

文字認識手法を用いた固有平面上での動作識別

大野 宏^{†*} 山本 正信^{††}

本論文では、文字認識手法を用いた固有平面上での動作識別方法を提案する。ロボットモデルに基づく動画解析で人体の姿勢データ列を求める。これら姿勢データ列を2次元固有空間（固有平面）上に射影すると、動作はその特徴を失わずに固有平面上の軌跡として表すことができる。この軌跡は連続した曲線になるため、動作の識別は文字認識と等価な問題となる。文字認識手法を固有平面上の曲線の識別に適用し、本手法が動作識別に有効であることを示す。

Gesture Recognition Using Character Recognition Techniques on Two-dimensional Eigenspace

HIROSHI OHNO^{†*} and MASANOBU YAMAMOTO^{††}

This paper describes a new method for gesture recognition using character recognition techniques on two-dimensional eigenspace. An image-based approach can calculate human body poses in motion from image sequences. The sequence of poses can be reduced into a trajectory on the two-dimensional eigenspace without any loss of gesture features, so that the gesture recognition equals the character recognition. Experiments for the gesture recognition using some character recognition techniques show our method is useful.

1. はじめに

人間の動作を追跡し、それを識別することができれば、人間と計算機のインタフェースや不審な動作をする人の判別などに利用することができる。また、同じ動作でもその人独自の特徴を抽出できれば、個性を取り入れたキャラクターアニメーションの生成や、個人の識別にも応用することができる。

人体の動作識別は計算機処理能力の飛躍的向上により、近年さかんに行われてきている。認識方法として、動画上に記録された動作を動画像どうしで比較する方法^{1)~4)}や、動作を運動パラメータで記述し運動パラメータどうしで比較する方法^{5),6)}などが提案されている。

動画像どうしを比較する方法として、Muraseらはパラメトリック固有空間法を発展させ、歩行動画像から人体を切り出し、その画像集合から固有空間を構成

し、この空間上の軌跡で歩行動作を表現し個人の識別を行っている⁴⁾。しかし、画像そのものの次元が大きいため、固有空間法で次元を圧縮しても識別には十数次元が必要になり、また動画像どうしを比較するためカメラと被測定者の位置関係に強い制約がある。

一方、Campbellらは、姿勢パラメータを比較する方法として、モーションキャプチャから得られた運動パラメータのうち、特徴ある2つのパラメータから構成される位相軌道で動作を表現し、パレーのステップの識別を行った⁶⁾。この方法は、得られたパラメータから特徴ある2つのパラメータをステップごとに選ばなければならない。また、姿勢パラメータの取得にセンサを人体に取り付ける必要があるため、自然な動作が得られず、セキュリティ関連には使えない。

そこで本論文では、動画像解析法⁷⁾により非接触で求めた人体の姿勢データ列を2次元の固有空間上に射影することで、動作の特徴を失うことなく平面上の軌跡で動作を表現する。さらに、この軌跡は連続した曲線となるため、動作の識別は文字認識と等価な問題となり、文字認識に有効な手法を利用することで、動作の識別を容易に行うことができる。実験では、よく知られている動作としてラジオ体操を取り上げ、個々の動作を2次元固有空間上に射影し、文字認識手法を適用して実際に動作の識別実験を行いその有効性を示す。

† 新潟県工業技術総合研究所

Industrial Research Institute of Niigata Prefecture

☆ 現在、新潟大学大学院自然科学研究科

Presently with Graduate School of Science and Technology, Niigata University

†† 新潟大学工学部情報工学科

Department of Information Engineering, Niigata University

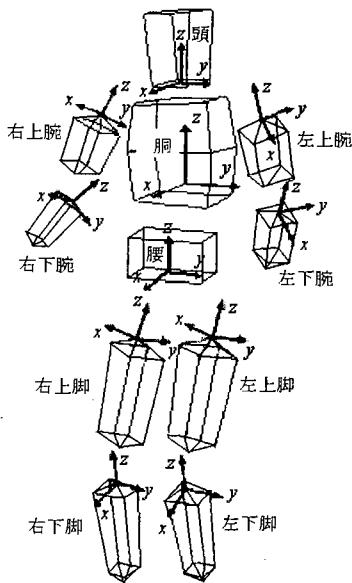


図1 人体モデル
Fig.1 Body model.

2章で人体の姿勢表現と動画像解析, 3章で2次元固有空間上に動作を表現する方法, 4章で動作の識別方法, 5章で実際に行った動作の識別実験, 6章で他の方法との比較について述べる.

2. 人体の姿勢表現と動画像解析

人体を図1のような多関節ロボットでモデル化し, CADモデリングシステムによりコンピュータ上に作成する. 人体は胸と頭, 腰, 左右上腕, 下腕, 上脚, 下脚の11パーツから成る階層構造で構成され, 各パーツは関節リンクで結ばれている. このモデルの姿勢は, きをつけの姿勢からの各関節角の回転角度で表される. また, このモデルの運動パラメータは, 各関節角の角速度とモデルの最上位親物体である胸の並進速度と角速度である. 1台のカメラで人体の追跡を行った場合, 腕や脚が胸で隠れたり, カメラの位置によって正確に人体の動きを追跡できない場合がある. これらの問題を解決するため, ここでは異なる位置に置かれた複数のカメラを利用して人体の位置と姿勢を求める.

