

# 加速度センサを用いたジェスチャ認識のための 筋活動量にもとづく重みづけ手法

梶原 大暉<sup>1</sup> 村尾 和哉<sup>1</sup>

概要：マイクロエレクトロニクス技術の発展によるコンピュータの小型化、高性能化によりコンピュータを常に身につけて生活するウェアラブルコンピューティングに注目が集まっている。ジェスチャ認識では、あらかじめ認識対象となるジェスチャのデータを学習データとして採取し、認識したい未知のジェスチャデータと学習データを比較することで認識している。これまでに提案されているジェスチャ認識システムでは、事前に認識対象となるジェスチャを数回繰り返してもらい、そのデータを学習データとして用いるが、ジェスチャ動作中に力が入り過ぎてしまったり、反対に力が入らな過ぎてしまうことで、同じジェスチャのデータでもばらつきがある。本研究では、学習データとして加速度と筋電位を採取し、入力データは加速度のみを用いてジェスチャを認識する手法を提案する。3種類の野球の投球フォーム「オーバースロー」、「サイドスロー」、「アンダースロー」のジェスチャを認識対象とし、認識には Dynamic time-warping (DTW) を用いる。同種のジェスチャの2つの加速度データの DTW 距離を計算し、筋活動量と DTW 距離の変化の関係を調査する。得られた関係から実際に未知のジェスチャを認識する際に DTW 距離計算において学習データの筋活動量に応じて距離計算方法を変え、動作のばらつきが大きい区間の DTW 距離への影響を小さくする手法を提案し、加速度のみで認識する手法を比較手法として提案手法の有効性を検証した。学習データサンプル数が各ジェスチャ1個ずつの時、被験者5人中に2人の認識精度が上がった。

## 1. はじめに

マイクロエレクトロニクス技術の発展によるコンピュータの小型化、高性能化によりコンピュータを常に身につけて生活するウェアラブルコンピューティングに注目が集まっている。ウェアラブルコンピューティングの発展により、加速度センサ、ジャイロセンサ、筋電位センサなどの装着型センサを用いたジェスチャ認識に関する研究が数多く行われている。今では、iPhone や Android 端末などの携帯電話や、PS4 や Wii などの家庭用ゲーム機のコントローラにも加速度センサが内蔵されており、そのセンサ値から端末の傾きや手の動きなどを検出してユーザの直観的な入力を実現している。

ジェスチャ認識では、あらかじめ認識対象となるジェスチャのデータを学習データとして採取し、認識したい未知のジェスチャデータと学習データを比較することで認識している。これまでに提案されているジェスチャ認識システムでは、事前に認識対象となるジェスチャを数回繰り返してもらい、そのデータを学習データとしているが、ジェスチャ動作中に力が入り過ぎてしまったり、逆に力が入らな過ぎてしまうことで、同じジェスチャのデータでもばらつ

きがある。このばらつきが大きいと、ばらつきを網羅できるほどの学習データサンプル数が必要であり、学習データサンプル数が不十分であると認識精度の低下を引き起こす可能性がある。認識精度を高くするために、従来の研究では大量の学習データを採取したり、加速度センサとジャイロセンサなどの他のセンサを組み合わせで認識している。しかし、学習データの数を増やすとすることで、より多くのデータを採取する必要がある。また、他のセンサと組み合わせで認識すると、装着するセンサの数が増えるため、ユーザの負担や金銭的コストが増える。

本研究では学習データとして加速度と筋電位を採取し、入力データは加速度のみを用いてジェスチャを認識する手法を提案する。認識には Dynamic time-warping (DTW[11]) を用いる。DTW は時系列データを波形の形状がもっとも類似するように非線形伸縮して距離を計算する手法である。まず、同種のジェスチャの2つの加速度データの DTW 距離を計算し、筋活動量と DTW 距離の変化の関係を調査する。得られた関係から実際に未知のジェスチャを認識する際に DTW 距離計算において学習データの筋活動量に応じて距離計算方法を変え、動作のばらつきが大きい区間の DTW 距離への影響を小さくする。これにより、加速度データのみで認識する手法と比較して同量の学習データサ

