

加速度センサを用いたジェスチャ認識における 連続動作の認識手法

渡邊 光¹ 村尾和哉² 望月祐洋³ 西尾信彦²

概要：加速度センサを用いてユーザの一回きりの動作であるジェスチャを認識する研究が盛んである。既存のジェスチャ認識の研究では認識したいジェスチャの加速度が正確に切り出されることを前提としており、ジェスチャが間隔を空けずに連続して行われている場合や、ジェスチャの前後で関係のない動作が行われている場合は認識精度が低下してしまうという問題点がある。そこで本研究では、加速度データストリームから教師データとの類似度が高い区間である部分シーケンスを検出し、検出した部分シーケンスの中から尤度の高いものをジェスチャ認識の結果として出力することで、ジェスチャが間隔を空けずに連続して行われている場合や、ジェスチャの前後で関係のない動作が行われている場合でも高精度でジェスチャを認識する手法を提案する。5名の被験者から観測した加速度データに対して提案手法を適用し、右手で記号を描くジェスチャではF値の平均値が0.78、右手で数字を描くジェスチャではF値の平均値が0.79となった。

1. はじめに

スマートフォンやウェアラブル端末の普及に伴い、加速度センサを用いてユーザの行動を認識し、認識した行動をインタフェースやライフログとして用いるサービスが登場しており、ユーザのジェスチャを認識する技術に注目が集まっています。加速度センサを用いたジェスチャ認識では一般的に、認識するジェスチャの加速度データ（教師データ）を事前に登録し、認識したいジェスチャの加速度データとすべての教師データの距離を計算し、最も距離が小さい教師データのジェスチャを認識結果とする方法が用いられる。しかしながら、この手法は認識したいジェスチャの加速度の前後で一定時間静止している場合は高い精度で認識できるが、ジェスチャが間隔を空けずに連続して行われる場合や、ジェスチャの前後に認識したいジェスチャとは無関係の動作が行われる場合は、認識したいジェスチャの境界を見つけることが困難であるため、認識精度が低下するという問題がある。ジェスチャをインタフェースとして利用する場合は、連続した入力や、ジェスチャの前後に学習して

いない動作が行われるため、ジェスチャの前後で静止するという制約を設けることは現実的ではない。

本研究ではジェスチャの前後に学習していない動作が含まれる場合や、ジェスチャを間隔を空けずに連続して行った場合でも、高精度に認識が可能なアルゴリズムを提案する。具体的には、入力となる加速度のストリームデータからジェスチャの教師データ（クエリ）に類似する区間を検出し、検出された区間の中からジェスチャの認識結果として尤度が高いものをジェスチャの認識結果として出力する。ストリームデータからクエリに類似する区間を検出する方法としては、櫻井らが提案した時系列データの中からクエリに類似する部分を検出するアルゴリズムであるSPRING [1]がある。しかしジェスチャ認識に対して単純にSPRINGを利用すると、間隔を空けずに連続したジェスチャを認識できない場合がある。また、ジェスチャ認識では通常、復習種類のジェスチャのクエリがあるが、SPRINGは一つのクエリしか設定できない。そのため、SPRINGを改良したアルゴリズムを用いて検出を行う。評価実験では5名の被験者が「○」「△」「×」の3種類のジェスチャを5種類の組合せで行った際の加速度と、「0」から「9」までの10種類の数字を描くジェスチャを行った際の加速度に対して提案手法を適用し、ジェスチャ認識の適合率および再現率、F値を算出した。

¹ 立命館大学大学院 情報理工学研究所
Graduate School of Information Science and Engineering,
Ritsumeikan University

² 立命館大学 情報理工学部
College of Information Science and Engineering, Rit-
sumeikan University

³ 立命館大学 総合科学技術研究機構
Research Organization of Science and Technology, Rit-
sumeikan University

2. 関連研究

「歩く」や「走る」などの一定時間継続される行動の認識は、通常、入力された加速度データを一定のサンプル数（ウィンドウ）で区切り、ウィンドウ内のデータの平均値や分散値などの特徴量に変換して、サポートベクタマシン（SVM）やランダムフォレスト（RF）などの認識器を用いて認識する。しかし、特徴量に変換するとウィンドウ内のデータの時系列での変化の情報が失われる。ジェスチャ認識では、認識対象とするジェスチャどうしが類似している場合が多いため、特徴量を用いた手法はジェスチャ認識には適さない。そのため、通常、ジェスチャ認識では、入力された加速度データと認識するジェスチャの加速度データ（教師データ、またはクエリ）の類似度を計算し、類似度が高いクエリのジェスチャを結果として出力する手法が用いられる [2], [3], [4]。類似度の計算方法としては、2つの時系列波形の時間的非線形圧縮距離を計算する Dynamic Time Warping (DTW) [5] や複数のクエリから生成した認識モデルから入力データが生成される尤度を計算する Hidden Markov Model (HMM) がある。HMM は音声認識で用いられていることが多いが、DTW と比較して大量の教師データが必要となり、ジェスチャのデータを大量に収集することは困難であるため、本研究では DTW を用いる。