動画像の追跡開始フレームにおいて, ロボットモデルを複数のカメラからの画像中の人体像に手でフィットさせる. このときモデルは人体の姿勢と位置を正しく表している. 時空間こう配法により運動パラメータを推定し, 初期フレームの位置と姿勢に累積することにより, 任意のフレームでの人体の位置と姿勢が得られる(付録). 例として図2に, ラジオ体操の「両足で飛ぶ運動」に対し4方向からのカメラ画像で追跡を行っているようすを示す. 動作識別では, このうち胸の位置と頭の姿勢を除いた胸, 腰, 左右上腕,

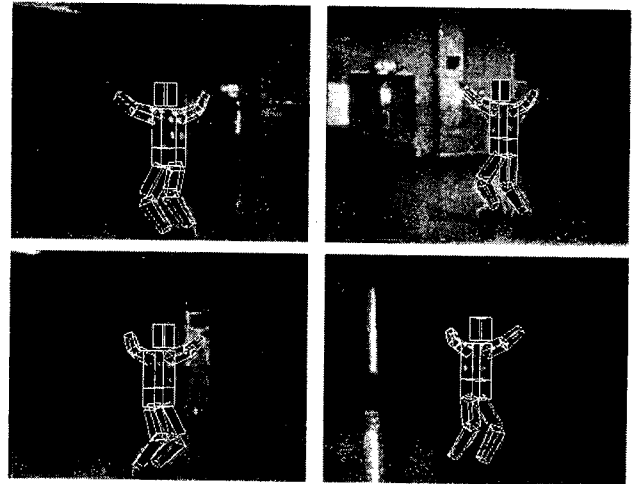


図2 4方向からのカメラ画像と人体モデル(両足で飛ぶ運動)
Fig.2 Images from four cameras, on which the body model overlaps.

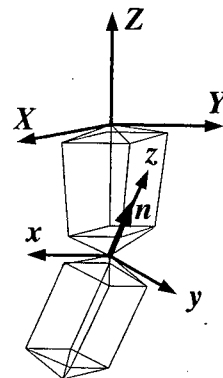


図3 パーツの姿勢
Fig.3 Part pose.

下腕, 上脚, 下脚10パーツの姿勢を使用する.

動画像解析から得られる姿勢は, ロボットモデルの腕や脚などの各パーツ関節リンクの親パーツに対する x, y, z 軸の回転角である. 各パーツの x, y, z 軸は図1に示すとおりで, 体軸方向を z 軸としている. 本追跡方法では, 画像上で腕, 脚のパーツが描かれている領域が小さいため, これらパーツのねじれ方向である z 軸回りの回転角を測定するのが難しく, 同じ姿勢でも異なる回転角で表される場合がある. この姿勢のあいまいさを回避するため, 図3のように親パーツ座標系 (X, Y, Z) に対する各パーツの z 軸のベクトル成分値 (n_x, n_y, n_z) で姿勢を表現する. φ, ψ をそれぞれ x, y 軸回りの回転角とすると, 回転角から z 軸のベクトル成分値への変換式は次のようになる.

$$\begin{cases} n_x = -\sin \psi \\ n_y = \cos \psi \sin \varphi \\ n_z = \cos \psi \cos \varphi \end{cases} \quad (1)$$

