

# 運動覚センサと筋電位センサを併用したジェスチャ認識

3B-3

坂口貴司\* \*\* 岡林孝志\*\* 金森務\* 井口征士\*\*  
 \*イメージ情報科学研究所 \*\*大阪大学基礎工学部

## 1 はじめに

人間の身体動作には豊かな情報が含まれており、マシンおよびマンマシンインタフェースの有力な手段と考えられる。身体ジェスチャは、シンボルで表現される意味の情報だけでなく、意味内容の程度を表す「とても」「やや」などのサブシンボリックな情報（以降、本論文ではそれぞれ「シンボル情報」、「バリュウ情報」と呼ぶ）を認識することが重要である。

しかし、一般的な位置・姿勢計測に基づいた方法ではバリュウ情報をうまく認識することができない。また画像処理による方法[1]は非接触式であるという利点はあるが、オクルージョンなどの問題がある。また磁気センサによって位置および姿勢を検出する方法[2]は周囲の磁界環境に影響されるなどの問題がある。さらに両手法とも計測範囲が限られるという欠点がある。

そこで、著者はバリュウ情報を抽出しやすい運動覚センサの中で、オクルージョンの心配がなく磁界環境の影響も受けないという特長を有するジャイロセンサと加速度センサを用いて人間動作を力学的に計測および認識する手法[3],[4]によって、高いバリュウ情報認識度をあげた。しかし、一部、動作（加）速度よりもむしろ力の入れ具合によって身体表現を行うジェスチャに関しては認識率が良くなかった。

そこで、本論文では上記手法に筋電位センサを用いて測定したEMG（皮膚表面筋電）データを加えた、より多様で高精度なバリュウ情報認識手法を提案する。

本手法は人間の骨格モデルに基づかないリアルタイムジェスチャ認識である。

## 2 計測および認識システム

図1にジェスチャ計測システム構成図を示す。3軸の加速度センサとジャイロセンサはたばこサイズのユニットに納め、ユーザの手の甲にサポーターで止める。筋電位センサは上/下腕に1ユニットずつ装着する。図2に使用した筋電位センサと装着方法を示す。

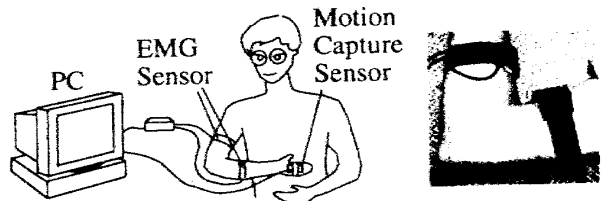


図1 システム構成

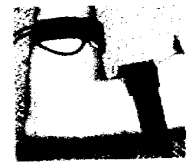


図2 筋電位センサ

認識には非線形パターン認識に適したニューラルネットワークを用いた。図3に認識システムを表す。

ジェスチャを行うとその間の3軸の加速度・角速度および筋電位データが約30Hzで得られる。そこから特徴量を抽出し、3層のニューラルネットワークを用いて、そのジェスチャのシンボル情報やバリュウ情報を出力させる。

特徴量にはジェスチャ動作中の動的情報を表すものとして加速度センサとジャイロセンサの出力の各軸毎の平均値およびピーク振幅、計算式(1),(2)で求めたバリュウ情報評価関数およびその自乗値のそれぞれ平均値（記号P）を使用し、これに上・下腕に装着した筋電位センサの出力のピーク振幅を入力ノードとして加えた。

$$P_a(t) = \sqrt{(a_x(t)^2 + a_y(t)^2 + a_z(t)^2)/3} \quad (1)$$

$$P_g(t) = \sqrt{(g_x(t)^2 + g_y(t)^2 + g_z(t)^2)/3} \quad (2)$$

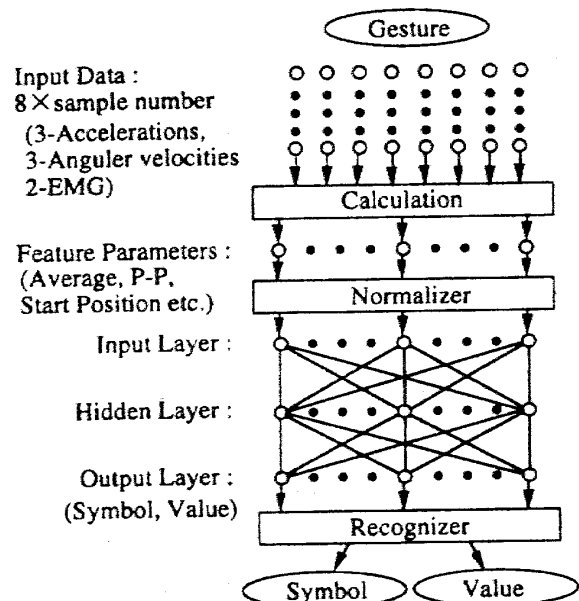


図3 認識システム

Gesture Recognition by integrating Motion Sensors and EMG Sensors

Takashi Sakaguchi \*\*, Takashi Okabayashi \*\*, Tsutomu Kanamori \*, Seiji Inokuchi \*\*

