

## 多数カメラを用いた手形状認識法と その仮想空間インタフェースへの応用

内海 章<sup>†</sup> 大谷 淳<sup>†</sup> 中津良平<sup>†</sup>

多数カメラを用いた手形状認識手法とそれを利用した仮想空間操作システムについて述べる。従来の手形状認識システムでは、カメラと手のひらの位置関係によって生じる自己オクルージョンについて考慮されていないため、安定した認識結果を得るには利用者がつねにカメラ位置を意識しながら手振り動作を行う必要があった。本システムでは、我々の提案した多視点情報を用いた姿勢推定法による姿勢推定結果に基づき、手形状認識に適した画像を選択して認識処理を行う。これにより自己オクルージョンを回避することが可能となり、利用者はカメラ位置を意識する必要がない。手形状の記述にはP型フーリエ記述子を用い、その低域成分により認識器を構成した。P型フーリエ記述子による形状記述は、形状の画像内の平行移動、拡大・縮小等に対して不変であり、前述の視点選択と組み合わせることで手の姿勢によらず安定した形状認識が実現できる。本手法の有効性を確かめるため、13種類の手形状について判別実験を行い、良好な結果を得た。また、本手法のマンマシンインタフェースへの応用例として仮想空間操作システムを構築し、手形状の認識結果が操作コマンドとして利用可能であることを確認した。同システムでは、操作者が手形状の遷移により定義された6種類のコマンドにより、コンピュータグラフィクスで表示された仮想物体を操作できる。本システムにより、従来装置の装着を必要としていたジェスチャ型のインタフェースを非接触で実現することができる。

### Multiple-camera-based Hand Shape Recognition Method and Its Application to Virtual Space Interaction

AKIRA UTSUMI,<sup>†</sup> JUN OHYA<sup>†</sup> and RYOHEI NAKATSU<sup>†</sup>

We propose a hand shape recognition system using multiple-viewpoint images. Most conventional hand shape recognition systems were not concerned about the self-occlusion problem due to the hand rotation and a user needed to pay attention to his hand's direction to avoid the problem. We employ multiple-viewpoint images to estimate the pose of a human hand. After a pose estimation, a "best view" for the hand shape recognition is selected based on the estimation result. Hand shape recognition is performed based on the shape representation using P-type Fourier descriptor that is not affected with image scaling and translation. Based on the shape recognition, we developed a system where a user can create virtual graphical scenes interactively. In the system, a user can change the virtual objects' positions, sizes, colors, etc. with hand gestures. This system can be used as a user interface device, replacing glove-type devices and overcoming most of the disadvantages of contact-type devices.

#### 1. はじめに

直感的で使いやすいインタフェースとして、手振りを利用するシステムがこれまで数多く提案されている。これらのシステムの大半は、データグローブ<sup>TM</sup>に代表される特殊な装置の着用を利用者に求めるものであり<sup>1)~3)</sup>、着脱の繁雑さ等使用者の負担が大きく、広く普及するには至っていない。同様のシステムを非接触の装置で実現するために、画像処理により手振りを検

出する手法が提案されてきたが、速度や安定性の面で問題があった。我々は、これらの問題点を解決するため、多数のカメラを用いた手の位置・姿勢の推定法を提案している<sup>4)</sup>。この方法では、複数の視点から得られるロバストな特徴量により、手の形状変化にかかわらず位置・姿勢を安定かつ高速に求めることができる。これにより、求められた位置・姿勢に基づきその後の処理に適した画像を選択することができる。

本論文では、姿勢推定結果による視点選択を用いたフーリエ記述子による手形状検出手法を提案する。上坂<sup>5)</sup>によって提案されたP型フーリエ記述子はパターンを少ないパラメータで記述できる、平行移動・

<sup>†</sup> ATR 知能映像通信研究所  
ATR Media Integration & Communications Research  
Laboratories

拡大縮小について不変である，開曲線に適用できるといった形状認識に優れた特徴を持ち，これまでも部品形状の認識<sup>6)</sup>，文字認識<sup>7)</sup>，人間の横顔認識<sup>8)</sup>等に広く利用されてきた．P型フーリエ記述子の持つこれらの特徴は，輪郭線による記述が可能な手形状の認識においても有効であると考えられる．

一方で，このように2次元特徴により手形状を記述する場合，カメラの光軸に対する手の姿勢変化による撮影像の変形（自己オクルージョン）が大きな問題となる．従来行われてきた2次元特徴による手形状判別には，エッジマップによるもの<sup>9)</sup>，伸ばされた指の本数によるもの<sup>10),11)</sup>等があったが，いずれもカメラの光軸方向と手の姿勢について一定の関係を仮定しており，手が回転する場合の動作に関しては対応できていなかった．これに対し，手指の3次元モデルを用いて形状復元を目指す研究も行われてきたが<sup>12)~14)</sup>，これらの手法は精緻な姿勢情報が得られることが期待できる一方で，手指の関節の自由度が大きく計算コストが膨大になるという問題がある．また，オクルージョンを回避する検討もなされておらず，処理の前提となる画像特徴がオクルージョンにより検出できない場合，やはり推定が困難になると考えられる．