例として図4に, 動画像追跡を行って求めた, ラジオ

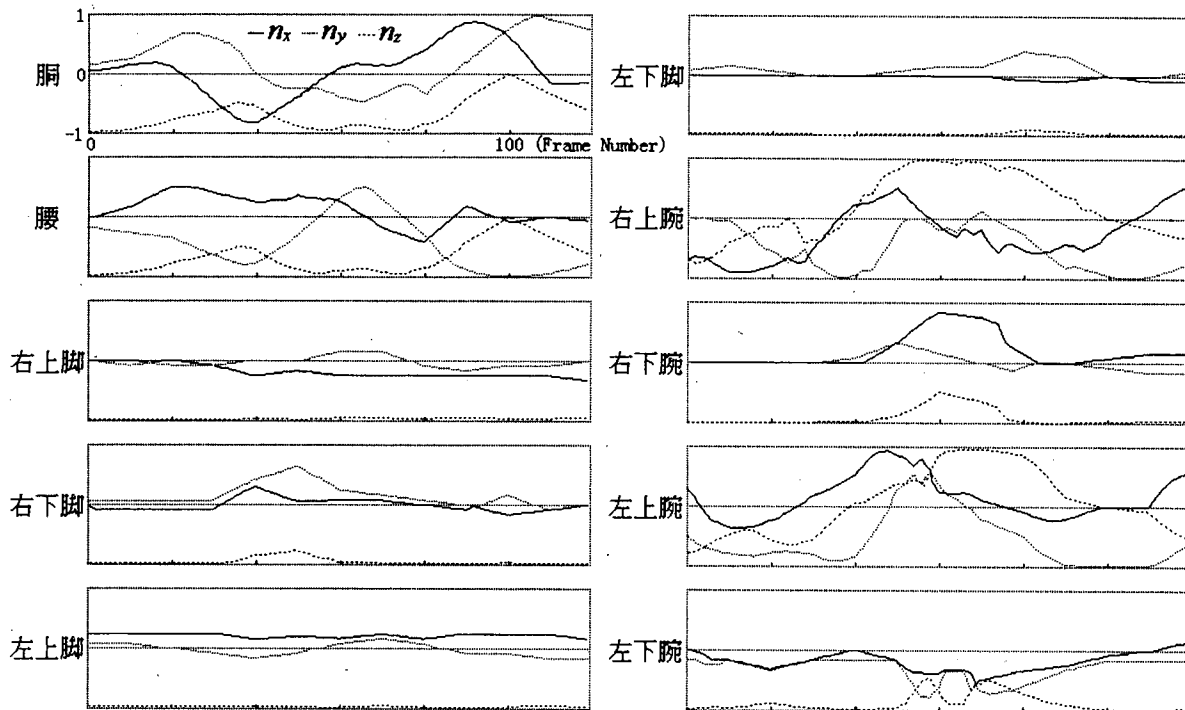


図4 姿勢を表す単位体軸ベクトルの成分値 (体を回す運動)
Fig. 4 Components of unit vectors representing parts pose.

体操の「体を回す運動」での胸、腰、左右上腕、下腕、上脚、下脚 10 パーツの z 軸のベクトル成分値を示す。

3. 動作の固有空間表現

人体は、胸、腰、腕、脚など多くのパーツから構成されるため、姿勢パラメータの数は多くなる。もし、この多くの姿勢パラメータを低次元の特徴空間上に表現することができれば、動作の識別は容易になる。ここでは、教示動作と未知の動作を、固有空間法により低次元空間上に表現する方法について述べる。

胸、腰、腕、脚は複雑な動きをするが、これらは互いに関連して動くと考えられる。そこでこの性質を利用し、KL 展開を使って姿勢パラメータ数を圧縮することを考える。

最初に教示動作から固有空間を計算する。各フレームでの姿勢パラメータを要素とするベクトルを

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T \quad (2)$$

と表す。ここで i は画像のフレーム番号を、 m は姿勢パラメータの総数を表す (図4の例では $m = 30$)。各値は式 (1) により $-1 \sim 1$ に正規化されている。教示動作における人体の姿勢の平均 \mathbf{c}

$$\mathbf{c} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad (3)$$

を計算する。 N は画像の総フレーム数を表す。教示動作の各姿勢から姿勢の平均を差し引いた行列 \mathbf{X}

$$\mathbf{X} = [\mathbf{x}_1 - \mathbf{c}, \mathbf{x}_2 - \mathbf{c}, \dots, \mathbf{x}_N - \mathbf{c}]^T \quad (4)$$

を作る。この姿勢データ列 \mathbf{X} の共分散行列 \mathbf{Q} は

$$\mathbf{Q} = \mathbf{X}\mathbf{X}^T \quad (5)$$

で与えられる。 k 次元の固有空間は次の固有方程式

$$\lambda_i \mathbf{e}_i = \mathbf{Q}\mathbf{e}_i \quad (6)$$

を解き、 m 個の固有値の中から大きい順に k 個選んだ固有値 ($\lambda_1 \geq \dots \geq \lambda_k$) に対する固有ベクトル ($\mathbf{e}_1, \dots, \mathbf{e}_k$) を基底ベクトルとすることにより得られる。

次に、教示動作をこの固有空間上に射影する。教示動作の姿勢パラメータから姿勢の平均 \mathbf{c} を引いたベクトルを次式により固有空間上に射影する。

$$\mathbf{g}_i = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]^T (\mathbf{x}_i - \mathbf{c}) \quad (7)$$

各フレームにおける人体の姿勢は、固有空間上の 1 点に射影される。前後のフレームにおける姿勢は、その変化が小さいとすれば、空間内でも近い点に射影される。そのため、連続して変化する姿勢は、空間上の軌跡で表される。

ここで固有空間の次元を考える。固有値 m 個の総和のうち大きい方から n 個の固有値の占める割合すなわち累積寄与率 SC_n は

$$SC_n = \frac{\sum_{j=1}^n \lambda_j}{\sum_{j=1}^m \lambda_j} \quad (8)$$

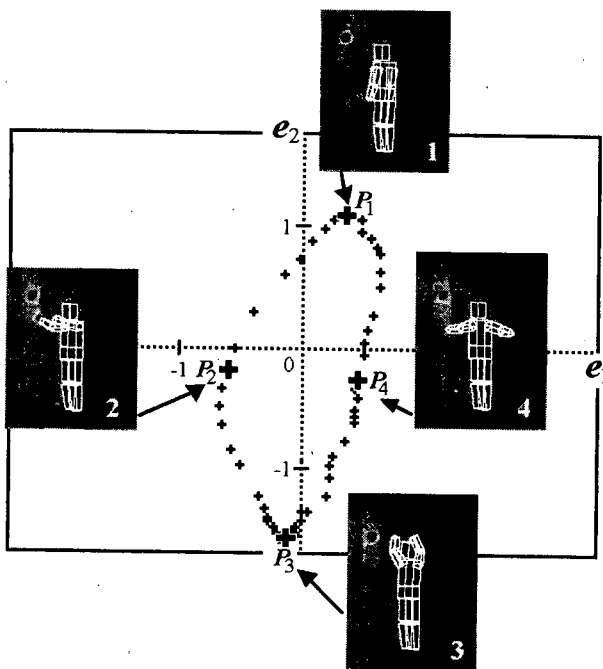


図5 固有平面上の動作(伸びの運動)
Fig. 5 Gesture on eigenplane.

