

形状行列からの特徴選択に基づく動きの分割

市村 直幸[†] 富田 文明[†]

Motion Segmentation Based on Feature Selection from Shape Matrix

Naoyuki ICHIMURA[†] and Fumiaki TOMITA[†]

あらまし 特徴対応データを用いる動きの分割は、与えられたデータをいくつかの部分集合に分ける組合せ問題である。本論文では、各物体から分割を行うのに最低限必要なデータを選択し、組合せ数を減少させる動きの分割方法を提案する。提案方法では、特徴選択を、複数の物体の形状を表す空間の基底の生成として行う。各基底を物体上の特徴と1対1に対応づ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ことを示す。特徴選択により組合せ数が減少するため、処理に数値最適化を必要としない。また、縮退した形状をもつ物体の識別や物体数の推定も行える。数値実験より提案方法の有効性について検討を行った。

キーワード 複数運動, 特徴選択, セグメンテーション, 因子分解法

1. まえがき

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、動画像圧縮におけるレイヤ表現 (layered representation) 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。

動きのセグメンテーションの方法は、これまでに数多く提案されている。それらの中には、ハフ変換を原理とする投票方法 [2], [3] や非階層クラスタリング [4], [5], 確率場モデルに基づく方法 [6], [7] などによりセグメンテーションを行うものがある。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、対応付けを通じて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る方法の中には、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及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 [8] や、因子分解に基づく方法 [9]~[12] がある。これらの方法では、運動及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 [11], [12] 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、運動及び形状パラメータそれぞれを要素とする運動行列 (motion matrix) 及び形状行列 (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) を構成する。セグメンテーションが正しく行われると、この形状相互作用行列がブロック対角化される。この性質から、セグメンテーションを形状相互作用行列の行と列の入替えとして行える。Costeira らは、セグメンテーションを行うため、すべての可能な行と列の入替えに関して、形状相互作用行列の非対角ブロック成分を最小にするように最適化を行うとしている [11], [12]。

しかし、物体数や特徴対応データの数が大きい場合、すべての可能な行と列の入替えの組合せ数は非常に大きくなる。このような組合せ最適化には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。

[†] 電子技術総合研究所, つくば市

Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba-shi, 305-8568 Japan

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 [13]~[15]。必要な数だけの特徴の選択により、組合せ数が全特徴を用いた場合に比べ減少する。よって、非線形最適化を用いずにセグメンテーションが行える。

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。これに対し、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、特徴選択を系統的に行う。一つひと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。

$$T_j^i = \begin{pmatrix} R_j^i & | & t_j^i \end{pmatrix} \quad (2)$$

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。

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

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 フレームにわたり追跡した P 個の特徴の対応データは、行列 D_s と T を

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

$$T_{3F \times 4N} = \left(T^1, T^2, \dots, T^N \right) \quad (5)$$

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

とすると、次の行列 W_s によって表現される。

$$W_s = \begin{matrix} T & D_s \\ 3F \times P & 3F \times 4N & 4N \times P \end{matrix} \quad (7)$$

この行列では、最初の 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 の列の入替えが生じる。その影響を入替行列 (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。その行の入替えを P_{row} で表す。入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。

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

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、入替行列 P_{col} を推定することといえる。

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、更に、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 運動及び形状行列

特異値分解により計測行列を次の形に分解する。

$$W_{2F(\text{or } 3F) \times P} = U_{2F(\text{or } 3F)} \Sigma_{r \times r} V^t_{r \times P}$$

$$= \begin{matrix} \mathbf{M} & \mathbf{S} \\ 2F(\text{or } 3F) \times r & r \times P \end{matrix} \quad (9)$$

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

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}^{1/2}\mathbf{A} \quad (11)$$

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

ここで、 σ_i ($i = 1, \dots, r$) は特異値であり、 r は計測行列のランク、 \mathbf{A} は任意の正則行列である。この行列 \mathbf{A} はカメラモデルより導出される計量拘束より定められる [16], [17].

