

運動特徴と形状特徴に基づいたジェスチャー認識と 手話認識への応用

澤田 秀之[†] 橋本周司[†] 松島俊明^{††}

意志や感情の伝達においては、言語以上に非言語的手段によるところが大きいといわれる。身振り手振りなどのジェスチャーはその代表的なものであり、日常の対話において自然に現れるばかりでなく、言語障害者あるいは言葉の通じない異国人の間の会話では主要なコミュニケーション手段となっている。ジェスチャーは、主に手腕部の運動形態と手の形状によって表されると考え、データグローブ、ポジションセンサおよび小型3次元加速度センサを用いて、5指の屈曲、空間的な手の配置、腕のダイナミカルな動きを検出し、それらを統合することによってジェスチャー認識を行った。各センサから得られる計測データをもとに特徴パラメータを抽出し、パターン間の距離を計算することによってジェスチャー認識を行う。ジェスチャーの形状ばかりでなく手腕部の運動のダイナミクスを考慮することにより、従来手法に比べてより多くのジェスチャー認識が可能となったばかりでなく、従来より問題となっていたジェスチャーの始まりと終わりの検出もロバストに行うことが可能となった。さらにジェスチャーによる柔軟なヒューマンインタフェース実現に向け、実時間手話単語認識システムへと拡張するとともに、手話単語データベースの構築を進めている。

A Study of Gesture Recognition Based on Motion and Hand Figure Primitives and Its Application to Sign Language Recognition

HIDEYUKI SAWADA,[†] SHUJI HASHIMOTO[†]
and TOSHIKI MATSUSHIMA^{††}

Gesture plays an important role in our daily life as nonverbal media for emotional human communication. And the gesticulation is a necessary method of communication among the people who are hearing impaired as sign language. The arm motion and the hand figure are considered to be the primitive information in understanding gestures. This paper proposes a gesture recognition algorithm based on motion primitives and hand figure primitives extracted from acceleration sensors, position sensors and datagloves, to understand human dynamic movements and hand postures. The sensor integration method realized the robust gesture recognition in spite of the simple algorithm. The algorithm is applied to a sign-language recognition system for the realization of new human-machine interface through gestures. We also propose a sign language database based on the motion primitives and hand figure primitives.

1. ま え が き

我々の日常生活における情報伝達においては、身体メディアによる非言語的手段が音声による言語的手段に劣らぬ機能を持っているといわれる^{1)~3)}。身振り手振りなどのジェスチャーはその代表的なものであり、特に意志や感情の伝達においては大きな役割を果たし

ている。ジェスチャーは日常の対話において自然に現れるばかりでなく、言語障害者あるいは言葉の通じない異国人の間の会話では主要なコミュニケーション手段となっている。通常、人間に対して用いられるこれらのジェスチャーが機械によって理解できれば、伝達者の感情や情緒などの微妙な表現にも対応できる自然な入力手段となりうるであろう。

筆者らはこれまでに画像処理やデータグローブ、加速度センサを用いたジェスチャー認識手法を検討してきた^{4)~9)}。前者の2つは、空間的な手の配置あるいは5指の屈曲などを基にしてジェスチャーを認識するものである。一方後者では、人間の意志や感情が、

[†] 早稲田大学理工学部応用物理学科

Department of Applied Physics, School of Science and Engineering, Waseda University

^{††} 東邦大学理学部情報科学科

Department of Information Sciences, Toho University

手などの位置よりもむしろ身体に加えらるる力に顕著に現れると考え、運動加速度の直接検出によるジェスチャー認識を試みた。

本論文では、これら 3次元加速度センサとデータグローブによる手法を統合したジェスチャー認識について述べる。ジェスチャーは、主に手腕部の運動形態と手の形状によって表されると考え、前者の認識に加速度センサを、後者にはデータグローブを用いて認識を行っている。さらに手話単語データベースの構築と実時間手話単語認識への拡張について報告する。ジェスチャーの形状ばかりでなく動作のダイナミクスを考慮することにより、従来手法に比べてより多くのジェスチャー認識が可能となったばかりでなく、従来より問題となっていたジェスチャーの始まりと終わりの検出も容易に行うことが可能となった。

2. ジェスチャー計測と特徴パラメータ

2.1 加速度センサと手腕部の運動記述パラメータ

加速度の計測には、ジャイロをはじめとして様々な手法が用いられているが、ジェスチャー認識に用いるには、手軽に扱えて、かつ手先などに装着した場合に自由な運動が行える必要がある。また、腕の運動加速度はきわめて大きく通常の運動でも最大値は 20 G を越えるため、測定レンジも大きくなければならない。ここではおもりの変位を測定する簡単な加速度センサを用いている（日本光電工業製三軸加速度センサ TA-513G）。

