

# 単一カメラによる動画像からの複数人物の自動検出に関する研究 -尤度比較による画像中人物の増減検出-

Research on automatic tracking of multiple people from dynamic image with single camera  
- Detection of increase or decrease the number of people in images by using likelihood ratio comparison -

小関 英剛\*1      森下 壮一郎\*2      浅間 一\*2  
KOSEKI, Hidetaka      MORISHITA, Soichiro.      ASAMA, Hajime

\*1東京大学 工学系研究科  
School of engineering The University of Tokyo

\*2東京大学人工物工学研究センター  
RACE, The University of Tokyo

This paper proposes a new method for positioning and counting moving objects in footage of single camera. We have proposed a method which detects moving objects in the dynamic picture image with a Gaussian mixture model. Although it does not accommodate the change of the number of moving objects automatically. We overcame the shortcoming with our new proposed method and confirmed the validity of the new method with one footage.

## 1. はじめに

公共空間等の広い場所における歩行中の人物への適切なサービス提供や不審者検出を行うには、人物の移動軌跡を解析し意図を推定することが有効である [1]。動画像から人物の移動軌跡を取得するには、(1) 動画像から移動物体領域を抽出し、(2) 人物の足元座標移動軌跡を取得する手続きが必要である。(1) については、適応的背景差分法を用いた照明条件に頑健な移動物体領域抽出手法 [2] が提案されている。(2) については、人物領域の重心からの足元位置推定手法 [3] が提案されている。また、広瀬らによって混合分布推定による複数人物検出 [4] が提案された。しかし、[4] の手法では、複数人物の人数を自動的に判別できるには至っていない。以上から、本研究では単一カメラより得た動画像を用いた複数人物移動軌跡の自動判別の手法提案及び、今後の展望について述べる。

## 2. 移動人数の自動判別アルゴリズム

移動人物の抽出には背景差分法 [2] を用いた。本研究では移動物体領域の重心と分散から人物領域を覆う楕円を得る。その楕円の長軸方向を足元とみなす [3]。動画像中の複数人物の人数が把握できている際の人数分の重心と分散を求めるには、EM アルゴリズムによる混合分布推定 [4] を行う。

混合分布推定において、EM アルゴリズムはデータ点に適当な初期値混合度  $\alpha$ 、重心  $\mu$ 、分散行列  $\Sigma$  が与えられたとき、最尤推定によってもっともらしい  $\alpha$ 、 $\mu$ 、 $\Sigma$  を得る [5]。また画素数を  $n$  とすると得られたパラメータを元に画素  $x_n$  について  $k$  番目のガウス分布に属する尤度  $p(x_n; \mu_k, \Sigma_k)$  を求める。

クラス数  $M_1 = 1$  とし、動画像に混合推定分布推定を適用した結果を Fig.1 に示す。クラスの変動がなければ Fig.1(a) のように 1 人の人物領域に対してひとつの楕円が追従していることが確認できる。また、クラス数が既知である場合にも正しい人物同定が行える。しかし、画面内に新たな人物が入ると、1 つのクラスに対し 2 人分の画素の重心と分散が適用され Fig.1(b) のように人物の誤検出が生じる。

### 2.1 提案手法

本研究で用いた動画像において、人物の増減に対してフレームレートは十分大きい。そこでフレームごとに画面内に入る人



(a) success      (b) failed

図 1: Calculation results of EM algorithm

数は、最大 1 人ずつであると仮定した。画面外へ出る場合も同様である。それに対応したクラスの追加と削除を、以下の手法を用いて自動的に行うことで複数人物の自動検出を行う。

### 2.2 クラスの追加

クラスの追加には以下の 2 つの手法を組み合わせることで対応する。

#### 2.2.1 学習データによる楕円領域の比較

Fig.1(b) からわかるように複数人物を 1 クラスと誤認識すると、移動物体領域の分散は適切に追従している状態より大きくなる。森下らは学習データと実際に求めた楕円のパラメータ比較を行い楕円領域内の人物の人数判定を行った [6]。

