

## 形状行列からの特徴選択に基づく動きの分割

市村 直幸<sup>†</sup> 石山 豊<sup>‡</sup> 角 保志<sup>†</sup> 富田 文明<sup>†</sup>

<sup>†</sup>電子技術総合研究所  
〒305 つくば市梅園 1-1-4  
ichimura@etl.go.jp

<sup>‡</sup>スタンレー電気(株)技術研究所  
〒225 横浜市青葉区荏田西 1-3-1

あらまし 本論文では、計測行列の特異値分解を通じ得られる形状行列からの、特徴の選択に基づく動きのセグメンテーション方法を提案する。提案方法では、特徴の選択を複数の物体の形状を表す線形空間の基底の生成として行う。各基底を物体上の特徴と対応付けて生成することにより、“セグメンテーション情報がない状態”で、“各物体”から特徴を選択できる。選択した特徴を用いることで計算量が減少し、セグメンテーションに必要な組み合わせ最適化が枚挙的に行える。その結果、数値計算的に極めて安定したセグメンテーションが行える。実画像を用いた実験により、提案方法の有効性について検討した。

キーワード 複数運動、セグメンテーション、特徴選択、因子分解法

## Motion Segmentation Based on Feature Selection from Shape Matrix

Naoyuki ICHIMURA<sup>†</sup>, Yutaka ISHIYAMA<sup>‡</sup>  
Yasushi SUMI<sup>†</sup>, and Fumiaki TOMITA<sup>†</sup>

<sup>†</sup>Electrotechnical Laboratory  
1-1-4, Umezono, Tsukuba  
Ibaraki, 305 Japan  
ichimura@etl.go.jp

<sup>‡</sup>Stanley Electric Co., Ltd., R&D  
1-3-1, Edanishi, Aobaku  
Yokohama, Kanagawa, 225 Japan

**Abstract** A motion segmentation algorithm based on the feature selection from the shape matrix is proposed in this paper. In the proposed method, the feature selection is carried out as a calculation of the bases of the linear space that represents the shape of the multiple objects. The features can be selected from “each” object “without segmentation information” by keeping the correspondence of the bases to the features. The computational cost is reduced by the feature selection; the combinatorial problem in the segmentation can be solved by the exhaustive manner. Thus the numerical computation of the proposed method is very stable. The experimental result for the real image is shown.

**key words** multiple motions, segmentation, feature selection, factorization method

## 1 はじめに

セグメンテーションは、コンピュータビジョンにおける基本的な処理の一つである。近年においては、動きからの3次元構造の復元 (structure from motion) を通じた物体モデリングや動画画像圧縮との関係から、動きを用いたセグメンテーションの必要性が高まっている。

動きを用いたセグメンテーションは、基本的には、対応付けにより得られる各時点における物体上の特徴の座標、つまり、特徴対応データを用いて行われ、これまでに数多くのアルゴリズムが提案されている。それらの中には、ハフ変換を原理とする投票方法 [2] や非階層クラスタリング [3]、確率場モデルに基づく方法 [4][5] などにより、動画画像から得られる運動パラメータの分割を行うものがある。これらの方法では、運動パラメータを用いてセグメンテーションを行うが、各物体に対する運動パラメータを精度良く求めるにはセグメンテーションを必要とする、という「鶏と卵」の問題がある [1]。

これらに対し、複数の物体の特徴対応データが混在したまま、直接セグメンテーションに必要な運動および形状パラメータを求める方法がある。このような方法として、線形重ね合せの原理から導かれる複数の物体のエピポーラ拘束方程式を用いる方法 [6] や、因子分解に基づく方法 [7] [8] [9] がある。これらの方法では、運動および形状パラメータを初期のセグメンテーションを行うことなく求められることから、上記の「鶏と卵」の問題が生じない。

本論文では、Costeira らによって提案された因子分解に基づく方法 [8][9] を取り扱う。この方法では、特徴対応データを要素とする計測行列の特異値分解を通じ、運動および形状パラメータそれぞれを要素とする運動行列 (motion matrix) および形状行列 (shape matrix) を求める。そして、形状行列より構成される形状相互作用行列 (shape interaction matrix) の行と列の入れ替えとして、セグメンテーションを行うとしている。

しかし、Costeira らの論文 [8][9] には形状相互作用行列の行と列の入れ替え方法、つまり、セグメンテーションの具体的な方法の詳細が示されていない。また、特徴の数が多くなった場合には、行と列の入れ替えの組合わせ数が大きくなるため、非線形最適化が必要と考えられる。非線形最適化を用いると、最終的なセグメンテーション結果が初期状態に依存することや、計算量が多いなどの問題が生じる場合がある。