ウェアラブルセンサを用いたジェスチャ認識の研究のひとつとして、出田ら [3] は途中までのジェスチャの加速度データと教師データとの距離を逐次的に計算することで、ジェスチャの入力途中でもジェスチャを認識できるアルゴリズムを提案している。入力データと教師データの距離を逐次的に計算できるように DTW を改良し、DTW 距離が最小となる教師データの距離と、次点の教師データの距離が相対的に大きくなった時点で、DTW 距離が最小の教師データに対応するジェスチャを認識結果として出力する。しかし、この手法ではジェスチャのデータが正確に切り出せることを前提としており、ジェスチャの前後に学習していない動作が含まれる場合や、間隔を空けずにジェスチャを連続して行った場合を考慮していない。

Joseph ら [2] は櫻井ら [1] は入力された時系列ストリームデータから、教師データと類似した区間（部分シーケンス）の抽出を行う手法として、DTW を改良した手法である SPRING を提案している。SPRING は、観測された時系列ストリームデータの中から、教師データとの距離が閾値未満となる部分シーケンスを網羅的に検出する。検出された部分シーケンスは区間が重複するものがあるため、その中から教師データとの距離が極小値をとる部分シーケンスのみを検出する。しかし、SPRING は観測された時系列ストリームデータの中に、部分シーケンスが間隔を空けずに出現するような場合を想定していないため、ジェスチャが間隔を空けずに連続した場合はジェスチャを部分シーケ

ンスとして検出できない場合がある。また、SPRING はひとつの時系列ストリームデータから、複数の教師データを用いて部分シーケンスを検出することを想定していないため、同時に部分シーケンスが複数検出される場合がある。そのため、SPRING を複数の教師データを利用するジェスチャ認識に用いるには、同時に部分シーケンスが複数検出された場合に、その中から適切なものを選択する手法が必要である。

3. 提案手法

本研究ではジェスチャが間隔を空けずに連続している場合や、ジェスチャの前後に学習していない動作が含まれる場合に、ユーザがジェスチャの前後で静止したり、ボタンを押したりして明示的にジェスチャの境界を入力することがなくても高い精度でジェスチャを認識するシステムの構築を目指す。提案手法では、観測された加速度のストリームデータに対してジェスチャの教師データごとに SPRING を適用し、教師データと類似する区間（部分シーケンス）を抽出する。しかし、SPRING を用いて教師データと類似する区間を抽出する際、間隔を空けずにジェスチャを行った場合の加速度から正しくジェスチャが検出されない場合がある。そのため本研究では、SPRING をジェスチャが連続して行われている場合を考慮したアルゴリズムに改良し、改良したアルゴリズムを用いてジェスチャ認識を行う。

また、SPRING は1つの入力された時系列ストリームデータに対して、複数の教師データから部分シーケンスを抽出することを想定していない。そのため観測された加速度に対してジェスチャの教師データごとに SPRING を適用し部分シーケンスを抽出した場合、同じ時刻に複数のジェスチャが認識結果として出力される場合があると考えられる。本研究では観測された加速度のストリームデータから、改良した SPRING を用いて教師データごとに部分シーケンスを抽出し、抽出された部分シーケンスの中から尤度の高いものをジェスチャ認識の認識結果として出力する。

3.1 SPRING アルゴリズムと問題点

SPRING は加速度センサや温度センサなどから入力される時系列データから、事前に与えられた教師データと類似した区間（部分シーケンス）を求めるアルゴリズムである。SPRING のアルゴリズムを図 1 に示す。観測された長さ n の時系列データを X_1, \dots, X_n 、事前に与えられた長さ m の教師データを Y_1, \dots, Y_m とする。ここで、 $d(t, i) (t = 1, \dots, n; i = 1, \dots, m)$ をタイムワーピング行列上の点 (t, i) における時系列データと教師データの距離、 $s(t, i)$ を $d(t, i)$ に対応する部分シーケンスの開始点とする。また、 $d(t, i)$ および $s(t, i)$ は式 1 のように求められる。

```

1: Algorithm SPRING
2: input: a new value  $x_t$  at  $t$ 
3: output: qualifying subsequence if any
4: for  $i = 1$  to  $m$  do
5:   Compute  $d_i$  and  $s_i$  by Equations 2 3
6: end for
7: if  $d_m \leq \varepsilon$  then
8:   if  $\forall_i d_i \geq d_{\min} \cup s_i < t_e$  then
9:     Report( $d_{\min}, t_s, t_e$ )
10:    for  $i = 1$  to  $m$  do
11:      if  $s_i \leq t_e$  then
12:         $d_i = \infty$ 
13:      end if
14:    end for
15:  end if
16: end if
17: if  $d_m \leq \varepsilon \cap d_m < d_{\min}$  then
18:    $d_{\min} = d_m; t_s = s_m; t_e = t$ 
19: end if
20: Substitute  $d'_i$  for  $d_i$ 
21: Substitute  $s'_i$  for  $s_i$ 

