

# IMUを用いた人体運動のモーション認識

Recognition of human body movement using Inertial Measurement Unit

博士前期課程 2年 青木崇 指導教員 堤正臣 教授

The objective of this study is the realization of recognition of human body movement using IMU (Inertial Measurement Unit). This paper shows segmentation and the recognition results of the movement of the right arm during Japanese drum performance to confirm the segmentation and recognition possibilities. In past studies the real-time recognition of human body movement using joint angles acquired by motion capture has been realized. The segmentation algorithm is now implemented for angular velocities. And to adjust the parameter settings in order to match the segmentation results in manual segmentation results. The similar movement patterns are classified into each model. Recognition is performed using these models. In addition recognition is performed using Wii remote instead of IMU as the measurement device.

## 1. 結言

これまで多くのロボットは産業分野などの特殊環境で用いられてきたが、近年は掃除用ロボットや警備用ロボットなどがオフィスや家庭における日常生活に広く使われ始めている。また、介護ロボットや医療ロボットの開発も行われており、人間とロボットの距離が近くなってきている。今後、ロボットが日常生活で人間と共存していくためには、人間とロボットで相互コミュニケーションを行うことが必要となる。そこで、ロボットが人間の動作を理解し、それに対応する行動を選択し実行することができれば、人間とロボットの相互コミュニケーションが可能となる。高野ら<sup>(1)(2)</sup>によって、人体各部位の位置や各関節角度情報から運動パターンのモデル化を行い、新たに入力された観測データがどの運動パターンに分類されるかを判定可能であることが報告されている。ここで、人体運動の観測データとして用いる位置や各関節角度情報の取得は主に光学式モーションキャプチャが用いられる。この装置は高額かつ大規模な設備が必要であり、観測範囲が設置されたカメラの撮影範囲内に限られることから、日常生活で用いるには適さない。また、関節角度を算出するには逆運動学の計算が必要となることから、計算量が増える。そこで本研究では、人体各部位にIMU(Inertial Measurement Unit)を取り付け、取得した加速度・角速度データを用いて運動パターンのモデルを作成し、モーション認識を行う。関節角度を用いた方法と比較して安価かつ小規模な設備によって、人体運動のモーション認識を実現することを目的とする。

## 2. 運動の分節化

人間は他者の行動を観察したとき、一連の動作を運動パターン毎に区切り動作認識を行っている。このため、運動パターンの分節化を行うことは人体運動のモーション認識を実現するにあたり、基本となる重要な情報処理となる。本研究では、人間が認識を行うプロセスを再現するため、音声認識の分野で多く用いられている隠れマルコフモデル(HMM:Hidden Markov Model)<sup>(3)(4)</sup>を用いた。

2.1 HMMによる記号列化 観測した時系列データ  $O$  を微小時間幅  $w_{span}$  で微小観測データ時系列  $o$  に区切り  $O = \{o(1), o(2), \dots, o(i), \dots, o(n)\}$  とし、

$i$  番目の微小観測時系列データ  $o(i)$  は式1のように定義する。

$$o(i) = [\hat{o}((i-1)w_{span} + 1) \dots \hat{o}(iw_{span})] \quad (1)$$

ただし、 $\hat{o}(t)$  は時刻  $t$  における観測データの列ベクトルとする。微小観測時系列データ  $o(t)$  に対してHMMを  $N$  個作成する。HMM( $\lambda$ ) については以下に示す。本研究では状態が左から右へと遷移、現在の状態から次の状態へ遷移すると前の状態へは戻らない Left-to-Right 型を用いてモデル化を行った。

$$\lambda = \{Q, A, B, \Pi\} \quad (2)$$

- $Q = \{q_1, \dots, q_n\}$ : 状態の集合
- $A = \{a_{ij}\}$ :  $q_i$  から  $q_j$  への遷移確率  $a_{ij}$  の集合である状態遷移確率行列
- $B$ : 状態  $q_i$  で記号列を出力する確率分布の集合
- $\Pi = \{\pi_i\}$ : 初期状態確率の集合