<sup>1</sup> 立命館大学 情報理工学研究所

ンプル数でも認識精度が向上する。筋電位のデータは学習データ採取時のみ必要であり、認識時の入力データは加速度のみであるため、システム利用時のユーザの負担は既存の加速度センサのみを用いるシステムと同等である。

以降、2節では関連研究を紹介する。3節では提案手法について述べ、4節では提案手法の性能を評価し、最後に5節で本研究をまとめる。

## 2. 関連研究

本節では加速度センサや角速度センサを用いたジェスチャ認識の研究および筋電位センサを用いたシステムの研究を紹介する。

### 2.1 ジェスチャ認識に関する研究

ジェスチャ認識に関する研究は、今までに数多く行われている。Muraoら[1]は、9個の加速度センサと9個の角速度センサをタブレット端末を想定した盤面上に格子状に配置し、27種類のジェスチャの認識を行い、センサ数やセンサの位置を変更することで認識精度の変化を調査している。盤面の保持位置から遠い加速度センサは有効であり、またジェスチャ動作の軸との交点にジャイロセンサを配置することが有効であると示している。また、分類するジェスチャの数が増えると、認識精度が低下することを示している。Agrawalら[2]は、加速度センサが内蔵されている携帯電話をペンのように持ち、空中に描かれたアルファベットを認識している。加速度データを文字の軌跡に変換して認識している。Junkeraら[3]は、「ボタンを押す」、「握手をする」、「受話器を取る」などの日常で行われる10種類の動作を、右腕上腕と右腕前腕に慣性センサを装着し、HMMを用いて認識している。

坂口[10]らは、手の甲に加速度センサと角速度センサを装着し12種類のジェスチャを認識している。加速度センサのみ、角速度センサのみで認識するよりも両方のセンサを組み合わせて認識する方が精度が高いことを示している。

Liu[4]らは、直線や四角形などの8種類の図形を描くジェスチャを、加速度センサを用いて認識している。日々のジェスチャ動作の変化に対応するために、4000試行以上のデータを1か月かけて採取している。認識精度は98.6%を示している。吉田[9]らは、ジェスチャの加速度データを7日分集めることで、ユーザの疲労やジェスチャフォームの忘却がジェスチャ認識に与える影響を調査している。また、リアルタイムでDTW距離の平均と分散を求めることで、ユーザの状態を考慮した適切な学習データ位置探索手法を提案している。高い認識精度を示しているが、大量の学習データを採取する必要がある。また、日々のジェスチャ変化に対応するために、長期間でデータを採取する必要がある。

### 2.2 筋電位に関する研究

Duenteら[8]は、前腕に筋電位の取得および筋肉刺激を与えられる電極を装着し、スマートウォッチの通知を筋肉刺激で伝え、筋電位の値から通知の受け取りと拒否を認識するシステムを提案している。通知の受け取りを手首を上にする動作、通知の拒否を手首を下す動作で認識し、エラー率は3.9%を示している。Zhangら[5]は、手首に加速度センサ、前腕に筋電位センサを装着して、手話の単語と文の認識および18種類のルービックキューブの操作の認識をしている。筋電位の値からジェスチャを切り出している。Ammarら[7]は、192個の電極が格子状に配置された筋電位センサを用いて、1本の指を曲げたり複数の指を曲げたりする27種類のジェスチャの認識を行い、90%の認識精度を示している。学習データとテストデータが異なる日に採取された環境では、センサの装着位置がわずかに変わり認識精度が59%となっているが、装着位置の補正アルゴリズムを適応することで、認識精度は75%に上昇することを示している。

