

形状特徴と動作方向のクラスタリングに基づく行動の種類獲得

瀧田 悠一 太原 育夫

東京理科大学 理工学部情報科学科

1 はじめに

現在広く使われている映像監視システムは、監視対象の行動を知ることが目的としている。しかし、複数の光景においてどんな行動が起きているかを常に把握し続けることは困難である。そこで移動体の行動の種類を生成し、自動識別に用いる研究が行われている。Tao[1]らは、位置や矩形サイズなど空間的な特徴を用いてクラスタリングすることで、位置に起因する行動を生成する手法を提案している。しかし、広場や歩道など、行動が変化する位置にまとまりがない環境では、期待する行動の種類が獲得できない。

そこで本稿では、動作軌跡の形状と方向の類似性に基づいて、行動の種類を生成し自動識別に用いる方法を提案する。動作領域の軌跡に注目することで、“動作の状態”に対応した種類を生成することができる。本稿では、提案手法の概要について述べ、最後に屋外環境における実験を通して提案手法の有効性を示す。

2 行動の種類を生成

2.1 移動体の抽出と追跡

まず、移動体のシルエット（前面領域）を生成する。本研究では実環境における雑音と影を背景とみなすために、背景モデルとして適応背景混合モデル [2] を採用している。次に入力画像に対して、Mean-Shift 法 [3] により前フレーム中の移動体矩形の対応付けを行うことで、識別番号 (ID) を引き継いだ移動体矩形を得る。ここで対応付けが行われなかった一定サイズ以上の連結領域は、新たな移動体として新しい ID を割り当てている。

2.2 ID に基づく動作履歴画像の生成

N フレーム間の移動体矩形の履歴から ID ごとに動作履歴画像 (MHI: Motion History Image) [4] を生成する (図 1)。時刻 t における識別番号 ID の移動体の MHI は以下の式により求める。

$$MHI_N(x, y) = \begin{cases} \tau & \text{if } D(x_{ID}, y_{ID}, t) = 1 \\ 0 & \text{else if } MHI_N(x, y) < (\tau - N) \end{cases} \quad (1)$$

$D(x_{ID}, y_{ID}, t)$ は移動体矩形中の画素値、 τ は経過フレーム数、 N は最大履歴フレーム数である。ここで ID は一意に割り当てられるため、移動体同士が交差しても互いに独立して MHI を生成することができる。

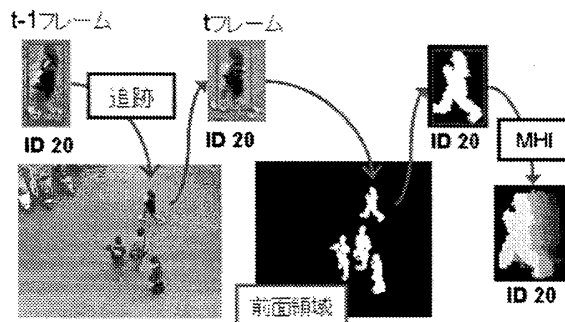


図 1: システムの流れ

2.3 形状特徴量の抽出

MHI の形状特徴として以下の特徴量を使用する。

(1) 矩形の幅と高さ

動作の大きさと方向に関係がある特徴量として MHI の幅と高さを用いる。MHI の連結領域を囲う矩形を求め、その幅 W と高さ H を使用する。

(2) 面積

動作の大きさと速さに関係がある特徴量として MHI の面積 A を用いる。

(3) 2 次重心モーメント

動作の進行方向に関係がある特徴量として X, Y 方向の 2 次重心モーメント M_{20}, M_{02} を用いる。

2.4 動作方向の抽出

MHI に対して Sobel フィルタ処理を水平・垂直方向に行うことで、方向ベクトル F_x, F_y を得ることができ、画素ごとに角度 $\phi = \arctan \frac{F_y}{F_x}$ が求まる。次に、MHI の M 個の局所領域から局所的な動作方向 $\bar{\phi}_i$ ($i = 1, \dots, M$) を生成する。これは以下の式で求まる。[4]

$$\bar{\phi}_i = \phi_{ref} + \frac{\sum_{x,y} |\phi(x, y) - \phi_{ref}| \times \alpha MHI_N(x, y)}{\sum_{x,y} \alpha MHI_N(x, y)} \quad (2)$$

ここで ϕ_{ref} は局所領域内で最頻出する角度、 $\alpha MHI_N(x, y)$ は MHI 上での値が最近のものほど大きな重み付け (0 から 1) を返す関数である。

Acquisition of Types of Actions based on Clustering of Shapes and Movement Directions.

