

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

7P-1

坂口貴司* ** 金森務* 片寄晴弘* 佐藤宏介*** 井口征士**

*イメージ情報科学研究所 **大阪大学基礎工学部 ***奈良先端科学技術大学院大学

1 はじめに

人間のジェスチャには豊かな情報が含まれており、今後のマンマシンおよびマンマンインタフェースの有力な手段と考えられる。そこで、近年、VR、手話認識などの分野において、人間動作の計測および認識を行う研究が盛んに行われている。しかし、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、計測範囲に限られるなどの問題がある。

さらに、人間の動作の中でもジェスチャでは、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）の認識が重要であるが、一般的な位置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。

そこで著者らは、ジャイロセンサと加速度センサを用いて、人間動作の力学的な計測および認識手法を検討している[3], [4].

本論文では最初に両センサを用いたジェスチャ認識について提案し、開発した3軸のジェスチャセンシングユニットについて説明する。また、これとニューロネットを用いて10種類のジェスチャ（手話動作）の認識実験を行い、シンボル情報認識およびバリュウ情報認識における有効性を検討したので報告する。

2 システム概要

2.1 3軸ジェスチャセンシングユニット

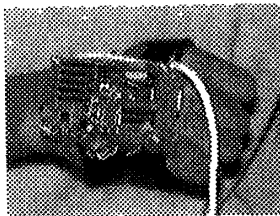


Photo.1 Gesture Sensing Unit (W72 D46 H55)

Photo.1に開発した3軸ジェスチャセンシングユニット

Gesture Recognition using Gyroscopes and Accelerometers

Takashi SAKAGUCHI* **, Tsutomu KANAMORI*,

Haruhiro KATAYOSE*, Kosuke Sato***, Seiji Inokuchi**

* Laboratories of Image Information Science and Technology,

** Osaka University, *** Nara Institute of Science and Technology

ニットと装着方法を示す。センシングデータはA/D変換器を通して有線パソコンへ送信され処理される。なお、計測データの周波数帯域は狭く、無線化することは容易である。

2.2 認識システム概要

Fig.1に認識システムを表す。本システムは開発した上記3次元センシングユニットを右手の甲にバンドで装着する。ジェスチャを行うとその間の3軸の加速度および角速度データが約30Hzで得られる。そこから特徴量を抽出し、バックプロパゲーションアルゴリズムによって求めた3層のニューラルネットを用いて、そのジェスチャのシンボルやバリュウを出力させる。

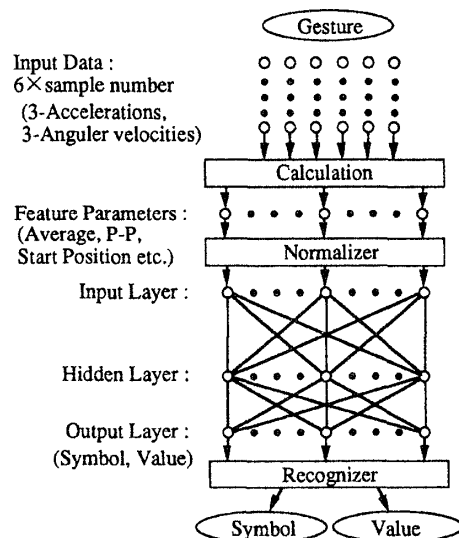


Fig.1 Recognition System

シンボル情報認識の入力ノードには両センサの時系列データから得られる各軸毎の平均値、ピーク振幅、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割り当てた。

バリュウ情報認識の入力ノードには前述の各センサの平均値およびピーク振幅（記号2）と計算式(1), (2)で求めたバリュウ情報評価関数およびその自乗値（記号P）のサンプル平均を入力ノード候補とした。式中Ax, Ay, Az, Gx, Gy, Gz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。

$$Pa(t) = \sqrt{((Ax(t)^2 + Ay(t)^2 + Az(t)^2) / 3)} \quad (1)$$

$$Pg(t) = \sqrt{((Gx(t)^2 + Gy(t)^2 + Gz(t)^2) / 3)} \quad (2)$$

また中間層のノード数は16とした。出力層は出力ノード数を認識ジェスチャ数に対応させて、教示時に対応ノードのみを1.0としその他を0.0とした。認識時には最大値を示すノードに対応するジェスチャを認識結果とし、その値(0.0~1.0)をバリュウ情報とした。

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/形容動詞など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10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」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た。

3.2 バリュウ情報認識

ジェスチャデータはバリュウ情報認識用として10種類×5程度×10回の計500データを収集し、半分を教示データ、残りの半分をテストデータとした。バリュウ情報は0.0~1.0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0.1, 0.3, 0.5, 0.7, 0.9の5段階を収集した。出力ノードはバリュウノード1つとした。なお、認識は特定者に限定した。テストデータに対するバリュウ情報認識のずれをTable 1に示す。

単独センサの場合、認識度はあまり良くないが、併用センサの場合、2Pまたは2の情報を用いるとバリュウ情報の認識が可能である。

	2P	2	P
Accelerometers	0.146	0.132	0.147
Gyroscopes	0.155	0.141	0.263
Both	0.063	0.069	0.145

Table 1 Deviation of quantitative information for input node (All gesture net)

3.3 意味(シンボル)情報認識

出力ノードはジェスチャ数である10を割り当てた。ここでは特定者認識(A), 不特定者認識(B), グループBの被験者のジェスチャを教示データに加えた場合の認識(C)の3種類の認識実験を行った。不特定者には手話熟練者1名を含む5名を選定し、ジェスチャデータとして5人×10種類×2回の計100データを収集し、半分を教示データ、残りの半

分をテストデータとした。

	unit : %		
	A	B	C
Accelerometers	99.6	70.0	92.0
Gyroscopes	100.0	72.5	94.0
Both	98.8	85.0	100.0

A : Particular (teaching)
B : Non-Particular (no teaching)
C : Member B, teaching

Table 2 Recognition rate of symbol information

Table 2に認識結果を表す。特定者ジェスチャ認識の場合、単独センサでもほぼ100%認識することができた。これに対して不特定者ジェスチャ認識の場合、個人差の影響で当然認識率は低下するが、単独センサを用いた場合に比べて、両センサを併用した場合の認識率はそれ程低下しない。また、不特定者のジェスチャも教示データに加えた場合、学習に幅が広がるため、認識率は上がり、特に両センサを併用した場合の認識率は100%となった。

4 おわりに

加速度センサとジャイロセンサを用いた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、3軸の加速度センサとジャイロセンサを組み込んだジェスチャセンシングユニットを開発した。また、これとニューロネットワークを用いて10種類のジェスチャ(手話動作)の認識実験を行い、シンボル情報に加えて、バリュウ情報の取得が可能であること、シンボル情報認識においてもバリュウ情報認識においても両センサの併用が有効であることを確認した。

参考文献

- 1) 高橋, 関, 小島, 岡: ジェスチャー動画のスポッティング認識, 信学論, J77-D-II-8, 1552/1561(1994).
- 2) 佐川, 酒匂, 大平, 崎山, 阿部: 圧縮連続DP照合を用いた手話認識方式, 信学論, J77-D-II-4, 753/763(1994).
- 3) 坂口, 金森, 片寄, 佐藤, 井口: ジャイロセンサを用いた人間の動作計測, 第50回情処全, 4, 329/330(1995).
- 4) 坂口, 金森, 片寄, 佐藤, 井口: 人間動作計測のための加速度センサとジャイロセンサを用いた回転動作計測手法の開発, 第52回情処全, 6, 181/182(1996).
- 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)