HMM( $\lambda_k$ ) の初期パラメータはランダムに設定するため、観測データ  $o(i)$  に対し最大尤度を出力するようにHMM( $\lambda_k$ ) のパラメータを変更する必要がある。また、最大尤度を出力するHMM( $\lambda_{max}$ ) を求め、パラメータを最適化する。HMMパラメータの最適化には Baum-Welch アルゴリズムを用いる。なお、Baum-Welch アルゴリズムの式は付録に記すものとする。観測データをHMMにて抽象化するイメージ図を図1に示す。

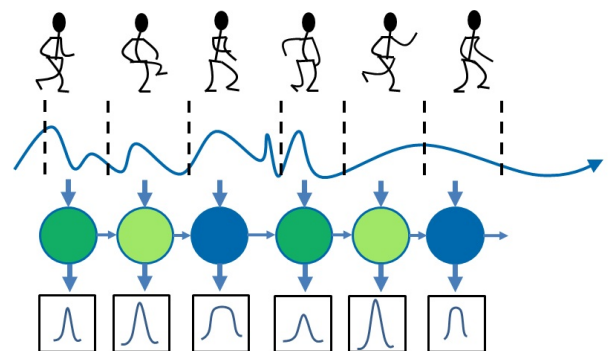


Fig.1 Hidden Markov Model

2.2 相関学習 観測データ  $o(i)$  が  $N$  個のHMMから生成される尤度をそれぞれ計算する。大きな尤度を算出した順にHMMを  $m$  個選択し、 $o(i)$  に対して大きな尤度を示したHMMの集合を  $S(i) = \{\lambda^1, \lambda^2, \dots, \lambda^m\}$  とする。 $k$  番目のHMM( $\lambda_k$ ) が

集合  $S(i)$  に含まれる場合、その要素を 1 とし、含まれない場合、その要素を 0 とする様な  $N$  次元の列ベクトル  $\hat{x}(i)$  を式 3 のように定義する。

$$\hat{x}(i) = [\hat{x}_{i1}, \hat{x}_{i2}, \dots, \hat{x}_{iN}]^T \quad (3)$$

$$\hat{x}_{ik} = \begin{cases} 1 & \lambda_k \in S(i) \\ 0 & \lambda_k \notin S(i) \end{cases} \quad (4)$$

$\hat{x}(i)$  は  $i$  での運動データ情報しか持たないため、過去の履歴情報も含めたベクトル  $x(i)$  を式 6 のように定義する。

$$x^*(i) = [\hat{x}(t-M+1)^T \dots \hat{x}(t)^T]^T \quad (5)$$

$$x(i) = \frac{x^*(i)}{\|x^*(i)\|} \quad (6)$$

$M$  は相関学習に用いる履歴情報の時間幅とする。これにより過去の履歴情報を用いることにより直前の特徴量との相関が強くなりすぎ運動パターンを細かく分節化しすぎることを防ぐことができる。このように作成された HMM と同数の要素を持つ列ベクトル  $x(i)$  を特徴ベクトルとする。 $x(i-1)$  を入力パターン、 $x(i)$  を出力パターンと見なすことで相関行列  $W_0 = x(i-1) * x(i)^T$  と表すことができる。入力パターンから予測される出力パターンは  $W_0 x(i-1)$  と表すことができ、この相関行列を用いることで次の記号列を想起できる。ここで、逐次学習を行うため相関行列を式 (7) のように定義する。

$$W(t) = \alpha W(i-1) + (1-\alpha)x(i)x(i-1)^T \quad (7)$$

$\alpha$  は安定化係数を表す。 $x(i)x(i-1)^T$  は直前の特徴ベクトル  $x(i-1)$  から現時刻の特徴ベクトル  $x(i)$  への写像を表している。過去の特徴ベクトルのダイナミクス  $W(i-1)$  を逐次足していくことにより、現時刻の特徴ベクトルのダイナミクス  $W(i)$  が求められる。