図 1 にその内部構造を示す。大きさは縦 20 mm、幅 15 mm、高さ 12 mm で、重量は 12.5 g と小型軽量なため、容易に体の一部に取り付けることができる。センサ内部には x , y , z 軸方向の加速度を独立に検出するために、各々の軸方向で金属おもりが圧電素子に接続してある。測定範囲は、 $-25 \sim +25$ G であり、加速度にともなう圧電素子の電圧が出力される。この出力は増幅され A/D 変換ボードを通じて 12 ビットのデジタル値として計算機に取り込まれる。感度は 1 G あたり約 5 mV である。なおドリフト除去のために、遮断周波数 3 Hz の低域除去フィルタをアンプ入力前に通している。

加速度センサからは、 x , y , z 方向の加速度値が $a_x(t)$, $a_y(t)$, $a_z(t)$ として得られる。なお本研究でセンサを手の甲に固定して使用する場合には、指先の向きがセンサ x 軸の負の向きとなるよう統一している。図 2 (a) に、 $y-z$ 面で円を描く動作を行ったときのセンサ出力値の例を、後述のポジションセンサからの出力値（同図 (b)）およびポジションセンサの出力の二

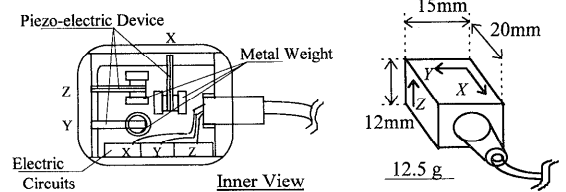
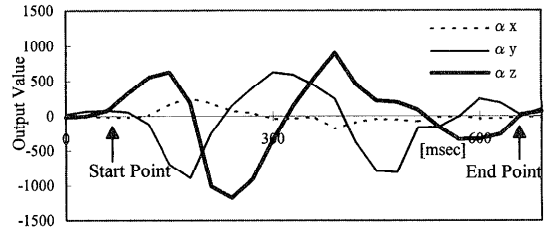
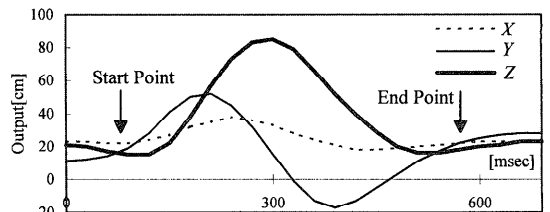


図 1 3次元加速度センサの構造
Fig. 1 3D acceleration sensor.



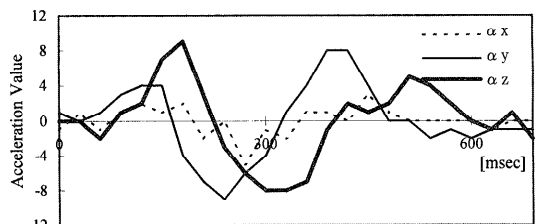
(a) 加速度センサの出力

(a) Output of acceleration sensor



(b) ポジションセンサの出力

(b) Output of position sensor



(c) ポジションデータより得られる加速度

(c) Acceleration obtained by position data

図 2 「円を描く」動作による各センサの出力例

Fig. 2 Examples of sensor outputs (gesture of circle).

階差分（同図 (c)）との比較で示す。各軸方向に取り付けられたおもりの変位を測定するという本センサの構造上、3軸の加速度値の分離が不十分で干渉が見られるが、ポジションセンサを用いた場合に比べ、遅れやノイズが少なく運動の検出を行えることが分かる。

この 3次元加速度ベクトルの時系列パターンから、ジェスチャーの認識を行うのであるが、ダイナミカルな動作から得られる加速度データの場合、単なるベク

表 1 加速度センサの特徴パラメータ

Table 1 Characteristic parameters of acceleration sensor.

P_d	動きの激しさの特性 (加速度ベクトルの変化)
P_g	回転方向特性 (ベクトル間の外積)
P_r	主動作の方向特性 (ベクトル外接四角形の縦横比)
$P_{a0} \sim P_{a7}$	方向の分布特性 (ベクトルの方向分布)

トル系列のパターンマッチングでは高い認識率が得られない。ジェスチャー運動の大域的な特徴を抽出する必要がある。ここでは運動の直感的な把握を容易にするために、次のような 3 つの 2 次元ベクトルに基づいて運動の特徴を抽出している。

$$A_1(t) = (a_y(t), a_z(t)), A_2(t) = (a_x(t), a_z(t))$$

$$A_3(t) = (a_x(t), a_y(t)) \quad (1)$$

ここで $A_1(t)$, $A_2(t)$, $A_3(t)$ は、それぞれ $y-z$ 平面, $z-x$ 平面, $x-y$ 平面へ射影した加速度ベクトルである。ここでは、データ取り込み周波数 f_{in} は 30 Hz としている。これは急激に方向や速度を変えるような運動の計測においては不十分なサンプリング周波数ではあるが、ここでは最小限の加速度データを用いて実時間でジェスチャー認識を行うことに重点をおいた。

