

因子分解法と判別基準を用いた複数運動の分割*

4 K - 3

市村 直幸

電子技術総合研究所 知能情報部

1 はじめに

動きの分割は、動きからの3次元構造の復元や、コンテンツベースのビデオ圧縮などで必要とされる。因子分解法に基づく動きの分割は、一度の行列の分解で複数の物体の運動情報が得られるという利点をもち、いくつかの方法が提案されている [2]-[4]。しかし、これらの方法は雑音に対する耐性が高いとは言えない。それは、シーン内には一つの動きしかない、という拘束を用いて計測行列のランクを規定できないためである [5]。また、計測行列のランク推定は、雑音や投影モデルの近似の影響で困難である。本稿では、雑音や外れ値を含む特徴対応データからの動きの分割を目的とし、因子分解法から得られる動きの類似度を表す行列と、判別基準を用いた方法を提案する。提案方法では、動きの類似度の分離度を最大にする特徴を選択し、それを用いて全特徴をグルーピングする。

2 因子分解法による動きの類似度の取得

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 はフレーム数と特徴数である。アフィン投影を仮定し、特異値分解により計測行列を次のように分解する [1]。

$$W = U \Sigma V^t \tag{1}$$

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

ここで、 $r = 1, \dots, P$ であり、非零の特異値の数に一致する場合には計測行列のランクを表す。この分解の結果より、次の形状相互作用行列を得る [2]。

$$X = VV^t = (x_1, \dots, x_P)^t \tag{2}$$

$P \times P$

データが雑音を含まない場合には、 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。つまり、この行列は特徴間の動きの類似度を表現していると言える。

3 判別基準による特徴の選択

従来、式 (2) の行列の要素全てを用いて、動きの分割が行われていた [2]。それに対し提案方法では、判別

基準に基づいて、動きの類似度の分離度を最大にする特徴を選択して用いる。この特徴選択により、雑音や外れ値の影響を自動的に軽減できる。

まず、式 (1) の r を適当に与え、行列 X を構成する。次に、 X の各行 x_k ($k = 1, \dots, P$) の要素をソートする。そして、その要素を次の判別基準に基づいて2つのグループに分割する。

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2} \tag{3}$$

$$\sigma_B^2 = N^1 N^2 (\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^2 \tag{4}$$

$$\sigma_W^2 = N^1 \sigma_1^2 + N^2 \sigma_2^2 \tag{5}$$

ここで $\bar{\varepsilon}_1, \bar{\varepsilon}_2, \sigma_1^2, \sigma_2^2, N^1, N^2$ は各グループの要素の平均、分散、および、個数である。また、 σ_B^2 と σ_W^2 は、グループ間分散およびグループ内分散である。行ベクトルの各要素をしきい値として用い、 λ を最大にする分割を求める。

λ_k を k 番目の行ベクトルに対する判別基準の値とする。最も大きな類似度の分離度を持つ特徴を、次式より求める。

$$k_{select} = \arg \max_k \lambda_k \tag{6}$$

行ベクトル $x_{k_{select}}$ の分割結果から得られる2つのグループのうち、より大きな平均類似度を持つグループを取り除く。

上記の処理を再帰的に繰り返し、他の動きに対応するグループを取り出す。選択された行ベクトルに含まれる要素の平均、または、 $\lambda_{k_{select}}$ があらかじめ与えたしきい値以下になったら、処理を終了する。

4 パラメータ r に対する処理

前節において、式 (1) の r は計測行列のランクがわかる場合には、その値にする。しかし、実データでは、ランクを推定することは困難である。本研究では、 r をある範囲で変化させ、前節のアルゴリズムを適用した。そして、抽出された g 個のグループの $\lambda_{k_{select}}$ の和

$$\lambda(r) = \sum_{i=1}^g \lambda_{k_{select}}(i, r) \tag{7}$$

を最大にする r を用いた際の分割を、最終的な結果とした。

*Motion Segmentation Using Factorization Method and Discriminant Criterion, Naoyuki Ichimura is with the Electrotechnical Laboratory, E-mail: ichimura@etl.go.jp

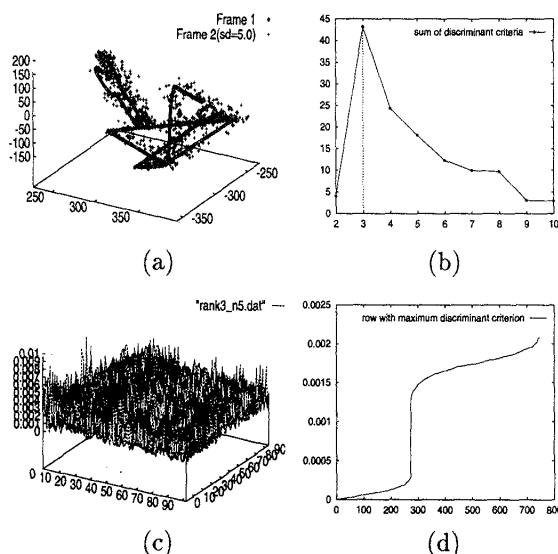


図 1: 人工的に動きを与えたデータの分割結果. (a) 標準偏差 5 のガウス雑音を加えたデータ. (b) r に対する判別基準の和の変化. (c) 形状相互作用行列. (d) 最大の分離度をもつ行の要素の変化.