式 (9) 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 (11) の行列 \mathbf{M} は運動パラメータ \mathbf{T} に対応し、式 (12) の行列 \mathbf{S} は特徴対応データ \mathbf{D}_s に対応する。この対応関係より、ここでは \mathbf{M} と \mathbf{S} を導くもとなる行列 \mathbf{U} と \mathbf{V} を、それぞれ、運動行列及び形状行列と呼ぶ。

また、計測行列のランク r は $4N$ に対応する。よって、物体数は形状行列のランクから求められる。しかし、この対応関係は物体形状が縮退している場合、つまり、シーンに平面や直線が含まれる場合には成立しなくなる。

2.3 形状相互作用行列

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替えと等価である。計測行列の列は形状行列の行と 1 対 1 対応なので、形状行列の行の入替えでもセグメンテーションは行える。

Costeira ら [11] は、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。

$$\mathbf{V}_{inter}^{P \times P} = \mathbf{V}\mathbf{V}^t \quad (13)$$

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、行と列がともに各特徴に対応する。よって、行と列の同時入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 (9) の対応関係より、式 (12) の行列 \mathbf{S} は式 (4) と同じブロック対角行列となる。その状

態を \mathbf{S}^* で表す。

式 (12) より、

$$\mathbf{V}^{*t} = \mathbf{\Sigma}^{*-1/2}\mathbf{A}^*\mathbf{S}^* \quad (14)$$

であることから、

$$\begin{aligned} \mathbf{V}_{inter}^* &= \mathbf{S}^{*t}\mathbf{A}^{*t}\mathbf{\Sigma}^{*-1}\mathbf{A}^*\mathbf{S}^* \\ &= \mathbf{S}^{*t} \left[\left(\mathbf{A}^{*-1}\mathbf{\Sigma}^{*1/2}\mathbf{V}^{*t} \right) \cdot \left(\mathbf{V}^*\mathbf{\Sigma}^{*1/2}\mathbf{A}^{*-t} \right) \right]^{-1} \mathbf{S}^* \\ &= \mathbf{S}^{*t} \left(\mathbf{S}^*\mathbf{S}^{*t} \right)^{-1} \mathbf{S}^* \end{aligned} \quad (15)$$

\mathbf{S}^* はブロック対角行列なので、明らかに式 (15) もブロック対角行列になる。式 (15) の要素 $v_{inter}^*(i, j)$ は次の性質をもつ。

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

以上から、形状相互作用行列を用いた場合、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. 従来の方法の問題点と提案方法の処理手順

3.1 従来の方法の問題点

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含まれる組合せ数

セグメンテーションを行う前には、各物体上の特徴数は不明である。よって、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各ブロック行列の大きさがわからない (図 1)。よって、すべての可能な行と列の入替えに関して、形状相互作用行列の非対角ブロック成分を最小にする最適化方法が必要とされる。

N 個の物体、 P 個の特徴に対し、セグメンテーションにおける組合せ数は次式で与えられる。

$$\frac{1}{N!} \sum_{k=1}^N (-1)^{N-k} {}_N C_k k^P \quad (16)$$

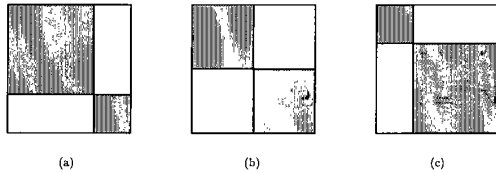


図1 形状相互作用行列におけるセグメンテーション。ブロック対角化の結果は各物体の特徴の数に依存して、例えば上記の(a), (b), (c)のように変わる

Fig.1 The shape interaction matrix corresponding to the segmentation result. The sizes of the block matrices in the shape interaction matrix vary with the number of the features of each object as shown in (a), (b) and (c).

これは、 P 個の特徴を N 個の物体に分割する組合せ数である。例えば、 $N = 4$, $P = 1230$ ^(注1) に対し、この組合せ数は近似的に 2×10^{740} となる。この組合せ数の枚挙的探索は非常に困難と考えられる。このような場合には、非線形最適化が適用されるが、局所解や計算量が問題になる場合がある。

(2) 縮退した形状の識別

形状相互作用行列を用いたセグメンテーションは、縮退した形状も取り扱える。しかし、どの物体が、つまり、どのブロック行列が縮退した形状に対応するかわからない。縮退した形状に対しては、例外的な処理を必要とする場合があり、その識別は重要である。

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、特徴選択に基づくセグメンテーション方法を提案する。提案方法の処理手順を図2に示す。まず、計測行列の特異値分解を行い、運動及び形状行列を求める。計測行列のランク r は特異値より推定する。

次に形状行列のQR分解を通じて、 r 個の特徴を選択する。このQR分解を用いる特徴選択は、“セグメンテーション情報が無い状態”で、“各物体”より特徴を選択することができる。よって、提案方法の特徴選択では「鶏と卵」の問題は生じない。

次の段階では、選択した特徴のセグメンテーションを行う。 r 個の特徴のみを用いるため、組合せ数は減少する。よって、物体数が極端に多くない限り、セグメンテーションでは最適化を必要としない。また、縮退した形状の識別と物体数の推定も、このセグメンテーションと同時にされる。

