

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

出田 怜^{1,a)} 村尾 和哉^{2,b)} 寺田 努^{1,3,c)} 塚本 昌彦^{1,d)}

概要：端末に搭載された加速度センサから得られる値を用いて端末の動きを検出することで従来のインタフェースでは困難であった入力の実現できる．一般的に加速度センサを用いたジェスチャ認識では，ジェスチャ終了後に認識処理を開始するため，ジェスチャ入力を行ってからフィードバックの出力までに遅延が生じ，インタフェースの操作性の低下につながる．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，途中までのジェスチャ入力データと教師データとの距離を逐次的に計算する早期認識アルゴリズムとその応用について述べる．提案手法では各教師データとの距離計算において，認識結果の候補が拮抗している場合には，認識結果を出力せずに次サンプルの入力を待ち，確度の高い唯一の候補が決定した時点で認識結果を出力する．提案手法を現場で用いた例として，参加型演劇における参加者の動き取得に早期認識を用いた事例を示し，早期認識の有用性を考察する．

1. はじめに

マイクロエレクトロニクスの発展によるコンピュータの小型化や高性能化，軽量化に伴い，携帯電話やタブレットパソコンといったコンピュータを持ち歩いて，いつでも利用する形態であるモバイルコンピューティングが人々の生活に浸透している．しかしながら，デバイスの小型化によるボタンの減少や表示領域の縮小によって操作性は低下しているといえる．これに対して，端末に加速度センサなどの小型センサを内蔵し，それらを応用することでより操作性を向上させたデバイスが数多く登場しており，その応用が注目されている．特に，iPhone や Android 端末などのスマートフォンや Wii や PS3 などの家庭用ゲーム機のコントローラには加速度センサが内蔵され，センサの値から端末の傾きや動きなどを検出して，ブラウジング中のスクロール操作や 3D オブジェクトの描画，キャラクタの操作などが直観的に行える．このことから今後，加速度センサによるジェスチャ認識を利用した入力インタフェースが数多く登場することが予想される．一般的に加速度センサを用いたジェスチャ認識では，ジェスチャ終了後に認識処理を開始するため，例えば，端末を傾けて 1 秒後に画面内の

表示が回転するといったように，インタフェースの操作性の低下につながる．単純にジェスチャの途中で認識するアプローチでは，途中までの動作が類似している複数のジェスチャを学習している場合，認識精度の低下を招く恐れがある．そこで筆者らは，未知のジェスチャデータとあらかじめ登録しておいたジェスチャデータとの距離計算を逐次的に行い，あるジェスチャの尤度が他のジェスチャの尤度に対して相対的に高くなった時点で認識結果を出力することで，認識精度を低下させることなく早期にジェスチャを認識する手法を提案している [1]．本稿では，提案手法の現場での応用例として，観客参加型演劇 YOUPLAY Vol.1 における参加者の動きを認識した事例について述べる．観客参加演劇とは，一般客が観衆ではなく演者として演劇を楽しむライブエンタテインメントである．

以降，2 章で関連研究を紹介し，3 章では早期認識のアルゴリズムを説明する．4 章で観客参加型演劇に提案手法を適用した事例について紹介する．最後に 5 章で本研究をまとめる．

2. 関連研究

2.1 ジェスチャ認識を用いたアプリケーション

富林らは両手に装着した加速度センサによりジェスチャを認識し，DJ パフォーマンスを行う DJ 支援システムを提案している [2]．数種類のジェスチャを音量の調整や，曲の再生，停止などに割り当て，認識結果に応じて操作を行い，見た目にもインパクトがあるパフォーマンスを実現し