McIntoshiら[6]は、前腕に筋電位センサと圧力センサを装着して、指を曲げるジェスチャや、手首を曲げるジェスチャ、拳を握るジェスチャなどの15種類のジェスチャの認識を行っている。筋電位センサのみ、または圧力センサのみで認識するよりも両方のセンサを組み合わせて認識する方が認識精度が高いことを示している。しかし、この手法は認識時も筋電位センサを装着する必要がある、現実的ではない。

## 3. 提案手法

本研究では、3種類の野球の投球フォーム「オーバースロー」、「サイドスロー」、「アンダースロー」のジェスチャを認識する。はじめに3種類の投球フォームの3軸の加速度と筋電位のデータを集め、同一のジェスチャの加速度データどうしてDynamic time-warping (DTW) 距離[11]を計算し、DTW距離のワーピングパス上での距離の増加量と筋活動量の関係を調査する。この時、筋活動量は筋電位から求める。そして、提案手法は未知のジェスチャの加速度データと学習データの加速度データのDTWを用いて認識する際に、学習データの筋電位データを重みとすることで、動作のばらつきの影響を小さくする手法を提案する。

### 3.1 DTW距離の変化と筋活動量の関係の調査

野球のピッチングフォームの「オーバースロー」、「サイドスロー」、「アンダースロー」の3種類のジェスチャデータを採取した。被験者は20代の男性5名で、5名とも右投げであり、内1名は野球経験があった。室内で3種類のジェスチャを10回ずつ行った。実際にボールを投げるのではなく、ボールを持たずに投球フォームだけを行った。被験者には右腕手首に加速度センサ、右腕前腕部と右腕

三角筋部に筋電位センサを装着してもらい、10回×3種類×5人=150個のデータを採取した。使用したセンサはATR-Promotions社製TSND151とAMP-151である。加速度センサのサンプリングレートは50Hz、筋電位センサのサンプリングレートは1000Hzで計測した。採取したデータはジェスチャ区間を目視で切り出した。採取した筋電位にウィンドウサイズ200[ms]でRoot mean square (RMS)を適用し筋活動量を得た。

同一ジェスチャの加速度データどうしのDTW距離を計算し、ワーピングパス上の距離の増加量と筋活動量の関係を調査する。DTWのアルゴリズムを以下に示す。長さ $m$ と $n$ の2つの離散波形 $X = (x_1, \dots, x_m)$ と $Y = (y_1, \dots, y_n)$ を比較する際、 $m \times n$ 行列 $d(i, j) = (x_i - y_j)^2$ を定義する。次に、 $X$ および $Y$ のインデックスのペアからなるワーピングパス $W = (w_1, \dots, w_k)$ を探索する。ワーピングパスは以下の3つの条件を満足する。

- 境界条件

$$w_1 = (1, 1), w_k = (m, n)$$

- 連続性

$$w_k = (a, b), w_{k-1} = (a', b') \Rightarrow a - a' \leq 1 \wedge b - b' \leq 1$$

- 単調性

$$w_k = (a, b), w_{k-1} = (a', b') \Rightarrow a - a' \geq 0 \wedge b - b' \geq 0$$

距離最小のパスを探索するために、以下の手順を適用する。

#### (1) Initialization:

$$DTW(0, 0) = 0$$

$$DTW(i, 0) = \infty \text{ for } i = 1, \dots, m$$

$$DTW(0, j) = \infty \text{ for } j = 1, \dots, n$$

#### (2) Do for $i = 1, 2, \dots, m$ :

Do for  $j = 1, 2, \dots, n$ :

$$DTW(i, j) = d(i, j) + \min \begin{cases} DTW(i-1, j-1) \\ DTW(i-1, j) \\ DTW(i, j-1) \end{cases}$$

#### (3) Output:

$$\text{Return } DTW(m, n)/(m+n)$$

上式によって得られたコスト $DTW(m, n)$ が $X$ と $Y$ の距離となる。ジェスチャを認識する際は、波形と正解ジェスチャのラベルの組である複数のテンプレートを保持しておく。入力データとすべてのテンプレートとの距離を計算し、最短距離のテンプレートに付与されているジェスチャラベルを結果として出力する。