```

図 1 SPRING のアルゴリズム

$$d(t, i) = |x_t - y_i| + d_{best} \quad (1)$$

$$d'(t, i) = d(t - 1, i)$$

$$d_{best} = \min\{d(t, i - 1), d'(t - 1, i), d'(t - 1, i - 1)\}$$

$$d(t, 0) = 0, d(0, i) = \infty$$

$$s(t, i) = \begin{cases} s(t, i - 1) & (d(t, i - 1) = d_{best}) \\ s'(t - 1, i) & (d'(t - 1, i) = d_{best}) \\ s'(t - 1, i - 1) & (d'(t - 1, i - 1) = d_{best}) \end{cases}$$

$$s'(t, i) = s(t - 1, i)$$

d_i と d'_i ($i = 1, \dots, m$) を m 個の距離値の配列, s_i と s'_i ($i = 1, \dots, m$) を m 個の整数値の配列とする. また, 時刻 t において, $d_i = d(t, i)$, $d'_i = d(t - 1, i)$, $s_i = s(t, i)$, $s'_i = s(t - 1, i)$ とする. 各部分シーケンスの距離については, 式 2, 式 3 のように計算する. SPRING は入力された時系列データと教師データの距離を算出し, 距離 d_m が $d_m \leq \varepsilon$ の条件を満たし, かつ $d_m \leq d_{\min}$ の条件を満たす場合に d_{\min} の値を d_m に更新する. このとき同時に部分シーケンスの開始点を $t_s = s_m$, 終了点を $t_e = t$ に更新することで, x_{t_s}, \dots, x_{t_e} の区間を部分シーケンスとして検出する. ここで ε を距離の閾値となる任意の正の実数とし, 本研究では $\varepsilon = 70$ とする. 検出された部分シーケンスは区間が重複するものがあるため, 区間が重複するものの中から教師データとの距離が最も小さい部分シーケンスを結果として出力する. 具体的には, $\forall_i d_i \geq d_{\min} \vee s_i < t_e$ の条件を満たす場合に部分シーケンスを出力することで, 区間が重複する部分シーケンスの中から教師データとの距離が最も小さい部分シーケンスを出力する.

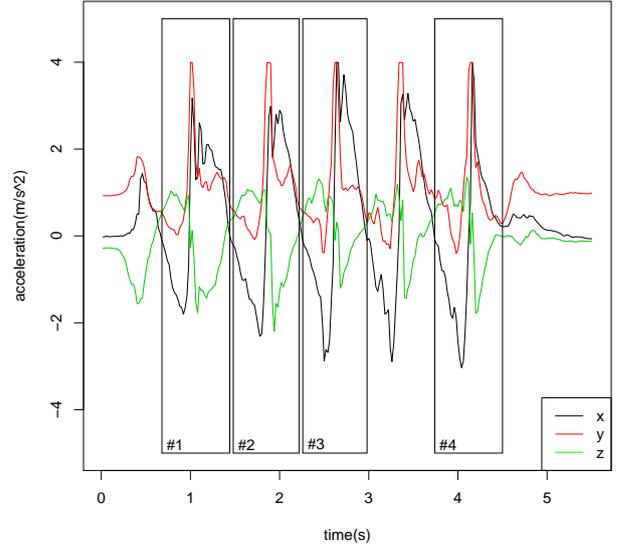


図 2 チョップを連続で 5 回行った際の加速度に対する SPRING の適用結果

$$d_i = |x_t - y_i| + d_{best} \quad (2)$$

$$d_{best} = \min\{d_{i-1}, d'_i, d'_{i-1}\}$$

$$d_0 = d'_0 = 0$$

$$s_i = \begin{cases} s_{i-1} (d_{i-1} = d_{best}) \\ s'_i (d'_i = d_{best}) \\ s'_{i-1} (d'_{i-1} = d_{best}) \end{cases} \quad (3)$$

SPRING には同じジェスチャが連続して行われると, ジェスチャを見逃す可能性があるという問題点がある. 一例として, 図 2 に「チョップ」のジェスチャを 5 回連続で行った際の 3 軸の加速度データから, SPRING を用いて「チョップ」の教師データと類似した部分を検出した結果を示す. 図中の黒線の矩形は部分シーケンスとして検出した区間を示している. 図 2 では, 5 回中 4 回 (#1, #2, #3, #4) の「チョップ」が正しく認識されているが, 4 回目の「チョップ」が認識されていない. SPRING では図 1 の 7 行目の $\forall_i d_i \geq d_{\min}$ では, 時系列データと教師データの距離 d_m が最小値 d_{\min} をとった後, 距離 d_i ($i = 1, \dots, m$) が上昇した場合に部分シーケンスを出力している. しかし, 同じジェスチャが連続して行われた場合は 1 回目のジェスチャの部分シーケンスを検出した後, 2 回目のジェスチャとの距離を計算するため, 教師データとの距離が 1 回目のジェスチャを検出した際の極小値よりも小さくなる場合は, 1 回目のジェスチャは部分シーケンスとして検出されるが, $\forall_i d_i \leq d_{\min}$ を満たさないため部分シーケンスとして出力されず, ジェスチャの見逃しが発生する.

```

1: Algorithm SPRING-improved
2: input: a new value  $x_t$  at  $t$ 
3: output: qualifying subsequence if any
4: for  $i = 1$  to  $m$  do
5:   Compute  $d_i$  and  $s_i$  by Equations 2 , 3
6: end for
7: if  $d_m \leq \varepsilon$  then
8:    $t_s = t_m; t_e = t$ 
9:   if  $\forall_i d_i \geq d_{\min} \cup s_i < t_e \cup$  Equation 4 then
10:     Report( $d_{\min}, t_s, t_e$ )
11:     for  $i = 1$  to  $m$  do
12:       if  $s_i \leq t_e$  then
13:          $d_i = \infty$ 
14:       end if
15:     end for
16:   end if
17: end if
18: if  $d_m \leq \varepsilon \cap d_m < d_{\min}$  then
19:    $d_{\min} = d_m; t_s = s_m; t_e = t$ 
20: end if
21: Substitute  $d'_i$  for  $d_i$ 
22: Substitute  $s'_i$  for  $s_i$ 