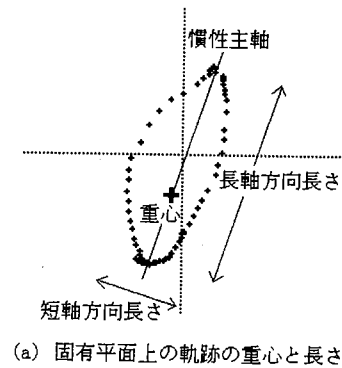
で与えられ、この値が固有空間の次元を決定する1つの目安となる。この値がある程度高ければ n 次元の固有空間でもとの動作を忠実に表すことができる。しかし、少ない次元であっても動作の特徴を失うことがなければ、識別計算を容易に行える。本論文では固有空間の次元をなるべく小さくし、かつ動作識別が正しく行える値として2次元固有空間上に動作を表現して識別を行った。この2次元空間を固有平面と呼ぶことにする。例として図5に固有平面上に表したラジオ体操の「伸びの運動」を示す。1~4の姿勢はそれぞれ $P_1 \sim P_4$ の点に射影され、動作は固有平面上の軌跡で表現される。

4. 識別方法

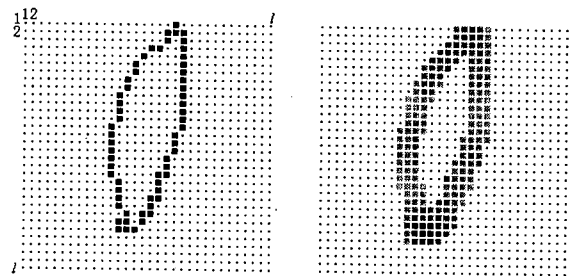
教示動作から得られた固有ベクトルを用いて、未知の動作を固有平面上に射影すれば、これも平面上の軌跡で表すことができる。すなわち、各フレームでの人体の姿勢を要素とするベクトルを x_i とし、先に求めた各姿勢の平均 c を引いたベクトルを次式により固有平面上に射影する。

$$g_i = [e_1, e_2]^T (x_i - c) \quad (9)$$

次に、教示動作がいくつかあるとき、これらの軌跡と未知の動作の軌跡を比較し、最も似ている教示動作を識別結果とする。ここでは、動作の軌跡は固有平面上での曲線となるため、動作の識別は文字認識の問題とまったく等価と考えられる。文字認識は古くから研



(a) 固有平面上の軌跡の重心と長さ



(b) 点列の曲線化と重みづけ

図6 固有平面上の軌跡の識別
Fig. 6 Gesture trajectory on eigenplane.

究され多くの手法が提案されており、そこでは特徴ベクトルの選択が重要となる。代表的な文字認識手法として次に示す2通りの方法で特徴ベクトルを選び、識別を試みる。

4.1 特徴ベクトルの選択

(1) 重心と長さを特徴ベクトルとする方法

図6(a)に示すように固有平面上の軌跡の特徴として、その重心の位置と慣性主軸の長軸方向の長さや短軸方向の長さを求め、これらを特徴ベクトルの要素とする。この特徴空間の各軸は、重心の位置と長さという種類の異なる要素となるため、前処理として各特徴軸を標準偏差で正規化し、平均回りの分散を等しくしておく。

(2) 部分空間の固有ベクトルを特徴ベクトルとする方法

文字認識の代表的な手法である部分空間法⁸⁾を適用し、この部分空間を構成する固有ベクトルを特徴ベクトルとする。具体的には、教示動作の曲線を図6(b)左図のように $l \times l$ のメッシュで区切り、 l^2 次元のベクトルで表現してもう一度KL展開し、固有値の大きい方からいくつかの固有ベクトルからなる部分空間上の1点に動作を射影し比較する。

部分空間法による文字認識では文字の大きさと位置の正規化を行うが、動作識別では固有平面上での軌跡の位置がある程度変われば姿勢も変わるため、大きさと位置の正規化を行わない。しかし、軌跡の位置が

微少に変わる場合はほぼ同じ姿勢となるため、正規化を行わないことで誤認識が起きてしまう。そこで、パターンの線を太くして比較する。また、固有平面上の点列を結んで曲線にし、単純に線が通るところを1、通らないところを0にするだけでは、点の分布のかたよりといった動作種類に対応する特徴の1つが失われてしまう。そのため、図6(b)の右図のように点の分布密度によって重みづけして線を太くする方法でも識別を行う。

4.2 線形判別法

固有平面上の軌跡として表された動作を、前述した特徴ベクトルからなる空間上の1点に射影し、動作の識別を行う。教示動作を複数人の動作から構成するため、次に述べる線形判別法により未知の動作の識別を行う。

特徴空間の次元を d とし、空間上に2クラスのパターンが分布していると仮定する。この2クラスを識別するための最適な1次元軸をフィッシャーの線形判別法で求める⁹⁾。