5人の被験者をA, B, C, D, Eとし、被験者Aの、オーバースローのデータのうち2つのデータのすべての組合せでDTW距離を計算したときのDTW距離の変化と三角筋の筋電位の関係を図1(左)に、同様にサイドスローとアンダースローについてもそれぞれ図1(中)と図1(右)に示す。

ワーピングパスの隣接する要素におけるDTW距離の差 $d_l = DTW(w_{l+1}) - DTW(w_l) (0 \leq l < k)$ を計算する。DTW距離を計算したときのワーピングパスの隣接する要素におけるDTW距離の差 $d_l (0 \leq l < k)$ とする。また、学習データと同時に採取した筋電位を $x_0, \dots, x_n$ とし、 $w_l$ の学習データのインデックス $j$ における筋活動量 $x_j (0 \leq j < k)$ を得る。図の縦軸は $d_l$ 、横軸は $x_j$ である。

同様に被験者B, C, D, Dについてはそれぞれ図2, 図3, 図4, 図5に示す。図より被験者Aと被験者Eは筋活動量が大きい時にDTW距離の増加が大きいことがわかる。また被験者Bは筋活動量が被験者Bの同一のジェスチャから得られた筋活動量の最大値の半分あたりでDTW距離が大きく増加することがわかる。同様に被験者Cのサイドスローとアンダースローのジェスチャと、被験者Dのオーバースローとサイドスローとアンダースローのジェスチャは、同一被験者の同一ジェスチャから得られた筋活動量の最大値の半分あたりでDTW距離が大きく増加することがわかる。この結果より、学習データの筋活動量が同一被験者の同一ジェスチャから得られた筋活動量の最大値の半分より大きい時は同じジェスチャを認識した時にDTW距離が大きく増加しているため、その時にDTW距離を小さくするように補正を掛けることで、同じジェスチャどうしでDTW距離計算を行った時にジェスチャの類似度を高めることができると考えられる。

### 3.2 提案手法のアルゴリズム

筋活動量の閾値を $d (0 < d)$ 、重み付け係数を $g (0 < g < 1)$ とすると、提案手法では、DTW距離を以下のように求める。

#### (2) Do for $i = 1, 2, \dots, n$

Do for  $j = 1, 2, \dots, m$

if  $x_j > d$

$$DTW(i, j) = |a_i - b_j| * g + \min \begin{cases} DTW(i-1, j-1) \\ DTW(i-1, j) \\ DTW(i, j-1) \end{cases}$$

else

$$DTW(i, j) = |a_i - b_j| + \min \begin{cases} DTW(i-1, j-1) \\ DTW(i-1, j) \\ DTW(i, j-1) \end{cases}$$

提案手法では、(2)の中の $|a_i - b_j|$ の計算を学習データの三角筋の筋活動量である $x_j$ によって変える。 $x_j$ が閾値 $d$ より大きければ $|a_i - b_j| * g$ とする。 $g$ は $0 < g < 1$ であるので $|a_i - b_j| * g < |a_i - b_j|$ となり、筋活動量が閾値より大きいとき $DTW(i, j)$ の評価を小さくすることができる。定数 $d$ は3.1節より学習データの筋活動量 $x$ から上位5%を外れ値として除き、残りの筋活動量の最大値の半分の値とした。また、 $g$ は0.5とした。