¹ 神戸大学大学院工学研究科

² 立命館大学情報理工学部

³ 科学技術振興機構さきがけ

a) izuta.r@stu.kobe-u.ac.jp

b) murao@cs.ritsumeit.ac.jp

c) tsutomu@eedept.kobe-u.ac.jp

d) tuka@kobe-u.ac.jp

ている。Liu らは Nokia 研究所が提唱する直線や円など 8 種類の図形を描く動作を 1 個の 3 軸加速度センサで認識するシステムを提案している [3]。また、提案システムを用いて、ビデオ共有サービスにおけるビデオリストのモバイル端末向け閲覧インターフェースを実装している。Agrawal らは携帯電話に搭載された加速度センサの値から空間に描いた文字を認識するシステムを提案している [4]。空間に文字を書く際に、文字を書くスピード等に制約を設けることによって 83% の認識精度を達成している。矢谷らはモバイル端末における情報の移動を直観的に実現する Toss-It というインターフェースを提案している [5]。Personal Digital Assistant (PDA) に加速度センサを取り付けてトスの投げ上げの角度や速度、振りの回転方向などを検出し、カメラを用いてユーザ同士の位置を測定すること、特定のユーザへ情報を投げ渡すといった直観的な操作でデータ送信ができる。しかしながら、これらの認識システムはジェスチャ終了後に認識処理を開始しており、認識結果に応じたフィードバックの出力に遅延が発生し、インターフェースの操作性が低下すると考えられる。出力の遅延を軽減するためには、ジェスチャ終了前に認識処理を開始しなくてはならない。Ammar らはデータグローブを用いて空中に書いた文字を認識し、ハンズフリーの入力インターフェースを実現している。データグローブには加速度センサと角速度センサが搭載されており、空中に文字を書いている動作を見分けるためにサポートベクターマシンの、動作認識には隠れマルコフモデルをそれぞれ用いている。このシステムではセンサデータを処理機能の高い他のデバイスに転送して、解析を行なっているので、リアルタイムで認識すると認識結果の出力に遅延が発生してしまう。結果の出力の遅延を軽減するためには、データの転送を行わず、ひとつのデバイスで認識処理も行なう必要がある。

2.2 ジェスチャの早期認識

管家らは実際に存在するドラム (実ドラム) と実際には存在しないドラム (仮想ドラム) の叩打の識別をスティックに取り付けた加速度センサを用いて行い、仮想ドラム叩打時のみ音を出力する Airstic Drum を提案している [7]。適切なドラム演奏を行うためには叩打と同時に音を出力する必要があり、直前に叩打の識別を行わなければならない。そこで叩打直前の特徴的なデータ波形に着目し、閾値を設けることによって叩打の識別を行っている。しかし、多数の動作に対応するためには個別に閾値を設定する必要がある。また、このシステムでは叩打といった比較的単純なジェスチャを対象としているが、より複雑なジェスチャを対象とすると認識は困難であると考えられる。藤本らはダンスステップを認識して BGM に合わせた効果音を再生する手法を提案している [8]。BGM の拍の前後の 2 段階でダンスステップの認識を行い、各段階で出力の修正を行うこ

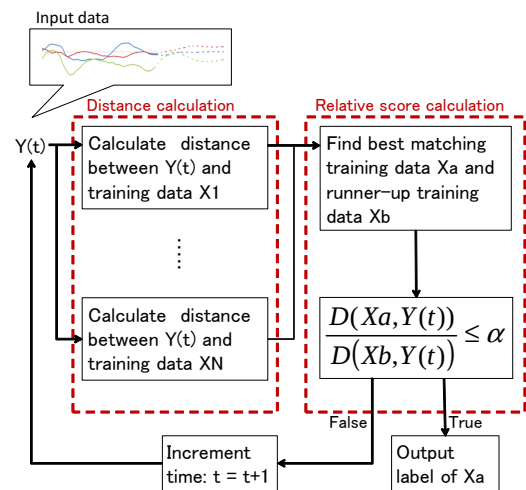


図 1 システム構成

とで、認識精度および効果音の出力のタイミングの改善を行っている。効果音を出力する場合は認識が間違っても即座に修正すれば聴衆への影響は小さいが、入力インターフェースにおいては認識結果に応じて出力したコマンドを取り消したり修正したりすることは困難である。

入力インターフェースにおけるジェスチャ認識では、認識結果を確定して出力するタイミングが重要となる。たとえば、認識対象となるジェスチャの種類が多いと認識結果が高い確度を得るまでに長い入力データが必要である。また、ジェスチャの途中段階まで類似したジェスチャが認識対象に含まれる場合は、類似した部分が終了するまで結果の出力を待たなければならない。村尾らは、認識するジェスチャの種類を変化させた時の認識精度について実験から評価を行っており、類似したジェスチャを認識対象から除外しながらジェスチャ数を減少させるに従い、認識精度が上昇することを確認している [9]。