これに対し我々のシステムでは，前述の手姿勢推定法により推定した手の姿勢に基づいて形状認識に用いるカメラを選択することで，オクルージョンを回避し，安定に手形状認識を行うことができる．前述のP型フーリエ記述子の持つ特徴とあわせ，手の姿勢変化によらず安定な形状認識が可能となる．

本手法のマンマシンインタフェースへの応用例として，我々は手振りにより仮想空間内の物体配置を行う仮想空間操作システムを構築した．利用者は，手形状の遷移により定義されたコマンドにより，把持・移動，拡大・縮小等の仮想物体操作を行うことができる．我々は，このような非接触型のインタフェースにより，仮想空間とのインタラクションがより容易に行えるようになると考えている．

本論文では，次章（2章）において我々の提案した手の位置・姿勢推定法について概略を述べた後，3章でP型フーリエ記述子を用いた手形状推定法について説明する．4章では手姿勢推定法とフーリエ記述子による手形状推定法を組み合わせることにより安定に形状推定が行えることを実験により示す．5章では，本システムを用いたアプリケーションの一例として，手振りを用いた仮想空間操作システムについて述べる．6章で本論文をまとめる．

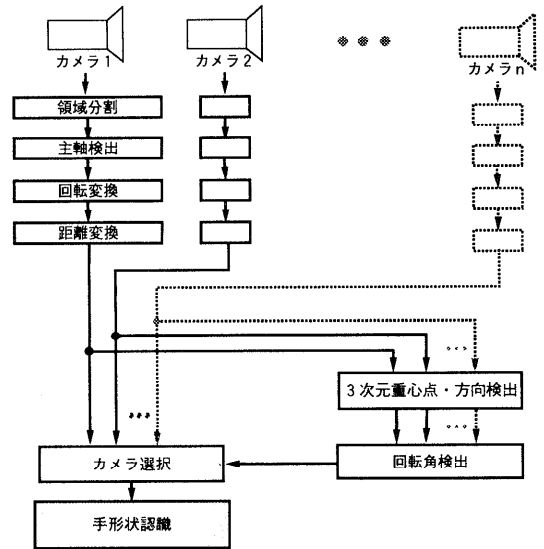


図1 システム構成図  
Fig. 1 System diagram.

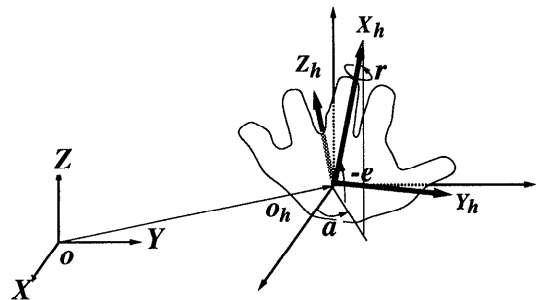


図2 手姿勢推定システムの座標系  
Fig. 2 Coordinate system for hand posture estimation.

## 2. 実時間手姿勢推定

本論文で検討する手振り認識装置の構成を図1に，姿勢推定に用いる座標系を図2にそれぞれ示す．図2において  $(X, Y, Z)$ ， $(X_h, Y_h, Z_h)$  はそれぞれ世界座標系，手のひら座標系を表す右手系直交座標で， $X_h$  は指先の方向（手方向）， $Z_h$  は手のひら法線方向をそれぞれ示す．図1に見られるように，人物の手を  $n$  台 ( $n \geq 2$ ) のカメラで観測することにより獲得される手の領域を含む各画像において，肌色領域を色情報および輝度情報を用いて手領域のシルエットとして抽出する．得られたシルエット画像にSobelフィルタを適用し，手領域の平均エッジ方向を求める（主軸検出部）．この平均エッジ方向を画像内の指先方向と見なす<sup>15)</sup>．以後の処理の安定化のため，得られた指先方向に基づき指先が上方を向くよう画像を回転変換した後で距離変換（各画素に領域境界からの最短距離を持たせる画

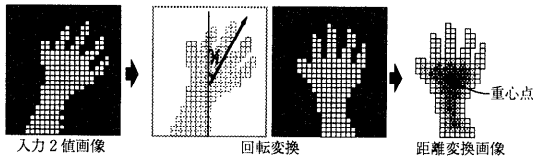


図3 重心点の検出(模式図)

Fig. 3 Detection of a center of gravity point.