各射影加速度ベクトルの時系列から、表 1 に示す 11 個の運動特徴パラメータの抽出を行う。ここでは、実時間ジェスチャー認識を可能とするために最低限必要と思われる運動の特性として、運動によって生じる加速度変化の大きさ、回転運動の方向、動作の向きを用いている。ここで得られる特徴は、主として動きの動的特性となる。

P_d は加速度ベクトルの時間差分によって与えられるパラメータであり、

$$P_d = \sum_{t=1}^{S-1} \|\Delta A(t)\| \quad (2)$$

$$\Delta A(t) = A(t+1) - A(t) \quad (3)$$

によって求められる。

P_g は動作の回転方向の情報を表すパラメータである。これは次式 (4), (5) に示すように $t, t+1$ のベクトル間の外積により計算され、たとえば $A(t)$ から $A(t+1)$ へ反時計回りに移動した場合、 P_g に 1 が加算される。

$$P_g = \sum_{t=1}^{S-1} u(A(t) \times A(t+1)) \quad (4)$$

$$u(m) = \begin{cases} 1 & : m \geq 0 \\ 0 & : m < 0 \end{cases} \quad (5)$$

P_r は S 個のベクトルによって各加速度平面に描かれる図形に対し、水平におかれた外接四角形の縦横比であり、絶対値の大きいベクトルによって与えられる方向の情報を持つパラメータである。たとえば $y-z$ 平面においては $A_1(t)$ より、

$$P_r = \frac{Vt^+ - Vt^-}{Hr^+ - Hr^-} \quad (6)$$

$$Hr^+ = \max_S \{a_y\}, Hr^- = \min_S \{a_y\}$$

$$Vt^+ = \max_S \{a_z\}, Vt^- = \min_S \{a_z\}$$

によって求められる。

P_{a0} から P_{a7} までのパラメータは、 $\pi/4$ ごとの 8 つの主方向への加速度データの偏りを表すものである。一連の加速度データに対して、一種のメンバーシップ関数を用いて各主方向への所属度合いをファジー集合的に決定し、各方向の加速度の密度を次式で求める。

$$P_{an} = \sum_{t=1}^S F_n(\theta_t) \quad (7)$$

$F_n(\theta_t)$: Membership Function
of n -th Main Direction

θ_t : Direction of $A(t)$

$n = 0, 1, 2, \dots, 7$

メンバーシップ関数は図 3 のように各主方向から隣接する主方向まで直線的に減少する孤立三角波関数とした。たとえば図 3 に示すように $\pi/4$ より少し大きい角度 θ 方向の加速度ベクトルが観測された場合は、 P_{a5} と P_{a6} に対してそれぞれ f_5 と f_6 の値が、その他の P_a には 0 が加えられる。このように、 P_a には運動の方向の特徴が現れることになる。

以上のすべての特徴パラメータが時系列ベクトル $A_1(t)$, $A_2(t)$, $A_3(t)$ からそれぞれ計算され、合計 33 個の運動特徴を用いてジェスチャー認識が実行される。

2.2 データグローブと手形状記述パラメータ

前述の加速度センサは身体のダイナミカルな動きを計測するには適しているが、ジェスチャー、特に手話

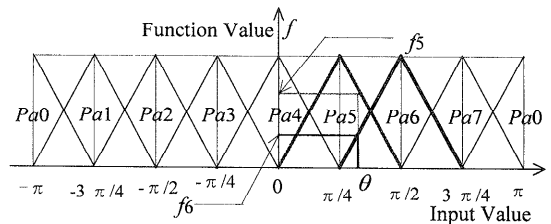


図 3 メンバーシップ関数

Fig. 3 Membership function.

表 2 データグローブの特徴パラメータ
Table 2 Characteristic parameters of dataglove.

R0-R2	手の 3 次元位置
R3, R4	$\sin(\theta_{pt}), \cos(\theta_{pt})$
R5, R6	$\sin(\theta_{rt}), \cos(\theta_{rt})$
R7, R8	$\sin(\theta_{yw}), \cos(\theta_{yw})$
R9-R13	5 指の曲がり度

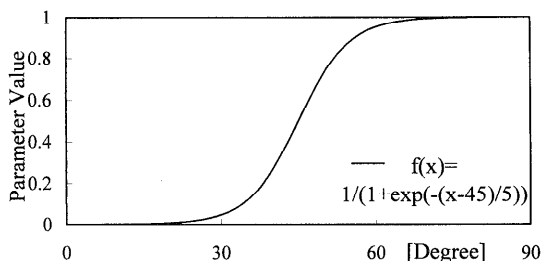


図 4 指の曲がり度の評価関数
Fig. 4 Logistic function.