Yuichi TAKITA and Ikuo HAHARA

Department of Information Sciences, Tokyo University of Science

2.5 特徴量のクラスタリング

実環境では行動ごとに発生する回数に差があり、かつクラスタの境界付近に位置する曖昧なデータが多いため、本研究では、特徴量の分布に正規混合分布を仮定した上で EM アルゴリズム [5] を用いる。

(1) 形状特徴を用いたクラスタリング

抽出した形状特徴量は測定単位が異なるため、標準化した後に 5 次元の特徴ベクトル $[W, H, A, M_{20}, M_{02}]$ としてクラスタリングを行う。この結果、動作軌跡の形状が似ている行動のクラスタが K 個生成される。

(2) 動作方向を用いたクラスタリング

MHI から得られた M 個の動作方向 ϕ_i を水平・垂直方向成分 F_x, F_y に分解することで、 $2M$ 個の特徴量 $[F_x^1, F_y^1, \dots, F_x^M, F_y^M]$ を得る。これを用いて、形状特徴クラスタに属するデータに対してそれぞれクラスタリングを行うことで、 C 個の動作方向が類似するクラスタが生成される。

3 実験

提案手法の有効性を検証するため実験を行った。カメラとしては Panasonic 製ネットワークカメラ BB-HCM581 を使用し、建物入り口前の T 字路を、15 分、毎秒 30 フレーム、解像度 320×240 で撮影したものを実験データとした。また MHI に用いる最大履歴フレーム数は $N = 10$ 、局所領域数 $M = 25$ とした。

3.1 実験結果と評価

クラスタ数を徐々に増やしていき、最も意味的整合性が取れているクラスタ数を最適な行動の種類数とした。実験では $K = 4$ 、 $C = 2$ とした。生成された行動の種類は、移動体の動作の大きさと進行方向を反映している (表 1)。また歩行者が縦に歩きはじめれば位置に関わらず縦と識別されていることから、動作の状態に対応した行動の種類が獲得できている (図 2)。

1 人の歩行は縦・横ともに動作軌跡の形状がよく類似しており識別率は良い。しかし、鞆を持つ歩行者と自転車の斜め走行は 2 人の併進：左として、3 人の密集団と 2 台の自転車の併進は自転車の走行：右として誤識別しており、対応するクラスタの正答率と検出率は低くなっている。他に静止している人やとても遅い歩行の誤検出が発生している。

行動クラスタの誤分類 (鞆を持つ歩行者や自転車の斜め走行など) は、動作軌跡の大きさが似ている場合に発生している。これは大きさの比重が高くなるような特徴量に注目したからである。よって MHI 上の動作方向の分布など、大きさに不変でより形状を表現できるような特徴量も取り入れる必要がある。

表 1: 行動クラスタごとの正答率と検出率

クラスタ		発生数	検出数	正答率 [%]	検出率 [%]
形状	方向				
1 人の歩行：縦	上	526	538	97.8	98.9
	下	108	112	96.4	100
1 人の歩行：横	左	1054	989	100	93.2
	右	1241	1241	100	99.1
2 人の併進：横	左	89	203	43.8	100
	右	285	270	100	86.7
自転車の走行	左	105	61	100	57.1
	右	157	244	64.3	86.0

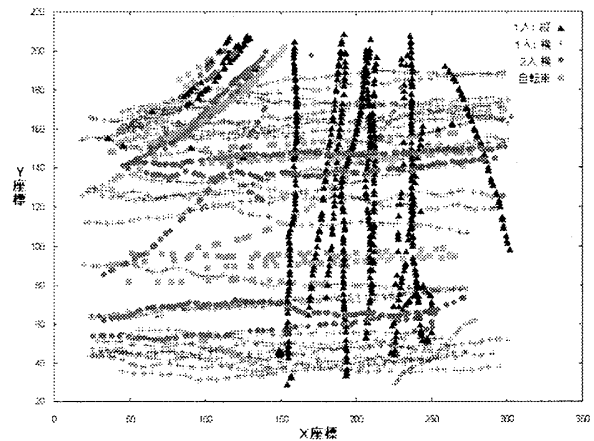


図 2: 光景上でのクラスタの分布

4 まとめ

本稿では、移動体の動作軌跡の形状と方向の類似性に着目することで、動作の状態に対応した行動の種類を生成する方法を提案した。今後の課題として、生成される行動種類の記述力を向上するために、動作方向の分布を形状特徴として用いることなどがあげられる。

参考文献

- [1] Tao Xiang, Shaogang Gong and Dennis Parkinson, "Autonomous visual events detection and classification without explicit objectcentred segmentation and tracking," BMVC, pp.233-242, 2002.
- [2] P.K.TraKulPong, R.Bowden, "An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection," 2nd European Workshop on Advanced Video Based Surveillance Systems, 2001.
- [3] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," IEEEconf.on Computer Vision and Pattern Recognition, pp.II:142-149, 2000.
- [4] Davis.J, Bobick.A, "The representation and recognition of human movement using temporal templates," Proceedings of Computer Vision and Pattern Recognition, pp.928-934, 1999.
- [5] C.M.Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.