像変換)を行う。このときのカメラ  $i$  の画像の回転角を  $\gamma_i$  とする(図3)。ここで、距離変換は2つの目的を持つ。1つは画像内の手のひら領域の重心点検出への利用であり、距離変換画像の極大点(スケルトン)を手の重心点として検出する。距離変換値の極大点は指曲げ等の形状変化に対して安定であるという、3次元復元にとって望ましい性質を持つ。3次元重心位置  $O_h$ (図2)は、各カメラで得られた重心点からステレオ法により決定する。

距離変換の持つもう1つの目的は3次元手方向まわりの回転角の推定である。3次元手方向  $X_h$  は、重心点同様に各カメラで得られた平均エッジ方向(主軸)よりステレオ法により決定する。これは、手の姿勢に関して手方向まわりの回転を除く2自由度(図2 a, e)を定めることを意味する。残る手方向まわりの回転角(図2 r)は、各画像の重心点(スケルトン)の持つ距離変換値から、あらかじめ定めた距離変換値の観測モデルを用いて最尤推定する<sup>4)</sup>。重心点の距離変換値は前述のとおり指曲げ等の形状変化の影響を受けにくく、回転角の推定も形状変化に対して安定である。以上により、手の位置・姿勢に関するすべての自由度が定まり、手方向  $X_h$  と回転角  $r$  から手のひらの法線ベクトル  $Z_h$  を定めることができる(図2)。以下の処理では、ここで求めた各カメラ画像に対する回転角  $\gamma_i$  と手のひら法線ベクトル  $Z_h$  を利用する。なお、3次元重心位置と3次元手方向を定めるためのステレオ対応には、前フレームでの手の法線ベクトル  $Z_h$  に近いカメラ軸を持つカメラ2台を用いる。これは、これらのカメラが正面に近い方向から手を観測しており、距離変換値を用いて安定した重心検出を行えることが期待できるためである。

今回実装したシステムの処理速度は手形状認識を含めて約10 Hzとなっている。

### 3. P型フーリエ記述子による形状認識

#### 3.1 P型フーリエ記述子による形状記述

手形状認識には、多数の視点で得られた画像のうちカメラの光軸が手のひらの法線軸に最も近い視点の画像を選択して用いる。すなわち、姿勢推定により求め

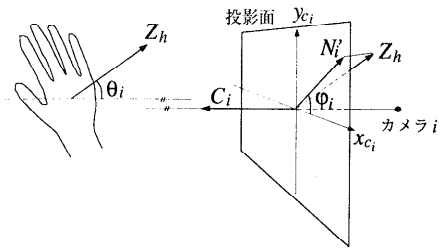


図4 選択視点への投影

Fig. 4 Hand image projection on a selected camera  $i$ .

られた手の法線ベクトルを  $Z_h$ 、 $i$  番目のカメラの光軸ベクトルを  $C_i$  とするとき(図4)、これらのなす角度  $\theta_i$  を最小にするカメラ  $i$  を選択する。

手指間のオクルージョンは、手のひらを正面から撮影する場合に最も起こりにくいと考えられるため(実験結果については次章で示す)、本視点選択によりオクルージョンの回避が期待できる。以下では、選択されたカメラの画像を用いた手形状認識手法について述べる。

2次元の曲線を複素平面上の点列と考え、長さの等しい線分からなる折線で近似する。折れ線の各頂点を  $z(j) = x(j) + iy(j)$  とするとき  $|z(j+1) - z(j)| = \delta, j = 0, \dots, n-1$ 、各折れ線を  $\delta$  で正規化して、折れ線のP表現  $w$  を得る。

$$w(j) = (z(j+1) - z(j))/\delta \quad (1)$$

$w$  の離散フーリエ展開により、フーリエ係数  $c(k)$  が次のように求まる。

$$c(k) = \frac{1}{n} \sum_{j=0}^{n-1} w(j) \exp\left(-2\pi i \frac{jk}{n}\right) \quad (2)$$

係数の集合  $\{c(k); k = -N, \dots, 0, \dots, N\}$  が  $N$  次のP型フーリエ記述子である。以下では、 $c(k)$  の大きさ  $|c(k)|$  を形状記述に利用する。

手形状をP型フーリエ記述子で記述するため、選択されたカメラ  $i$  の画像から手のひらの輪郭線を抽出する。ここでは、特に手形状を特徴づける手指部分について記述することを考える。前章で述べたように、入力2値画像は重心点検出の前処理として指先が上方を向くよう角度  $\gamma_i$  の回転変換を受けているため、手の領域内で重心点よりも上方部分のみの輪郭線を抽出することで手指部分の形状が得られる(図5)。ここで輪郭線の抽出は右回りとし、抽出した各画素の座標を  $(a(t), b(t))$  とする ( $t = 0, \dots, m-1$ ;  $m$  は輪郭線を構成する全画素数)。