においては手指の静的な形状も大きな意味を持っている。そこで本システムでは、データグローブとポジションセンサを用いて手の形状および 3 次元位置と姿勢の実時間計測を行っている。本データグローブ（日商エレクトロニクス SuperGlove）は、曲げセンサを指裏側部分に配置することによって指の曲げ角に応じた出力が得られる。さらに手の甲部分に固定したポジションセンサ（Polhemus Sensor）によって、手の 3 次元位置 x, y, z と、姿勢としてピッチ θ_{pt} 、ロール θ_{rt} 、ヨー角 θ_{yw} を得ることができる。ポジションセンサによる 3 次元位置座標は、地上を基準とする人間の座標系と一致し、鉛直上方を z 軸正の向き、手前側を x 軸正の向き、左を y 軸正の向きと設定した。

データグローブおよびポジションセンサからの出力をもとに、表 2 に示す 14 個の特徴パラメータが抽出される。手の姿勢 R3-R8 は、得られたピッチ、ロール、ヨー角の正弦、余弦を取ったものであり、 $\pm(n\pi)/2 [n = 0, 1, 2]$ 付近の手の姿勢の選択性を高めるばかりでなく、 $-\pi$ から $+\pi$ への変化点での不連続性をなくすることができる。また、各指の曲がり度は、図 4 に示したロジスティック関数によって屈曲角度より変換される。伸びている状態を 0 として、完全に曲がった状態の 1 までの間の値をとるが、これによって曲がっている状態と伸ばしている状態を容易に弁別することができる。

なお、これらのパラメータは、あらかじめ使用者の体型に合わせてジェスチャー空間を設定することによって、規格化されたものである。使用者はまず、前後、

左右、上下の位置限界値と最大運動速度の入力を行ったうえで、手の初期位置の設定を行う。ポジションセンサから得られる位置パラメータ R0-R2 は、この初期位置を基準として算出される。さらに手形状のパラメータ R9-R13 は、はじめに指の曲げ伸ばし動作を行うことによって規格化が行われる。

3. ジェスチャー認識

3.1 ジェスチャー認識の流れ

ジェスチャー認識には、加速度ベクトルおよびデータグローブから得られる特徴パラメータを利用し、認識アルゴリズムは個人対応型としている。システムを使用する前にあらかじめ学習モードで個人の運動特徴の標準パターンを作成してから認識モードに入る。これにより特定動作者の認識が実現できるばかりでなく、認識させたいジェスチャーを自由に選ぶことができる。

時系列データから特定のジェスチャーを認識するためには、その始まりを検出しなければならないが、従来のジェスチャー認識手法ではこれが大きな問題となっていた。加速度センサを用いた本手法では、動作の開始を加速度の立ち上がりの検出によって容易に検出可能であるという利点を持つ。ここではジェスチャーの開始は、加速度の絶対値の大きさで判断している（図 2(a)~(c) 参照）。使用者の非動作時の手ぶれの大きさをあらかじめ測定しておき、これよりも大きな加速度入力があった場合にジェスチャー開始と見なしている。

前述したように、ジェスチャーは主に手腕部の運動形態と手の形状によって表されるが、人間の視覚的な認知特性を考えた場合、運動中の細かい手形状変化はあまり意味を持たないであろうと考えられる¹⁰⁾。運動による手部の移動形態と、その前後の静止時の手形状が大きな意味を持つと考え、ここでは、

—手形状—運動—手形状—

をジェスチャーの基本単位として認識を行っている。

まず、加速度データによりジェスチャー開始を認識した時点で、データグローブからのデータを記録をする。同時に加速度データの記録を開始し、加速度パターンからジェスチャー終了が認識された時点で、再度手形状、ポジションデータの取得を行う。これらをジェスチャー計測データとして認識アルゴリズムに渡す。この流れを図 5 に示す。認識部では、計測データをもとに各特徴パラメータを抽出し、ジェスチャー認識を行う。ジェスチャー認識は、手腕部の運動認識、手形状、姿勢について別々に認識を行い、それらを統合することによって意味が決定される。