このように、認識結果を出力するタイミングは認識対象とするジェスチャの種類および認識したいジェスチャの数に依存するため、高い精度で可能な限り早期に認識結果を出力するためにはそのタイミングを動的に変更する必要がある。本研究では、入力波形と教師データとの距離を逐次的に計算し、最近傍の教師データと次点の教師データとの距離の差が相対的に大きくなった時点で認識結果の確度が高いと判断し、最近傍の教師データに付与されているジェスチャを認識結果として出力する手法を提案する。

3. 早期認識手法

筆者らはこれまでに、ジェスチャの早期認識手法を提案している。本章では早期認識アルゴリズムについて説明する。

3.1 システム構成

一般的に加速度センサを用いたジェスチャ認識では、認

識したいジェスチャのデータとジェスチャのラベルのペア (教師データ) を事前に登録しておく必要がある。登録したすべての教師データと認識したい未知のデータ (入力データ) との距離を比較し、距離が最も小さい教師データのラベルを入力データの認識結果として出力する。一般的には、ジェスチャ終了までの入力データと教師データを用いて距離計算を行うが、本研究では図 1 に示すように、入力データが途中までであることを考慮した認識アルゴリズムを提案する。提案する認識アルゴリズムでは 2 段階の計算処理から成る。第一段階では、入力データとすべての教師データとの距離を計算し、距離の最も近いものと 2 番目に近いものを検索する。第二段階では、最も近い距離と 2 番目に近い距離を比較することで認識結果の候補が拮抗している場合には、認識結果を出力せずに、より長い入力データを用いて判断し、確度の高い結果を出力する。

3.2 距離計算アルゴリズム

提案手法では、距離計算のアルゴリズムとして Dynamic Time Warping (DTW)[10] を用いる。DTW は時間的非線形圧縮距離を計算することで、長さや部分的に速度が異なるデータの距離を計算でき、ジェスチャ認識の距離計算において主要なアルゴリズムである。しかしながら、DTW は入力データと教師データがともにジェスチャの開始から終了までの完全なデータであることを想定しており、途中までの入力データに対して DTW をそのまま適用すると、入力したデータが完全なジェスチャであるとして処理されるため正しく認識できない。そのため、本研究では、途中までの入力データにも対応した距離計算を行うように DTW のアルゴリズムを改良する。

具体的なアルゴリズムを以下に示す。ある時点における入力データの長さが t であるとする、その入力データに合う教師データの部分波形の長さも t に近いと想定し、距離計算の対象と教師データの長さに制約を設け、 $t - \varepsilon$ から $t + \varepsilon$ とする。この時の ε を検索幅と呼ぶ。教師データ $X = (x_1, \dots, x_m)$ と長さ t の入力データ $Y(t) = (y_1, \dots, y_t)$ の距離を計算する際、 $m \times t$ 行列の $d(x_i, y_j) = (x_i - y_j)^2$ を定義する。次に、 X と $Y(t)$ のインデックスのペアからなるワーピングパス $F = (f_1, \dots, f_k)$ を検索する。ここで、ワーピングパスは次の 3 つの条件を満たす。

- 境界条件
 $w_1 = (1, 1), w_k = (t + \varepsilon, t)$
 - 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
 $w_k = (a, b), w_{k-1} = (a', b') \Rightarrow a - a' \geq 0 \wedge b - b' \geq 0$
- 上記の条件を満たす最小のパスを検索するために、以下の手順を適用する。

Initialization:

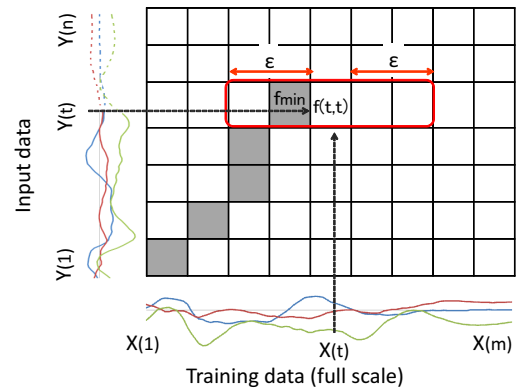


図 2 距離計算アルゴリズム

$$\begin{aligned} f(0, 0) &= 0 \\ f(i, 0) &= \infty \text{ for } i = 1, \dots, t + \varepsilon \\ f(0, j) &= \infty \text{ for } j = 1, \dots, t \end{aligned}$$