さて、こうして抽出された輪郭線は、手の法線ベクトル  $Z_h$  とカメラの光軸ベクトル  $C_i$  のなす角  $\theta_i$  の

変化に応じて、投影による変形を受けており、このままでは形状記述にとって不都合である。そこで、得られた輪郭線を以下のように補正する。

視点選択の結果  $\theta_i \ll 90^\circ$  であると考え、弱透視変換を仮定すると、観測される画像は、手の法線ベクトル  $Z_h$  を撮像面に投影した  $N'_i$  の方向（画像内の水平軸  $x_{c_i}$  となす角を  $\varphi_i$  とする）に  $\cos\theta$  倍の縮小を受けていることになる。そこで、抽出した輪郭線の各画素の座標  $(a(t), b(t))$  を次式で補正し  $(a'(t), b'(t))$ ,  $t = 0, \dots, m-1$  とする。

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = R(\varphi_i + \gamma_i) \begin{bmatrix} 1/\cos\theta & 0 \\ 0 & 1 \end{bmatrix} R(-\varphi_i - \gamma_i) \begin{bmatrix} a \\ b \end{bmatrix} \quad (3)$$

補正後の輪郭線を折線近似し、式 (1) の P 表現で表す。なお、ここでは手の表裏の違いを特に考慮していないが、これについては次節で述べる識別処理部で対応する。図 6 に後述の実験に用いた入力画像の例とそれぞれの画像について検出された輪郭線を示す。

### 3.2 識別手法

式 (2) で得られるフーリエ係数のうち低次の  $2N+1$  個により次の特徴ベクトル  $V$  を定義する。

$$V = \frac{[|c(-N)| \cdots |c(0)| \cdots |c(N)|]'}{\sqrt{\sum_{k=-N}^{-N} c(k)^2}} \quad (4)$$

サンプル画像に基づき、各形状  $s$  ( $s$  は認識対象となる各形状を表す) について  $V$  の平均ベクトル  $\bar{V}_s$ 、共

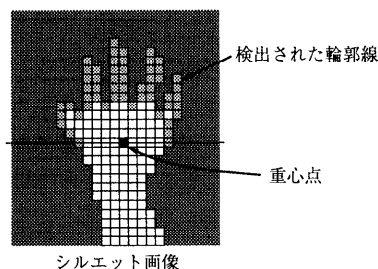


図 5 輪郭線抽出

Fig. 5 Contour extraction.

分散行列  $\Sigma_s$  を求める。形状  $s$  について  $V$  が観測される確率  $P(V|s)$  を次の正規分布で表現する。

$$P(V|s) = (2\pi)^{-(N+\frac{1}{2})} |\Sigma_s|^{-\frac{1}{2}} \exp\left(\frac{-(V - \bar{V}_s)\Sigma_s^{-1}(V - \bar{V}_s)}{2}\right) \quad (5)$$

これにより、与えられた画像に対する特徴ベクトル  $V$  を用いた形状の識別が上記確率  $P(V|s)$  を最大にする  $s$  を選択することにより可能となる。

ところで、本手法による画像選択においては手のひらの表裏の違いが考慮されておらず、そのため同じ手形状について表裏 2 種類の画像入力がありうる。このような輪郭線の左右の反転に対し、得られる P 型フーリエ記述子は次のようになる。

$$c'(k) = -\overline{c(-k)} \quad (6)$$

ここで、 $c'(k)$  は反転後の係数を示す。

そこで、手の表裏の反転を考慮し、特徴ベクトル  $V$  の成分を並べ替えた  $V^*$  を定義する。

$$V^* = \frac{[|c(N)| \cdots |c(0)| \cdots |c(-N)|]'}{\sqrt{\sum_{k=-N}^{-N} c(k)^2}} \quad (7)$$

形状識別においては、 $P(V|s)$ 、 $P(V^*|s)$  の両方を評価することで、入力画像の表裏の違いに対応する。すなわち、選択画像から抽出された輪郭線から  $V$ 、 $V^*$  を計算し、すべての形状  $s$  に対する  $P(V|s)$ 、 $P(V^*|s)$  のうち最大値をとる分布を選択、対応する  $s$  を識別結果とする。

## 4. 実験

### 4.1 実験画像の撮影と形状パラメータの決定

以下の実験のために、図 6 に示した手形状それぞれについて約 600 フレーム分の画像を撮影した。人間の手は変形の自由度が大きく、表現できる手形状の種類も多いため、それらすべてについて検討することは困難である。そこで、今回の実験では、独立に運動させることが難しい中指と薬指を除く、3 本の指曲げにより定義される 8 種類の手形状 (shape0~7, 中指・薬指は常時曲げる) に加え、指を  $n$  本伸ばした形状

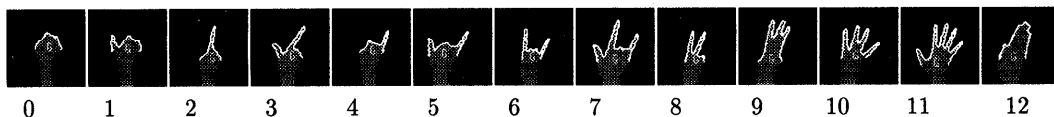


図 6 実験に用いた手形状についての輪郭線抽出例 (白線部分が輪郭線)

Fig. 6 Examples of extracted contours for 13 hand shapes used for experiments.