本論文では、セグメンテーションに必要な少数の

特徴を形状行列を用いて選択し、それらより構成される形状相互作用行列を用いる方法を提案する。必要な数だけの特徴の選択により、形状相互作用行列の行と列の入れ替えの組合わせ数が、全特徴を用いた場合に比べ大幅に減少する。よって、非線形最適化は必要がなく、計算量が少なく、かつ、数値計算的に極めて安定したセグメンテーションを行える。

全データから少数のデータを選択することは、他の方法でも計算量削減の観点から有効である。例えば、クラスタリングを用いる方法では、ランダムサンプリングにより選択を行う。一方、提案方法では、複数の物体の形状を表す線形空間の基底の生成を通じ、特徴の選択を系統的に行う。一つ一つの基底ベクトルを、物体上の特徴に対応するように構成すれば、“複数の物体の特徴が混在した状態”においても、“各物体”から特徴を選択できる。つまり、特徴の選択においても「鶏と卵」の問題は生じない、という点でも提案方法は従来の方法と異なる。

以下、2. において複数の物体に対する因子分解法について述べる。次に、3. において提案方法の処理手順を述べ、4. においてその手順の中心である特徴の選択に基づくセグメンテーションについて説明する。そして、5. において実験結果を示す。最後に、6. において本研究のまとめを示す。

## 2 複数の物体に対する因子分解法

複数の物体に対する因子分解法について述べる。物体数を  $N$ 、各物体に属する特徴の数を  $P^i$ 、全特徴数を  $P = \sum_{i=1}^N P^i$ 、フレーム数を  $F$  とする。

### 2.1 計測行列の構成とセグメンテーション

各物体上の特徴の座標を行列  $D_s^i$  ( $i = 1, \dots, N$ ) で表す。

$$D_s^i = \begin{pmatrix} x_1^i & x_2^i & \dots & x_{P^i}^i \\ y_1^i & y_2^i & \dots & y_{P^i}^i \\ z_1^i & z_2^i & \dots & z_{P^i}^i \\ 1 & 1 & \dots & 1 \end{pmatrix} \quad (1)$$

$i$  番目の物体の  $j$  番目 ( $j = 1, \dots, F$ ) のフレームにおける運動パラメータを行列  $T_j^i$  で表す。行列  $T_j^i$  は、回転行列  $R_j^i$  と並行移動ベクトル  $t_j^i$  より構成される。

$$T_j^i = \begin{pmatrix} R_j^i & t_j^i \end{pmatrix}_{3 \times 4} \quad (2)$$

$i$  番目の物体の全フレームを通じての運動パラメータを、次の行列  $T^i$  で表す。

$$T^i = \left( T_1^i, T_2^i, \dots, T_F^i \right)^t_{3F \times 4} \quad (3)$$

この行列で表される特徴の運動を追跡し、特徴対応データを得る。

$F$  フレームに渡り追跡した  $P$  個の特徴の対応データは、行列  $D_s$  と  $T$  を

$$D_s = \text{diag}(D_s^1, D_s^2, \dots, D_s^N) \quad (4)$$

$$T = (T^1, T^2, \dots, T^N) \quad (5)$$

$$3F, P \geq 4N \quad (6)$$

とすると、次の行列  $W_s$  によって表現される。

$$W_s = T D_s \quad (7)$$

この行列では、最初から  $P^1$  個の列は物体 1 の特徴対応データ、次の  $P^2$  個の列は物体 2 の特徴対応データというように、セグメンテーション情報が入っている。

しかし、通常、このセグメンテーション情報は得られない。つまり、特徴の画像上での走査順番などの計測の状況に依存して、 $W_s$  の列の入れ替えが生じる。その影響を入れ替え行列 (permutation matrix)  $P_{col}$  で表す。また、便宜上、データの順番を  $x_1 y_1 z_1 x_2 y_2 z_2$  の形から  $x_1 x_2 y_1 y_2 z_1 z_2$  の形に変更する。その行の入れ替えを  $P_{row}$  で表す。入れ替えを行った特徴対応データを要素とする行列を  $W$  で表す。

$$W = P_{col} P_{row} W_s \quad (8)$$

この行列  $W$  を、計測行列と呼ぶ。式 (8) から、本研究で目的とする複数の物体のセグメンテーションは、入れ替え行列  $P_{col}$  を推定することと言える。