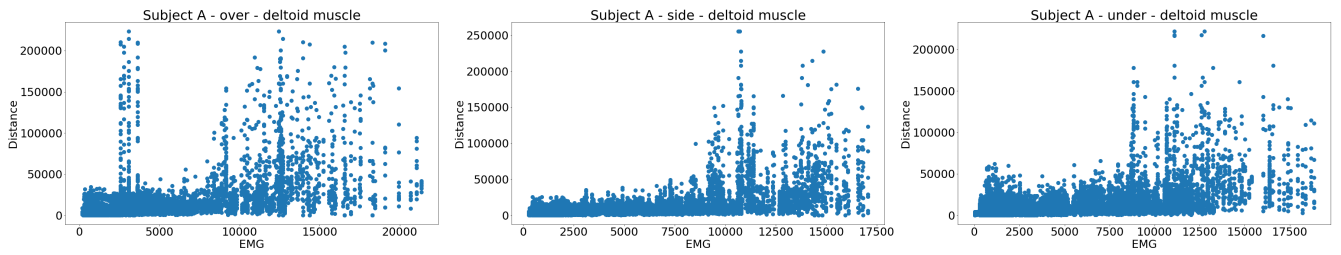


図 1 被験者 A のオーバースロー (左), サイドスロー (中), アンダースロー (右) の DTW 距離と筋活動量

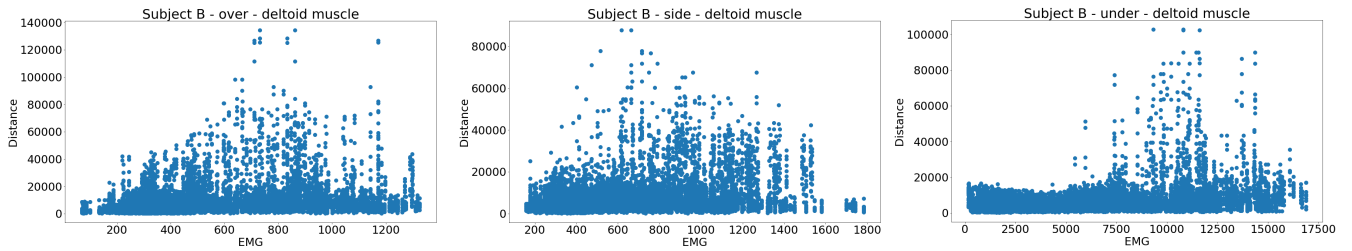


図 2 被験者 B のオーバースロー (左), サイドスロー (中), アンダースロー (右) の DTW 距離と筋活動量

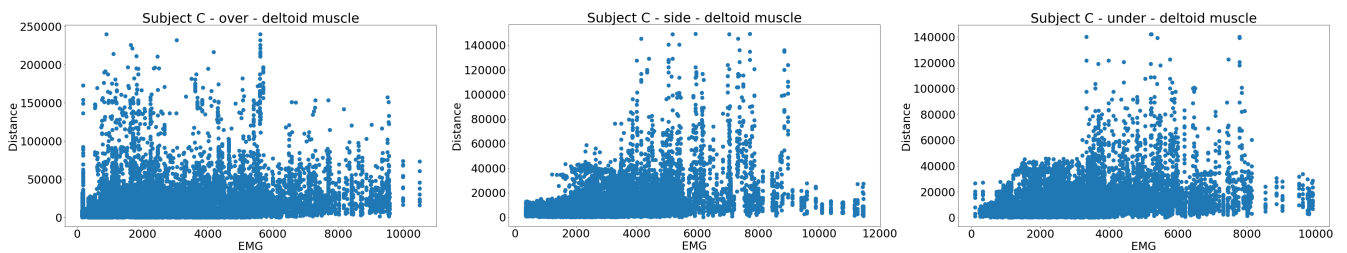


図 3 被験者 C のオーバースロー (左), サイドスロー (中), アンダースロー (右) の DTW 距離と筋活動量

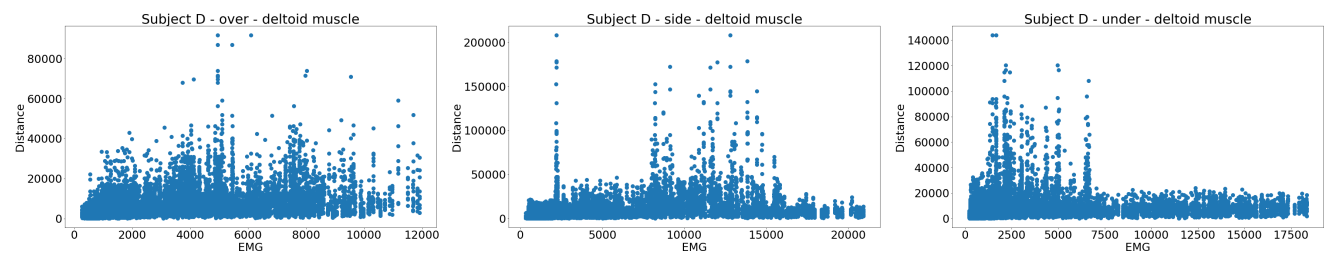


図 4 被験者 D のオーバースロー (左), サイドスロー (中), アンダースロー (右) の DTW 距離と筋活動量

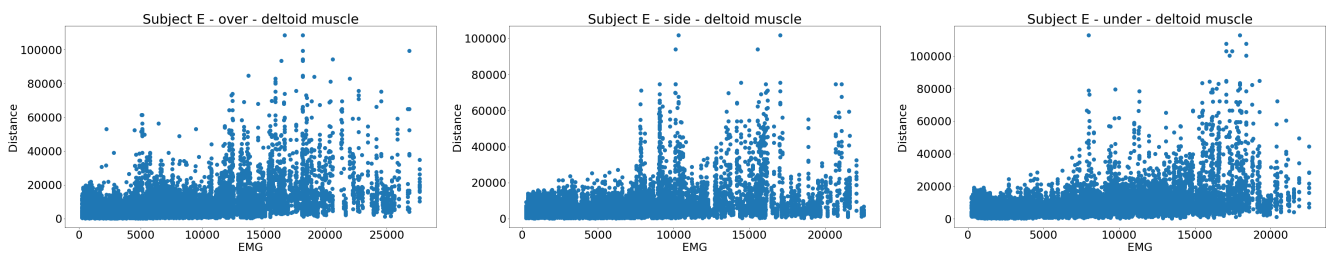


図 5 被験者 E のオーバースロー (左), サイドスロー (中), アンダースロー (右) の DTW 距離と筋活動量

## 4. 評価

### 4.1 評価環境

提案手法の有効性を検証するために3節で採取した10回×3種類×5人=150個のデータを使用し、提案手法と、比較手法として筋電位で重みづけせずに加速度データのDTW距離で認識する手法でジェスチャ認識精度を計測した。