最後の段階では、選択されずに残った $P - r$ 個の特徴のセグメンテーションを行う。

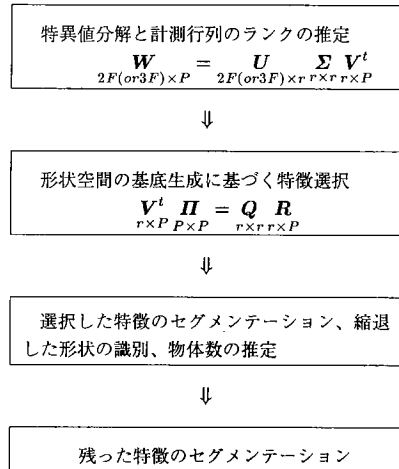


図2 提案方法の処理手順

Fig.2 The overview of the proposed algorithm.

各段階のより詳しい処理は4.で説明する。

4. 特徴選択に基づくセグメンテーション

4.1 形状空間の基底生成による特徴選択

特徴選択は、物体の特徴と1対1に対応している形状行列 V の行の選択により行う。

まず、いくつの特徴を選択するかを述べる。一つの物体から選択すべき特徴の数は、物体の形状に依存する。3次元物体ならば、形状表現のために同一平面上にない4点が最低必要である。また、平面なら3点、直線なら2点が最低必要である。複数の物体がある場合には、各物体の形状表現に必要な特徴数の合計が、全体として必要な特徴数となる。

セグメンテーションを行う前には、各物体の形状も物体数も不明である。よって、各物体にいくつの特徴が必要かはわからない。しかし、必要な特徴の総数は、計測行列のランク r よりわかる。なぜならランクは、 P 個の特徴の中で r 個の特徴のみが線形独立であり、複数の物体の形状表現には r 個の特徴があればよいことを示しているからである。よって、 P 個の特徴より r 個の特徴を選択すればよい。

N 個の物体に対し、それぞれ、 r_i ($i = 1, \dots, N$) 個の特徴が物体の形状表現に必要とする。この場合、セグメンテーションを通じて r 個の特徴を次式のように分解しなければならない。

$$r = r_1 + r_2 + \dots + r_N, \quad 1 \leq r_i \leq 4 \quad (17)$$

(注1)：これは5.における実験の条件である。

この r_i は、 i 番目の物体の形状を表す空間（以後、形状空間と呼ぶ）の次元に等しい。

特徴選択の最も大きな問題は、セグメンテーション情報が無い状態で、どのようにして i 番目の物体から形状に応じた r_i 個の特徴を選ぶかである。本論文では、複数の物体の形状空間の基底生成を通じた特徴選択を提案する。提案する特徴選択方法では、一つひと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（図 3）。基底は冗長性なしに各形状空間に割り当てられる。このことは、 r 個の基底ベクトル、つまり、それに対応する r 個の特徴が、式 (17) のように分解されることを意味する。なぜなら、 r_i 個以上のベクトルが i 番目の物体に割り当てられれば、その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からである。その状況では、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り、基底の条件に矛盾する。よって、物体上の特徴と基底ベクトルの対応をとる提案方法により、セグメンテーション情報が無い状態で、 i 番目の物体から r_i 個の特徴を選択できる。

物体上の特徴と基底ベクトルの対応を保つために、 r 個の特徴と 1 対 1 に対応する V^t の r 個の列より、基底を生成する。この処理を、形状行列の QR 分解 [18] により実現した。

QR 分解は次式で表される。

$$V^t \Pi = Q R \tag{18}$$

$r \times P$ $P \times P$ $r \times r$ $r \times P$

ここで、行列 Q と Π はそれぞれ、 V^t の列の正規直

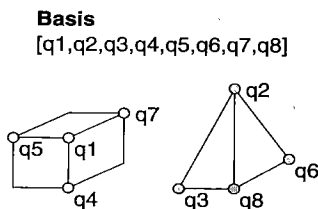


図 3 特徴に対応した基底の生成。物体上の特徴と対応するように、複数の物体の形状を表す線形空間の基底を生成する。各基底ベクトルは各物体に対して冗長なく割り当てられる。よって、各物体の形状を表すのに必要な特徴が、基底の生成を通じて選択される

Fig.3 The correspondence of the basis vectors to the features. Each basis vector of the shape space corresponds to a feature of the object. The basis vectors are assigned to the objects without redundancy. Thus the features needed to represent the shape of the objects can be selected through the construction of the basis.