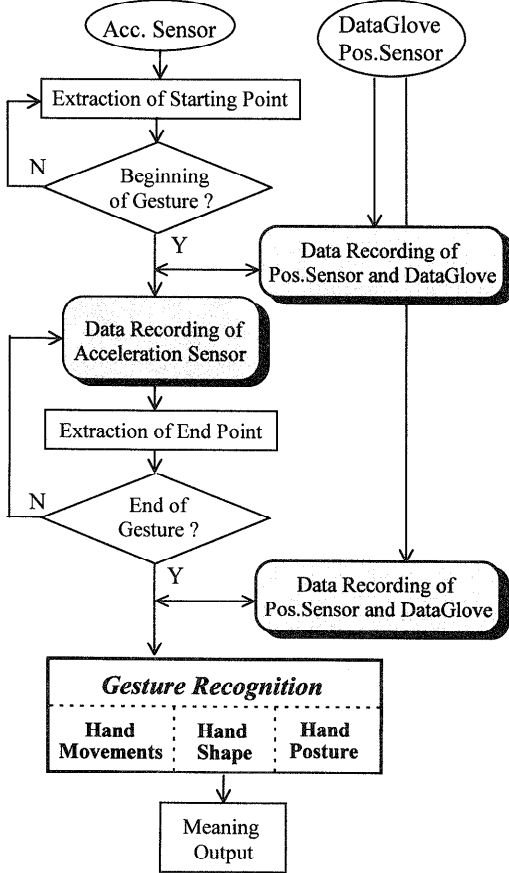


図5 ジェスチャー計測データ取得の流れ
Fig. 5 Diagram of data acquisition.

3.2 認識アルゴリズム

まず、学習モードで認識させたい動作のサンプルを M 回ずつ入力し、前記の各特徴パラメータについての平均値 E_α^G および標準偏差 μ_α^G を次式で求め、各ジェスチャー G に対する標準パターンデータを作成する。なお、動作認識は手腕部の運動認識、手形状、姿勢について別々に行われるため、運動特徴パラメータ P 's, 手形状 R9-R13, 姿勢 R1-R8 についてそれぞれ標準パターンが作成されることになる。またここでは試行回数 M は5とした。

$$E_\alpha^G = \frac{1}{M} \sum_{i=1}^M V_\alpha^{g_i} \tag{8}$$

$$\mu_\alpha^G = \sqrt{\frac{1}{M} \sum_{i=1}^M (V_\alpha^{g_i} - E_\alpha^G)^2} \tag{9}$$

V : Parameter Values of Learning Pattern
 g_i : i -th Sample of Gesture g
 α : P 's or R 's

表3 ジェスチャー認識実験の結果

Table 3 Experimental results of gesture recognition.

	本人のジェスチャー 正解率%	別人のジェスチャー 正解率% (誤認パターン)
1. 縦振り	100	100
2. 横振り	100	100
3. 斜め振り	100	75 (三角形)
4. 時計回り	100	100
5. 反時計回り	100	80 (三角形)
6. ひと振り	100	100
7. 星型	100	100
8. 三角形	100	85 (反時計回)
9. ハート型	100	100
10. 方向指が	100	85 (ひと振り)

認識モードでは未知動作を行うたびに各パラメータについて標準パターンの平均値 E_α^G との二乗誤差を計算し、それぞれ標準偏差 μ_α^G の二乗つまり分散値で割って規格化したものを重み付き誤差とする。これを各標準パターンごとに特徴パラメータ α について加え合わせたものをパターンの相違度 e_G とする。

$$e_G = \sum_\alpha \varepsilon_\alpha^G = \sum_\alpha \frac{(V'_\alpha - E_\alpha^G)^2}{(\mu_\alpha^G)^2} \tag{10}$$

V' : Parameter Values of Input Pattern

相違度が最小で、かつ下式 (11) に従って決めるしきい値 T_h よりも小さいものを標準パターンから選び認識結果とする。 T_h は、対象とするジェスチャーの標準パターン間の最小距離に相当する。

$$T_h = \min_{G \neq H} \left\{ \sum_\alpha \frac{(E_\alpha^H - E_\alpha^G)^2}{(\mu_\alpha^G)^2} \right\} \tag{11}$$

for all G, H

G, H : Gesture G, H

3.3 認識実験

手腕部運動形態の認識実験結果を表3に示す。縦振り、横振り、斜め振り、時計回り、反時計回り、ひと振り、星を描く、三角形を描く、ハートを描く、方向指示、胸の位置に静止の、11種類の手振り運動について20回ずつの認識実験を行った結果、100%の認識が可能であった。さらに、別人のジェスチャーを認識させた場合にも比較的良好な認識結果を得た。誤認識したのは、移動回転方向が同じである“斜め振り”、“反時計回り”と“三角形”、および動作方向のみが異なる“方向指示”と“強いひと振り”であった。このことから使用者は事前に個人の標準パラメータを作成してから認識動作を行ったほうがよいことは明らかであるが、ここで使用しているパラメータによって他人のジェス

チャーであってもある程度の認識が可能であることが分かった。日常生活でよく用いられる手振りはこちらのうち、縦振り、横振り、斜め振り、回転、強いひと振り、方向指示といった単体の動作、あるいは、それらいくつかの組合せが多いと考えられる。一方で、画像処理などを用いた場合に実時間認識が困難と思われる、星形、三角形、ハート型といったジェスチャー運動が3次元加速度パターンから認識可能であることが分かった。

