も高く、次にパー、最も正答率が低いものがチョキとなっている。

各特徴量、各識別器についてそれぞれ違った特徴を捉えられている。以上のことよりアンサンブル学習を行うことでさらなる正答率の向上が期待される。

(2) AdaBoost.M1 の結果と考察

実験 2 では、実験 1 で正答率の高い弱識別器の重みを大きくし、且つアンサンブル学習で多くの弱識別器が用いられるように他の弱識別器の重みを改良する。改良方法を以下に示す。

- SVM を用いた特徴量 *f* の重みを 2 倍に設定
- 他の弱判別器の重みを 1 に設定

実験 1 の結果から、SVM を用いた特徴 *f* の正答率が最も高いため、その重みを 2 倍にする。他の弱判別器の重みを 1 に設定する。そうすることで、実験 1 で識別結果の正答率が最も高かった弱識別器の結果をよりアンサンブル学習に強く反映させる。それ以外の弱識別器の重みを同じにすることで、アンサンブル学習が多くの弱識別器によって行われる。改良を行った AdaBoost.M1 の結果を図 5 に示す。図 5 は横軸が boosting 回数で、縦軸が各手形状の正答率を表している。boosting 回数が 15 回以上で各手形状の正答率が安定している。3 手形状の平均の正答率は boosting 回数 15 回目のとき最も高い。その時の正答率はグーの手形状が 86%、チョキの手形状が 75.6%、パーの手形状 81.3%、平均 81%となる。

boosting 回数を増やすごとに各手形状の正答率が向上していることは、多くの弱識別器が最終的な結果に影響を与えていることが推測される。実験 1 で最も正答率が高い識別器の結果より約 2.3%も正答率が高くなっており、

AdaBoost.M1 の改良は適切であると判断できる。boosting 回数については、図 5 より、boosting 回数が 15 以降のとき正答率が安定しているため boosting 回数は 20 で十分である。

5. おわりに

本稿では、上肢の sEMG を用いた手形状の予測の一手法を提案する。手形状の識別には、4 つの特徴量と 2 つの識別器を用いたアンサンブル学習を用いている。特徴量については、FFT から得られる特徴量 1 つと、wavelet 変換から得られる特徴量 3 つを用いている。FFT から得られる特徴量は EMG のスペクトルの強さに注目している。wavelet 変換から得られる特徴量は、EMG の時間的変化を考慮している。識別器は SVM と改良を行った k-NN である。複数の識別器を用いることでアンサンブル学習の弱識別器候補を増やしている。対象手形状はじゃんけんを用いるグー、チョキ、パーの 3 種類としている。被験者を 30 人とすることで、汎用性のある手法であることを示した。単一特徴量の識別結果では、FFT から得られる特徴量を用いた SVM の結果が 78.7%で最も正答率が高くなっている。アンサンブル学習の結果、正答率が 81%である。その正答率は単一の特徴量を用いる場合より向上した。以上の結果より、本稿で提案する手法は、じゃんけんの 3 つの手形状を判別する際、有効なものである。

参考文献

- [1] “google glass”. <https://www.google.com/glass/start/>,(参照 2017-06-05).
- [2] “Hoko Lens”. <https://www.microsoft.com/microsoft-hololens>, (参照 2017-06-05).
- [3] Shafivulla Mohammad, G. Vijay Kumar. Development of sEMG based human machine interface control system for robotic watch. International Conference on Research Advances in Integrated Navigation Systems (RAINS), 2016, pp.1 - 5.
- [4] Meredith J. Cler and Cara E. Stepp. Discrete Versus Continuous Mapping of Facial Electromyography for Human-Machine Interface Control: Performance and Training Effects. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2015, Vol. 23, No. 4
- [5] John V. Basmajian and Carlo J. De Luca. Muscles alive: their functions revealed by electromyography 5th ed. Williams & Wilkins, 1985
- [6] Agamemnon Krasoulis, Sethu Vijayakumar and Kianoush Nazarpour. Evaluation of regression methods for the continuous decoding of fingermovement from surface EMG and accelerometry. 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER) pp.631 - 634
- [7] Ali-Akbar Samadani, Dana Kulic. Hand gesture recognition based on surface electromyography. 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014, pp. 4196 - 4199.
- [8] “Myo”. <https://www.myo.com/>,(参照 2017-06-05).
- [9] I. A. Kapandji. Anatomia funzionale. Monduzzi, 2011