```

図 3 SPRING の改良手法のアルゴリズム

3.2 SPRING アルゴリズムの改良

3.1 で示したように、SPRING には同じジェスチャが連続して行われた場合、ジェスチャを見逃す問題点があるため、部分シーケンスの尤度が高い場合には認識結果として出力するよう SPRING のアルゴリズムを改良する。改良した SPRING の手法を 図 3 に示す。図 1 では、同じジェスチャが連続して行われた場合、2 回目のジェスチャを検出した際の極小値が 1 回目のジェスチャを検出した際の極小値よりも小さくなる場合に、1 回目のジェスチャを見逃してしまう問題点があった。そのため、部分シーケンスの始点と終点が一定サンプルの間変化しない場合に、その部分シーケンスの尤度が高いとみなし、 $\forall_i d_i \geq d_{\min}$ を満たさない場合でも認識結果として出力する。具体的には、式 4 の条件を満たす場合に部分シーケンスを認識結果として出力する。 t_s は部分シーケンスの開始時刻を、 t_e は部分シーケンスの終了時刻を示している。 $t_s(t)$ は時刻 t における t_s の値、 $t_e(t)$ を時刻 t における t_e の値、 ε_f を任意の整数とする。式 4 を追加することによって、部分シーケンスの始点 t_s と部分シーケンスの終点 t_e が ε_f サンプルの間変化しなかった場合に、部分シーケンスを認識結果として出力する。本研究では、 $\varepsilon_f = 10$ とする。

$$\forall_j t_s(j) = t_s(t) \cap t_e(j) = t_e(t) \quad (4)$$

$$(j = t - \varepsilon_f, \dots, t - 1)$$

3.3 ジェスチャ認識のアルゴリズム

3.2 節で説明した改良された SPRING のアルゴリズムを用いて、ジェスチャ認識の結果を出力するアルゴリズムについて説明する。SPRING のアルゴリズムを用いてジェ

```

1: input: a new acceleration  $x_t$  at  $t$ 
2: output: detected gesture
3:  $S =$  SubSequence output by SPRING(Fig.1)
4: if  $S! = null$  then
5:   Add  $S$  to  $P$ 
6: end if
7: if  $P \neq null$  then
8:   for each  $p \in P$  do
9:     if Satisfying equation 5 then
10:       Add  $p$  to  $C$ 
11:       Remove  $p$  from  $P$ 
12:     else
13:        $P' = P$ 
14:       Delete  $p$  from  $P'$ 
15:       for each  $p' \in P'$  do
16:         if  $M < D(p, p') \cap Dist(p') \leq Dist(p)$  then
17:           Delete  $P'$  from  $P$ 
18:         end if
19:       end for
20:       for each  $r \in R$  do
21:         if  $M < D(p, r) \cap Dist(r) \leq Dist(p)$  then
22:           Delete  $P'$  from  $P$ 
23:         end if
24:       end for
25:       if  $T \leq S_{time}$  then
26:         Report  $p$  as recognition result
27:         Add  $p$  to  $C$ 
28:         Remove  $p$  from  $P$ 
29:       end if
30:     end if
31:   end for
32: end if
33: if  $C \neq null$  then
34:    $min = \infty$ 
35:   for each  $c \in C$  do
36:     if  $Dist(C) < min$  then
37:        $Result = c$ 
38:     end if
39:   end for
40:   Report  $Result$  as gesture
41: end if

```

図 4 ジェスチャ認識のアルゴリズム

スチャ認識を行う場合は、観測された加速度に対して教師データごとに SPRING を適用し、SPRING によって得られた部分シーケンスに対応するジェスチャを認識結果として出力する方法が考えられる。しかしこの手法では同じ時刻に複数の部分シーケンスが検出されることが考えられる。そのため本研究では、入力された加速度に対して SPRING を適用し、検出された部分シーケンスの中から尤度が高いものをジェスチャの認識結果として出力する。