\* Laboratories of Image Information Science and Technology, \*\* Osaka University

式中  $a_x, a_y, a_z, g_x, g_y, g_z$  はそれぞれ加速度センサ, ジャイロセンサのサンプリング時間毎の3軸方向の出力である. またシンボル情報認識を行う場合にはジェスチャ前後の静的情報としてジェスチャ開始・終了時点の加速度データとジャイロデータ積分値を追加した. 中間層のノード数は16とした.

### 3 ジェスチャ認識実験

#### 3.1 実験概要

認識対象ジェスチャは日本手話(JSLs)[5]とし, バリュウ情報を多く含んでいる形容詞/形容動詞/動詞を表す手話の中から, 主に上肢運動によってその意味/程度を表現しているものを14種類選択した. 選択した単語は「寒い」「暑い」「嬉しい」「涼しい」「悲しい」「難しい」「美味しい」「明るい」「強い」「話す」「(気)楽な」「安心な」「怒る」「苦しい」であり, 並進/回転動作, 繰り返し/非繰り返し動作を含んでいる. なお, 実験は単独のジェスチャを認識対象とし, 認識は特定者に限定した.

ジェスチャデータはシンボル情報およびバリュウ情報認識用として14種類×5程度×10回の計700データを収集し, 半分を教示データ, 残りをテストデータとした. バリュウ情報は0~100までの値で表現し, 被験者の心理スケールを基準にして程度の小さい方から順に10, 30, 50, 70, 90の5段階を収集した. ここでは被験者の意図したバリュウ情報を真値とした.

#### 3.2 バリュウ情報認識

バリュウ専用ニューラルネットワークによる認識実験を行った. ニューラルネットワークはジェスチャ個別のものと全ジェスチャを認識するものを作成した.

表1に結果を表す. 表内の数字は被験者の意図したバリュウ情報(真値)と認識結果との偏差の平均値である. 縦の列は左から順に加速度センサとジャイロセンサからの情報のみ, EMGセンサからの情報のみ, 両センサからの情報を用いた場合のバリュウ情報認識誤差である. 上, 下段はそれぞれジェスチャ個別用, 全ジェスチャ用のニューラルネットワークで認識させた結果を表す.

ジェスチャ個別用のニューラルネットワークを用いた場合, 全ジェスチャのバリュウ情報認識誤差の平均

表1 バリュウ情報認識誤差(14ジェスチャ)

	Motion sensors	EMG sensors	Both
Average	10.0	12.6	8.4
All net	12.5	16.3	10.7

表2 シンボル情報認識率

Motion sensors	Motion & EMG
97.7%	90.9%

をとった結果, 両センサ情報の併用が有効であることを確認した. また, 全14ジェスチャを1つのニューラルネットワークで認識させた場合, 運動覚系や筋電位系のジェスチャを同一ネットで認識するため認識誤差は大きくなるが, やはり両センサを併用すれば認識誤差を少なくすることができることを確認した.

#### 3.3 シンボル情報認識

表2にシンボル情報認識の結果を示す. 運動覚センサのみを用いた場合, 97.7%認識することができた. これに対して筋電位センサ情報を入力ノードに加えた場合, 認識率は90.9%に下がった. これはジェスチャのシンボル情報と相関が少ないEMGのピーク振幅情報が入力ノードに加わったため, パターン認識の精度が悪化したためと考えられる.

### 4 おわりに

運動覚センサ(加速度センサとジャイロセンサ)および筋電位センサで計測したデータをニューロネットワークで統合化するシステムでジェスチャ計測および認識を行う手法を提案した. また14種類のジェスチャに関して認識実験を行い, 両タイプのセンサの情報統合(センサフュージョン)が有効であること, シンボル情報認識には運動覚センサのみの情報を用いる方法が有効であることを確認した.

### 参考文献

- [1]高橋, 関, 小島, 岡: ジェスチャー動画像のスポッティング認識, 信学論, J77-D-II-8, pp.1552-1561, 1994.
- [2]佐川, 酒匂, 大平, 崎山, 阿部: 圧縮連続D P 照合を用いた手話認識方式, 信学論, J77-D-II-4, pp.753-763, 1994.
- [3]坂口, 金森, 片寄, 佐藤, 井口: 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.
- [4]坂口, 金森, 片寄, 佐藤, 井口: 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, 第54回情報処, 4, pp.37-38, 1997.
- [5]井崎, 冬野: てことば「手話の本」うごきときもちの手話, あすなろ書房, Vol. 3, 1993.