交基底を列とする行列及び、 V^t の列の入替行列である。また行列 R は、演算の結果得られる行列である。 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 Π により、例えば V^t の列で次のような入替えが生じる。

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

$$\downarrow$$

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

そして、上記の例でいえば、140, 11, 30, 259, 182 番目の特徴に対応して、 Q の最初の五つの基底となる列が生成される。つまり、入替行列 Π により基底と特徴との対応はとられる。よって、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、 Π により入替えが行われた r 個の列に対応する特徴は、必ず r_i 個ずつ各物体に対応することだけはわかる。

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が無いに、各物体からその形状に応じた数の特徴を選ぶことができる。

4.2 選択した特徴の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。

$$V^t \Pi = \begin{pmatrix} V_{11}^t & V_{12}^t \\ \vdots & \vdots \end{pmatrix} \tag{19}$$

$r \times P$ $P \times P$ $r \times r$ $r \times (P-r)$

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。

$$X_{11} = V_{11} V_{11}^t = \{x_{ij}\} \quad i, j = 1, \dots, r$$

この行列の大きさは元の特徴数に依存せず、常に $r \times r$ である。

2.3 で述べた形状相互作用行列の特徴と、各物体の特徴数が 4 以下であることを利用し、以下のアルゴリズムでセグメンテーションを行う。

[Step 1] 初期化: $i = 0, k = 0, S = \{1, 2, \dots, r\}$.

[Step 2] 形状相互作用行列の各行の要素を大きい順にソートする。

[Step 3] S の最初の要素で示される X_{11} の行の、ソートされた要素を $\{x_1, x_2, \dots, x_r\}$ とおく。そして、次式の条件を満たす要素を見出す。

$$|x_{j+1}| < \theta \quad \text{and} \quad x_{(j+1)}/x_j < \theta$$

with the smallest $j + 1, 1 \leq j \leq r$ (20)

ここで θ はしきい値である。

[Step 4] 部分集合 $S_k = (s_{k1}, \dots, s_{kr_k})$ を作る。ここで、 $s_{kj}, j = 1, \dots, r_k$ は、Step 3 で抽出された要素より大きな値をもつ要素の、ソートされる前の列の番号である。つまり、 S_k は Step 2 のソートの入替結果を保持している。

[Step 5] S_k と同じ要素をもつ r_k 個の行を見つける。

[Step 6] S より Step 5 で見つけた r_k 個の行の番号を取り除く。そして、それらを集合 S_{seg} に加える。 $i := i + r_k$ とする。

[Step 7] もし $i < r$ ならば $k := k + 1$ とし、Step 3 へいく。そうでなければ、Step 8 へいく。

[Step 8] セグメンテーション結果は、 $S_{seg} = \{S_1, \dots, S_k\}$ であり、物体数は k である。各物体の形状空間の次元は r_k である。 $r_k < 4$ の場合、その物体の形状は縮退している。 S_{seg} による形状相互作用行列の行と列の入替えにより、セグメンテーション結果に対応するブロック対角行列を得る。

Step 3 の式 (20) の条件は、0 とみなされる最初の要素の位置を見出すためのものである。よって、しきい値 θ は、 10^{-5} など 0 に近い値にする。この条件を満たす要素は、各列の最初から 5 番目以内に存在する。なぜなら、各物体の最大の特徴数が 4 であるからである。すべての物体が 3 次元形状を有する場合には、特徴数は常に 4 である。よって、上記の条件は不要で、0 とみなされる最初の要素は常に 5 番目にある。

上記のアルゴリズムでは、 r_k により表される各ブロック行列の大きさにより、そのブロック行列に対応する物体の形状が縮退しているかどうかかわかる。また、 k はブロック行列の数であり、最終的には物体数を表す。このように、セグメンテーションと同時に、縮退した形状の識別と物体数の推定も行われる。

図 4 に選択した特徴のセグメンテーション結果の例を示す。図 4(a) で、白い部分が 0 に対応する要素を表す。まず、この例では $S = \{1, 2, 3, 4, 5, 6, 7\}$ となる。Step 2 により、各行の要素がソートされる。 S の最初の要素が 1 であるので、行列の 1 行目を調べる。この例では、非零の要素は三つである。よって、Step 3 と Step 4 より、 $S_1 = \{1, 2, 4\}$ を得る。 $r_1 = 3$ であり、これに対応する形状空間の次元は 3 で、その空間は平面を表すことがわかる。2 行目と 4 行目が S_1 と同じ要素をもつ。よって、Step 5 と Step 6 より、 $S = \{3, 5, 6, 7\}$ 及び $S_{seg} = \{1, 2, 4\}$ を得る。この更新された S に対し、Step 3 から Step 6 の処理を行

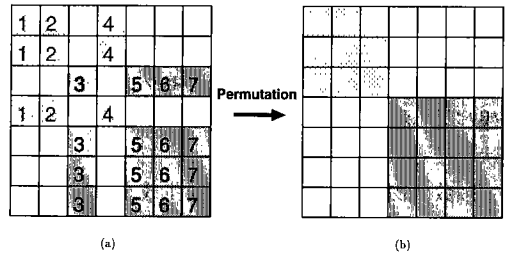


図 4 選択された特徴のセグメンテーションの例
Fig. 4 An example of the segmentation of the selected features.