各クラス内の変動が小さく2クラス間の変動が大きいことを、2クラスをよく分離するための評価基準とする。クラス ω_i ($i = 1, 2$) の変動を表す行列、変動行列を S_i

$$S_i = \sum_{y \in y_i} (y - m_i)(y - m_i)^T \quad (10)$$

と定義する。ここで、 y はクラス ω_i に属する特徴ベクトル、 m_i はその平均である。次に、クラス内変動行列 S_W とクラス間変動行列 S_B を

$$S_W = \sum_{i=1,2} \sum_{y \in y_i} (y - m_i)(y - m_i)^T \quad (11)$$

$$S_B = \sum_{i=1,2} n_i(m_i - m)(m_i - m)^T \quad (12)$$

とする。ここで m は全パターンの平均、 n_i はパターン数を表す。 d 次元特徴空間から1次元空間への変換を表す $d \times 1$ 行列を A とする。変換後のクラス内変動 \tilde{S}_W とクラス間変動行列 \tilde{S}_B の比を $J_S(A)$ と表すと

$$J_S(A) = \frac{\tilde{S}_B}{\tilde{S}_W} = \frac{A^T S_B A}{A^T S_W A} \quad (13)$$

となり、これを最大にする A を求める問題は、

$$\tilde{S}_W = A^T S_W A = 1 \quad (14)$$

という制約条件の下で

$$\tilde{S}_B = A^T S_B A \quad (15)$$

を最大化する変分問題に帰着する。これを解くと、行列 $S_W^{-1} S_B$ の最大固有値に対応する固有ベクトルが J_S を最大にする A となる。以上の計算で得られる

のは2クラスの境界の法線ベクトルのみである。判別のための境界は、求めた法線ベクトルに垂直で、変換後のクラス平均の中点を通り、2クラスの重心を結ぶ直線とのなす角が最も大きくなる平面とする。多クラスの場合は、その数だけ境界を求め、どのクラスに属するか判別する。

5. 識別実験

5.1 実験条件

本手法の有効性を検証するため、実際の動作を使って識別実験を行った。識別対象の動作は、ラジオ体操第一の9種類の運動(1. 伸びの運動, 2. 腕を振って足を曲げ伸ばす運動, 3. 腕を回す運動, 4. 体を横に曲げる運動, 5. 体を前後に曲げる運動, 6. 腕を上下に伸ばす運動, 7. 体を斜め下に下げる運動, 8. 体を回す運動, 9. 両足で飛ぶ運動¹⁰⁾)で、被験者は、年齢が20~40代の8人の男性A~Hであった。

画像は8ビットの白黒、画像の大きさは 320×240 ピクセルで、4方向に設置したカメラで撮影した画像上で追跡行う。ラジオ体操での各動作の1サイクル時間は運動の種類と被験者によって異なるが、48~200フレームであった。実験に使用した動作の数は88で、これら動作の画像総数は約8,000フレームであった。9種類の動作を追跡した結果を図7に示す。部分空間法による識別では、メッシュの大きさを 32×32 とした。

5.2 実験結果と考察

被験者8人を教示動作グループ(被験者A~F)と未知動作グループ(被験者G, H)の6人対2人に分ける。そして、教示動作の人数を、(1)被験者A~Dの4人、(2)被験者A~Eの5人、(3)被験者A~Fの6人の場合で、未知動作グループ2人による9種類の18動作を識別した。図8は、教示動作の人数が5人のときの、被験者Aによる固有平面上の9種類の動作を表している。また、図9は、全被験者8人による動作1「伸びの運動」の固有平面上の動作を表している。表1には固有平面上の軌跡の重心と慣性主軸の長軸方向長さ、短軸方向長さを特徴ベクトルとして識別した結果を示し、表2には部分空間法で重みなしとありによる識別結果を示す。これらの表で、たとえば認識率88.9%とは、18動作のうち16が正しく認識されたことを意味する。

表1と表2の実験結果より、固有平面上に表された動作の軌跡を比較することで、高い認識率を得ることができた。また、どの方法も教示動作を構成する人数を増やすと認識率は上がった。しかし、教示動作と未



図7 ラジオ体操の追跡結果 (1行目:伸びの運動, 2行目:腕を振って足を曲げ伸ばす運動, 3行目:腕を回す運動, 4行目:体を横に曲げる運動, 5行目:体を前後に曲げる運動, 6行目:腕を上下に伸ばす運動, 7行目:体を斜め下に下げる運動, 8行目:体を回す運動, 9行目:両足で飛ぶ運動)

Fig. 7 Tracking results of the radio gymnastic exercises.

知の動作の人を入れ替え様々な組合せで識別を行ったところ, 教示動作を構成する人数を増やしても認識率が上がらない場合もあった. たとえば, 非常に個性の強い被験者 D を教示動作の 5 人目あるいは 6 人目とすると, 人数を増やしても少しだけ認識率が下がるという結果になった.

