

# カメラ動きパラメータのクラスタリングによる 動視点からの動物体検出

原 聡<sup>†</sup> 三枝 亮<sup>‡</sup> 橋本 周司<sup>†</sup>

早稲田大学理工学部応用物理学科<sup>†</sup>

The Italian Institute of Technology<sup>‡</sup>

## 1. はじめに

動物体検出はロボットビジョンやセキュリティシステムなどの基礎となる技術である。特に、動視点からの動物体検出は、移動ロボットなどの環境認識に必要不可欠である。動視点からの動物体検出手法は、Costeria[1]らにより「作用行列」を用いた手法や金谷[2,3]らの一連の研究がある。また、Vidal[4]らによりエピポラ幾何を利用した手法も提案されている。

本研究では透視射影カメラモデルに基づき、カメラ動きパラメータのクラスタリングを用いた逐次処理手法を提案し、動画像からの動物体検出を行う。

## 2. 提案手法の流れ

提案手法では、透視射影カメラモデルを用いてカメラの動きを推定する。カメラの動き推定には、クラスタリングを用いる。得られたカメラの動きから物体の動きを推定し、そこから導かれる拘束条件からのずれが大きいものを動物体として検出する。なお、前提として画像中の背景の割合が動物体の割合よりも高いものとする。

### 2.1 透視射影カメラモデル

静止物に対してカメラが動いている場合、画像座標  $(X_i, Y_i)$  における特徴点のフロー  $(dX_i, dY_i)$  について、透視射影カメラモデルのもとで以下の式が成り立つ。

$$\vec{d} = z_i^{-1} F_i \vec{v} + G_i \vec{\omega} \quad (1)$$

$$\vec{d} = (dX_i, dY_i)^T$$

$$F_i = \begin{pmatrix} -f & 0 & X_i \\ 0 & -f & Y_i \end{pmatrix}$$

$$G_i = \begin{pmatrix} -X_i Y_i / f & -f - X_i^2 / f & Y_i \\ -f - Y_i^2 / f & -X_i Y_i / f & -X_i \end{pmatrix}$$

ただし、 $z_i$  は特徴点の奥行き、 $\vec{v} = (v_x, v_y, v_z)^T$

は、カメラ座標系におけるカメラの各軸方向への速度、 $\vec{\omega} = (\omega_x, \omega_y, \omega_z)^T$  は各軸まわりの角速度、 $f$  は焦点距離とする。以降、 $\{\vec{v}, \vec{\omega}\}$  をカメラ動きパラメータと呼ぶ。また、焦点距離  $f$  は既知とする。

### 2.2 カメラ動きパラメータ推定

動画像には動物体上のフローや、誤追跡フローが含まれる。これらの影響を除くためにクラスタリングによるロバスト推定を用いる。手順は以下の通りである。

1. フローの集合  $B$  から要素数  $n$  の  $m$  個の部分集合  $\{B_j; |B_j| = n\}_{j=1}^m$  をランダムに生成する。
2. 各  $B_j$  についてエピポラ幾何の手法[5]を用いて  $E$  行列を求め、そこからカメラの平行移動成分  $\vec{v}_j$  を推定する。

$$z_j^{-1} \vec{v}_j^T (\hat{F}_{jk}^T \vec{d}_{jk} + \hat{F}_{jk}^T G_{jk} \vec{\omega}_j) = 0 \quad (2)$$

から  $\hat{F}_{jk}^T G_{jk} \vec{\omega}_j + \vec{v}_j^\perp = -\hat{F}_{jk}^T \vec{d}_{jk}$  となるので、式(3)により最小二乗推定パラメータ  $\vec{\omega}_j$  を求め、 $\{\vec{v}_j, \vec{\omega}_j\}$  を  $B_j$  の推定パラメータとする。

$$\begin{bmatrix} \vec{\omega}_j \\ \vec{v}_j^\perp \end{bmatrix} = - \begin{bmatrix} \hat{F}_{j1}^T G_{j1} & I \\ \vdots & \vdots \\ \hat{F}_{jn}^T G_{jn} & I \end{bmatrix}^+ \begin{bmatrix} \hat{F}_{j1}^T \vec{d}_{j1} \\ \vdots \\ \hat{F}_{jn}^T \vec{d}_{jn} \end{bmatrix} \quad (3)$$

ただし、添え字  $jk$  は  $B_j$  の  $k$  番目の要素を表し、 $\hat{F}_i$  は  $F_i = [\vec{f}_x \ \vec{f}_y]^T$  としたときに  $\hat{F}_i = [\vec{f}_x \ -\vec{f}_y]^T$  で表される行列である。

3. 各  $B_j$  より得られた推定パラメータ  $\{\vec{v}_j, \vec{\omega}_j\}_{j=1}^m$  に対してクラスタリングを行い、最大クラスタの平均パラメータ  $\{\vec{v}, \vec{\omega}\}$  を最終的な推定パラメータとする。
4. 最後に、各点の奥行き  $z_i^{-1}$  を推定する。

$$\tilde{z}_i^{-1} = \max \left( \frac{\vec{v}^T F_i^T (\vec{d}_i - G_i \vec{\omega})}{\vec{v}^T F_i^T F_i \vec{v}}, 0 \right) \quad (4)$$

Moving object detection with a moving camera by camera-parameter clustering

<sup>†</sup> Satoshi Hara, Shuji Hashimoto, School of Science and Engineering, Waseda University

