

ロバスト統計を用いた背景位置合わせに基づく 移動カメラ画像からの移動物体の抽出

5N-5

朝岡 忠 横矢 直和 竹村 治雄 山澤 一誠
奈良先端科学技術大学院大学

1. はじめに

本報告ではロボットの移動中に出現する人物や自動車などの移動物体の検出・識別を目的とした移動カメラ画像からの移動物体の抽出手法について述べる。

移動カメラ画像では移動物体以外の背景領域にも見かけの動きが生じるので移動物体の抽出が困難である。本研究ではフレーム間での対応点を用いて背景領域の位置合わせを行い、それらのフレーム間差分に基づいて移動物体を抽出する。

画像内に移動物体が存在する場合、上記の対応点には背景以外の移動物体に属するものが不可避免的に含まれる。対応点から背景位置合わせのための画像変換パラメータを推定する場合、これが外れ値 (outlier) となるので最小2乗法などの最適化法では安定した結果が得られない (inlier は背景領域に属する対応点である)。そこで本研究ではロバスト統計 [1] を用いて背景位置合わせを行う。

2. 前提条件

本研究では以下の画像を対象とする。

- 1) 移動物体領域は背景領域よりも小さい。
- 2) 背景の奥行き方向の広がり、カメラと背景との距離に比べて小さい。
- 3) フレーム間でのカメラの視点移動の大きさは、カメラと背景との距離に比べて小さい。

1) より outlier の割合は 50% 未満であると仮定して (対応点探索については 3. で説明) 背景位置合わせにロバスト統計を用いる。2) および 3) よりカメラと背景との相対的な動きをパン・チルト・ズームで近似し、射影変換 $V'(x', y') = f(a, V(x, y))$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{a_7x + a_8y + 1} \begin{pmatrix} a_1x + a_2y + a_3 \\ a_4x + a_5y + a_6 \end{pmatrix} \quad (1)$$

を用いて背景位置合わせを行う。ここで、 $a = \{a_1, \dots, a_8\}$ は変換パラメータである。この式は一般に視点を変えずに視軸の向きを変える場合に使われる変換式である。

3. フレーム間の対応点探索

フレーム間での対応点探索は以下のように行う。まず一方のフレーム上で Moravec のインタレスト・オペレータ [2] を用いて特徴点を求める。背景位置合わせを信頼性良く行う

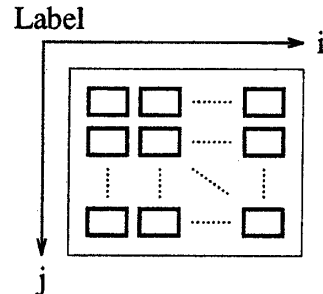


Fig. 1 特徴点抽出のためのウィンドウの設定

ためには、対応点は画面全体に様々かつ多数分布していることが望ましい。また outlier の割合を 50% 未満にするためにもその方が望ましい。そこで Fig. 1 に示すようにあらかじめ多数のウィンドウを等間隔に配置し (各ウィンドウのラベルを (i, j) とする)、各ウィンドウごとにインタレスト・オペレータの出力値が最大である点 $P_{ij}(x_p, y_p)$ を検出する。つぎに他方のフレーム上で P_{ij} と対応する点 $Q_{ij}(x_q, y_q)$ を相関法を用いて求める。

4. ロバスト統計による背景位置合わせ

得られた対応点からロバスト統計を用いて背景位置合わせのための変換パラメータを推定する。

ロバスト統計には連続的な手法である M-推定法と離散的な手法である最小2乗中央値 (LMS) 法がある。M-推定法は繰り返し法により解を推定するため計算量が小さいという利点があるが、適切な初期値を与える必要がある。一方 LMS 法は初期値を必要としないが、解法が乱数発生的な手法に基づくため一般に計算量が大きい。

そこで本研究ではまず局所的に分布した対応点からなる集合のみを用いて LMS 法を実行する。そしてこの推定値を初期値として M-推定法を行うことにより最適なパラメータを推定する。

4.1 LMS 法による初期値推定

LMS 法が適切に働くためには inlier のみからなる対応点の集合を少なくとも 1 つ選択できれば良い。ある対応点が inlier の場合その近傍の対応点も inlier である可能性が非常に高い。またパラメータ a は 4 つの対応点から求めることができる。

そこでつぎの 4 つの対応点からなる集合

$$(C_{i,j}, C_{i+1,j}, C_{i,j+1}, C_{i+1,j+1}) \quad (2)$$

ごとにパラメータ a を求める。ここで $C_{i,j}$ は対応点対 (P_{ij}, Q_{ij}) を表す。対応点が画面全体に様々かつ多数分布している場合、これらの集合の内少なくとも 1 つが inlier