上記の計測行列の大きさは、3次元空間での追跡結果が得られる場合には  $3F \times P$  となる。この場合、各特徴の運動と位置は3次元計測の段階でわかるので、セグメンテーション情報のみを求める。

また、画像空間での追跡結果が得られる場合には、式 (1) の  $x_j^i$  と  $y_j^i$  に対応する  $W$  の最初の  $2F$  個の行が得られる。よって、計測行列の大きさは  $2F \times P$  となる。この場合、各特徴の運動と位置、さらに、セグメンテーション情報を求める。また、式 (6) の条件は、 $2F, P \geq 4N$  となる。

## 2.2 因子分解と運動および形状パラメータ

特異値分解により計測行列を次の形に分解する。

$$W = U \Sigma V^t \quad (9)$$

ここで、

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_P) \quad (10)$$

であり、その成分  $\sigma_i$  ( $i = 1, \dots, P$ ) は特異値である。計測行列のランクが  $r$  の場合、特異値の大きい順から  $r$  個以降は全て 0 となる。よって、計測行列は、

$$W = U \Sigma V^t \quad (11)$$

と表せる。さらに、上式は以下のように表せる。

$$W = U \Sigma^{1/2} A A^{-1} \Sigma^{1/2} V^t \quad (12)$$

$$= M S \quad (13)$$

$$M = U \Sigma^{1/2} A \quad (14)$$

$$S = A^{-1} \Sigma^{1/2} V^t \quad (15)$$

ここで、行列  $A$  は任意の正則行列である。

式 (13) は、計測行列を二つの行列に分解した結果であるので、セグメンテーション情報が消失していることを除いて、式 (7) と対応させることができる。両式より、式 (14) の行列  $M$  は運動パラメータ  $T$  に対応し、式 (15) の行列  $S$  は特徴対応データ  $D_s$  に対応する。この対応関係より、ここでは  $M$  と  $S$  を導く基となる行列  $U$  と  $V$  を、それぞれ、運動行列および形状行列と呼ぶ。

また、計測行列のランク  $r$  は  $4N$  に対応する。以後この対応関係が成り立つ場合を仮定し、計測行列のランク  $r$  を  $4N$  で表す。このことは、

物体数  $N$  は計測行列のランクより決定できる

という仮定と等価である。

## 2.3 因子分解とセグメンテーション

2.1 で述べたように、計測行列を用いた場合、複数の物体のセグメンテーションは、特徴に対応する計測行列の列の入れ替えと等価である。しかし、計測行列の列は形状行列の行と一対一対応なので、形状行列の行の入れ替えでもセグメンテーションは行える。

Costeira らは [8]、形状行列から構成される次の形状相互作用行列 (shape interaction matrix) を用いて、セグメンテーションを行った。

$$V_{P \times P}^{inter} = V V^t \quad (16)$$

形状行列の行は各特徴と対応する。それをかけた結果の  $P \times P$  の大きさの形状相互作用行列では、行と列が共に各特徴に対応する。よって、行と列の同時入れ替えがセグメンテーションに対応する。

形状相互作用行列を用いた場合の、正しいセグメンテーション結果の表現を調べておかなければ、セグメンテーションはできない。この説明のために、

まず、セグメンテーション情報が得られていると仮定する。この場合、式 (7) と式 (13) の対応関係より、式 (15) の行列  $S$  は式 (4) と同じブロック対角行列となる。その状態を  $S^*$  で表す。

式 (15) より、

$$V^{*t} = \Sigma^{*-1/2} A^* S^* \quad (17)$$

であることから、

$$\begin{aligned} V_{inter}^* &= S^{*t} A^{*t} \Sigma^{*-1} A^* S^* \\ &= S^{*t} \left[ \left( A^{*-1} \Sigma^{*1/2} V^{*t} \right) \cdot \right. \\ &\quad \left. \left( V^* \Sigma^{*1/2} A^{*-t} \right) \right]^{-1} S^* \\ &= S^{*t} \left( S^* S^{*t} \right)^{-1} S^* \end{aligned} \quad (18)$$

$S^*$  はブロック対角行列なので、明らかに式 (18) もブロック対角行列になる。式 (18) の要素  $v_{inter}^*(i, j)$  は次の性質を持つ。

$$v_{inter}^*(i, j) = \begin{cases} \neq 0, & i \text{ 番目の行と } j \text{ 番目の列に} \\ & \text{対応する特徴が同一の物体} \\ & \text{に属する} \\ = 0, & i \text{ 番目の行と } j \text{ 番目の列に} \\ & \text{対応する特徴が同一の物体} \\ & \text{に属さない} \end{cases}$$