Do for $i = 1, 2, \dots, t + \varepsilon$

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

$$f(i, j) = d(x_i, y_j) + \min \begin{cases} f(i-1, j-1) \\ f(i-1, j) \\ f(i, j-1) \end{cases} \quad (1)$$

Output:

Return $D(X, Y(t)) =$

$$\min \begin{cases} f(t - \varepsilon, t) / (t - \varepsilon + t) \\ f(t - \varepsilon + 1, t) / (t - \varepsilon + 1 + t) \\ \vdots \\ f(t + \varepsilon - 1, t) / (t + \varepsilon - 1 + t) \\ f(t + \varepsilon, t) / (t + \varepsilon + t) \end{cases} \quad (2)$$

以上の手順によって、図 2 のような f の行列が作成される。これをタイムワーピング行列と呼ぶ。 $f(t - \varepsilon, t)$ から $f(t + \varepsilon, t)$ までの値は時刻 t における入力データと時刻 $t - \varepsilon$ から $t + \varepsilon$ までのそれぞれの教師データとの距離を表している。図では、行列内の影のつけられた部分が最小のパスである。したがって、 $D(X, Y(t)) = f(t - 1, t)$ が最小の距離であり、時刻 t における X と $Y(t)$ との距離となる。検索幅を採用して $t - \varepsilon \leq i \leq t + \varepsilon$ の中から最小値を検索することによって教師データとして登録したジェスチャのスピードと入力データのジェスチャのスピードが異なる場合でも適切な距離計算ができる。

この距離計算をすべての教師データについて行い、距離の最も近いジェスチャ g_1 と 2 番目に近いジェスチャ g_2 を導き出す。つまり、 $D(X_{g_1}, Y(t))$ と $D(X_{g_2}, Y(t))$ を導き出す。ただし、 X_{g_1} はジェスチャ $Y(t)$ と最も距離に近いジェスチャの教師データであり、 X_{g_2} はジェスチャ $Y(t)$ と 2 番目に距離に近いジェスチャの教師データである。上記の計算過程を入力データが追加されるたびに行う。ただ

し、タイムワーピング行列の計算は追加された入力データに相当する部分のみを計算するだけであるため、入力データが追加されるたびに行う計算処理の負荷は小さい。

3.3 相対スコア計算アルゴリズム

距離計算アルゴリズムによって得られた $D(X_{g1}, Y(t))$ と $D(X_{g2}, Y(t))$ を用いて相対スコアを計算する。相対スコアを用いることで入力データのジェスチャの冒頭部が複数のジェスチャと類似している場合、結果の出力を猶予して、より長い入力データを用いて確度の高い結果を出力する。具体的には、時刻 t における最小の距離 $D(X_{g1}, Y(t))$ と 2 番目に小さい距離 $D(X_{g2}, Y(t))$ の比が次式を満たすとき、認識結果を出力する。

$$\frac{D(X_{g1}, Y(t))}{D(X_{g2}, Y(t))} < \alpha \quad (3)$$

本研究では式 (3) の左辺を相対スコアと呼ぶ。 α ($0 < \alpha \leq 1$) は閾値であり、閾値が小さいほど結果を出力するために 2 番目の距離との差が大きくなければならず、より確度の高い結果が得られることを意味する。式 (3) を満たさなければ、認識結果の出力を行わず、入力データを $Y(t+1)$ に更新し、再度 3.2 節の距離計算を行う。また、極端に短い入力データではジェスチャの判別が困難であるため、入力データ数には下限を設定して、入力データが 10 サンプル (0.2 秒) 以下では認識処理を行わない。提案手法では認識結果の確度が高い時点で認識結果を出力するため、ジェスチャ終了まで待っても確度が高くない可能性がある。その場合は、ジェスチャ終了と同時に最も確度が高い認識結果を出力する。