手形状においては、各指の屈伸を考慮した場合 $2^5 = 32$ のパターンが考えられ、本アルゴリズムによってすべてのパターンの認識が可能であった。しかし実際のジェスチャーにおいて用いられる手形状パターンはその半数程度であると考えられる。たとえば、薬指のみ伸ばす、あるいは中指と小指のみ伸ばす、といった手形状が日常のコミュニケーションにおいて使われることは非常にまれである。

姿勢は、甲部の向きとして上 (Up)、下 (Down)、前 (Front)、後ろ (Back)、外向き (Outside)、内向き (Inside) の6パターンの分類および認識を行い、100%の認識結果が得られている。

4. 手話認識への応用

4.1 運動パターンに基づく手話データベース

現在日本語手話には、約3,000単語あるといわれている。しかし多くの手話は、いくつかの単純なジェスチャーの組合せで表すことができる^{9),11)}。手話で用いられる代表的な動作方向を図6に、また手形状を表4に示した。手形状パターンは、親指から小指までの5指の屈伸状態が、b(曲げ)、s(伸び)として表される。今回はこの11の動作方向を基本動作として取り上げ、さらに14の手形状と、姿勢として6つの手の甲の向き(Up, Down, Front, Back, Outside, Inside)を加えることにより手話単語の認識を行っている。

さらに我々は、基本動作の組合せをもとにした手話データベースの構築を試みている。手話においても、運動の形態とその前後の静止時の手形状が大きな意味を持つと考え、手形状と手腕部の運動をジェスチャーの基本単位として考え、手話の記述を行っている。データベースには手話単語が、図6の基本動作とその前後に付随する表4に示した手形状パターンの列によって記述されている。図7に、本データベースにおける手話単語および、手話で使われる日本語指文字の記述例を示す。なお指文字については、右肩の横で提示することを想定し、手形状および姿勢のみについて記述さ

れている。また、動作あるいは手形状が、手話の意味に対して影響を与えないものは印で表している。

4.2 手話単語認識システム

本システムは、加速度データ取得および、データグローブとポジションデータ取得用にそれぞれPCを用い、RS-232Cを通じてホストコンピュータ(SGI Indy)にデータを転送して認識を行っている。使用者は、それぞれデータグローブ、手の甲部分に固定したポジションセンサおよび加速度センサを装着することによって、手話のデータベース登録、手話単語の認識、データベースからの手話単語の意味検索を行う。ここで認識対象とする手話および指文字は、図6で示した11個の基本動作の列および、表4で示した手形状で記述できるものである。手形状については、データベースに記述された曲がった状態(bend)を1、伸ばした状態(straight)を0として、前述の認識アルゴリズムに従って形状が判定される。

図8に手話単語認識システムの流れを示す。加速度を検出した時点を手話動作開始と見なし、認識アルゴ

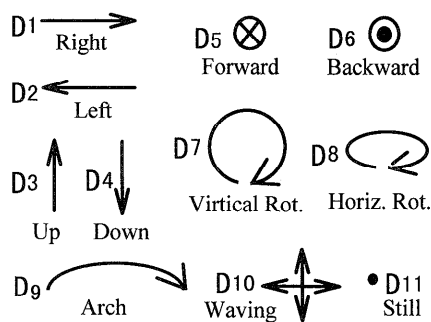


図6 手話の代表的な基本動作方向
Fig. 6 Hand movement primitives.

表4 手話の代表的な手形状

Table 4 Hand figures in sign language.

形状パターン	各指の屈伸 Straight/Bend
F1	bbbb
F2	sbbbb
F3	bsbbb
F4	bbsbb
F5	bbbs
F6	ssbbb
F7	bssbb
F8	sbbbs
F9	bsbbs
F10	sssb
F11	bsssb
F12	bssss
F13	bssss
F14	sssss

右手	左手	: 補足
あなた	You	: 手話意味
S F3/O	—	: 形状/姿勢
D5	D11	: 動作
E F3/O	—	: 形状/姿勢
楽しい	Happy	: 手話意味
S F3/B	F3/B	: 形状/姿勢
D3	D4	: 動作
F3/B	F3/B	: 形状/姿勢
D4	D3	: 動作
E F3/B	F3/B	: 形状/姿勢
北	North	: 手話意味
S F7/O	F7/O	: 形状/姿勢
D4	D4	: 動作
F7/O	F7/O	: 形状/姿勢
D2	D1	: 動作
E F7/O	F7/O	: 形状/姿勢
あ		: 指文字
F2/B	—	: 形状/姿勢
た		: 指文字
F2/O	—	: 形状/姿勢
や		: 指文字
F8/B	—	: 形状/姿勢

図7 手話データベースの記述例

Fig. 7 Examples of sign language database.

リズムが実行される。手話開始時点でまず、データグローブからのデータの記録を行う。同時に加速度データの記録を開始し、加速度パターンによる手腕部運動の認識アルゴリズムが実行される。認識結果は図6に示した基本動作方向の列として順次得られる。この動作方向認識結果の系列を監視することによって、動作方向パターンの変化が検出されたときを動作の変化点と見なし、データグローブからデータの取得を行う。手話動作終了が認識された時点で、再度手形状および姿勢データの取得を行う。この結果、

開始:手形状—運動—手形状—運動—...