尤度の高い部分シーケンスをジェスチャの認識結果として出力するためのアルゴリズムを 図 4 に示す。図 4 では、加速度データを 図 3 に示した SPRING に対して入力し、ジェスチャの教師データと類似する区間である部分シーケンス S を求める。SPRING により部分シーケンス S が検出された場合は、式 5 を用いて部分シーケンス S をジェス

チャの認識結果として出力するか、保留中のシーケンスに追加するかどうかを判定する。式 5 において、 d_1 は距離 $d[m]$ が最小となるジェスチャの距離、 d_2 は距離 $d[m]$ が 2 番目に小さいジェスチャの距離を示している。ここでは式 5 を満たす場合に部分シーケンス S を認識結果として出力候補となる部分シーケンス群 C に追加し、式 5 を満たさない場合は S を保留中のシーケンスとして P に追加する。式 5 の左辺の値は d_1 と d_2 の差が大きいほど小さい値をとる。そのため、式 5 の左辺の値が小さい場合は d_1 をとるジェスチャの距離が他のジェスチャの距離と比較して小さいため、他のジェスチャの部分シーケンスが d_1 をとるジェスチャと区間が重複する確率が低い。そのため、式 5 の左辺が閾値より低い場合は部分シーケンス S の尤度が高いとみなし、出力候補となるシーケンス群 C に追加する。閾値 α は本研究では 0.50 に設定した。

$$\frac{d_1}{d_2} < \alpha \quad (5)$$

3.3.1 尤度の低い部分シーケンスの削除

次に、保留中の部分シーケンスを、他の保留中の部分シーケンスや既にジェスチャの認識結果として出力された部分シーケンスとの、重複率や尤度の高低を比較することで、保留中の部分シーケンスの中から尤度が低いものを削除する処理を行う。尤度の比較には、式 6 により求められる部分シーケンス A に対する部分シーケンス B の重複率と、部分シーケンスの距離に関する評価値である式 7 を用いる。

$$D(A, B) = \begin{cases} 0 & (A_e < B_s \cup B_e < A_s) \\ \frac{B_e - B_s}{A_e - A_s} & (A_s < B_s \cap B_e < A_e) \\ \frac{A_e - A_s}{A_e - A_s} & (A_s < B_s) \\ \frac{B_e - A_s}{A_e - A_s} & (A_s > B_s) \\ 1 & (\text{otherwise}) \end{cases} \quad (6)$$

$$Dist(A) = \frac{d_{min}(A)}{length(A) + length(Y)} \quad (7)$$

式 6 により求められる部分シーケンス A に対する部分シーケンス B の重複率 $D(A, B)$ は $0 \leq D(A, B) \leq 1$ の範囲で部分シーケンス A の中で、部分シーケンス B と重複している区間の割合を表している。例えば部分シーケンス $A[1:10]$ と $B[5:10]$ に対しては $D(A, B) = 0.5$ となる。

式 7 により求められる部分シーケンスの距離に関する評価値は部分シーケンスとそれに対応する教師データの距離を表しており、値が小さいほど教師データとの類似度が高く、尤度が高いことを表している。式 7 において、 $d_{min}(A)$ は部分シーケンス A の距離の最小値を、 $length(A)$ は部分シーケンス A の長さを、 $length(T)$ は部分シーケンス A に対応する教師データ T の長さを表している。

まず保留中の部分シーケンスの尤度を他の保留中の部分

シーケンスと比較し、保留中の部分シーケンスの尤度が低い場合はこれを削除する。保留中の部分シーケンス P から部分シーケンスを 1 つ選択した物を p とし、 P から p を削除した部分シーケンス群を P' とする。ここで、 P' の部分シーケンス群の中に式 8 の条件を満たす部分シーケンス p' が存在する場合は P から p を削除する。 M は重複率の閾値であり、本研究では $M = 0.33$ に設定した。これを P に含まれる全ての部分シーケンスに対してくり返すことで、 P の中から尤度の低い保留中のシーケンスを除外する。

$$M < D(p, p') \cap Dist(p') < Dist(p) \quad (8)$$

式 8 の条件を満たす場合、 p と区間の重複率が高く、 p よりも式 7 で表される評価値が小さい部分シーケンスが存在する。したがって p はジェスチャの認識結果としての尤度が低いため、 p を削除する。

次に、保留中の部分シーケンスを、既にジェスチャの認識結果として出力された部分シーケンスと比較し、保留中の部分シーケンスの尤度が低い場合はこれを削除する。保留中の部分シーケンス P から部分シーケンスを 1 つ選択し、これを p 、既にジェスチャの認識結果として出力された部分シーケンス群を R とする。ここで、 R の部分シーケンス群の中に式 9 の条件を満たす部分シーケンス r が存在する場合は P から p を削除する。これを P に含まれる全ての部分シーケンスに対してくり返す。

$$M < D(p, r) \cap Dist(r) < Dist(p) \quad (9)$$

式 9 の条件を満たす場合、 p と区間の重複率が高く、 p よりも式 7 で表される評価値が小さい部分シーケンスが既にジェスチャの認識結果として出力されている。したがって p はジェスチャの認識結果としての尤度が低いため、 p を除外する。保留中の部分シーケンスのうち、保留中のシーケンスとして P に追加されてからの時間 S_{time} が閾値 T 以上となる部分シーケンスに関しては、出力候補の部分シーケンスとして C に追加し、 P から削除する。部分シーケンスが保留中の部分シーケンスとして P に追加されてから十分な時間が経過している場合、その部分シーケンスと重複する尤度の低い部分シーケンスは 3.3.1 における処理により除外される。そのため、保留中の部分シーケンスとして P に追加されてから十分な時間が経過した部分シーケンスは尤度の高い部分シーケンスとみなせるため、出力から T 以上の時間が経過した部分シーケンスは出力候補となる部分シーケンス群 C に追加する。