ここで、具体的な数値を用いた距離計算の例を図 3 に示す。実際は加速度センサのデータは 3 軸であるが、ここでは簡単化のために 1 軸で説明する。ジェスチャ開始から終了までの教師データを $X = \{5, 12, 6, 10, 6, 5, 18, 20, 10, 7\}$ 、与えられる入力データを $Y(t) = \{11, 6, 9, 4, 2, \dots\}$ とする。 $t = 1$ から順に 1 サンプルずつ入力されるたびに入力データに対して図 3 に示すようにタイムワーピング行列 $f(i, j)$ を作成していく。ただし、ここでは入力データ数の下限は設定せずに説明する。まず、 $Y(1) = 11$ が入力されると、タイムワーピング行列の 1 行目が計算される。検索幅を $\varepsilon = 1$ 設定すると、この時刻における距離となる候補は $f(1, 1)$ と $f(2, 1)$ であり、最小値である $f(1, 1) = 6$ が時刻 $t = 1$ における X と $Y(t)$ の距離となる。さらに、 $Y(2) = \{11, 6\}$ が入力されると、2 行目のタイムワーピング行列が作成され、この時の距離の候補は $f(1, 2)$ と $f(2, 2)$ と $f(3, 2)$ であり、その中で最小値である $f(3, 2) = 7$ が時刻 $t = 2$ における距離となる。 $f(1, 2) = f(3, 2)$ のように、候補の中で同じ値をもつ要素が存在する場合、教師データのインデックスが大きい方を距離として選択する。このように、入力データが更新されるたびにタイムワーピング行列を計算し、相対ス



図 3 距離計算の例

コアが閾値 α を下回ったときに認識結果を出力する。相対スコアが閾値 α を下回らなければ認識結果を出力せずに、入力データを更新して再度距離計算を行う。

3.4 認識精度

提案手法を用いた場合の認識精度への影響を示す。タブレット端末使用時に行われると想定する 27 種類のジェスチャのデータを 3 軸加速度センサを 1 個配置した盤面を用いて 1 人の被験者から採取し、認識精度と認識結果の出力からジェスチャ終了までの残り時間の関係を比較手法と提案手法で調査した。比較手法は、ジェスチャの開始から一定時間までのデータ波形を用いて距離計算を行う。具体的には、長さ t のデータ波形が入力されたときに認識結果を出力するとし、長さ t の教師データと距離計算を行い、距離が最も小さい教師データのラベルを認識結果とする。

各ジェスチャ 8 回ずつのデータを採取し、1 回分を教師データ、残りの 7 回分をテストデータとした。使用した加速度センサはワイヤレステクノロジー社製 WAA-006[11] で、サンプリング周波数は 50Hz でデータを採取した。27 種類のジェスチャの動作時間の平均は 1.40 秒であり、最短は 0.66 秒、最長は 2.34 秒である。被験者の解釈の違いによる影響を軽減するために、ジェスチャを被験者に教示する際は言葉や図ではなく、実際の動作を示して行った。センサデータは東芝社製 PC (dynabook, RX3 TM240E/3HD) に Bluetooth で送信され、PC 上で計算処理を行った。

比較手法と提案手法それぞれにおいて、入力データの長さを変えた時の、ジェスチャの残り時間と認識精度を図 4 に示す。ジェスチャの残り時間とは入力されるジェスチャの全長から入力されたジェスチャの長さを引いたものである。議論を簡単にするため、入力されてから結果が出力されるまでの処理時間は考慮していない。計算処理時間については次節で述べる。比較手法では、入力データの長さを 1 サンプルずつ変化させて認識精度の計算を行った。また、教師データが入力データより短いとき、教師データの終端にセンサ値 0 を追加して長さをそろえた。提案手法では、

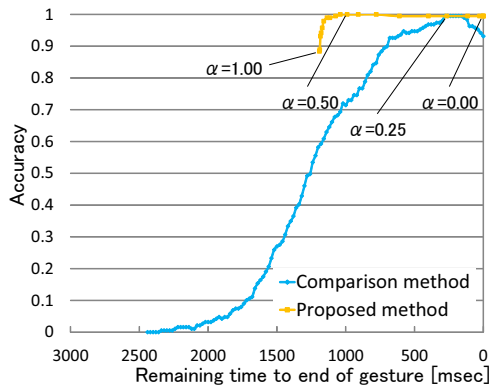


図 4 ジェスチャ終了までの残り時間に対する認識精度の変化

閾値 α の値を 0 から 1 まで 0.05 ずつ変化させて認識精度の計算を行った。図中の縦軸はジェスチャの平均認識精度を表しており、横軸はジェスチャの残り時間を表している。また、検索幅 ε の大きさは 10 である。