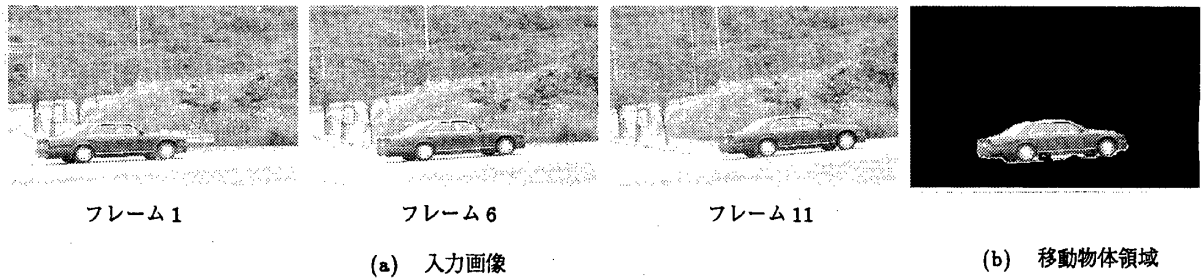


Fig. 2 移動物体の抽出結果

のみからなる確率は十分 1 に近い。そして対応点 P_{ij} と Q_{ij} の誤差 E_{ij} を

$$E_{ij} = |P_{ij} - f(a, Q_{ij})| = \sqrt{(x_p - x_q')^2 + (y_p - y_q')^2} \quad (3)$$

とし、各 a について E_{ij} の中間値

$$\text{Median}\{E_{ij}\} \quad (4)$$

を求める。それが最小となるときの a を推定値 $\hat{a}$ として選択する。

背景の位置合わせを信頼性良く行うには大局的に分布した多数の対応点から変換パラメータを求める必要がある。そこで $\hat{a}$ を初期値とした M-推定法により最適なパラメータを求める。

4.2 M-推定法による変換パラメータ推定

M-推定法を用いて以下のようにパラメータ a を推定する。

- 1) a の初期値として $\hat{a}$ を与える。

$$a \leftarrow \hat{a} \quad (5)$$

- 2) 各対応点について式 (3) により誤差 E_{ij} を計算し、その中間値 med を求める。

$$med = \text{Median}\{E_{ij}\} \quad (6)$$

- 3) E_{ij} に基づいて各対応点ごとの重み w_{ij} を求める。

$$w_{ij} = \begin{cases} (1 - (E_{ij}/med)^2)^2 & \text{if } E_{ij} \leq med \\ 0 & \text{if } E_{ij} > med \end{cases} \quad (7)$$

- 4) つぎの非線形最小 2 乗法を実行し a の値を更新する。

$$S(a) = \sum_{ij} w_{ij} E_{ij}^2 \rightarrow \min \quad (8)$$

- 5) w_{ij} が変化しなくなるまで 2) から 4) の処理を繰り返す。

5. 移動物体の抽出

3 つのフレーム t_1 、 t_2 、 t_3 を用いてフレーム t_2 における移動物体を抽出する。

まずフレーム t_1 とフレーム t_2 との間で背景位置合わせを行い、それらの差分画像 $S(t_1)$ を求める。フレーム t_2 とフ

レーム t_3 についても同様に差分画像 $S(t_2)$ を求める。得られた差分画像 $S(t_1)$ 、 $S(t_2)$ には背景位置合わせの近似誤差などに起因する雑音成分が多く存在する。そこであらかじめメディアンフィルタにより雑音成分を除去したのちに $S(t_1)$ 、 $S(t_2)$ を 2 値化し、マスク画像 $M(t_1)$ 、 $M(t_2)$ を求める。さらにマスク画像 $M(t_1)$ 、 $M(t_2)$ において一定面積以下の微小な領域を雑音成分として除去する。

これらのマスク画像 $M(t_1)$ と $M(t_2)$ の論理積をとるだけでは移動物体領域に欠損が生じる場合がある。そこで嶋田らの手法 [3] を適用する。 $M(t_1)$ と $M(t_2)$ をそれぞれ膨張し、これらの論理積をとる。これにより移動物体の輪郭領域が得られるので、この領域内部を塗りつぶし、膨張した分だけ収縮する。このようにして得られたマスク画像を用いて移動物体を切り出す。

6. 実験結果

移動物体の抽出実験に用いた 3 枚の入力画像 (ビデオレートで撮影) を Fig.2(a) に示す。画像の大きさはそれぞれ 320×200 である。背景位置合わせには 6×6 の 36 個の対応点を使用した。なお対応点探索を容易にするためにフレーム 2~5 および 7~10 を用いて対応づけを行った。抽出された移動物体領域を同図 (b) に示す。

7. むすび

提案手法では背景位置合わせに近似による誤差が原理的に生じる。そのためフレーム間差分によって得られる移動物体領域が小さい場合には雑音成分に埋もれて正しく移動物体を抽出することができない。今後は移動物体の抽出方法および適切なフレーム (タイムスケール) の選択方法について検討を行う予定である。

謝辞

本研究を行うにあたり御協力を頂いた住友電気工業株式会社の寺久保敏氏並びに有益な助言を頂いた新情報処理開発機構の高橋裕信氏に感謝します。

参考文献

- [1] 佐藤, "画像処理における統計モデル利用に関する最近の動向: ロバスト推定法と MDL 基準", Medical Imaging Technology, Vol.12, No.1, pp.48 - 58 (1994).
- [2] H.P.Moravec, "Visual mapping by a robot rover", Proc. 6th IJCAI, pp.598 - 600 (1979).
- [3] 嶋田, 安達, "逐次マッチングによる移動物体の切出し", 電子情報通信学会論文誌 (D-II), Vol.J76-D-II, No.10, pp.2196 - 2203 (1993).