以上から、形状相互作用行列に用いた場合、正しいセグメンテーション結果は、ブロック対角行列として表現される。

### 3 従来方法の問題点と提案方法の処理手順

#### 3.1 従来方法の問題点

2.3より、ある形状相互作用行列が与えられた場合、それをブロック対角化するような行と列の入れ替えにより、セグメンテーションが行える。しかし、それを実行する場合、次の2点に問題がある。

##### (1) セグメンテーションに用いる評価関数の定義

セグメンテーション情報が無いため、形状相互作用行列がブロック対角行列になった場合の各ブロックの大きさはわからない。この状態で、例えば、形状相互作用行列のブロック対角成分の和を評価関数とし、行と列の入れ替えを行うとする。しかし、図1に示すように、ブロック対角化の結果は各物体の特徴の数に依存して変わるため、評価関数が定義できない。

(2) 形状相互作用行列の行と列の入れ替えにおける数値計算

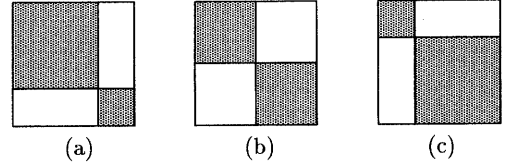


図1: 形状相互作用行列におけるセグメンテーション。ブロック対角化の結果は各物体の特徴の数に依存して、例えば上記の(a),(b),(c)のように変わる。

形状相互作用行列の大きさは  $P \times P$  であるので、行と列の入れ替えの組み合わせ数は、次式で求められる。

$$\binom{P}{2} = \frac{P(P-1)}{2} \quad (19)$$

例えば  $P$  を 1000 とした場合、この数は 499500 となる。この程度の数ならば、全ての組み合わせを尽くす、つまり、枚挙的に計算できる。しかし、枚挙的に計算を行うと、入れ替えの必要のない行と列に対しても評価関数を計算しなければならず、無駄な計算時間が多くなる。

また、枚挙的に計算ができないほど組み合わせ数が増えた場合、非線形最適化が必要となる。その場合、初期の形状相互作用行列の要素配置への依存性など、数値計算の安定性が問題となる。

#### 3.2 提案方法の処理手順

本論文では、上記の問題点を踏まえ、“セグメンテーション情報が無い状態”で、“各物体”から特徴を選択することを提案する。

各物体から既知の数の特徴を選択できれば、形状相互作用行列に含まれる各ブロックの大きさがわかる。よって、正しいセグメンテーション結果に対応する形状相互作用行列の要素配置が明確になる。よって、セグメンテーションの評価関数を定義できる。

また、少数の特徴のみからセグメンテーションを行うため、形状相互作用行列の行と列の入れ替えの組み合わせ数が、全特徴を用いた場合に比べ大幅に減少する。よって、枚挙的に計算を行っても、入れ替えの必要のない行と列に対する評価関数の計算量が減り、無駄な計算時間は減少する。また、枚挙的に計算が行える可能性が大きくなる。そのため、数値計算的に安定したセグメンテーションが行える。

提案するセグメンテーション方法の処理手順を図2に示す。提案方法では、まず、計測行列の特異値分解から、そのランクを推定する。そして、そのランクが  $4N$  に等しいことから物体数を推定する。その

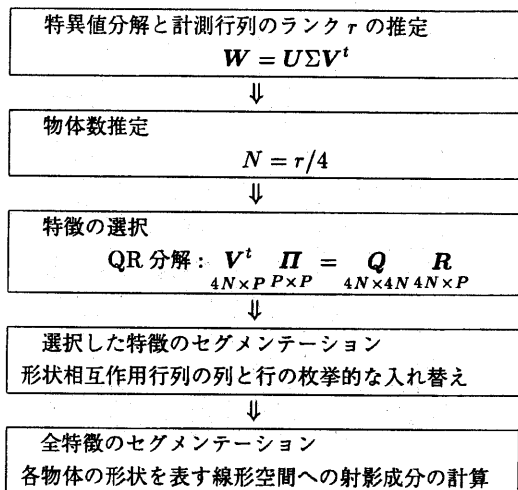


図 2: 提案方法の処理手順

後に、特徴の選択に基づくセグメンテーションを行う。特徴の選択以降の処理については、4. で詳しく説明する。

## 4 特徴の選択に基づくセグメンテーション

### 4.1 特徴の選択

大きさが  $P \times 4N$  の形状行列の列は、計測行列の行の正規直交基底である。つまり、複数の物体の形状をまとめて表現する線形空間の基底である。

計測行列のランクは  $4N$  になると仮定しているので、形状行列のランクも  $4N$  となる。よって、複数の物体の形状を表現する線形空間の基底の構成には、各物体 4 つ、合計で  $4N$  個の特徴があればよい。