図 4 より、比較手法において、ジェスチャ終了 800 ミリ秒前に平均認識精度が 0.8 を越えており、ほとんどすべてのジェスチャがジェスチャ終了 800 ミリ秒前に認識できている。

4. 演劇での活用

早期認識の応用例として、本研究では観客参加型演劇を対象とする。演者は剣や銃などの小道具を使用し、動きに合わせて効果音を出すことで迫力のある演技をすることができる。例えば演者が剣を振り下ろすと同時に風切り音のような効果音を出すことで、力強さを表現することができる。このような演出を行なうためには、動きに合わせて効果音を出力する必要があるが、そのためにはあらかじめ決められたタイミングで出力される効果音に合わせてその動作を行うなど、演者にアドリブを許さない構成となってしまう。そこで、小道具にセンサを搭載し演者の動きを認識することによって、事前の打合せや練習なく演者の動きに合わせて効果音を出力することが可能になる。このとき、演者の動きに合わせてリアルタイムで効果音を出力するためには、早期認識が有効に働くと考えられる。

4.1 観客参加型演劇 YOUPLAY

一般客が観衆ではなく演者として演劇を楽しむライブエンタテインメントがある。これを観客参加型演劇という。この観客参加型演劇では演技を参加者自身に任せるため、演者の動作に応じてオペレータが効果音を手動で出すことは困難である。そこで、観客参加型演劇「YOUPLAY Vol. 1 スペースレンジャーの不思議な惑星」[12]において、小道具の虫取り網の竿に加速度センサを搭載して「上から振り下ろす」と「右から左へ振る」の 2 種類の振りの動作を認識し、振りに応じて風を切るような効果音を出力するシ

ステムを実装した。YOUPLAY とは図 5 に示す最大 10 人のキャラクタを初対面同士あるいは友人同士で演じることによって、演劇を楽しむ約 30 分のライブエンタテインメントである。YOUPLAY の実施の様子を図 6 に示す。演者が装着するヘルメットの頭頂部に複数の赤外線 LED、内部に加速度センサやヘッドセットなどが組み込まれている。会場の天井に設置された赤外線カメラにより、演者の位置をトラッキングし、加速度センサを用いて動きを認識する。例えば、演者の周りに現在の酸素量などのパラメータを追従させたり、特定の歩き方をしたときにヘッドホンから音を出す。各キャラクタはセンサが内蔵されている銃や虫取り網などのアイテムを所持しており、それらを駆使して物語を進めていく演出となっている。網を所持している「タカミネさん」というキャラクタは宇宙の生物を捕まえることが得意である設定となっており、物語中で網を使って生物を捕まえる場面がある。作・演出はノンバーバルパフォーマンス集団「オリジナルテンポ」のプロデューサーでもあるウォーリー木下氏、映像関連ハードウェアとしては大阪万博公演プロジェクトマッピングを行なっている株式会社タケナカが担当している。Vol. 1 は 2013 年 11 月 16 日から 24 日まで開催し、計 40 公演を行った。

4.2 システム構成

YOUPLAY Vol.1 では加速度センサを搭載した虫取り網を作成し、提案する早期認識技術を用いて網を振る動作を認識している。図 7 に示すように、竿の中央部付近に加速度センサ、Arduino Fio、XBee、リチウムイオン電池を搭載している。演劇では多くの無線機器を用いるため、加速度センサのデータを PC に送信して、PC 側で認識処理を行うと、通信量が多くなり他の無線通信の妨げとなる恐れがある。そこで図 8 に示すように、作成した虫取り網の Arduino Fio 上で認識処理を行ない、認識結果のみを PC に XBee を用いて ZigBee で送信し、舞台装置のスピーカから効果音を発生させる構成としている。また、Arduino Fio のみでは、SRAM のメモリ容量が不足しているため Microchip Technology 社のシリアル SRAM チップを用いてメモリ容量の拡張を行なった。動作認識を行うためには、あらかじめユーザ虫取り網を振る動作の加速度データを教師データとして登録しておく必要があるが、今回は筆者らのうち 1 名の「上から振り下ろす」と「右から左へ振る」の 2 種類のデータを教師データとした。

4.3 早期認識の使用場面と考察

ストーリー中で動作認識を使う場面を図 9 に示す。虫取り網を使う場面は序盤のキャラクタの紹介の場面(図 9(a))と床にプロジェクタで投影されたねずみの映像を追いかけて網で捕まえる場面(図 9(b))の 2 箇所のみであり、不要な場面での効果音の出力を防ぐために、その他の場面では