表1 最適カメラによる手形状認識結果  
Table 1 Results of shape recognition with best-view selection.

		入力形状												
		shape0	shape1	shape2	shape3	shape4	shape5	shape6	shape7	shape8	shape9	shape10	shape11	shape12
フレーム数		298	299	299	298	297	295	298	297	299	293	286	299	300
認識結果	shape0	297	6	0	0	26	0	0	0	0	0	0	0	0
	shape1	1	270	0	0	36	7	0	0	0	1	0	0	0
	shape2	0	0	288	6	4	0	1	0	45	9	1	0	0
	shape3	0	0	1	285	1	0	2	1	14	9	0	0	0
	shape4	0	23	2	1	215	6	0	0	6	0	0	0	1
	shape5	0	0	0	0	10	280	0	2	0	1	0	1	0
	shape6	0	0	2	1	0	0	279	3	1	5	0	0	0
	shape7	0	0	0	0	0	1	0	277	0	1	0	0	0
	shape8	0	0	5	1	0	0	1	0	213	26	1	0	0
	shape9	0	0	1	4	5	1	14	11	26	202	32	1	7
	shape10	0	0	0	0	0	0	1	3	0	33	250	13	0
	shape11	0	0	0	0	0	0	0	0	0	0	2	284	0
	shape12	0	0	0	0	0	0	0	0	0	0	0	0	292
認識率 (%)		99.7	90.3	96.3	95.6	72.4	94.9	93.6	93.3	71.2	68.9	87.4	95.0	97.3

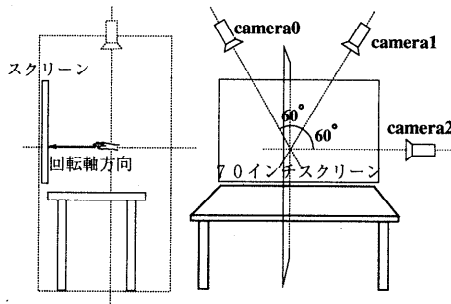


図7 実験環境  
Fig. 7 Experimental environment.

( $n = 0 \sim 5$ , shape0,2,8,9,10,11), すべての指を揃えた形状 (shape12) の13種類を認識対象とした。

実験に用いた環境を図7に示す。ここでは、同心円上に中央を向けて約60度間隔で3台のカメラを配置した。ここでカメラの配置を同心円上60度間隔としたのは、次章で述べるようなアプリケーションにおいて、手方向(回転軸)が図6のようにスクリーン方向を向くことが多いと考えられるためである。その場合、手の回転に対して有効に視点選択を行うにはカメラがスクリーンと平行な同心円上に等間隔に配置されていることが望ましく、今回の実験で用意した3台のカメラではその間隔は60度となる。

被験者には撮影中、検出範囲内でつねに手のひらの回転と移動を続けるよう求めた。このため撮影された画像は広い範囲の回転角と位置の変化を持つようになる。4.3節では、この角度変化と認識結果の関係を示す。なお、前述のとおり本システムの処理速度は約10 Hzであるため、本章の実験では撮影した全画像について形状パラメータを求められるよう、3台のカメラ

で得られた画像をいったん3台の追記型ビデオディスクに記録した後に、コマ送りにより1フレームごとに処理した。全フレームのうち300フレームを学習用に、残りを認識実験用に用いた。

パラメータ学習時には、各形状について初期フレームの画像から前章で述べた方法により特徴ベクトル  $V_0$  を抽出した。2フレーム以降については、 $V$ ,  $V^*$  のうち、ユークリッド距離が  $V_0$  に近いものを選んで記録した。形状ごとに得られた300サンプルの特徴ベクトルを基に前章で述べた確率分布を決定した。