**2.3 境界判定** ある運動パターンから異なる運動パターンに遷移する時点では、過去の情報から次の動きを予測することは困難である。そこで現在の特徴ベクトル  $x(i)$  と、相関学習から得られる履歴情報から予測された特徴ベクトル  $W(i-1)x(i-1)$  との誤差  $E$  を算出する。

$$E(i) = \|x(i) - W(i-1)x(i-1)\| \quad (8)$$

誤差  $E$  が大きな値を示したとき、過去の運動と現在の運動との間に類似性が低下しており、運動パターンが変更または新たな運動に変化したものとする。逆に、誤差が小さな値を示した場合は、現在の運動は過去の運動と同じ運動パターンであり、運動パターンの境界は閾値を設けることで判定する。この判定方法を用いることで、観測データがゼロや停止状態になる時点も運動パターンの境界とするのではなく、連続した運動でも境界を判定することが可能となる。

**2.4 HMM の学習と競合学習** 稲邑ら<sup>(5)</sup>により見まね学習を通じて原始シンボルを獲得し、高

次の知能処理を実現するミメシスモデルが提案されている。運動パターンのダイナミクスの抽象化と原始シンボル化、原始シンボルを用いた他者の運動認識と原始シンボルからの自然な動作パターンの生成を行う学習、認識、生成を一つのシステムで実現でき自律的にデータベースを作成できることが報告されている<sup>(6)</sup>。運動パターンを原始シンボルとして抽象化するため、原始シンボルの学習と HMM と強化学習を行う。分節化のときと同様に、 $N_U$  個の  $HMM(\Lambda_k : k = 1, 2, 3, \dots, N_U)$  を用意し、初期パラメータはランダムに設定する。 $j$  番目に分節化された運動パターンデータ  $o_{segment}(j)$  に対して、最大尤度を出力する  $HMM(\Lambda_{max})$  を求める。

$$\Lambda_{max} = \arg \max_{\Lambda_k} P(o_{segment}(i) | \Lambda_k) \quad (9)$$

$HMM(\Lambda_{max})$  を最適化するため、訓練データとして  $o_{segment}(j)$  を保存し用いる。最適化のために保存したパラメータ数が閾値に達すれば、保存した複数のパターンデータを削除する。パターンデータが閾値未満の場合は  $HMM(\Lambda_{max})$  のパラメータ最適化は実行しない。上記の競合学習処理回数が閾値回数に達したときを十分に学習が行われたとみなし、 $HMM(\Lambda_k)$  の学習を終了する。閾値に達しなかった場合は次のパターンデータ  $o_{segment}(j+1)$  を訓練データとして処理を行う。分節化された運動パターンを  $HMM(\Lambda_k)$  に分類し、 $o_{segment}(j)$  に対し最大尤度を出力する  $HMM(\Lambda_{max})$  のパラメータを上記の方法で最適化することによって  $HMM(\Lambda_{max})$  は類似した運動パターンを表現した原始シンボル HMM となる。認識は、 $o_{segment}$  を出力する尤度  $P(o_{segment} | \Lambda_k)$  が最大値を出力する原子シンボル  $HMM(\Lambda_k)$  を  $o_{segment}$  と同じ運動パターンであると判定する。

$$\Lambda_{max} = \arg \max_{\Lambda_k} P(o_{segment} | \Lambda_k) \quad (10)$$

なお、尤度の計算はフォワードアルゴリズムを用いて行う。フォワードアルゴリズムの詳細については付録に記す。

### 3. 実験概要

本研究の実験は、初めに太鼓を叩く動作を IMU によって計測し、得られたデータによって動作認識を行った。IMU による認識が確認された後、IMU より身近で日常生活に用いることができると考え、IMU と同様に加速度・角速度センサを内蔵した Wii リモコンを用いて認識を行った。

**3.1 IMU と Wii リモコン** 本実験で用いた IMU と Wii リモコンの軸設定を Fig.2 に示す。IMU にはジースポート社製の Pocket-IMU を用いた。

**3.2 実験方法 1** 実験では和太鼓を叩く運動を被験者に行わせ、実際に叩く動作を行った右腕の動きを用いて分節化・認識を行った。Fig.2 に示す様に被験者の右手首に IMU を取り付け、右手で持った撥を使って太鼓を叩く。被験者には初めにモデル作成