図 5 YOUPLAY VOL.1 のキャラクタ



図 6 YOUPLAY VOL.1 の様子

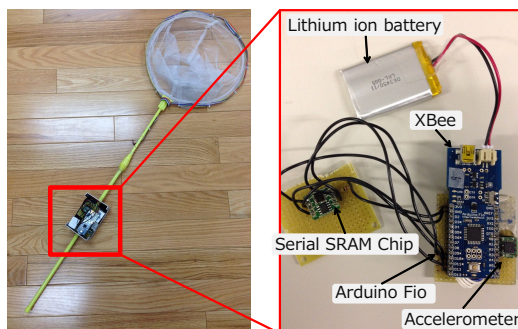


図 7 センサの位置

音声出力されないようにした。キャラクタ紹介の場面では虫取り網を使って虫を捕まえる演技をさせるなど網に注目が集まる演出をした。

キャラクタ紹介の場面では、網を振り効果音が発生すると演者らが驚きを示したことから演者に振りと言が連動していることを認識させることができたと考えられる。しかし、他のキャラクタの紹介の時に網を持ったまま拍手をすると誤って効果音が流れてしまい、紹介の邪魔になった。このように予期せぬ場面で効果音が流れてしまうことがあったため、それ以降の公演では網を持ったキャラクタの

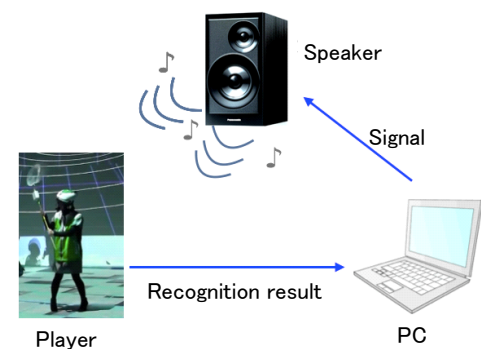


図 8 システム構成

紹介が終了した時点で効果音の出力をオフにした。

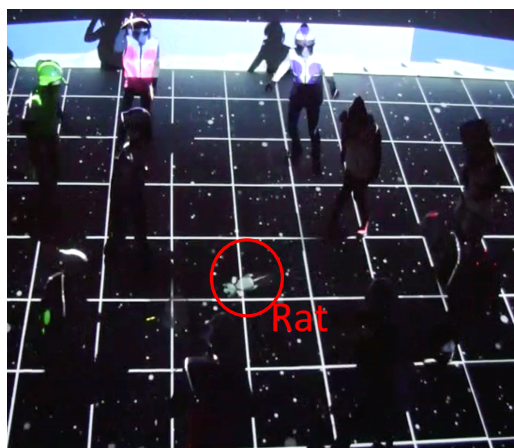
床に投影されたねずみを追いかける場面では、床に向かって網を上から振り下ろす動作が多く見受けられた。この時、振り下ろす途中で風を切る効果音出力されており、動作と効果音のずれによる違和感が解消されていたことから、提案するアルゴリズムは有効であると考えられる。また網を持った手を大きく振りながら追いかける演者の場合、追いかけている最中に効果音の誤出力が目立った。この原因として、今回は2種類の振りを認識したが、網の動きを検出した場合、必ずどちらかの認識結果を出力するようにしていたため、誤出力してしまったと考えられる。従って、入力データと教師データとの距離が大きい場合は効果音を出力しないことによって、正しい振り以外の動作に対する誤出力を削減できると考えられる。

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ける早期認識手法を紹介し、その応用として観客参加型演劇における舞台アイテムへの応用例を示した。提案手法では、早期かつ適切なタイミングで認識するために、入力データと教師データの距離を逐次的に計算し、最も近い距離と2番目に近い距離を相対的に判断することで確度の高い認



(a) キャラクタ紹介場面



(b) ねずみを捕まえる場面

図 9 動作認識の使用場面

識結果を得ることができる。参加型演劇での活用では、小道具の虫取り網に加速度センサを搭載して、早期認識技術を用いて網の振りを振り終える前に認識することで、動作と効果音とのずれが軽減されており、早期認識の有効性を確かめることができた。しかしながら、予期せぬ場面で効果音が発生する場面が多く見受けられた。今後は、効果音を出力しない教師データを登録しておくなど、効果音の誤出力を削減する手法を検討する予定である。また、動作の途中で効果音を出力するが、演技に与える影響を調査する予定である。