以上より、選択する特徴の数の条件は「各物体から 4 つ以上」である。以下では、最低限必要な 4 つを選ぶとして話を進める。

選択する際の最も大きな問題は、セグメンテーション情報がない状態で、どのようにして各物体から 4 つの特徴を選ぶかである。

複数の物体の形状を表す線形空間の  $4N$  個の基底は、基底の条件を満たす限り、どのように構成してもよい。本論文では、一つ一つの基底ベクトルが物体上の特徴に対応するように構成することを提案する (図 3)。この場合、構成された基底は複数の物体の形状を表す線形空間の基底であり、それぞれが物体上の特徴に対応することから、それらは必ず 4 つずつ各物体に対応している。何故なら、もし、5 つ以上のベクトルが物体上の特徴に対応すれば、その

### Basis

[q1,q2,q3,q4,q5,q6,q7,q8]

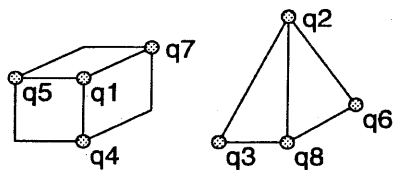


図 3: 特徴に対応した基底の生成。物体上の特徴と対応するように、複数の物体の形状を表す線形空間の基底を生成する。

物体の表現は冗長となり、他のいずれかの物体の表現が損なわれる。よって、全体として複数の物体の形状を正しく表現しないことになる。これは、基底の条件に矛盾する。

以上から、複数の物体の形状を表す線形空間の基底ベクトル一つ一つを、物体上の特徴に対応するように構成できれば、セグメンテーション情報がない状態で各物体から特徴を選択できる。

本論文では上記の処理を、形状行列の QR 分解 [10] により行った。QR 分解は次式で表される。

$$V^t \Pi = Q R \quad (20)$$

$4N \times P \quad P \times P \quad 4N \times 4N \quad 4N \times P$

ここで、行列  $Q$  と  $\Pi$  はそれぞれ、 $V^t$  の列の正規直交基底を列とする行列、および、 $V^t$  の列の入れ替え行列である。また行列  $R$  は、演算の結果得られる上三角行列である。 $Q$  は  $V^t$  の特徴に対応する列の正規直交基底であるので、複数の物体の形状を表す線形空間の基底である。

行列  $\Pi$  により、例えば  $V^t$  の列で次のような入れ替えが生じる。

$$V^t = (v_1, v_2, v_3, v_4, v_5, \dots)$$

↓

$$V^t = (v_{140}, v_{11}, v_{30}, v_{259}, v_{182}, \dots)$$

そして、上記の例で言えば、140,11,30,259,182 番目の特徴に対応して、 $Q$  の最初の 5 つの基底となる列が生成される。この段階では、これらの各基底がどの物体に対応するものかはわからないが、入れ替え行列  $\Pi$  により基底と特徴との対応はとられる。よって、 $\Pi$  により入れ替えが行われた  $4N$  個の列に対応する特徴は、必ず 4 つずつ各物体に対応する。

以上より、形状行列の QR 分解を通じて、セグメンテーション情報なしに、各物体から 4 つの特徴を選ぶことができる。

## 4.2 選択した特徴から構成された形状相互作用行列を用いたセグメンテーション

QR分解の結果、形状行列は次のように分割される。

$$V^t \Pi = \left( \begin{array}{c|c} V_{11}^t & V_{12}^t \\ \hline 4N \times 4N & 4N \times (P-4N) \end{array} \right) \quad (21)$$

選択した特徴に対応する  $V_{11}$  から形状相互作用行列を作る。

$$V_{11inter} = V_{11} V_{11}^t \quad (22)$$

この行列の大きさは元の特徴の数に依存せず、常に  $4N \times 4N$  である。よって、正しいセグメンテーション結果に対応する形状相互作用行列の要素配置は、図1(b)のように、 $4 \times 4$ の均等なブロックに分けられた形になる。よって、セグメンテーションの評価関数として、例えば次式のブロック対角成分の和を定義できる。

$$J = \sum_{i=0}^{N-1} \sum_{j=4i+1}^{4(i+1)} \sum_{k=4i+1}^{4(i+1)} v_{11inter}(j, k) \quad (23)$$