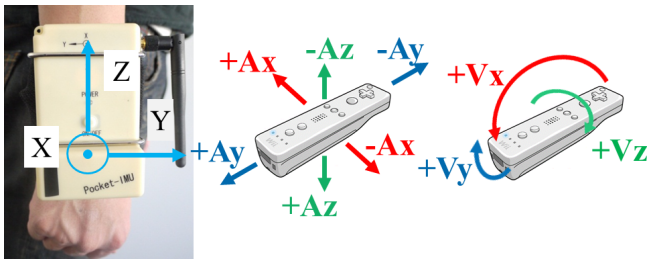


Fig.2 IMU and Wii Remote

用の実験として、同じ叩き方を連続して10回繰り返し行い、その間の角速度を取得した。叩き方はそれぞれ「上から強く叩く (Strongly)」、「上から弱く叩く (Weakly)」、「横から叩く (Sideways)」の三種類を10回ずつ繰り返して行う。

続けて認識用の実験を行った。この実験では叩く角度は全て上からとし、叩く強さを三段階に分け、それぞれの強さで4回ずつ叩いた際の腕の角速度を取得した。

3.3 実験方法2 実験では被験者が右手でWiiリモコンを持ち異なる振り方で計測した結果の分節化・認識を行った。被験者には同じ振り方を連続して10回繰り返し行い、その間の角速度を取得した。振り方はそれぞれ「大きく振る (Widely)」、「小さく振る (Small)」、「横に振る (Sideways)」の三種類を10回ずつ繰り返して行った。

#### 4. 自動分節化とモデルの作成

4.1 自動分節化の結果 「上から強く叩く」という動作で得られた角速度  $V_x$ ,  $V_y$  を用いて分節化を行った結果を Fig.3 に示す。グラフには手動で分割した区間として、人間が動画と波形を見て一つの運動パターンだと判断した位置を示す。具体的には、腕が停止している状態から撥を振り上げ、太鼓を叩いて腕が停止するまでを一つの運動パターンとした。Fig.3より自動で行った分節区間が手動の分節区間と一致したことがわかり、HMMを用いて相関学習を行う方法によって運動パターン毎に分節化を行えることが確認できた。角速度では  $X$ ,  $Y$  が共にゼロ付近となる。これは叩いた後、次に振上げる運動を行う直前が運動の切れ目となるからである。しかし、自動での分節化では10回繰り返した波形の1回目の分節化はうまくいっていない。これは1回目の動作だけ他と異なる姿勢から叩き始めたため以降の波形と異なったことが影響したものと考えられる。

4.2 モデルの作成 IMU, Wii リモコンそれぞれのモデルを作成した。各モデルについて Table1 に示すし、このモデルを用いて認識を行う。

Table 1. Model

Name	Type	Name	Type
IMU Model 1	Strongly	Wii Model 1	Widely
IMU Model 2	Weakly	Wii Model 2	Small
IMU Model 3	Sidely	Wii Model 3	Sidely

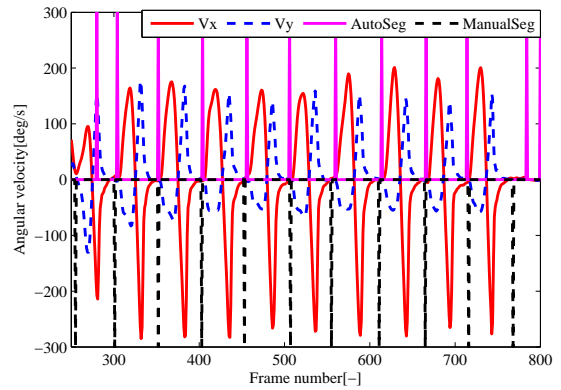


Fig.3 Manual and Auto sementation results

#### 5. 運動パターンの認識

5.1 認識結果1 作成したモデルを用いて、新たに入力された観測時系列データの認識を行う。叩き方を3段階変化させ各4回ずつ叩いた計測データを自動分節化した結果を Fig.4 に示す。図4より、運動パターンごとに区切られ運動パターン から はIMU Model 1に対応する「上から強く叩く」、運動パターン から の運動はどの運動パターンにも対応していない「上から中程度で叩く」、運動パターン , はIMU Model 2に対応する「上から弱く叩く」となっている。運動パターン では2回太鼓を叩いているが分割化ができていない。これは値が他の動作と比較して小さかったため分割化できなかったものと考えられる。これは分節化に加速度を加えることで解消されると考えている。分節化された各運動パターンを各