3.3.2 認識結果として出力する部分シーケンスの選択

出力候補の部分シーケンス群 C の中に複数の部分シーケンスが存在する場合は、 C の中から最も尤度の高い部分シーケンスを 1 つ選択してジェスチャの認識結果として出力する。具体的には、 C の中から $Dist(c)$ の値が最小とな

る部分シーケンス c をジェスチャの認識結果として出力する。また、ジェスチャの認識結果として出力された c は R に追加する。

4. 評価

提案手法による連続して発生したジェスチャの認識精度を調査するために、「○」「△」「×」の3種類のジェスチャを5種類の組合せで行った際の加速度と「0」から「9」までの10種類の数字を手で空中に描くジェスチャを行った際の加速度データに対して提案手法を適用し、ジェスチャ認識の適合率、再現率およびF値を算出した。

評価実験では右手首に加速度センサを装着した20代の被験者5人（被験者AからE）から観測した加速度データを用い、オフラインで解析した。被験者らは右手首に無線加速度センサ（ATR-Promotions社製TSND121^{*1}）を装着し、ジェスチャを行った際の加速度データをサンプリング周波数50Hzで観測した。適合率、再現率およびF値は式10により算出した。

$$\begin{aligned} \text{適合率} &= \frac{\text{正しく認識されたジェスチャの出力回数}}{\text{実際にジェスチャを行った回数}} \\ \text{再現率} &= \frac{\text{正しく認識されたジェスチャの出力回数}}{\text{ジェスチャ認識結果の出力回数}} \\ \text{F値} &= \frac{2 \cdot \text{適合率} \cdot \text{再現率}}{\text{適合率} + \text{再現率}} \end{aligned} \quad (10)$$

4.1 記号のジェスチャを用いた評価

3種類のジェスチャ（手で「○」を描く、手で「△」を描く、手で「×」を描く）に対して、表1に示すように各ジェスチャを5種類の組合せ（同じジェスチャを連続して行う・異なるジェスチャを交互に行う・ジェスチャの直前に非学習動作を行う・ジェスチャの直後に非学習動作を行う・ジェスチャの直前と直後に非学習動作を行う）で、計15種類の連続したジェスチャの組合せを被験者AからEが行った際の加速度データを観測し、観測した加速度データに対して提案手法を適用した。表1の組合せ7から15における非学習動作としては歩行時のように腕を振る動作を行った。

被験者AからEが表1に示した15通りの組合せでジェスチャを行った際の適合率、再現率およびF値を表2に示す。表2に示した組合せの番号は、表1の番号に対応している。表2において、全てのジェスチャの組合せのF値の平均は0.78となった。しかし、被験者Aの組合せ12、14、15では再現率が低下し、被験者Aの組合せ11と被験者D、Eの組合せ15ではF値が0となっている。理由として、被験者Aの組合せ12、14、15では被験者Aから観測した加速度データを用いて作成した「△」の教師データ

^{*1} 小型無線多機能センサ「TSND121/151」 | ATR-Promotions
<http://www.atr-p.com/products/TSND121.html>

表1 ジェスチャの組合せの一覧

組合せ	ジェスチャの内容
1	「○」を連続5回
2	「△」を連続5回
3	「×」を連続5回
4	「○」と「×」を交互に5回
5	「○」と「△」を交互に5回
6	「△」と「×」を交互に5回
7	「○」の直前に非学習動作
8	「△」の直前に非学習動作
9	「×」の直前に非学習動作
10	「○」の直後に非学習動作
11	「△」の直後に非学習動作
12	「×」の直後に非学習動作
13	「○」の前後に非学習動作
14	「△」の前後に非学習動作
15	「×」の前後に非学習動作

表2 記号のジェスチャを行った際の被験者5名の評価結果

組合せ	被験者A			被験者B			被験者C		
	適合率	再現率	F値	適合率	再現率	F値	適合率	再現率	F値
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.83	0.91
2	0.80	0.80	0.80	1.00	0.83	0.91	1.00	0.83	0.91
3	0.40	0.67	0.50	0.60	1.00	0.75	0.80	0.80	0.80
4	0.80	1.00	0.89	0.60	0.75	0.67	0.80	0.89	0.84
5	0.70	0.88	0.78	0.90	0.90	0.90	0.90	1.00	0.95
6	0.60	0.67	0.63	0.50	0.63	0.56	0.60	0.75	0.67
7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	1.00	0.33	0.50	1.00	1.00	1.00	1.00	1.00	1.00
9	1.00	0.25	0.40	1.00	1.00	1.00	1.00	0.50	0.67
10	1.00	1.00	1.00	1.00	0.50	0.67	1.00	1.00	1.00
11	0.00	0.00	0.00	1.00	0.50	0.67	1.00	1.00	1.00
12	1.00	0.20	0.33	1.00	1.00	1.00	1.00	1.00	1.00
13	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
14	1.00	0.13	0.22	1.00	1.00	1.00	1.00	0.33	0.50
15	1.00	0.25	0.40	1.00	1.00	1.00	1.00	1.00	1.00