部分空間法による方法では, 識別に使用する空間の次

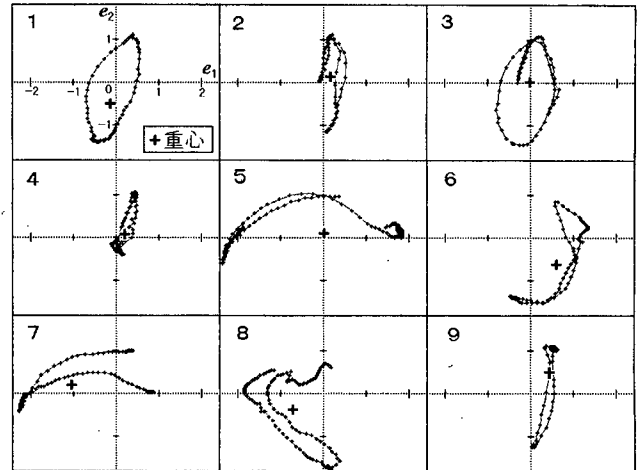


図8 固有平面上の9種類の動作 (被験者 A)

Fig. 8 Nine classes of gesture on the eigenplane in the radio gymnastic exercises.

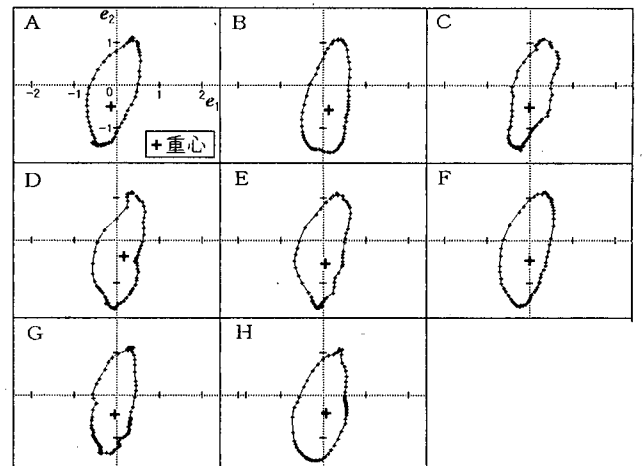


図9 固有平面上の8人の動作 (伸びの運動)

Fig. 9 One gesture of eight persons on the eigenplane.

元を, 累積寄与率がおおよそ 60%となる 4次元, 80%となる 13次元, 90%となる 22次元でそれぞれ認識率を求めた. 次元数を上げると認識率も上がるが, 空間の次元を増やす場合は, 教示動作の数もそれに伴って増やす必要がある. 今回, 次元が 13と 20の場合は教示動作の数がこれら次元数を上まわらないため, その認識率は参考値とした.

重みづけの有無で認識率を比較すると, 重みづけを行った方が認識率は高くなった. 動作 1と動作 3, 動作 2と動作 4と動作 9, 動作 5と動作 7は, 図 8に示すように固有平面上での軌跡の形が似ており, 形だけで識別を行うと誤認識がおきる. しかし, 動作によって点列の分布密度が異なるため, 分布密度で重みづけを行ってから識別を行うと認識率は改善された.

表 1に示す実験 (軌跡の重心と長さで比較) では, 教示動作が 6人のとき 100%という認識率が得られた.

表1 認識率 (%) (重心と長さ)

Table 1 Recognition result (gravity and length).

教示動作を構成する人数 (人)		
4	5	6
88.9	94.4	100.0

表2 認識率 (%) (部分空間法)

Table 2 Recognition result (subspace method).

	次元	教示動作を構成する人数 (人)		
		4	5	6
重みなし	4	72.2	77.8	83.3
	13	83.3	83.3	83.3
	22	83.3	83.3	83.3
重みあり	4	83.3	83.3	88.9
	13	83.3	88.9	88.9
	22	88.9	88.9	94.4

しかし、ここでは9種類の動作の1試行で識別を行っており、同一被験者が同じ動作を複数試行した場合、人体の動きがまったく同じになるとは限らない。そこで、同一被験者が同じ動作を複数回行った場合でも正しく識別されるか確認する実験を行った。対象動作は、固有平面上での軌跡の形が似て誤って認識されやすい動作1と動作3を選び、被験者G, Hがそれぞれ4試行し計16動作を識別した。図10に、被験者Gが動作1と動作3をそれぞれ5回行ったときの固有平面上の動作を示す。軌跡の形はまったく同じとはならないが、非常に似ていることが分かる。また、同じ動作を異なる被験者が行ったときにくらべ、同じ被験者が複数回行ったときの動きのばらつきは小さいことが予測される。教示動作を構成する人数を6人とし、軌跡の重心と長さで識別したところ16動作ですべて正しく識別された。

人体モデルの各パーツの姿勢パラメータをKL展開を使って圧縮し、動作を表す固有空間の次元を2次元とした。大きい方から2つの固有値の累積寄与率は、教示動作の人数が4人、5人、6人ではほぼ58%とそれほど高くなかった。しかし、認識率の高さは固有平面上でも十分に動作識別が行えるものと考えられる。

6. 他の方法との比較

本手法と同じく姿勢パラメータを使って動作識別を行うCampbellら⁶⁾の方法と比較を行った。Campbellらは、姿勢パラメータのうち特徴ある2つのパラメータから位相軌道と呼ばれる2次元平面上の軌跡で動作を表し、これに多項式をあてはめ教示動作との相関で識別を行った。一方、本手法では全身の姿勢パラメータを圧縮して固有平面上に動作を表し、文字認識手法