上式のような評価関数に従い、式(22)の行と列の入れ替えを、ブロック対角行列になるまで行う。行と列の入れ替えの組み合わせの数は、

$$\binom{4N}{2} = \frac{4N(4N-1)}{2} \quad (24)$$

となる。物体数を10とすると、この数は780となる。これは、十分に枚挙的に探索することのできる数である。仮に、枚挙的に探索できる限界が、式(19)で特徴の数を1000とした場合の数とする。すると、取り扱える物体数は  $1000/4 = 250$  個となる。このことから、特徴の選択により、物体数が実用上十分と思われる数百程度まで、枚挙的な行と列の入れ替えによりセグメンテーションが行えると考えられる。

実際の処理では、全ての可能な入れ替え方について評価関数を計算し、値が増加した組み合わせに対し入れ替えを行う。それを、ブロック対角行列になるまで繰り返す。この処理では、単に配列要素の和である評価関数を計算するだけなので、数値計算的に極めて安定したセグメンテーションが行える。

## 4.3 全特徴のセグメンテーション

選択した特徴を用いたセグメンテーションにより、式(21)の  $V_{11}^t$  の列が、各物体に対応する4つの特徴ごとに並ぶように入れ替えられる。その入れ替え

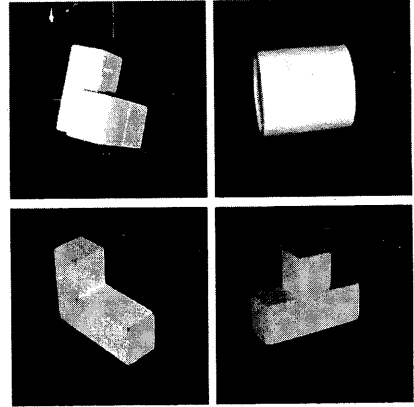


図4: 実験に用いた4つの物体。

に対応し、式(20)の行列  $Q$  の列も、次式のように各物体に対応した形に入れ替えることができる。

$$\tilde{Q} = \left[ \begin{array}{c|c|c|c} \tilde{Q}^1 & \tilde{Q}^2 & \dots & \tilde{Q}^N \\ \hline 4N \times 4N & 4N \times 4N & & 4N \times 4N \end{array} \right] \quad (25)$$

行列  $\tilde{Q}^i$  ( $i = 1, \dots, N$ ) の列は、各物体の形状を表す線形空間の基底である。この基底を用いて、次式より、全特徴のセグメンテーションを行う。

$$ObjNo(j) = \underset{i}{\operatorname{argmax}} \left( \left\| \tilde{Q}^i v_j \right\|^2 \right) \quad (26)$$

$$i = 1, \dots, N, j = 1, \dots, P$$

ここで  $ObjNo(j)$  は  $j$  番目の特徴へ割り当てられる物体番号、 $v_j$  は  $V^t$  の各列である。上式は、射影成分の最も大きな部分空間に対応する物体に、特徴を割り振ることを意味している。

以上で、特徴の選択に基づくセグメンテーションは終了する。

## 5 実験

提案手法の有効性を確認するために、物体数、各物体の特徴の数、および、物体の3次元形状が既知の場合について、実験を行った。

### 5.1 画像入力と追跡

画像入力は、ステレオカメラシステム、パタパタ[11]により行った。対象とした物体は、図4に示す4つである。これらの物体の追跡には、幾何モデルとの照合に基づく方法を用いた[12][13]。追跡の一例を図5に示す。物体の移動は人間の手により行った。追跡の結果、幾何モデル上のサンプル点(図7参照)を特徴とする特徴対応データが得られる。一つ一つの物体に対し図5のように追跡を行い、得られたデータ

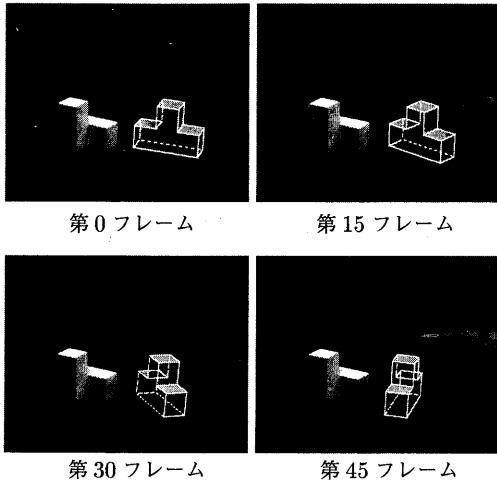


図 5: 物体追跡の一例. 文献 [13] より抜粋.