うと、 $S_2 = \{3, 5, 6, 7\}$ を得る。この段階で処理は終了する。 $k = 2$ で処理が終了したことから、物体数は 2 となる。また、セグメンテーション結果は、 $S_{seg} = \{1, 2, 4, 3, 5, 6, 7\}$ となる。この集合により表される入替えを行うと、形状相互作用行列はブロック対角化される (図 4(b))。

4.3 残った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 (19) の V_{11}^t の列が、各物体に対応する r_i 個の特徴ごとに並ぶように入れ替えられる。その入替えに対応し、式 (18) 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。

$$\tilde{Q} = \begin{bmatrix} \tilde{Q}^1 & \tilde{Q}^2 & \dots & \tilde{Q}^N \\ r \times r_1 & r \times r_2 & & r \times r_N \end{bmatrix} \quad (21)$$

行列 \tilde{Q}^i ($i = 1, \dots, N$) の列は、各物体の形状空間の基底である。この基底は、選択されずに残った $P - r$ 個の特徴のセグメンテーションに利用できる。

本論文では、Noguchi [19] によって提案された射影行列 G^i を用いセグメンテーションを行った。

$$ObjNo(j) = \underset{i}{\operatorname{argmax}} \left(\|G^i v_j\|^2 \right) \quad (22)$$

$$G^i = \tilde{Q}^i \left\{ \tilde{Q}^i \left(\tilde{Q}^{it} \tilde{Q}^i \right)^{-1} \right\}^t \quad (23)$$

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

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。射影行列 G^i は、次の性質をもつ。

$$\begin{cases} G^i v_j \neq 0, & v_j \in \operatorname{span}(\tilde{Q}^i) \\ G^i v_j = 0, & v_j \notin \operatorname{span}(\tilde{Q}^i) \end{cases} \quad (24)$$

単に各形状空間への正射影を計算すると、形状空間の

次元の違いが射影成分の大きさに影響を与える。つまり、縮退した形状に対しては、基底ベクトルの数が少なくなるので、射影成分も小さくなってしまふ。これは、セグメンテーションの誤りの原因になる。これに対し、ここで用いた射影行列は、式(24)の性質より次元の違いにかかわらず、どの特徴がどの形状空間に属するかを識別できる。

以上で、特徴選択に基づくセグメンテーションは終了する。

5. 実験

提案方法の動作を確認するために、物体数、各物体の特徴の数、及び物体の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[20]により行った。対象とした物体は、図5に示す四つである。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[21],[22]。追跡の一例を図6に示す。物体の移動は人間の手により行った^(注2)。追跡の結果、幾何モデル上のサンプル点(図9)を特徴とする特徴対応データが得られる。一つひとつの物体に対し図6のように追跡を行い、得られたデータ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、特徴数は図5の各物体それぞれ、328、280、296、326で、合計1230である。

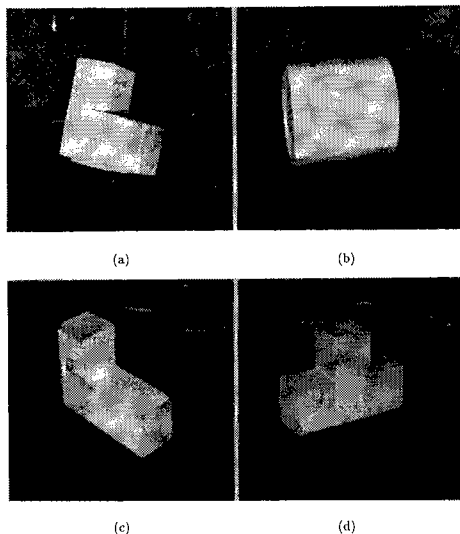


図5 実験に用いた四つの物体
Fig. 5 Four objects used in the experiments.

