

## 加速度センサを用いたHMMによるジェスチャー認識

3B-5

野津拓人、澤田秀之、橋本周司

早稲田大学

## 1. はじめに

人間が日常行っているコミュニケーションの手段として、音声言語などのバーバルコミュニケーションと、ジェスチャーなどのノンバーバルコミュニケーションがある。ジェスチャーは、音声によるコミュニケーションよりも、人間の感情などの感性的な情報を伝えることができると考えられている [1]。ジェスチャーをマン・マシンインターフェースとして用いるために、現在さまざまな手法によるジェスチャー認識の研究がおこなわれている。

ここでは、3次元加速度データを入力としてHMMを用いてジェスチャー認識をおこなう試みについて報告する。

## 2. 本手法の概要

## 2.1 処理の流れ

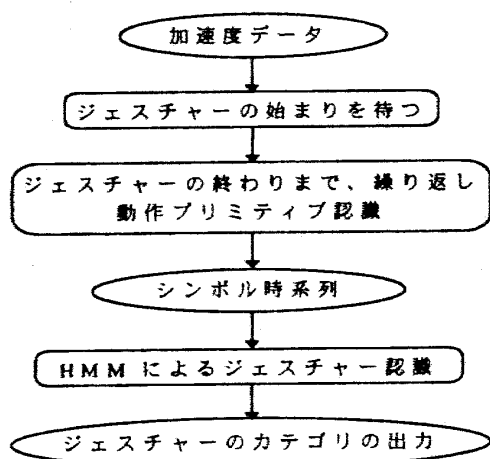


図1. 処理の流れ

HMM Based Gesture Recognition Using Acceleration Sensor  
Takuto NOTSU, Hideyuki SAWADA  
and Shuji HASHIMOTO  
Department of Applied Physics, WASEDA University  
3-4-1, Okubo, Shinjuku-ku, Tokyo, 169-8555, Japan

処理は大きく2つの部分に分けることができる。まずセンサから得られる加速度データをシンボル化するシンボル系列生成部と、これを入力としたHMMによるジェスチャー認識部である。シンボル系列生成部では、動きの要素単位（動作プリミティブ）認識を行い、各シンボルに割り当てる。これを繰り返し行うことにより生成されるシンボルの時系列に基づき、ジェスチャーの認識を行う。

## 2.2 加速度センサと動作プリミティブ認識

加速度データには、人間の体にかかる力が直接的に表われるので、ジェスチャーに込められた感情などをとらえるのにむいていると考えられる。また、図2に示すように、ジェスチャーの切り出しの問題についても、動きの始まりや終わりを敏感に検出することが可能である。加速度があらかじめ設定した閾値をこえたところをジェスチャーの始まりとして、その時点から動作プリミティブの認識が開始され、シンボル時系列が生成されていく。加速度値が閾値以下になる期間がある程度続いたところを、ジェスチャーの終わりとし、シンボルの生成を終了する。

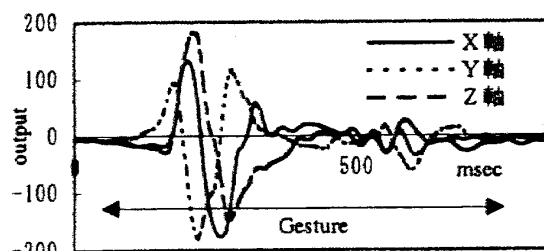


図2. 加速度センサーの出力

動作プリミティブの認識法は以下のとおりである。加速度ベクトルから加速度の変化の大きさ、動作の回転方向、主な動きの方向などを表現する特徴パラメータを抽出し、これに基づいて動作プリミティブの認識を行う。まず、学習モードで各動作プリミ

ィブを数回ずつ入力し、各パラメータの平均と分散を計算することにより標準パターンを作成しておく。認識時には、各特徴パラメータごとに入力パターンと標準パターンとの二乗誤差を分散値で割った、重みつき誤差を計算する。これをすべての特徴パラメータについて和をとったものを標準パターンからの距離とする。最も距離の近い標準パターンを認識結果とし、その動作プリミティブに割り当てられたシンボルを出力する。この方法で、動作プリミティブ単独のジェスチャーでは、ほぼ 100% の認識が可能である [2]。

### 2.3 HMM認識

HMM (Hidden Markov Model) は、確率的にシンボル時系列を生成するモデルで、主に不特定話者の音声認識の分野で用いられている。これを認識に用いる場合には、あらかじめカテゴリ毎の HMM を用意し、標準パターンのサンプル時系列の出力確率が最大になるように、HMM のパラメータ (状態遷移確率、出力確率) をカテゴリごとに学習する。そして、このパラメータを用いて、観測されたシンボル時系列が各 HMM から出力される確率を算出する。最も確率の高い HMM のカテゴリを認識結果とする。

本手法では、以下のような HMM を用いた。

- ・ 状態数 4
- ・ left-to-right
- ・ 飛び遷移あり

学習には Baum-Welch アルゴリズムを用い、認識には Forward アルゴリズムを用いた。

## 3. 実験

本手法を用いて、比較的簡単なジェスチャーの認識実験を行った。

### 3.1 実験方法

図 3 に示すように、多角形を描く、ジグザグに振るなど 20 種類のジェスチャーを認識対象とした。なお、3次元加速度センサを右手の甲に装着しジェスチャーを行う。動作プリミティブを上・下・左・右・前・後・静止の 7 個とし、それをもとに認識を行った。

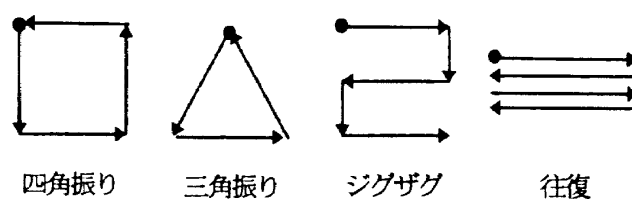


図3. ジェスチャーの例

### 3.2 結果と考察

認識結果の例を表 1 に示す。括弧内のアルファベットはジェスチャーの運動面を表わし、A は床と平行な面、B は体と平行な面を表わしている。

ジェスチャー	正答率 (%)	主な誤認パターン
四角振り (A 時計回り)	100	—
四角振り (B 時計回り)	94	三角振り
三角振り (B 時計回り)	75	逆三角振り
三角振り (B 反時計回り)	68	四角振り
ジグザグ (B 左右)	85	四角振り

表 1 認識結果の例

認識結果は平均で 正答率が 83.6% となった。

ジェスチャーの中に動作プリミティブに含まれていない動きを含んでいるジェスチャーは認識率が相対的に低下している。これは、動作プリミティブを増やすことによって解決できると思われる。ただしその際に、動作プリミティブの選択の仕方を慎重にする必要があり、効率の良い動作プリミティブの選び方の検討が今後の課題の一つである。

また、今回の手法では動作プリミティブを一定時間間隔で切り出していた。この方法では動作プリミティブを正確に切り出すことが難しく、動作速度に対するロバスト性が低下すると考えられる。これは加速度パターンから動作プリミティブの切れ目を検出し、自動的に切り出しを行うことにより解決可能であると考えられる。

これらの点を改善し、よりロバストなジェスチャー認識を今後目指していきたい。

### 参考文献

- [1] 黒川隆夫, "ノンバーバルインターフェース" オーム社 (1994)
- [2] 澤田秀之, 橋本周司, 松島俊明, 「ヒューマンインターフェースとしてのジェスチャー認識」インタラクション '97 論文集, 情報処理学会, pp. 25-32, 1997