表 1: 特異値

順番	特異値	順番	特異値
1	130739	10	15.9511
2	35807.2	11	13.3373
3	313.306	12	11.8829
4	201.673	13	11.1419
5	47.1147	14	9.89105
6	37.6064	15	7.08395
7	23.8859	16	5.34026
8	21.7266	17	8.30671e-11
9	18.9427	18	3.5091e-11

をランダムに並べ替えることにより、計測行列を構成した。フレーム数は 50、特徴の数は各物体それぞれ、328、326、296、280 で、合計 1230 である。

3次元計測データ  $(x, y, z)$  の奥行き情報を使わず、 $(x, y)$  のみを用いた。これは、正射影モデルで投影された特徴を、画像平面上で追跡したことに等価である。計測行列の大きさは  $100 \times 1230$  となる。

## 5.2 セグメンテーションと形状復元の結果

計測行列の特異値分解の結果、表 1 の特異値を得た。最大特異値に  $10^{-6}$  をかけた数をしきい値として、それ以下の特異値は 0 と見做した。その結果、計測行列のランクは 16 となり、物体数は 4 と推定された。図 6 にセグメンテーション結果を示す。図 6(a) は、物体の位置関係と運動を表す。選択した 16 個の特徴から構成された形状相互作用行列 (図 6(b)) に対し、120 通りの組み合わせの行と列の入れ替えを行った結果、

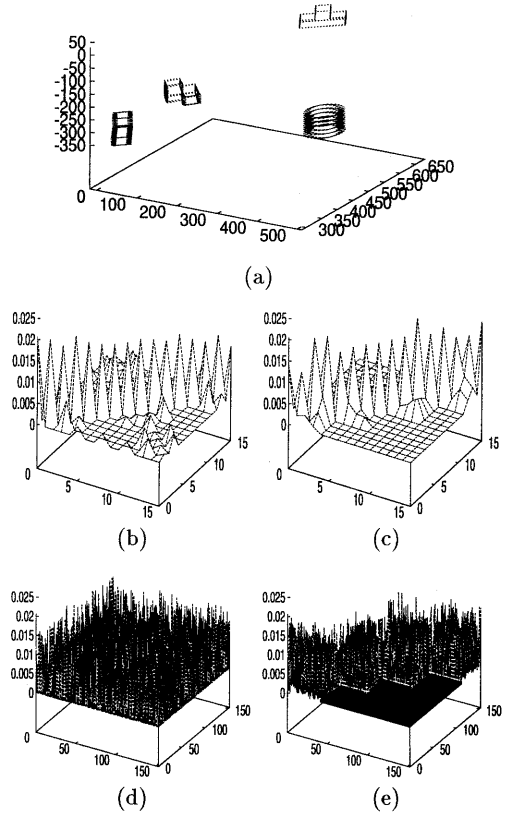


図 6: 2次元特徴を用いた場合のセグメンテーション結果。(a) 物体の位置関係と運動。(b) 選択した特徴の形状相互作用行列。(c) (b) のセグメンテーション結果。(d) 全特徴の形状相互作用行列。(e) (d) のセグメンテーション結果。(d) と (e) では行列の様子がわかりやすくなるように、8つごとに特徴を間引いて表示した。

ブロック対角行列となった (図 6(c))。このセグメンテーション結果を用いて、式 (26) より全特徴のセグメンテーションを行うと、その形状相互作用行列もブロック対角化された (図 6(d),(e))。

セグメンテーション結果を用いて、単一物体に対する因子分解法 [14] により物体形状を復元した結果を図 7 に示す。図 4 に示した 4 つの物体が正しくセグメンテーションされ、その形状が得られている。

図 7 の “Selected Points” は、選択した 4 つの特徴を結んだ直線である。ブロック状の物体では、ほとんどコーナーから、また、円筒ではエッジとなる上面と底面から、特徴が選択された。このような特徴が自動的に選択できることは興味深く、選択のメカニズムをより詳しく調べたいと考えている。

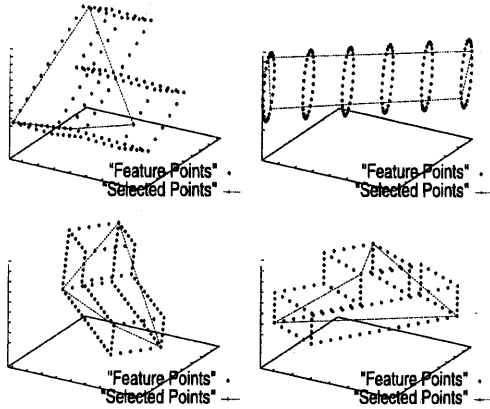


図 7: 物体形状の復元結果. 図 4 の各物体に対応している. "Selected Points" は, セグメンテーションで選択した 4 点を結んだものである.