3次元計測データ (x, y, z) の奥行情報を使わず、 (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 100×1230 となる。

追跡結果を用い、すべての物体が3次元形状をもつ場合と、縮退した形状が含まれる場合の二つの実験を行った。

5.1.1 すべての物体が3次元形状をもつ場合

この実験では、図5に示した四つの物体を用いた。図7は、1230個の特徴の追跡結果を示す。

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} を掛けた数をしきい値として、それ以下の特異値は0とみなした。その結果、計測行

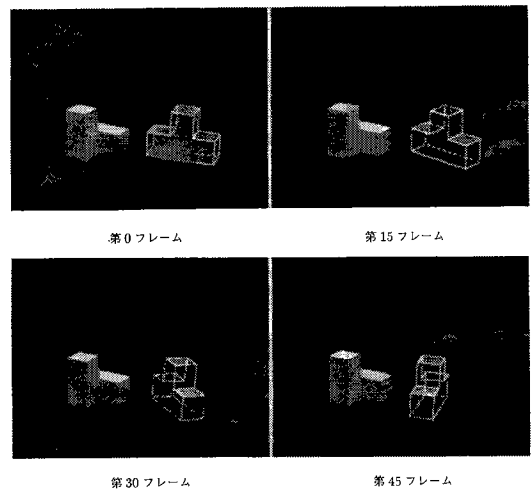


図6 物体追跡の一例 [22]
Fig. 6 An example of the model-based object tracking [22].

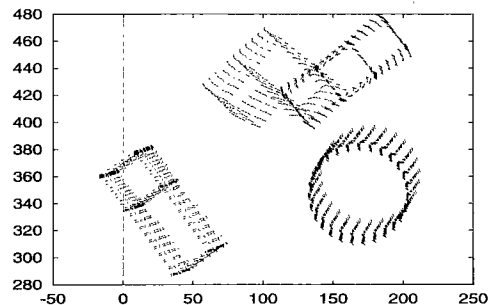


図7 四つの物体の特徴対応
Fig. 7 Feature correspondences of the four objects.

(注2) : ロボット等を用いて系統的な運動をさせていないため、各物体の運動の独立性は事前には保証されていなかった。

列のランクは 16 となった。よって、1230 個の特徴から 16 個の特徴が選択された。

図 8 にセグメンテーション結果を示す。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 8(a)) に対し、4.2 で述べたアルゴリズムを適用した結果、各形状空間の次元と物体数は正しく推定された。よって、セグメンテーション結果より、形状相互作用行列はブ

表 1 特異値 (すべての物体が 3 次元形状をもつ場合)
Table 1 Singular values of the full-3D case.

Order	Singular Value	Order	Singular Value
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

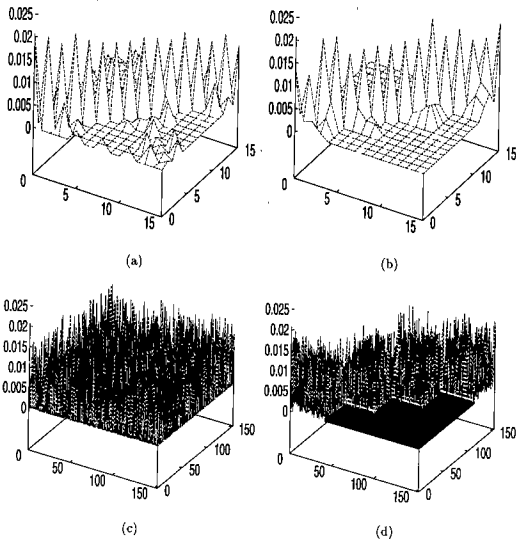


図 8 セグメンテーション結果 (すべての物体が 3 次元形状をもつ場合)。(a) 選択した特徴の形状相互作用行列。(b) (a) のセグメンテーション結果。(c) 全特徴の形状相互作用行列。(d) (c) のセグメンテーション結果。(c) と (d) では行列の様子がわかりやすくなるように、八つごとに特徴を間引いて表示した

Fig. 8 The segmentation result of the full-3D case. (a) the shape interaction matrix constructed by the selected features. (b) segmentation result of (a). (c) the shape interaction matrix constructed by the all features. (d) segmentation result of (c). In (c) and (d), every 8 features are sampled to display the shape interaction matrix clearly.

ロック対角化された (図 8(b))。このセグメンテーション結果を用いて、式 (22) より残った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 8(c), (d))。この結果は、四つの物体が正しくセグメンテーションされたことを示す。

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 [16] により物体形状を復元した結果を図 9 に示す。図 5 に示した四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 9 の “Selected Points” は、選択した四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。

5.1.2 縮退した形状を含む場合

この実験では、図 5 の中の三つの物体を用いた。マグカップに関しては、底面の部分のみのデータを使用した。この部分は平面であるため、シーンに縮退した形状が含まれることになる。図 10 に、772 個の特徴点の追跡結果を示す。

表 2 に、14 個の特異値を示す。これらの特異値より、計測行列のランクは 11 となった。図 11 にセグメンテーション結果を示す。4.2 に示したアルゴリズム