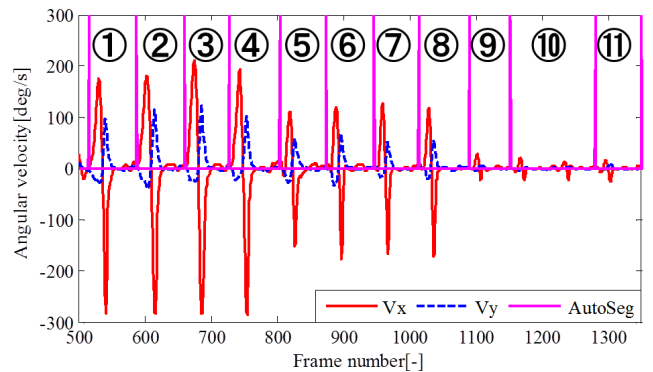


Fig.4 IMU Recognition results

モデルに対しフォワードアルゴリズムを用いて対数尤度  $\log P(o | \lambda)$  を算出し、絶対値が小さな数値を示しているモデルをその運動パターンと近いモデルであると認識する。結果は運動パターン から をIMU Model 1と認識し、 から , をIMU Model 2と認識した。認識した動作にIMU Model 3と認識した動作はなく、各モデルに対応する運動パターンは認識できている。しかし、中程度の強さで叩く動作の認識はIMU Model1, IMU Model2両方で認識されていることから波形が似ており、大きさのみが異なる場合には認識がうまくいかない場合があることが分かった。

5.2 認識結果2 Wiiによって作成したモデルを用いて、「大きく振る動作」を行った結果の認識を行った。自動分節化を行った結果を Fig.5 に示す。図5より、運動パターンごとに区切られ運動パターンは9つに分かれた。最後の1回は続く動作がなかったため分節化されなかったため認識が行われなかった。

Table 2. Wii Recognition results

Wii Model 1	-129	-27.4	-92.0	-118	-110
Wii Model 2	-150	-34.8	-119	-147	-146
Wii Model 3	-143	-32.1	-113	-139	-77.6
Wii Model 1	-169	-89.0	-153	-137	
Wii Model 2	-205	-118	-177	-151	
Wii Model 3	-194	-93.5	-168	-144	

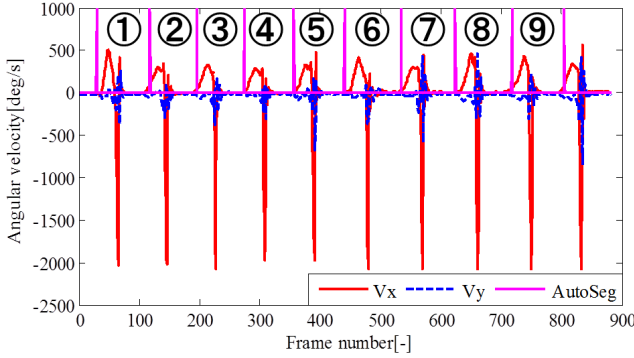


Fig.5 Wii Recognition results

結果は運動パターンを除く全てが Wii Model 1 と認識された。運動パターンは Wii Model 3 と認識された。これは振り上げたときの角速度よりも振った後角速度の方が大きい点で他の動作とは異なっていたため認識ができなかったものと考えている。