認識するジェスチャデータと学習データは同一被験者のものを使用する。例えば、被験者Aのジェスチャを認識する時は、被験者Aのオーバースロー、サイドスローとアンダースロー、アンダースローを学習データとし、学習データサンプル数を各ジェスチャ1個から9個まで1個ずつ変化させたときの認識精度の変化を検証した。学習データ以外のデータをテストデータとし、学習データとテストデータの組合せを変える交差検証を行った。

### 4.2 結果

被験者A, B, C, D, Eの認識結果をそれぞれ図6, 図7, 図8, 図9, 図10に示す。被験者Aは学習データサンプル数が各ジェスチャ1個ずつの時、提案手法の方が比較手法よりも0.02認識精度が高かった。学習データサンプル数が各ジェスチャ2個ずつ以上になると比較の方が認識精度が高かった。被験者Bは比較手法の方が提案手法よりも常に認識精度が高かった。また、比較手法では学習データサンプル数が各ジェスチャ4個ずつ以上、提案手法では各ジェスチャ6個ずつ以上の時認識精度が1.0となっていた。被験者Cは学習データサンプル数が各ジェスチャ1個ずつの時、提案手法の方が比較手法より認識精度が0.02認識精度が高かった。学習データサンプル数が各ジェスチャ2個ずつ以上になると比較の方が認識精度が高かった。被験者Eは常に比較手法の方が認識精度が高かった。被験者Eは提案手法と比較手法ともに、どの学習データサンプル数でも認識精度が1.0であった。学習データサンプル数が各ジェスチャ3個ずつ以上になると、比較手法の認識精度は全ての被験者で0.9以上となっていた。