組合せ	被験者D			被験者E		
	適合率	再現率	F値	適合率	再現率	F値
1	1.00	0.83	0.91	1.00	0.83	0.91
2	1.00	1.00	1.00	0.80	0.80	0.80
3	0.60	1.00	0.75	0.60	1.00	0.75
4	0.80	0.80	0.80	0.70	0.88	0.78
5	1.00	0.83	0.91	0.70	0.78	0.74
6	1.00	0.91	0.95	0.90	0.82	0.86
7	0.00	0.00	0.00	1.00	0.33	0.50
8	1.00	1.00	1.00	1.00	1.00	1.00
9	0.00	0.00	0.00	1.00	1.00	1.00
10	1.00	1.00	1.00	1.00	1.00	1.00
11	1.00	1.00	1.00	1.00	1.00	1.00
12	1.00	1.00	1.00	1.00	1.00	1.00
13	1.00	1.00	1.00	1.00	1.00	1.00
14	1.00	1.00	1.00	1.00	0.50	0.67
15	0.00	0.00	0.00	0.00	0.00	0.00

と被験者Aが非学習動作を行っている際の加速度データの距離が低下し、非学習動作が誤って「△」のジェスチャと誤認識されたため、再現率が低下したと考えられる。一例として、被験者Aが組合せ12のジェスチャを行った際の加速度とジェスチャの認識結果を図5に示す。図5において、縦軸は加速度 (m/s^2)、横軸は時間 (s)、黒、赤、青の線はそれぞれ加速度のx軸、y軸、z軸を表している。

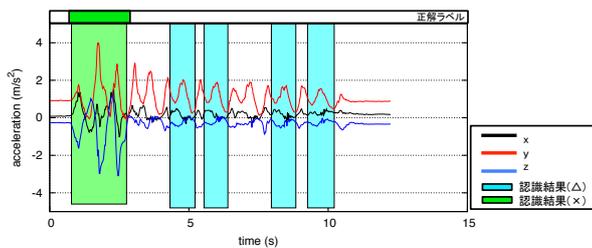


図 5 被験者 A が組合せ 12 を行った際の加速度とジェスチャの認識結果

また、水色の矩形はジェスチャ「△」の認識結果として出力された区間を、緑色の矩形はジェスチャ「×」の認識結果として出力された区間を表している。図 5 では「×」のジェスチャが正しく認識されているが、非学習動作が誤って「△」と検出されたため、再現率が低下している。被験者 A の組合せ 11 では「△」のジェスチャが、被験者 D, E の組合せ 15 では「×」のジェスチャが行われているが、これらのジェスチャが認識されなかったため F 値が 0 となっている。

4.2 数字のジェスチャを用いた評価

「0」から「9」までの 10 種類の数字を手で空中に描くジェスチャを、「0」から「9」までのジェスチャをそれぞれ 5 回行う場合と、「0」から「9」までのジェスチャを順番に行う計 11 通りの組合せで被験者 A から E がジェスチャを行った際の加速度データを採取し、提案手法を適用した。ただし、「0」から「9」までのジェスチャをそれぞれ 5 回行う組合せは 1 名の被験者が 2 種類、5 名で計 10 種類行い、「0」から「9」までのジェスチャを各 1 回ずつ行う組合せは全ての被験者に対して行った。

被験者 A から E から観測した加速度データに対して提案手法を適用した際の適合率、再現率および F 値を表 3 に示す。

表 3 において、全てのジェスチャの組合せの F 値の平均は 0.79 となった。しかし、「0」を 5 回行った場合や「6」を 5 回行った場合の F 値はそれぞれ 0.36, 0.46 と低い値となった。これは「6」のジェスチャの一部に「0」のジェスチャと同じ動作が含まれているため、加速度データと教師データの距離を算出した際に一方のジェスチャの距離が下がった際、同時に他方のジェスチャの距離も低下してしまったことが原因である。一例として、「0」を 5 回行った場合の加速度とジェスチャの認識結果を図 6 に示す。図 6 において、縦軸は加速度 (m/s^2)、横軸は時間 (s)、黒、赤、青の線はそれぞれ加速度の x 軸、y 軸、z 軸を表している。また、赤色の矩形はジェスチャ「0」の認識結果として出力された区間を、青色の矩形はジェスチャ「6」の認識結果として出力された区間を表している。図 6 では 1 から 3 回目の「0」は正しく認識されているが、3 回目の「0」の認