## 6 むすび

本論文では, 形状行列からの特徴選択に基づくセグメンテーション方法を提案した. 提案方法では, 物体上の特徴と対応するように, 複数の物体の形状を表す線形空間の基底を構成した. この方法により, セグメンテーション情報がない状態で, 各物体から特徴を選択できた. そのため, 少数の特徴のみでセグメンテーションができ, 計算量が少なく, かつ, 数値計算的に安定した処理が実現できた.

本論文では, 物体数  $N$  は計測行列のランクより決定できると仮定した. この仮定は, 対象物体に平面や直線が含まれる場合には成り立たない. 例えば, 2つの物体が平面と 3次元物体であるなら, ランクは 7 になる. この場合でも, 特徴の選択には問題はない. しかし, 式 (23) の評価関数の定義が不明確となってしまう. もし, 平面と 3次元物体の組み合わせとわかるならば, 評価関数の定義の問題もなくなるが, 現実にはこの組み合わせがわかるとは限らない.

また, 本論文では, 特徴の選択に QR 分解を用いた. 実験により, QR 分解により選択される特徴が, 物体のコーナーやエッジに対応することを示した. 特徴の選択方法と選択される特徴の関係をより詳しく調べることは, 提案方法の適用範囲や特徴抽出そのものの意味に関わり重要である.

上記の 2 点が課題であり, 現在それについて検討中である.

謝辞 研究の機会を与えて下さいました電総研情報情報部, 大津展之部長に感謝致します. また, 日頃御討論頂く VVV プロジェクトの関係諸氏に感謝致します.

## 参考文献

- [1] 志沢 雅彦: "CVCV-WG 特別報告: コンピュータビジョンにおける技術評論と将来展望 (VIII) 運動からの 3 次元復元 - 複数運動の扱いを中心に -, " 情処学 CV 研報 99-9, pp.59-66, 1996
- [2] G. Adiv: "Determining three-dimensional motion and structure from optical flow generated by several moving objects," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.7, No.4, pp.384-401, 1985
- [3] B. G. Schunck: "Image flow segmentation and estimation by constraint line clustering," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.11, No. 10, pp. 1010 - 1027, 1989
- [4] D. W. Murray and B. F. Buxton: "Scene segmentation from visual motion using global optimization," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.9, No.2, pp.220-228, 1987
- [5] J. Konrad and E. Dubois: "Bayesian estimation of motion vector fields," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.14, No.9, pp.910-927, 1992
- [6] M. Shizawa: "Transparent 3D motions and structures from point correspondences in two frames: A quasi-optimal, closed-form, linear algorithm and degeneracy analysis," Proc. First Asian Conference on Computer Vision, pp.329-334, 1993
- [7] T. E. Boulton and L. G. Brown: "Factorization-based segmentation of motions," Proc. IEEE Workshop on Visual Motion, pp.179-186, 1991
- [8] J. Costeira and T. Kanade: "A multi-body factorization method for motion analysis," Technical report, School of computer science, Carnegie Mellon University, CMU-CS-TR-94-220, 1994
- [9] J. Costeira and T. Kanade: "A multi-body factorization method for motion analysis," Proc. IEEE Fifth International Conference on Computer Vision, pp.1071-1076, 1995
- [10] C. W. Ueberhuber: "Numerical Computation 2, Methods, Software and Analysis," Springer-Verlag, 1997
- [11] 河合 良浩, 石山 豊, 植芝 俊夫, 角 保志, 高橋 裕信, 富田 文明: "ステレオカメラシステム - バタバタ -, " 画像の認識・理解シンポジウム (MIRU'94) 予稿集, II, PP.127-134, 1994
- [12] 角 保志, 富田 文明: "ステレオビジョンによる 3次元物体の認識," 信学論 D-II, Vol.J80-D-II, No.5, pp. 1105-1112, 1997
- [13] 石山 豊, 角 保志, 富田 文明: "ステレオ視による位置姿勢の変化する 3次元物体の追跡," 情処学 CV 研報 95-3, pp.17-24, 1995
- [14] C. Tomasi and T. Kanade: "Shape and motion from image streams under orthography: a factorization method," International Journal of Computer Vision, 9, 2, pp.137-154, 1992