## 6. 結論

本稿では以下のことを行った。

- 角速度データによる運動パターン毎の自動分節化
- IMU による計測データによる認識
- Wii リモコンによる計測データによる認識

分割化では角速度の絶対値が小さい場合、分割化されないことがあることがわかった。また、ある程度の認識は行うことができたが、動作の大小変化のみの場合認識がうまく行えないことがわかった。これは分割化・認識共に加速度データを用いることで解消されるのではないかと考えている。しかし、加速度データを付加して分節化を行うとパラメータ設定が難しくなることから、加速度と角速度の分節化を別々に行った結果をもとに行うことを試みようと考えている。

## 謝辞

プログラムを提供して下さった東京大学中村研究室、カナダ Waterloo Univ. の Dana Kulic 博士、フランス Lead Lab の J.J. ホクテュリエ博士、和太鼓奏者の影山伊作様へここに感謝の意を表します。

- (1) W. Takano and Y. Nakamura, "Humanoid robot's autonomous acquisition of proto-symbols through motion segmentation," In IEEE Int. Conf. on Humanoid Robots, pp.425-431, 2006.
- (2) クリッチ・ダナ, 高野 渉, 中村 仁彦, "逐次記憶化および分類による全身運動知識の自己組織化", JSME Conf. on Robotics and Mechatronics, 1P1-G21, 2008.
- (3) 吉井, "音声情報処理", 森北出版株式会社, 1998.
- (4) C.M. ビショップほか, "パターン認識と機械学習上・下", シュプリンガー・ジャパン, 2007.
- (5) 稲邑 哲也, 中村 仁彦, 戸嶋 巖樹, 江崎 英明, "ミメシス理論に基づく見まね学習とシンボル創発の統合モデル", 日本ロボット学会誌, Vol.22, No.2, pp.256-263, 2004.
- (6) 高野 渉, 中村 仁彦, "統計的相関に基づく動作パターンのリアルタイム教師なし分節化と原始シンボルの自律的獲得", 日本ロボット学会誌, Vol.27, No.9, pp.1046-1057, 2009.

## A 付録 尤度計算アルゴリズム

尤度計算に用いたフォワードアルゴリズムを以下に示す。

$$\alpha_i^*(t) = \begin{cases} \pi_i & t = 0 \\ \sum_{j=1}^N \alpha_j(t-1) a_{ij} b_j(o(t)) & t > 0 \end{cases} \quad (A.1)$$

$$\alpha_i(t) = C(t) \alpha_i^*(t) \quad (A.2)$$

$$C(t) = \left[ \sum_{i=1}^N \alpha_i^*(t) \right] \quad (A.3)$$

$$\log P(O | \lambda) = \log \left[ - \sum_i C(T) \right] \quad (A.4)$$

ここで、 $b_i(o(t))$  は状態  $i$  で  $o(t)$  を出力する確率、 $C(t)$  はアンダーフローを回避するためのスケール係数である。このフォワードアルゴリズムを用いて尤度の算出を行う。

## B 付録 HMM パラメータ学習アルゴリズム

今回用いた出力記号列  $O$  がに対する HMM パラメータ  $A = a_{ij}$ ,  $B = b_{ij}$ ,  $\Pi$  を求める Baum-Welch アルゴリズムについて以下に示す。

$$\gamma_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \alpha_T(i)} \quad (B.1)$$

$$\gamma_t(i) = \sum_{j=1}^N \gamma_t(i, j) \quad (B.2)$$

を定義し、次式の EM アルゴリズムによって新しいパラメータの推定を行う。

$$\hat{\pi}_i = \gamma_1(i) \quad (B.3)$$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \gamma_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \quad (B.4)$$

$$\hat{b}_i(k) = \frac{\sum_{t: o_t=k} \gamma_t(i)}{\sum_{t=1}^T \gamma_t(i)} \quad (B.5)$$

この後、パラメータの更新を行い、この式 (B.3) ~ (B.5) が収束するまで推定を続け、得られたパラメータを下式とする。

$$\pi = \hat{\pi} \quad (B.6)$$

$$a_{ij} = \hat{a}_{ij} \quad (B.7)$$

$$b_i(k) = \hat{b}_i(k) \quad (B.8)$$

このようにして最大尤度を出力する  $A, B, \Pi$  を得る。