...—運動—手形状:終了

の順で手話動作シーケンスが得られることになる。これは、それぞれ基本動作パターンおよび、手形状パターン列としてジェスチャーテーブルにシーケンシャルに書き込まれていき、手話データベースを参照することによって手話単語が認識される。手話単語認識には有限オートマトンを用いており、図9に「あなた」と「北」の認識を行った例を示す。

一方、指文字については、使用者の右肩の横で提示することとし、あらかじめここを動作エリアとして登録しておく。加速度センサからジェスチャー開始が認

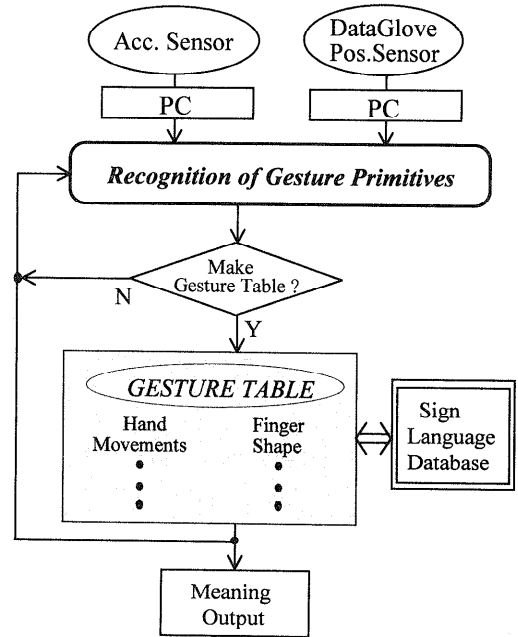
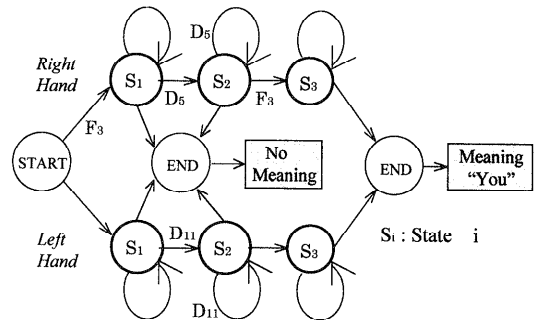


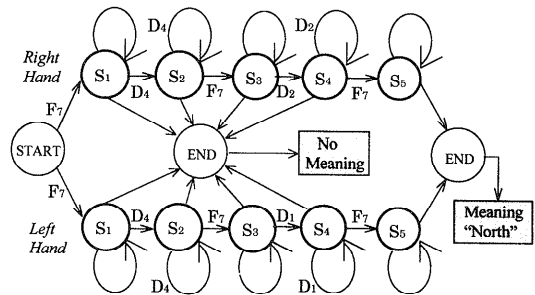
図8 手話単語認識の流れ

Fig. 8 Diagram of sign language recognition.



(a) 「あなた」の認識

(a) Recognition of "You"



(b) 「北」の認識

(b) Recognition of "North"

図9 手話単語認識オートマトン

Fig. 9 Automata in sign language recognition.

表5 手話認識実験結果の例
Table 5 Experimental results of sign language recognition.

手話単語	正解率 (%)	誤認パターン
あなた	100	-
好き	100	-
さようなら	93	意味なし
楽しい	87	意味なし
北	83	意味なし
あ (指文字)	100	-
た (指文字)	100	-

識され、指文字提示エリアにて動作が行われたときには、指文字として手形状判定が行われるようにした。

現在までに、自己紹介や日常生活でよく使われる30程度の手話単語と日本語指文字の登録を行い、その実時間認識を実現している。表5に、30回ずつの認識実験を行った結果の例を示す。

11の基本動作単体で表されるワンアクションの手話については、100%の認識が得られている。一方で、たとえば図7に示した「北」の認識において、D4→D2と1つの手話単語内で動作の方向が変わる場合、その方向変化点において動作パターンの誤認識をおこし、オートマトンが「意味なし」で終了してしまうことがみられた。これが認識率低下の原因となっており、方向変化点における動作パターン誤認識の排除が現在の課題となっている。その一方、オートマトンを用いた本手法では、手話動作が不明瞭な場合に認識途中で意味のパスが切れ「意味なし」となってしまうが、この場合は単語単位の認識であるため、ユーザーに再度入力をうながすなどの対話処理の手法を考慮することができる。