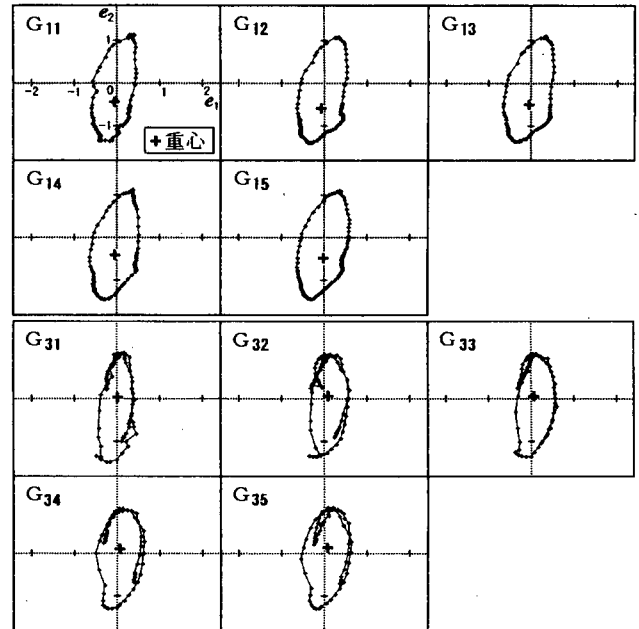


図10 被験者Gが同じ動作を5回行ったときの固有平面上の動作 (上が動作1「伸びの運動」で下が動作3「腕を回す運動」)
Fig. 10 Five gestures of one person on the eigenplane.

表3 認識率 (%)

Table 3 Recognition result.

	教示動作を構成する人数 (人)		
	4	5	6
右上腕	83.3	83.3	83.3
右上脚	33.3	33.3	33.3

を使って識別を行っている。両者の本質的な違いは前半の動作の表現方法であるため、後半の識別に関しては、文字認識手法(軌跡の重心と長さ)を使って識別を行った。ここでは、値の変化が最も大きい右上腕と最も小さい右上脚の体軸方向のベクトルのxとz成分値から位相軌道を構成し、それぞれについて識別実験を行った。表3に実験結果を示す。

右上腕から位相軌道を構成する場合、高い認識率が得られたが、右上脚から位相軌道を構成する場合の認識率は低い。位相軌道を構成する特徴パラメータの選び方によって認識率がかなり異なることが分かる。Campbellらの方法は一部の姿勢パラメータのみで識別を行うため、パラメータの選び方によって認識率が大きく変わってしまう。一方、本論文で提案した手法は全身の姿勢パラメータを反映した2次元平面上の軌跡で動作を表し識別を行うため、高い認識率が得られる。

7. おわりに

本論文では、人間の動作を固有平面上に表現し、動作の識別を行う手法について述べた。この手法を用い

て動作の識別実験を行い、その有効性を確認した。人間の動作は複雑ではあるが、2次元固有平面上の軌跡として表すことで識別は容易となった。また、古くから研究されている文字認識の手法を使って、高い認識率を得ることができた。

今後の課題としては、本手法を使いより多くの動作の識別を行うこと、同一動作を本手法で解析し、個人識別への応用を検討することである。また、動作の識別には、文字認識のような基準データがないため、動作識別のための基準データ作成も重要な課題である。

参考文献

- 1) 高橋勝彦, 関進, 小島浩, 岡隆一: ジェスチャー動画像のスポッティング認識, 信学論 (D-II), Vol.J77-D-II, No.8, pp.1552-1561 (1994).
- 2) 大和淳司, 倉掛正治, 伴野明, 石井健一郎: カテゴリー別VQを用いたHMMによる動作認識法, 信学論 (D-II), Vol.J77-D-II, No.7, pp.1311-1318 (1994).
- 3) 畠直志, 岩井儀雄, 谷内田正彦: 動き情報と情報圧縮を用いたロバストなジェスチャー認識手法, 信学論 (D-II), Vol.J81-D-II, No.9, pp.1983-1992 (1998).
- 4) Murase, H. and Sakai, R.: Moving Object Recognition in Eigenspace Representation: Gaint Analysis and Lip Reading, *Pattern Recognition Letters*, Vol.17, No.2, pp.155-162 (1996).
- 5) Boblick, A.F. and Wilson, A.D.: A State-Based Approach to the Representation and Recognition of Gesture, *IEEE Trans. Pattern Anal. and Mach. Intell.*, Vol.19, No.12, pp.1325-1337 (1997).
- 6) Campbell, L.W. and Bobick, A.F.: Recognition of Human Body Motion using Phase Space Constraints, *Proc. 5th ICCV*, pp.624-630 (1995).
- 7) 佐藤明和, 川田聡, 大崎喜彦, 山本正信: 多視点動画像からの人間動作の追跡と再構成, 信学論 (D-II), Vol.J80-D-II, No.6, pp.1581-1589 (1996).
- 8) Oja, E. (著), 小川英光, 佐藤誠 (訳): パターン認識と部分空間法, pp.5-23, 産業図書, 東京 (1986).
- 9) 石井健一郎, 上田修功, 前田英作, 村瀬洋: パターン認識, pp.114-121, オーム社, 東京 (1998).
- 10) 浜田靖一: ラジオ体操ハンドブック, ビネバル出版, 東京 (1994).
- 11) 山本正信, 川田聡, 近藤拓也, 越川和忠: ロボットモデルに基づく人間動作の3次元動画像追跡, 信学論 (D-II), Vol.J79-D-II, No.1, pp.71-83 (1996).