### 4.3 考察

実験結果から被験者Aと被験者Cは学習データサンプル数が各ジェスチャ1個ずつの時、提案手法の方が比較手法よりも認識精度が高く有効であることが分かった。しかし、被験者A, 被験者Cともに学習データサンプル数が各ジェスチャ2個ずつ以上になると比較手法の方が提案手法よりも認識精度が高かった。被験者Eは学習データサンプル数が各ジェスチャ1個ずつと最も少ないときに既に認識精度が1.0となっていたため、提案手法の有効性を確認することができなかった。提案手法であまり認識精度を上げることができなかった原因は、被験者ごとにDTW距離の増加量と筋活動量の関係が違ふことだと考えられる。3.1

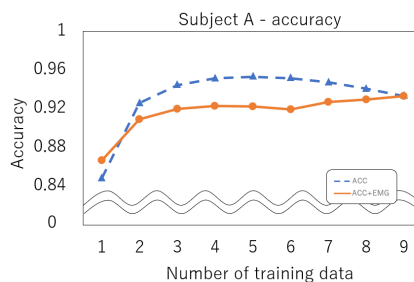


図6 被験者Aの認識結果

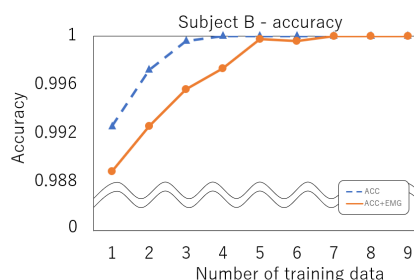


図7 被験者Bの認識結果

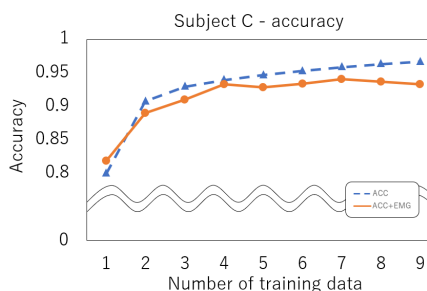


図8 被験者Cの認識結果

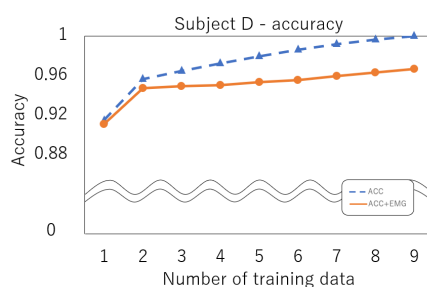


図9 被験者Dの認識結果

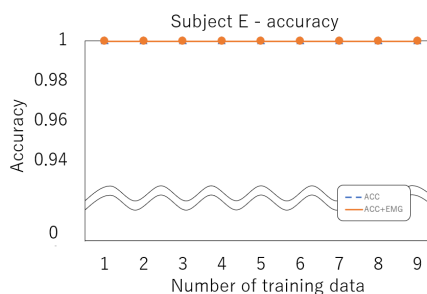


図10 被験者Eの認識結果

節より、被験者やジェスチャが違うことが確認でき、対象とするジェスチャや対象とするユーザごとに、提案手法の閾値や重み付け係数を変える必要があると考えられる。

## 5. おわりに

本研究では、同種のジェスチャの2つの加速度データのDTW距離を計算し、筋活動量とDTW距離の変化の関係を調査し、得られた関係から実際に未知のジェスチャを認識する際にDTW距離計算において学習データの筋活動量に応じて距離計算方法を変え、動作のばらつきが大きい区間のDTW距離への影響を小さくする手法を提案した。3種類の野球の投球フォーム「オーバースロー」、「サイドスロー」、「アンダースロー」のジェスチャを認識対象とし、実験結果から学習データの筋活動量が同一被験者の同一ジェスチャから得られた筋活動量の最大値の半分より大きい時は同じジェスチャを認識した時にDTW距離が大きく増加していることが分かり、その時にDTW距離を小さくするように補正を掛けてDTW距離計算を行った。被験者5人から3種類のジェスチャを各10回ずつ、5人×3種類×10回=150個のデータを採取し、筋電位で重みづけせずに加速度データのDTW距離で認識する手法と比較し、提案手法の有効性を検証した。学習データサンプル数が各ジェスチャ1個ずつの時、被験者5人中に2人の認識精度が上がった。今後は対象とするジェスチャや対象とするユーザごとに、提案手法の閾値や重み付け係数を変える方法を検討する。

## 参考文献

- [1] K. Murao, T. Terada, A. Yano, and R. Matsukura, "Evaluating Gesture Recognition by Multiple-Sensor-Containing Mobile Devices," in Proceedings of International Symposium on Wearable Computers (ISWC 2011), PP. 55–58 (2011).
- [2] S. Agrawal, I. Constandache, S. Gaonkar, R. Choudhury, K. Caves, and F. DeRuyter, "Using Mobile Phones to Write in Air," in Proceedings of the International Conference on Mobile Systems, Applications, and Services (Mobisys 2011), PP. 15–28 (2011).
