

QR 分解を用いた特徴選択に基づく複数運動の分割*

2K-4

市村 直幸[†] 石山 豊[‡] 角 保志[†] 富田 文明[†][†]電子技術総合研究所 [‡]スタンレー電気(株)技術研究所

1 はじめに

動きに基づくセグメンテーションは、動きからの3次元構造の復元を通じたモデリングや、コンテンツベースのビデオ圧縮などで、その必要性が高まっている。本稿では、因子分解に基づくセグメンテーション [1] における組み合わせ問題を、特徴選択により解決する方法を提案する。

2 因子分解に基づくセグメンテーション

時系列画像において物体の特徴を追跡し、特徴対応データを得る。特徴対応データを要素とする計測行列を W ($2F \times P$) とする。ここで、 F と P はフレーム数と全特徴数である。特異値分解により、計測行列を次のように分解する [2]。

$$W = U \Sigma V^t \quad (1)$$

$2F \times P$ $2F \times r$ $r \times r$ $r \times P$

ここで、 r は計測行列のランクである。

Costeira ら [1] は、次式で定義される形状相互作用行列を用いてセグメンテーションを行った。

$$X = VV^t \quad (2)$$

$P \times P$

$V^t = [v_1, v_2, \dots, v_P]$ の各列は、 P 個の特徴に一つ一つに対応する。よって、 X では、その行と列が共に特徴に一つ一つに対応する。このことから、 X の行と列の入れ替えにより、セグメンテーションが行える。

X の要素 x_{ij} は、 i 番目の行と j 番目の列に対応する特徴が同一物体に属する場合には 0 でなく、そうでない場合には 0 となる。この性質から、正しくセグメンテーションを行うと、 X はブロック対角化される。

3 特徴選択に基づくセグメンテーション

各物体の特徴数は未知であるため、 X をブロック対角化するには、基本的に、 P 個の特徴を N 個の物体に分割する全ての組み合わせを試す必要がある。通常、その組み合わせ数は非常に大きな数になるため、非線形最適化などがセグメンテーションに用いられる。

提案方法では、組み合わせ問題の解決のために、最低限必要な数だけ特徴を選択する。必要な特徴の数は、

各物体最大 4 つで、その総数は計測行列のランク r に等しい。 r 個の特徴しか用いないため、組み合わせ数が削減される。その結果、最適化を用いずにセグメンテーションが行える。以下に、特徴選択およびその後の処理を説明する。

3.1 QR 分解による特徴選択

特徴選択での問題は、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”から特徴を選択するかである。本稿では、 V^t の列の張る空間、つまり、複数の物体の形状を表す空間の、基底の生成を通じた特徴選択を提案する。基底は、各物体に対し、その形状の表現に最低限必要な数が割り当てられる。よって、基底を物体の特徴と一つ一つに対応させ生成すると、各物体から必要な数の特徴が選択できる。

基底と特徴の対応を保つため、特徴と一つ一つに対応している V^t の列 $\{v_j\}$ から r 個取り出し、それらから基底を生成する。この基底の生成を QR 分解 [3] により行った。QR 分解は次式で表される。

$$V^t \Pi = Q R \quad (3)$$

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

基底は行列 Q に得られる。基底の生成に使われた r 個の V^t の列は、 V^t の先頭 r 列に入れ替えられる。その入れ替え結果は行列 Π に保持され、何番目の特徴が選択されたかがわかる。行列 R は演算の結果得られる上三角行列である。

3.2 選択された特徴のセグメンテーション

選択された r 個の V^t の列より、形状相互作用行列 X_0 を作り、以下の手順でセグメンテーションを行う。

まず、 X_0 の各行で、要素を大きい順にソートする。そして、次の条件を満たす要素を見つける。これは、いくつの要素が 0 以外の値かを見いだす処理である。

$$|x_{ij}| < \theta \quad \text{and} \quad 1/(x_{ij} - x_{i(j+1)}) < \theta \quad (4)$$

ここで θ はしきい値で、 10^{-5} など 0 に近い値とする。条件を満たす要素は必ず上位 5 つ以内にある。もし、縮退した形状をもつ物体がなければ、しきい値は不要で、5 つ目の要素を選べば良い。

見つけた要素より大きな要素の番号の集合 S_i ($i = 1, \dots, r$) を作る。そして、 S_i の中で、同一の要素番号をもつものを見つけ、それらをまとめる。最大 $\frac{4(4+1)}{2}$ 回比較で 2 つの集合がまとめられる。

*Motion Segmentation Based on Feature Selection Using QR Decomposition. Naoyuki Ichimura, Yasushi Sumi and Fumiaki Tomita are with the Electrotechnical Laboratory. Yutaka Ishiyama is with the Stanley Electric Co., Ltd., R & D. E-mail: ichimura@etl.go.jp

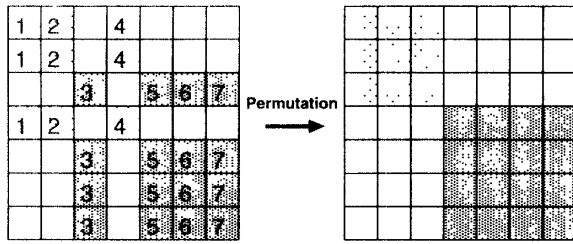


図 1: 選択された特徴のセグメンテーションの例.