5. む す び

加速度センサ、データグローブおよびボジションセンサを統合したジェスチャー認識と手話単語認識システムへの応用について述べた。ここでは、シンプルなハードウェア構成と単純な認識アルゴリズムによる実時間ジェスチャー認識を提案するとともに、加速度データをジェスチャー認識に用いることの有用性について報告した。これを手話認識に用いることにより、ジェスチャーを用いたマン・マシンインタフェースの入力部としての手がかりが得られたと考えている。

ジェスチャーは人間の微妙なニュアンスを表すのに適したメディアである。その一方でジェスチャーが果たす役割の大きい微妙な感情表現などのパターン化は困難である。ここで述べた加速度によるジェスチャー認識手法では容易に取得データの時間分解能を上げ

ることが可能であり、身体に及ぼされる細かな力のパターンと感情表現との関係を詳細に調べていくことが今後の課題であると考えている。また、本手話認識システムを用いて手話動作を解析、分類することによって、カパターンと手形状データをもとにした手話辞典の構築を目指している。

このようなジェスチャー認識の適用により、手話ばかりでなく多くの分野に利用できる感性豊かなマン・マシンインタフェースが実現できると思われる。たとえば、ジェスチャー認識を用いてジェスチャーとコマンドを対応付け、マウスでポインティングしながら片手のジェスチャーで操作するCADシステムなどが考えられる。また、指先に小型加速度センサを取り付けることにより、指先の運動パターンを利用した入力デバイスの実現も可能であろう。

本研究の一部は、文部省科学研究費（重点(2)「感性情報」課題番号06212225 および「人工現実感」課題番号07244221）によって行われている。

参 考 文 献

- 1) 黒川隆夫：ノンバーバルインタフェース，電子情報通信学会（編），オーム社（1994）。
- 2) 金山宣夫：世界20ヶ国ノンバーバル辞典，研究社出版（1983）。
- 3) Rokeby, D.: Body Language, *ACM Siggraph Art Show*, Atlanta (1988).
- 4) Morita, H., Ohteru, S. and Hashimoto, S.: Computer Music System that Follows a Human Conductor, *IEEE Computer*, Vol.24, No.7, pp.45-53 (1991).
- 5) Sato, A., Harada, T., Hashimoto, S. and Ohteru, S.: Singing and Playing in Musical Virtual Space, *Proc. International Computer Music Conference (ICMC)*, pp.289-292 (1991).
- 6) Sawada, H. and Hashimoto, S.: Musical Performance System Using 3D Acceleration Sensor, *Multimedia Modeling Towards Information Superhighway*, pp.293-306, World Scientific (1995).
- 7) 澤田秀之，橋本周司：加速度センサを用いたジェスチャー認識と音楽制御への応用，電子情報通信学会論文誌，Vol.J79-A, No.2, pp.452-459 (1996).
- 8) Sawada, H., Onoe, N. and Hashimoto, S.: Acceleration Sensor as an Input Device for Musical Environment, *Proc. ICMC*, pp.421-424 (1996).
- 9) 澤田秀之，橋本周司，松島俊明：ヒューマンインタフェースとしてのジェスチャー認識，情報処理学会インタラクシオン'97, pp.25-32 (1997).
- 10) 黒川隆夫，神戸大樹：身振り言語の記述とコン

ピュータ化身振り辞書の構築, 第5回ヒューマンインタフェースシンポジウム論文集, pp.347-352 (1989).

- 11) Kim, J., Jang, W. and Bien, Z.: A Dynamic Gesture Recognition System for the Korean Sign Language (KSL), *IEEE Trans. Systems, Man and Cybernetics*-Part B: Cybernetics, Vol.26, No.2 (1996).

(平成9年7月2日受付)

(平成10年1月16日採録)



澤田 秀之 (正会員)

1967年生。1990年早稲田大学理工学部応用物理学科卒業。1992年同大学院修士課程修了。1995年同大学院博士課程入学。1995～1997年日本学術振興会特別研究員。現在、早稲田大学理工学部助手。音楽情報処理、ニューラルネットワーク、ヒューマンインタフェースの研究に従事。電子情報通信学会、日本光学会、日本顔学会、ICMA各会員。



橋本 周司 (正会員)

1948年生。1970年早稲田大学理工学部応用物理学科卒業。東邦大学講師、助教授、早稲田大学助教授を経て、現在、早稲田大学理工学部教授。ロボティクス、画像処理、音楽情報処理、ヒューマンコミュニケーションの研究に従事。工学博士。電子情報通信学会、計測自動制御学会、日本VR学会、日本顔学会、ICMA各会員。



松島 俊明 (正会員)

1959年生。1983年早稲田大学理工学部応用物理学科卒業。1985年同大学院修士課程修了。1988年同大学院博士課程単位取得退学。東邦大学講師を経て、現在、同大学助教授。画像理解、パターン認識、音楽情報処理の研究に従事。工学博士。電子情報通信学会、日本顔学会、日本VR学会、ICMA、各会員。本会MUS研究会幹事。