#### 4.2 認識結果

認識実験用の各形状約300フレームの画像に対する形状認識結果を表1に示す。ここにみられるように、形状4, 8, 9, 10を除いていずれも90%以上、特に形状0, 2, 3, 12については95%以上という安定した認識結果が得られた。これにより、本手法による形状認識の有効性が示されたといえる。一方で、形状4および形状8, 9, 10について他と比べて認識率が低くなっているのは、これらの形状が回転による自己オクルージョンによる誤判別を起こしやすかったためと考えられる。誤判別の内訳をみると、形状4では形状0, 1, 形状8, 9, 10では形状2, 8, 9, 10というように、いずれも手の回転によって見た目の形状が近くなる手形状との誤判別が大部分を占めることが分かる。この点については、次節の実験結果とともに改めて考察する。なお、このようなオクルージョンによる誤判別は、視点数の増加により視点選択の機会を増やすことで回避できると考えられる。なお、このような個々の形状によるオクルージョンの増減と必要な視点数の関係については今後さらに検討していきたい。

### 4.3 視点選択の効果

ここでは、本手法で用いた視点選択の効果を確認するため、視点選択を行わず同一視点の画像を用いて形状認識を試みた。図8に13種類の手形状それぞれについての認識結果を示す。ここで、横軸は手姿勢推定で得られた手の法線軸とカメラの光軸のなす角度 $\theta$ を、縦軸はそれぞれのフレームについて得られた認識結果をそれぞれ表している。図8から明らかなように、すべての形状について、角度 $\theta$ の増加にともなって誤認識が増加している。前節の実験結果とあわせて考えることで、形状認識において視点選択が有効に作用していることが分かる。なお、形状8, 9, 10は他と比べてより小さい角度で誤認識を起す傾向が見られ、このことは前節の実験結果を裏付けている。このように形状認識が困難になる角度は形状によって異なるため、安定した認識に必要な視点数は一概には定めることができないが、今回用いた60度間隔の配置では、ほとんどの形状について90%以上の高い認識率を得ており、次に述べる仮想空間操作システムの操作も問題なく行うことができた。

### 5. 仮想空間操作システム

前章までに述べた手形状認識法を利用し、対話的に仮想空間を操作できる実時間システムを構築した。利用者は、手振りによってコマンドを発し、仮想物体の配置・削除・拡大縮小等の操作を行うことができる。コマンドは図6に示した13種類のうち7種類の手形状に対応させた。具体的には、各コマンドは表2に示すように2種類の手形状間の「形状遷移」を基本としており、形状11の提示後、他の形状(0, 2, 3, 6, 8, 12)を提示することでコマンドが実行あるいは開始される。

コンピュータグラフィクスで表示される仮想空間には、利用者の手の3次元位置を表す指示ポイントが描画されており、利用者は自分の手を動かすことにより操作対象となる仮想物体の内部にポイントを移動させることで、操作対象を直接指定することができる。

表2 コマンド一覧  
Table 2 Command list.

コマンド	形状遷移
把持・移動	11 → 0
拡大・縮小	11 → 2
削除	11 → 3
色・テクスチャ変更	11 → 6
分割	11 → 8
結合	11 → 12