図 1 の例では、 $S_1, S_2, S_4 = (1, 2, 4)$ および $S_3, S_5, S_6, S_7 = (3, 5, 6, 7)$ が同一の要素番号もち、まとめられる。得られた 2 つの集合それぞれが、1 つの物体に対応する特徴を表す。そして、この集合の数が物体数である。また、各集合に含まれる要素の数が 3 以下ならば、形状が縮退していることを表す。このように、物体数の推定と縮退した形状の識別も同時に行われる。結果から入れ替え行列 $(1, 2, 4, 3, 5, 6, 7)$ が得られ、これに従い X_j がブロック対角化できる。

3.3 残った特徴のセグメンテーション

選択された特徴のセグメンテーション結果に応じて、行列 Q の列の入れ替えを行う: $Q = [Q^1, \dots, Q^N]$ 。ここで、 Q^i は i 番目の物体の形状を表す基底である。この基底から、Noguchi[4] によって提案された射影行列 G^i を構成し、次式より選択されずに残った特徴のセグメンテーションを行う。

$$ObjNo(j) = \underset{i}{\operatorname{argmax}} (\|G^i v_j\|^2) \quad (5)$$

$$G_{r \times r}^i = Q^i \left\{ Q^{i^t} (Q^{i^t} Q^i)^{-1} \right\}^t \quad (6)$$

ここで、 $ObjNo(j)$ は j 番目の特徴に割り当てられる物体番号である。

4 実験結果

2 つの 3 次元物体および 1 つの平面を 50 フレーム追跡し、特徴対応データを得た (図 2(a))。全特徴数は 772 である。特異値より、計測行列のランクは 11 と推定された。11 個の特徴を選択し、セグメンテーションを行った。その結果、全特徴の形状相互作用行列がブロック対角化された (図 2(b),(c))。

各物体の形状の復元結果を図 2(d),(e),(f) に示す。各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。“Selected Points” は、各物体で選択された特徴を結んだ直線である。

5 まとめ

特徴選択に基づく複数運動のセグメンテーション方法を提案した。提案方法では、複数の物体の形状を表す空間の、基底の生成として特徴選択を行った。特徴数の削減により、数値最適化を用いることなく、セグ

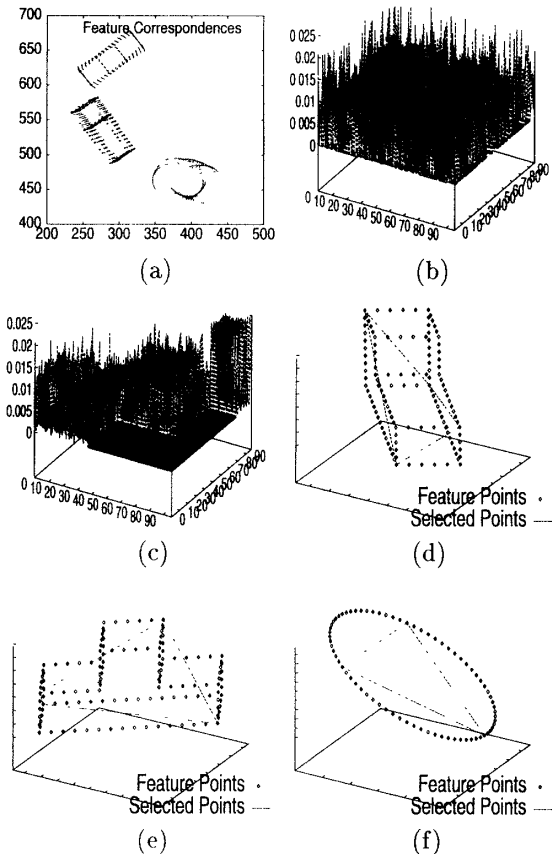


図 2: 実験結果. (a) 特徴対応データ. (b) 全特徴の形状相互作用行列. (c) セグメンテーション結果に対応する形状相互作用行列. (d),(e),(f) セグメンテーションされた物体の形状復元結果. (b),(c) では、表示のため、8 つずつデータを間引いた。

メンテーションが行えた。また、セグメンテーションと共に、物体数の推定と縮退した形状の識別も行えた。

謝辞 研究の機会を与えて下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] 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
- [2] 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
- [3] C. W. Ueberhuber: “Numerical Computation 2,” Springer-Verlag, 1997
- [4] Y. Noguchi: “Subspace Method of Feature Extraction Using Non-symmetric Projection Operators,” Bulletin of the Electrotechnical Laboratory, 40, 7, pp.11-27, 1976