付 録

A.1 動作の姿勢データ列を求めるアルゴリズム

ロボットモデルの全関節の位置と姿勢を表すベクトル $q = (q_1, q_2, \dots, q_n)^T$ を用いて、人体モデル上の任意の位置ベクトル $p = (x, y, z)^T$ を、次のように表すことができる。

$$p = f(q) \quad (16)$$

f は一般に非線型であるが q の時間変化率 \dot{q} に対する速度ベクトル \dot{p} はヤコビ行列 $J(q)$ を使い、次の線形式で関係づけられる。

$$\dot{p} = J(q)\dot{q} \quad (17)$$

ここで \dot{q} はロボットモデルの動きを表すもので、運動パラメータと呼ぶ。ヤコビ行列 $J(q)$ はロボットモデルの幾何学的構造から求めることができる¹¹⁾。

時空間こう配法を用いた運動パラメータの推定は次のようにして行う。画面上の点 (X, Y) における時刻 t での明るさを $E(X, Y, t)$ とすると、画像上での各点の速度ベクトル (U, V) は、次の時空間こう配法の方程式に拘束されることが知られている。

$$E_X U + E_Y V + E_T = 0 \quad (18)$$

ここで、 (E_X, E_Y) は明るさの空間こう配で、 E_t は時間こう配である。

さて、3次元空間内の直交座標系 (x, y, z) に対し、視点を原点、光軸を z 軸、焦点距離を f とした中心投影カメラモデルを設定する。このとき、空間内の点 $p = (x, y, z)$ とその投影面上の点 (X, Y) との関係は次の式で与えられる。

$$\begin{cases} X = f \frac{x}{z} \\ Y = f \frac{y}{z} \end{cases} \quad (19)$$

点 p の3次元速度ベクトル $\dot{p} = (u, v, w)^T$ と、その投影点の速度ベクトル (U, V) との関係は、式 (19) を時刻 t で微分することで得られる。

$$\begin{cases} U = \frac{fu - Xw}{z} \\ V = \frac{fv - Yw}{z} \end{cases} \quad (20)$$

式 (20) を式 (18) に代入すれば、

$$G\dot{p} = -E_t \quad (21)$$

を得る。ここで、

$$G = \left(f \frac{E_X}{z}, f \frac{E_Y}{z}, \frac{XE_X + YE_Y}{z} \right)^T \quad (22)$$

とする。さらに式 (17) を式 (21) に代入すれば、

$$G^T J \dot{q} = -E_t \quad (23)$$

を得、この式が時空間こう配法による運動パラメータ \dot{q} の拘束式になる。

次に、ロボットモデル上の複数の点から運動パラメータの推定式(式(23))を求め、それらを連立させた次式

$$\begin{pmatrix} G_1^T J_1 \\ G_2^T J_2 \\ \vdots \\ G_n^T J_n \end{pmatrix} \dot{q} = \begin{pmatrix} -E_{t1} \\ -E_{t2} \\ \vdots \\ -E_{tn} \end{pmatrix} \quad (24)$$

を最小二乗法で解き運動パラメータ \dot{q} を推定する。カメラの数が複数の場合は、式(24)をカメラの数だけ導出し、これらをさらに連立させて推定する。そして、推定したパラメータの値だけロボットモデルを移動させたときの位置と姿勢が、次のフレームでの人体の位置と姿勢になる。すなわち、人体のあるパーツの推定した x, y, z 軸回りの運動パラメータを q_{x1}, q_{y1}, q_{z1} 、初期フレームのフィッティングで求めた姿勢を $\varphi_0, \psi_0, \phi_0$ とし、回転行列 $R(\varphi, \psi, \phi)$ を

$$R(\varphi, \psi, \phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \\ \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \\ 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \quad (25)$$

とすると、次の第1フレームでのパーツの姿勢 $\varphi_1,$

ψ_1, ϕ_1 との間に

$$R(\varphi_1, \psi_1, \phi_1) = R(\varphi_0, \psi_0, \phi_0)R(q_{x1}, q_{y1}, q_{z1}) \quad (26)$$

という関係が成り立つ。この式(26)を $\varphi_1, \psi_1, \phi_1$ について解くと、パーツの姿勢が得られる。また、この計算を繰り返すことによって、 i 番目のフレームにおける人体のパーツの姿勢が得られる。

(平成10年12月28日受付)

(平成11年6月3日採録)



大野 宏 (学生会員)

昭和60年新潟大学工学部電子工学科卒業。昭和62年同大学大学院工学研究科電子工学専攻修士課程修了。同年新潟県庁に入庁。平成元年より新潟県工業技術総合研究所(旧新潟県工業技術センター)に勤務し、画像処理、メカトロニクス等の研究開発に従事。現在、新潟大学大学院自然科学研究科博士後期課程在学中。電子情報通信学会会員。



山本 正信 (正会員)

昭和48年九州工業大学工学部制御工学科卒業。昭和50年東京工業大学大学院修士課程修了。同年、電子技術総合研究所入所。動画像処理、コンピュータビジョン等の研究に従事。平成元~2年カナダ国立研究協議会招聘研究員。昭和62年情報処理学会研究賞受賞。平成4年より新潟大学工学部情報工学科教授。工学博士。電子情報通信学会、IEEE各会員。