表 3 数字のジェスチャを行った際の被験者 5 人の評価結果

組合せ	被験者	適合率	再現率	F 値
「0」を 5 回	A	0.40	0.33	0.36
「1」を 5 回	A	0.80	0.80	0.80
「2」を 5 回	B	0.80	1.00	0.89
「3」を 5 回	B	1.00	1.00	1.00
「4」を 5 回	C	1.00	0.83	0.91
「5」を 5 回	C	1.00	1.00	1.00
「6」を 5 回	D	0.60	0.38	0.46
「7」を 5 回	D	0.80	0.80	0.80
「8」を 5 回	E	0.60	0.75	0.67
「9」を 5 回	E	0.60	0.50	0.55
「0」から「9」を順番に行う	A	1.00	1.00	1.00
	B	1.00	0.77	0.87
	C	0.90	0.90	0.90
	D	0.80	0.73	0.76
	E	0.80	0.89	0.84

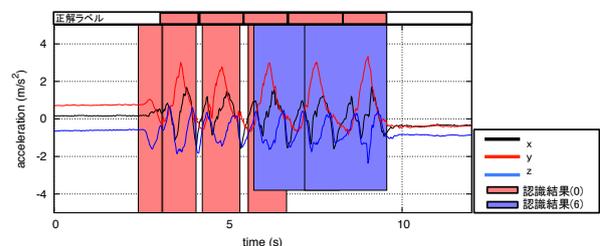


図 6 被験者 A が「0」のジェスチャを 5 回行った場合の加速度とジェスチャの認識結果

識結果と重複している認識結果が出力されていることや、4 回目と 5 回目の「0」のジェスチャが誤って「6」として認識されているため、F 値が低下している。また、「9」を 5 回行った場合に関しても「9」のジェスチャを行った際に誤って「0」と認識され、F 値が低下している。

5. まとめ

本研究では、加速度センサを用いたジェスチャ認識において、ジェスチャが間隔を空けずに連続して行われる場合や、ジェスチャの前後に無関係の動作が行われる場合でもジェスチャ認識を行う手法を提案した。入力された加速度データから SPRING を用いて教師データとの類似度が高い区間である部分シーケンスを検出し、検出した部分シーケンス中から尤度の高い部分シーケンスのみをジェスチャの認識結果として出力することで、ジェスチャが間隔を空けずに連続して行われる場合や、ジェスチャの前後に無関係の動作が行われる場合におけるジェスチャ認識を実現した。評価から提案手法を用いることで、「○」「△」「×」の 3 種類のジェスチャに関しては F 値の平均値が 0.78、「0」から「9」までの 10 種類のジェスチャに関しては F 値の平均値が 0.79 であったが、表 2 では被験者 A の非学習動作を伴うジェスチャや、表 3 の「0」「6」「9」をそれぞれ 5 回

行った際に F 値が低下する結果となった。

現時点では、SPRING により検出された部分シーケンスの中から尤もらしい部分シーケンスを選択する際、部分シーケンスの重複率と距離の最小値のみを用いて選択を行っているが、部分シーケンスが出力された際に、他のジェスチャの教師データとの距離などを用いて尤もらしい部分シーケンスの選択方法を向上させることを検討している。

また提案手法ではジェスチャが完了した時間から、ジェスチャが認識結果として検出されるまでに若干の遅延がある。ジェスチャ認識においては、ジェスチャが完了した後できるだけ短い遅延時間でジェスチャを出力するか、ジェスチャの完了前でもジェスチャの確度が高い場合は早期に認識結果を出力することが望ましい。そのため、提案手法を改良し遅延時間の短縮や早期認識への対応を行う予定である。

謝辞

本研究の一部は、科学研究費補助金基盤 (B) (15H02698) および公益財団法人栢森情報科学振興財団の支援によるものである。ここに記して謝意を表す。

参考文献

- [1] 櫻井保志, Faloutsos Christos, 山室雅司. ダイナミックタイムワーピング距離に基づくストリーム処理. 電子情報通信学会研究誌. D, 情報・システム, Vol. 92, No. 3, pp. 338-350 (Mar. 2009).
- [2] 村尾和哉, 寺田 努. 加速度センサの定常性判定による動作認識手法. 情報処理学会論文誌, Vol. 52, No. 6, pp. 1968-1979 (Jun. 2011).
- [3] 出田 怜, 村尾和哉, 寺田 努, 塚本昌彦. 加速度センサによるジェスチャの早期認識手法の提案とその応用. 情報処理学会研究報告 (モバイルコンピューティングとユビキタス通信研究会), Vol. 2014-MBL-71, No. 15, pp. 1-7 (May 2014).
- [4] J. Liu, L. Zhong, J. Wickramasuriya, V. Vasudevan. uWave: Accelerometer-based personalized gesture recognition and its applications. In proc. of International Conference on Pervasive and Mobile Computing (PerCom 2009), Vol. 5, No. 6, pp. 657-675 (Mar. 2009).
- [5] C.S. Myers, L.R. Rabiner. A comparative study of several dynamic time-warping algorithms for connected-word recognition. The Bell System Technical Journal, Vol. 60, No. 7, pp. 1389-1409 (Sep. 1981).