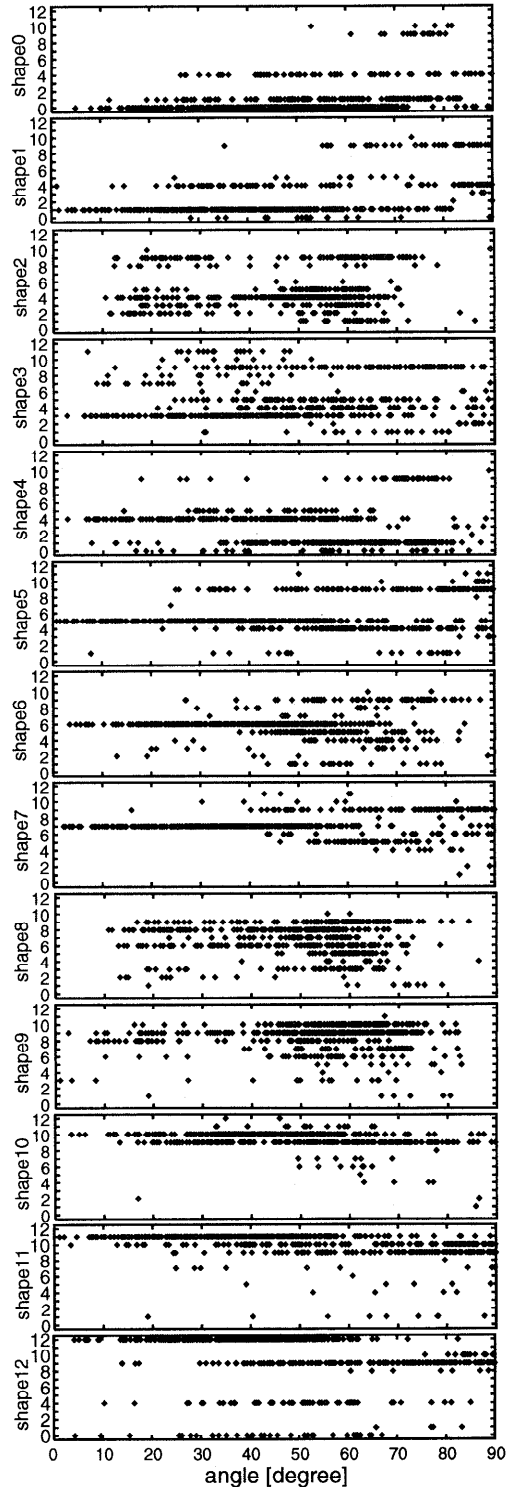


図8 手の回転角に対する手形状検出の安定性  
Fig. 8 Shape recognition stability versus hand rotation.



図9 システム動作例 (拡大・縮小)  
Fig. 9 System behavior example ('resize' command).



図10 システム動作例 (削除)  
Fig. 10 System behavior example ('delete' command).

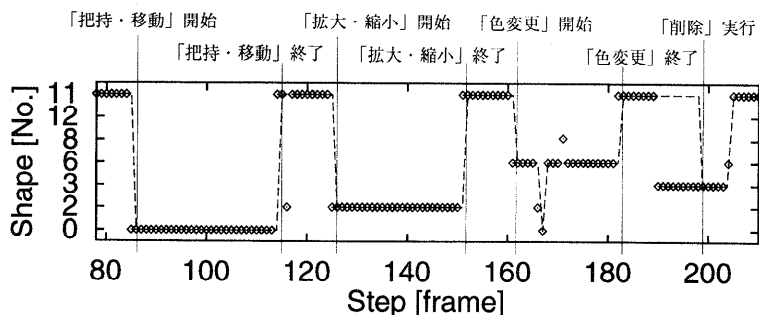


図11 コマンド認識例  
Fig. 11 Command recognition sequence (example).

「削除」(仮想物体を消去する),「結合」(近接する複数の仮想物体を結合する),「分割」(仮想物体の結合を解除する)の3コマンドについては,それぞれ操作対象内にポインタを移動させた後,各コマンドに対応する形状遷移を提示することで,当該コマンドが実行される。

「把持・移動」(仮想物体を掴んで移動させる),「拡大・縮小」(仮想物体の大きさを変更する),「色・テクスチャ変更」(仮想物体表面の色・テクスチャを変化させる)の3コマンドについては,ポインタによる指示,対応する形状遷移に続いて,さらに手を移動させることで,変更対象となる物体パラメータ(位置,大きさ,色等)が移動量に応じて変化する。パラメータの変更は,手形状を形状7に再び遷移させることで終了する。

図9, 図10にシステムの動作例を示す。図9では操作者が仮想物体の拡大を, 図10では消去を行って

いるところをそれぞれ示している。いずれの動作も形状7で物体内にポインタを移動し, コマンドに対応する手形状変化を与えることで処理が行われる。

コマンド実行時の形状認識の例(約150フレーム)を図11に示す。ここでは, 被験者が仮想物体の把持・移動, 拡大・縮小, 色・テクスチャ変更, 削除の各コマンドを順に実行したときの様子を示している。本システムでは形状の遷移時等に生じる誤認識を防ぐため, 過去3フレームのうち2フレーム以上で同一形状と認識されないと新しい手形状に遷移しないようフィルタ処理を組み込んでいる。また削除コマンドについては, 特に10フレーム連続の形状入力がないと遷移しないように実装した。図11で, 各点はそのフレームで認識された形状の番号を, 点線はフィルタ処理後の形状番号を示す。図内には, コマンドが実行されたフレームを記入している。一連の動作が安定して行われている様子が分かる。本システムにより, 形状認識による

インタフェースの実現が可能であることが示された。

## 6. ま と め

多数カメラを用いた手形状認識手法とそれを利用した仮想シーン生成システムについて述べた。本システムでは多視点情報による手姿勢推定結果に基づき、手形状認識に適した画像を選択し、手の輪郭線より得られるP型フーリエ記述子の低域成分により形状を判別した。検証実験により、P型フーリエ記述子による形状認識の安定性、視点選択の有効性を示した。本手法を組み入れた仮想空間操作システムでは、手形状の遷移により定義された6種類のコマンドにより、仮想物体の配置、連結、分割、削除、拡大縮小、色・テクスチャ指定を行え、本手法を用いた対話型システムの有効性を確認することができた。今後は、時系列の形状遷移による多様なジェスチャの認識について検討していく予定である。