- [3] H. Junkera, O. Amfta, P. Lukowiczb, and G. Tröster, "Gesture Spotting with Body-worn Inertial Sensors to Detect User Activities," Pattern Recognition, pp. 2010–2024 (2008).
- [4] J. Liu, L. Zhonga, J. Wickramasuriya, and V. Vasudevanb, "uWave: Accelerometer-based Personalized Gesture Recognition and its Applications," in Proceedings of the IEEE International Conference on Pervasive Computing and Communication (PerCom 2009), pp. 1–9 (2009).
- [5] 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 Transactions on Systems, Man, and Cybernetics, Vol. 41, Issue. 6, pp. 1064–1076 (2011).
- [6] J. McIntosh, C. McNeill, M. Fraser, F. Kerber, M. Löchtfeld, and A. Krüger, "EMPress: Practical Hand Gesture Classification with Wrist-Mounted EMG and Pressure Sensing," in Proceedings of the ACM SIGCHI International Conference on Human Factors in Computing Systems (CHI2015), pp. 2332–2342 (2015).
- [7] C. Amma, T. Krings, J. Böer, and T. Schultz, "Advancing Muscle-Computer Interfaces with High-Density Electromyography," in Proceedings of the 33rd Annual ACM SIGCHI International Conference on Human Factors in Computing Systems (CHI2015), pp. 929–938 (2015).
- [8] T. Duent, J. Schulte, M. Pfeiffer, and M. Rohs:MuscleIO, "Muscle-Based Input and Output for Casual Notifications," Journal of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT), Vol. 2, Issue 2, No. 64, pp. 1–21 (2018).
- [9] 吉田 稜, 村尾和哉, 寺田 努, 塚本昌彦, "加速度センサを用いたジェスチャ認識に疲労および忘却が与える影響の評価," 情報処理学会研究報告(ユビキタスコンピューティングシステム), Vol. 2012-UBI-33, No. 20, pp. 1–8 (2012).
- [10] 坂口貴司, 金森 務, 片寄晴弘, 佐藤宏介, 井口征士, "加速度センサとジャイロセンサを用いたジェスチャ認識," 計測自動制御学会論文集, Vol. 33, No. 12, pp. 1171–1177 (1997).
- [11] C. S. Myers and L. R. Rabiner, "A comparative study of several dynamic timewarping algorithms for connected word recognition," Journal of the The Bell System Technical Journal, Vol.60, pp. 1389–1409 (1981).