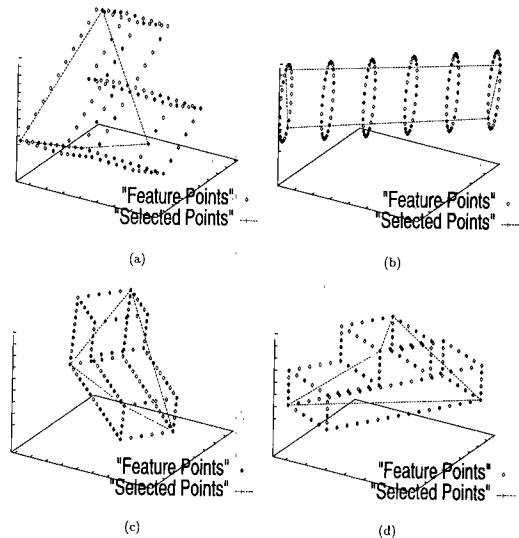


図 9 物体形状の復元結果 (すべての物体が 3 次元形状をもつ場合)。図 5 の各物体に対応している。“Selected Points” は、セグメンテーションで選択した特徴点を結んだものである

Fig. 9 Results of the shape reconstruction in the full-3D case. The “Selected Points” in the figures are the features selected in segmentation.

表2 特異値 (縮退した形状を含む場合)
Table 2 Singular values of the degenerate case.

Order	Singular Value	Order	Singular Value
1	110092	8	14.0002
2	28609.3	9	11.3941
3	229.167	10	8.44566
4	132.309	11	6.61284
5	41.211	12	7.45128e-11
6	27.9543	13	2.05956e-11
7	19.3066	14	1.30156e-11

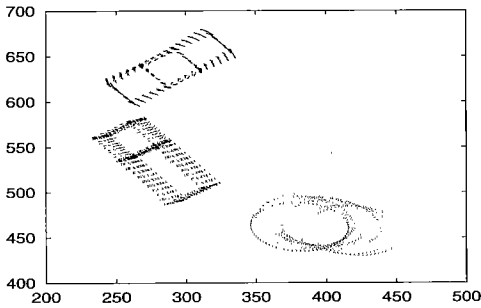


図10 縮退した形状を含む三つの物体の特徴対応
Fig.10 Feature correspondences of the three objects. The object with the degenerate shape, plane, is included in this data.

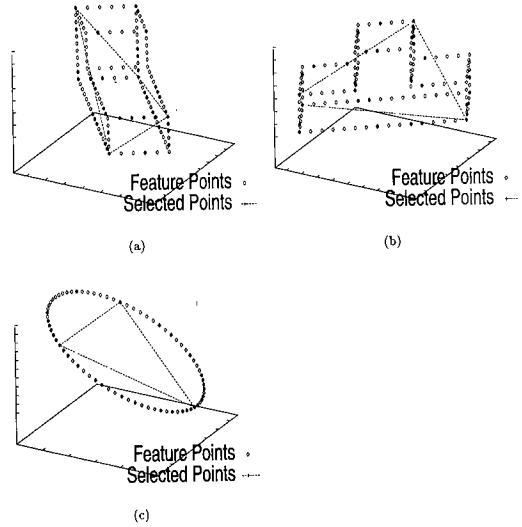


図12 物体形状の復元結果 (縮退した形状を含む場合). “Selected Points” は、セグメンテーションで選択した特徴点を結んだものである

Fig.12 Results of the shape reconstruction in the degenerate case. The “Selected Points” in the figures are the features selected in segmentation.

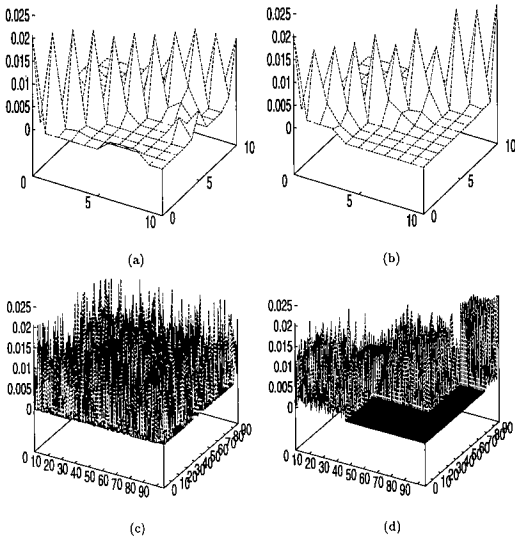


図11 セグメンテーション結果 (縮退した形状を含む場合). (a) 選択した特徴の形状相互作用行列. (b) (a) のセグメンテーション結果. (c) 全特徴の形状相互作用行列. (d) (c) のセグメンテーション結果

Fig.11 The segmentation result of the degenerate case. (a) the shape interaction matrix constructed by the selected features. (b) segmentation result of (a). (c) the shape interaction matrix constructed by the all features. (d) segmentation result of (c).