<sup>‡</sup> Ryo Saegusa, The Italian Institute of Technology

### 2.3 動物体候補の検出

カメラ動きパラメータ  $\{\tilde{v}, \tilde{\omega}\}$  と  $F_i, G_i$  の間には式(2)の関係が成り立つので、式(2)のカメラ動きパラメータ及び  $z_i^{-1}$  を推定値に置き換えた次式を用いて動物体候補の検出を行う。

$$e_i = \tilde{z}_i^{-1} \tilde{v}^T (\hat{F}_i^T \tilde{d}_i + \hat{F}_i^T G_i \tilde{\omega}) \quad (5)$$

各フローについて  $e_i$  が閾値  $\hat{e}$  を超えたものを動物体の候補として検出する。

### 2.4 履歴情報を用いた動物体検出

一般に、観測フローには誤追跡フローが含まれ、またパラメータのクラスタリングが失敗し適切な推定パラメータが求まらないことも起こりうる。そこで、動物体の検出候補に関する履歴情報を用いて動物体検出を行くこととする。

1. 過去  $s$  フレームにおいて動物体の候補となった割合が  $\alpha$  以上のフローを動物体として検出する。
2.  $a$  フレーム連続で動物体の候補でなかったフローは非動物体とする。
3.  $b$  フレーム連続で動物体の候補となったフローは動物体とする。

### 3. 実験

実験では、カメラに対して右方向に移動する人物を、ビデオカメラを手で把持して前進しながら撮影した動画像 (720 x 480[pixels], 30fps, 68 フレーム) を用いた。特徴点の検出および追跡は Kanade-Lucas-Tomasi の方法[6]により、毎フレームごとに新たな特徴点を追加しながら行った。

パラメータ推定は  $n=12, m=100$  として行った。クラスタリングの手法には群平均法を採用し、パラメータ空間上での距離尺度としてマハラノビス距離を用いた。クラスタ間の距離が 0.5 より大きくなったときの最大クラスタの平均パラメータを推定パラメータとした。また、動物体検出のパラメータは  $\hat{e}=10, s=10, \alpha=0.6, a=b=3$  とした。

上記の条件で推定を行った結果を図 1 及び図 2 に示す。白円が動物体として検出されたフローであり、白円の直径は過去  $s$  フレームで動物体として推定された割合を表す。図 1 では、動物体 (人物) の周辺に白円が観測され、動物体を適切に検出していることが確認できる。計算時間は 80(ms/frame) である。

### 4. おわりに

透視射影カメラモデルを用いた動画像からの動物体検出手法について述べた。提案手法では、

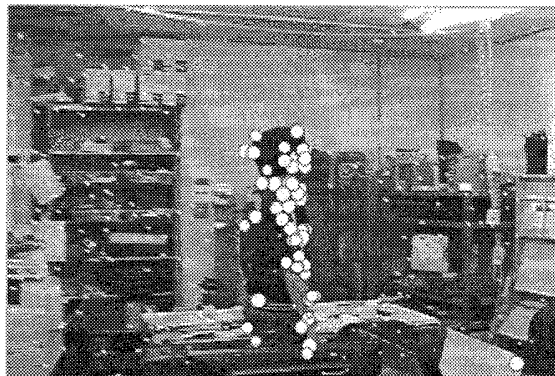


図 1. 動物体検出結果

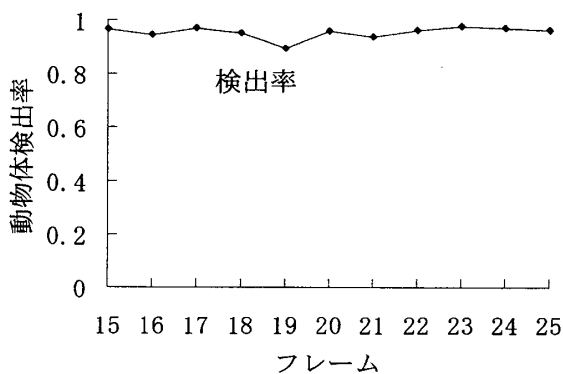


図 2. 動物体検出率

カメラ動きパラメータのクラスタリングを用いたロバスト推定によりカメラの動きを推定し、そこから導かれる拘束条件からのずれが大きい点を動物体として検出する。実動画を用いた実験により、本手法の有効性を確認した。現在、本手法のロボットへの搭載を検討している。

### 参考文献

- [1] J.P.Costeria, T.Kanade, "A multibody factorization method for independently moving objects", Int. J. Computer Vision, 29(3), pp.159-179, Sept.1998.
- [2] 讓田賢治, 坪内貴之, 菅谷保之, 金谷健一, "移動ビデオカメラ画像からの運動物体の抽出", 情報処理学会研究報告 2004-CVIM-143-6, pp.41-48, March 2004.
- [3] 菅谷保之, 金谷健一, "複数運動分離の教師なし学習による多段階最適化", 情報処理学会研究報告 2003-CVIM-138, pp.185-192, May 2003
- [4] R.Vidal, Y.Ma "A Unified Algebraic Approach to 2-D and 3-D Motion Segmentation", J. Mathematical Imaging and Vision, 25(3), pp.403-421, Oct.2006
- [5] 徐剛, 辻三郎, "3次元ビジョン", 共立出版社, 1998
- [6] C.Tomasi, T.Kanade, "Detection and Tracking of Point Features", CMU Tech. Rep. CMU-CS-91-132, Apr.1991.