謝辞 仮想オブジェクトをデザインして下さった榎沢順氏、システムの実装にご尽力下さった裏隠居宏、井村茂雄の両氏に深く感謝します。

## 参 考 文 献

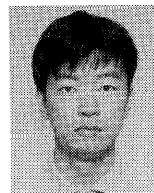
- 1) Takahashi, T. and Kishino, F.: Hand Gesture Coding Based on Experiments Using a Hand Gesture Interface Device, *SIGCHI Bulletin*, Vol.23, No.2, pp.67-74 (1991).
- 2) 吉田美寸夫, ジュリ A. ティヘリノ, 宮里 勉, 岸野文郎: 手振りと言語による仮想物体形状生成インタフェース, *テレビ誌*, Vol.50, No.10, pp.1482-1488 (1996).
- 3) Hanqiu, S.: Hand Interface in Traditional Modeling and Animation Tasks, *J. of Comput. Sci. & Technol.*, Vol.11, No.3, pp.286-295 (1996).
- 4) 内海 章, 宮里 勉, 岸野文郎, 大谷 淳, 中津良平: 距離変換処理を用いた多視点画像による手姿勢推定法, *映像情報メディア学会誌*, Vol.51, No.12, pp.2116-2125 (1997).
- 5) 上坂吉則: 開曲線にも適用できる新しいフーリエ記述子, *信学論 (A)*, Vol.J67-A, No.3, pp.166-173 (1984).
- 6) 伊藤豪俊, 平田達也, 石井直宏: フーリエ記述子を用いた部品の形状認識と処理, *信学論 (D)*, Vol.J71-D, No.6, pp.1065-1073 (1988).
- 7) 大友照彦, 原 健一: P型フーリエ記述子を用いたオンライン手書き漢字認識, *情報処理学会論文誌*, Vol.34, No.2, pp.281-288 (1993).
- 8) 相原恒博, 大上健二, 松岡 靖: 人間の横顔認識におけるP型フーリエ記述子の有効成分の個数について, *信学論 (D-II)*, Vol.J74-DII, No.10,

pp.1486-1487 (1991).

- 9) Uras, C. and Verri, A.: Hand Gesture Recognition from Edge Maps, *Proc. International Workshop on Automatic Face- and Gesture-Recognition*, pp.116-121 (1995).
- 10) 平川幹和子, 松尾英明, 高田雄二, 手嶋光隆: シルエット画像からの手形状認識, *テレビジョン学会技術報告*, VIS96-49, pp.15-20 (1996).
- 11) 長嶋ほか: 手の機能的な構造特徴量に基づく指文字認識, *テレビ誌*, Vol.48, No.8, pp.1503-1507 (1994).
- 12) Rehg, J.M. and Kanade, T.: Visual Tracking of High DOF Articulated Structures: An Application to Human Hand Tracking, *Computer Vision-ECCV '94, LNCS*, Vol.801, pp.35-46 (1994).
- 13) 中嶋正之, 柴 広有: 仮想現実世界構築のための指の動き検出法, *情報処理学会研究報告*, CG67-6, pp.41-46 (1994).
- 14) 岩井儀雄, 八木康史, 谷内田正彦: 単眼動画からの手の3次元運動と位置の推定, *信学論 (D-II)*, Vol.J80-D-II, No.1, pp.44-55 (1997).
- 15) 石淵耕一, 岩崎圭介, 竹村治雄, 岸野文郎: 画像処理を用いた実時間手振り推定とヒューマンインタフェースへの応用, *信学論 (D-II)*, Vol.J79-D-II, No.7, pp.1218-1229 (1996).

(平成10年6月5日受付)

(平成10年12月7日採録)



内海 章 (正会員)

1991年大阪府立大学工学部金属工学科卒業。1993年大阪大学基礎工学研究科情報工学修士課程修了。同年、ATR通信システム研究所に入社。画像処理、ヒューマンインタフェースの研究に従事。現在、ATR知能映像通信研究所研究員。



大谷 淳 (正会員)

1979年東京大学大学院精密機械工学専攻修士課程修了。同年電電公社(現NTT)電気通信研究所入所。以来、画像処理、カラー記録等の研究に従事。1988年より1年間、米国Maryland大学客員研究員。1992年(株)ATR通信システム研究所に出向。現在、(株)ATR知能映像通信研究所第一研究室長。仮想空間を介したコミュニケーション方式の研究に従事。工学博士。



**中津 良平（正会員）**

1969年京都大学工学部電子工学科卒業。1971年同大学大学院修士課程修了。同年日本電信電話公社（現NTT）武蔵野電気通信研究所入所。1980年横須賀電気通信研究所。主として音声認識の基礎研究，応用研究に従事。1990年NTT基礎研究所研究企画部長，1991年NTT基礎研究所情報科学研究部長。1994年よりATRに移り，現在（株）ATR知能映像通信研究所代表取締役社長。マルチメディア要素技術の研究およびマルチメディア技術を応用した通信方式の研究等に従事。工学博士（京大）。

---