現段階では、確度の観点のみからジェスチャの認識結果のタイミングを決定しているため、高い確度が最後まで得られないと従来手法と同様にジェスチャ終了後に認識処理を開始し、認識結果の出力が動作終了後になる場合がある。したがって今後は、拮抗している候補からジェスチャの残り時間を予測し、結果の出力が遅れるのであれば、高い確度が得られなかった場合でも、その時点において最も距離の近いジェスチャを出力する手法を提案する予定である。また、YOUPLAYの参加型演劇のような動作認識の技術を用いたインタフェースの利用を今後増やしていくためには、動作認識の専門的な知識がなくても容易に利用するこ

とができる環境を整えていく必要がある。従って、今後の課題として様々な場で早期認識を利用してもらうために早期認識アルゴリズムのライブラリ化を行ない、動作認識技術の普及を目指す。

参考文献

- [1] 出田 怜, 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操作の早期認識手法の提案, 情報処理学会マルチメディア・分散・協調とモバイル DICOMO2013 シンポジウム, Vol. 2013, No. 1, pp.655–662 (July 2013).
- [2] Tomibayashi, Y., Takegawa, Y., Terada, T., and Tsukamoto, M.: Wearable DJ System: a New Motion-Controlled DJ System, *Proc. The International Conference on Advance in Computer Entertainment Tecnology (ACE 2009)*, pp. 132–139 (Oct. 2009).
- [3] Liu, J., Wang, Z., Zhong, L., Wickramasuriya, J., and Vasudevan, V.: uWave: Accelerometer-based Personalized Gesture Recognition and its Applications, *Proc. IEEE International Conference on Pervasive Computing and Communication (PerCom 2009)*, pp. 1–9 (Mar. 2009).
- [4] Agrawal, S., Constandache, I., Gaonkar, S., Choudhury, R., Caves, K., and Deruyter, F.: Using Mobile Phones to Write in Air, *Proc. The International Conference on Mobile Systems, Applications, and Services (Mobisys 2011)*, pp. 15–28 (Jun. 2011).
- [5] Yatani, K., Tamura, K., Hiroki, K., Sugimoto, M., and Hashizume, H.: Toss-it: Intuitive Information Transfer Techniques for Mobile Devices, *Proc. The International Conference on Human Factors in Computing Systems (CHI 2005)*, pp. 1881–1884 (Apr. 2005).
- [6] Amma, C., Georgi, M., and Schultz, T.: Hands-Free Mobile Text Input by Spotting and Continuous Recognition of 3d-Space, *Proc. The International Symposium on Wearable Computers (ISWC2012)*, pp. 52–59 (Nov. 2012).
- [7] Kanke, H., Takegawa, Y., Terada, T., and Tsukamoto, M.: Airstic Drum: a Drumstick for Integration of Real and Virtual Drums, *Proc. The International Conference on Advance in Computer Entertainment Technology (ACE2012)*, pp. 57–69 (Nov. 2012).
- [8] Fujimoto, M., Fujita, N., Takegawa, Y., Terada, T., and Tsukamoto, M.: A Motion Recognition Method for a Wearable Dancing Musical Instrument, *Proc. The International Symposium on Wearable Computers (ISWC 2009)*, pp. 11–18 (Oct. 2009).
- [9] Murao, K., Terada, T., Yano, A., and Matsukura, R.: Evaluation Study on Sensor Placement and Gesture Selection for Mobile Devices, *Proc. The International Conference on Mobile and Ubiquitous Multimedia (MUM 2012)*, No. 7 (Dec. 2012).
- [10] Myers, C. and Rabiner, L.: A Comparative Study of Several Dynamic Time Warping Algorithms for Connected Word Recognition, *The Bell System Technical Journal*, Vol. 60, pp. 1389–1409 (Sep. 1981).
- [11] Wireless Technologies, Inc., available from (<http://www.wireless-t.jp/>).
- [12] YOUPLAY Vol. 1 スペースレンジャーの不思議な惑星, available from (<http://youplay.jp>).