により、各形状空間の次元と物体数は正しく推定された。よって、11個の特徴より構成された形状相互作用行列はブロック対角化された (図11(a), (b))。全特徴より構成された形状相互作用行列も、同様にブロック対角化された (図11(c), (d))。この結果は、三つの物体が正しくセグメンテーションされたことを示す。各物体の形状も正しく復元された (図12)。

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、数値最適化を必要としない処理が実現できた。また、特徴数が形状空間の次元を表すことから縮退した形状をもつ物体の識別及び、物体数の推定もセグメンテーションと同時にできた。

本論文の実験では、モデルベースの追跡を用いたため、誤対応や雑音の影響は加味されていない。その意味で、実験はアルゴリズムの原理的な動作の確認にとどまっている。誤対応や雑音を含むデータに対するア

ルゴリズムの適用が、今後の大きな課題の一つである。

謝辞 研究の機会を与えてくださいました電総研知能情報部、大津展之部長に感謝致します。また、日ごろ御討論頂く VVV プロジェクトの関係諸氏に感謝致します。

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] T.Y. Tian and M. Shah, "Recovering 3D motion of multiple objects using adaptive Hough transformation," IEEE Trans. Pattern Anal. & Mach. Intell., vol.19, no.10, pp.1178-1183, 1997.
- [4] 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.
- [5] T. Weiss and E.H. Adelson, "A unified mixture framework for motion segmentation: Incorporating spatial coherence and estimating the number of models," Proc. Computer Vision and Pattern Recognition, pp.321-326, 1996.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] T.E. Boulton and L.G. Brown, "Factorization-based segmentation of motions," Proc. IEEE Workshop on Visual Motion, pp.179-186, 1991.
- [10] C.W. Gear, "Feature grouping in moving objects," Proc. IEEE Workshop on Motion of Non-rigid and Articulated Objects, pp.214-219, 1994.
- [11] 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.
- [12] 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.
- [13] 市村直幸, 石山 豊, 角 保志, 富田文明, "形状行列からの特徴選択に基づく動きの分割," 情処学 CV 研報 106-12, pp.83-90, 1997.
- [14] N. Ichimura and F. Tomita, "Motion Segmentation Based on Feature Selection from Shape Matrix," Technical Report of the Electrotechnical Laboratory, TR-97-22, 1997.
- [15] 市村直幸, 石山 豊, 角 保志, 富田文明, "QR 分解を用いた特徴選択に基づく複数運動の分割," 第 55 回情処全大, vol.II, pp.120-121, 1997.
- [16] C. Tomasi and T.Kanade, "Shape and motion from image streams under orthography: A factorization method," International Journal of Computer Vision, vol.9, no.2, pp.137-154, 1992.
- [17] C.J. Poelman and T. Kanade, "A paraperspective factorization method for shape and motion recovery," IEEE Trans. Pattern Anal. & Mach. Intell., vol.19, no.3, pp.206-218, 1997.
- [18] G.H. Golub and C.F. Van Loan, "Matrix Computations," The Johns Hopkins University Press, 1996.
- [19] Y. Noguchi, "Subspace method of feature extraction using non-symmetric projection operators," Bulletin of the Electrotechnical Laboratory, vol.40, no.7, pp.11-27, 1976.
- [20] 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—パタパタ," 画像の認識・理解シンポジウム (MIRU '94) 予稿集, II, pp.127-134, 1994.
- [21] 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, March 1997.
- [22] 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.

(平成 10 年 2 月 18 日受付, 7 月 9 日再受付)



市村 直幸 (正員)

平 1 電通大・電気通信・通信卒。平 6 同大大学院電気通信学研究科博士後期課程単位取得退学。同年電子技術総合研究所入所。現在知能情報部主任研究官。コンピュータビジョン, データ圧縮等の研究に従事。工博, 情報処理学会, IEEE 各会員。



富田 文明 (正員)

昭 47 阪大・基礎工・制御卒。昭 53 同大大学院博士課程了。工博。同年電子技術総合研究所入所。昭 58~59 カーネギーメロン大学計算機科学科客員研究員。現在知能システム部主任研究官。画像処理, コンピュータビジョン, 知能ロボットの研究に従事。平 2 情報処理学会創立 30 周年記念論文賞受賞。著書「Computer Analysis of Visual Textures」(共著)など。情報処理学会, 日本ロボット学会各会員。