5 実験結果

5.1 人工データ

二つの物体の 3 次元データに、人工的な動きを与えた。画像の雑音や追跡の誤りを模擬するため、移動後のデータに標準偏差 5 のガウス雑音を加えた (図 1(a))。データの投影には、焦点距離が 8[mm] のレンズで 2[m] の距離から物体を観測した場合の透視投影行列を用いた。分割の終了条件のしきい値は、 1.5×10^{-3} と 6.0 である。式 (7) の値の変化より、 $r = 3$ の結果を採用した (図 1(b))。判別基準に従い選択した特徴は、雑音下においても分割に必要な情報を有しており (図 1(c),(d))、二つの物体は正しく分割できた。

5.2 実データ

人間の動きを 2 秒間追跡したデータを用いた。角検出器により得られた特徴点 (図 2 の四角) を、ブロックマッチングで追跡した。隠れ等でマッチング時の誤差が大きくなった特徴点は追跡の途中で除去したが、それ以外の雑音の軽減や外れ値除去等の処理は行っていない。分割の終了条件のしきい値は、 1.5×10^{-3} と 3.0 である。 $r = 4$ の場合の結果を採用した (図 3(a))。分割結果は、左手、顔と肩、および、背景に対する特徴点にはほぼ対応していた (図 3(b)-(d))。この結果より、二つの動きのある部分に対して、それぞれ、エピポラ拘束を求めることができた (図 2 の線)。

6 まとめ

因子分解法より得られる動きの類似度と、判別基準を用いた動きの分割方法を提案した。因子分解に基づ

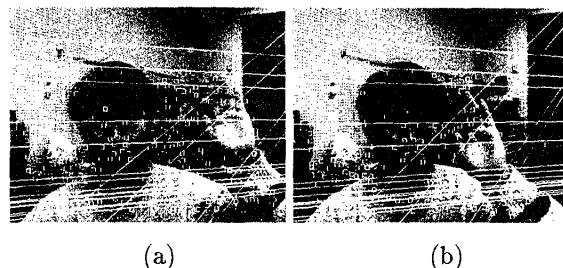


図 2: 人間の動きの追跡結果と分割結果より計算した各部分のエピポラ拘束. (a) 第 1 フレーム. (b) 第 60 フレーム.

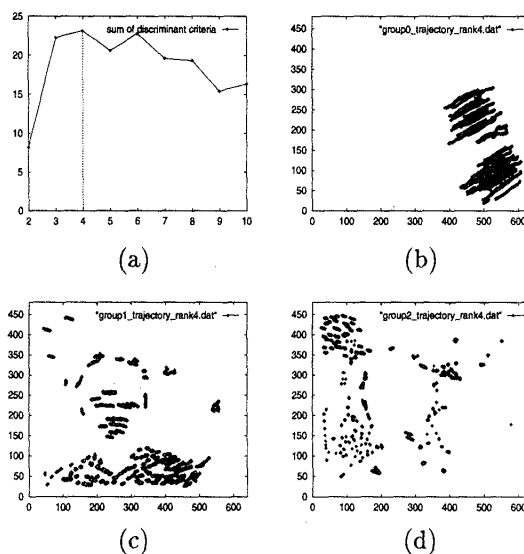


図 3: 人間の動きの分割結果. (a) 判別基準の和の変化. (b)-(d) 各グループに対する動きの軌跡.

く他の方法との関連付け等が、今後の課題である。

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。

参考文献

- [1] C. Tomasi and T. Kanade: "Shape and motion from image streams under orthography: a factorization method," *Internat. J. Comp. Vis.*, 9, 2, pp.137-154, 1992
- [2] J. P. Costeira and T. Kanade: "A multi-body factorization method for independently moving objects," *Internat. J. Comp. Vis.*, 29, 3, pp.159-179, 1998
- [3] C. W. Gear: "Multibody grouping from motion images," *Internat. J. of Comp. Vis.*, 29, 2, pp.133-150, 1998
- [4] 市村直幸、富田文明: "形状行列からの特徴選択に基づく動きの分割," *信学論 D-II*, Vol.J81-D-II, No.12, pp. 2757-2766, 1998
- [5] T. Kanade and D. D. Morris: "Factorization methods for structure from motion," *Phil. Trans. R. Soc. Lond. A*, 356, pp.1153-1173, 1998