実測値による楕円の長軸短軸長をそれぞれ  $S_1$ 、 $S_2$ 、学習データの推測値による楕円の長軸短軸長をそれぞれ  $S'_1$ 、 $S'_2$  とし、以下の式を用いて評価を行う。

$$\alpha = (S_1 S_2) / (S'_1 S'_2) \quad (1)$$

式 (1) の閾値を  $\theta_1$  とし、人数増加の判定を行う。本研究で用いた学習データは [6] を利用した。

#### 2.2.2 クラス増減時における尤度の変化

画素  $x_n$  がクラス  $k$  に属する尤度  $p$  は  $x_n$ 、 $\mu$ 、 $\Sigma$  を用いると、以下の式で表せる。

$$p(x_n; \mu_k, \Sigma_k) = \frac{\exp(-(x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k) / 2)}{2\pi |\Sigma_k|^{1/2}} \quad (2)$$

楕円領域がクラスを追跡しているとき尤度の和は

$$\beta_n = \sum_{k=1}^K p(x_n; \mu_k, \Sigma_k) \quad (3)$$

で表せる。適当な閾値  $\theta_2$  を設定すると、人物画素に楕円領域が適切に割り当てられている場合、 $\beta_n > \theta_2$  となる。しかし動画像中に新たな人物の画素が入ると、それらの画素は外れ値と

して  $\theta_2$  以下の値となる。以上より  $\alpha > \theta_1$  及び、 $\beta_n < \theta_2$  となる画素数の合計  $\gamma$  が一定割合 ( $\theta_3$ ) 以上存在するとき、すなわち  $\theta_3 < \gamma/n$  のとき新たなクラスと判断できる。

### 2.3 クラスの削除

一方で、画面内の人数が減少した場合、Fig.2のように1人の画素に複数の楕円領域が割り当てられる。このとき  $\alpha < \theta_1$ 、 $\gamma \approx 0$  となり外れ値が存在しない。そのため1種類のクラス数の計算では人物の減少に対応できない。そこで通常の計算に用いている  $M_1$  とは別に、人数の減少を確認するためのクラス数  $M_2 = M_1 - 1$  で人物の追跡を行う。 $M_1$ 、 $M_2$  において  $\alpha$ 、 $\beta_n$  の外れ値が観測できない場合、 $M_1$  のクラス数が画面内の人数よりも多いと判断でき、クラスの削除を行なう。



図 2: Calculation results of EM algorithm

### 2.4 提案手法の流れ

画面内に人がいない状態から検出を行う。また人物を検出し始めた  $t$  フレーム目のクラス数を、 $M_1(t) = M_2(t) = 1$  とする。毎フレーム EM アルゴリズムを適用し、以下の条件に従いクラスの増減を行う。

1. クラス数が  $M_1(t) = M_2(t) = 1$  のとき
  - $\theta_1 < M_1(t)$  の  $\alpha$  かつ  $\theta_3 < M_1(t)$  の  $\gamma/n$   
→  $M_1(t+1)$  を 1 追加
2. クラス数が  $M_1(t) = M_2(t) + 1$  のとき
  - $\theta_1 < M_1(t)$  の  $\alpha$  かつ  $\theta_3 < M_1(t)$  の  $\gamma/n$   
→  $M_1(t+1), M_2(t+1)$  を 1 追加
  - $\theta_1 < M_1(t), M_2(t)$  の  $\alpha$  かつ  $M_2(t) > 1$   
→  $M_1(t+1), M_2(t+1)$  を 1 削除
  - $\theta_1 < M_1(t), M_2(t)$  の  $\alpha$  かつ  $M_2(t) = 1$   
→  $M_1(t+1)$  を 1 削除

クラス数は必ず 1 以上でなければならないので、上記のように場合分けして  $M_1 = 1$  のときに  $M_2 = 1$  とする。また画面内人数が 1 の場合には、クラスの削除が必要なくなるため  $M_1 = M_2 = 1$  として検出を行う。

## 3. 実証実験

### 3.1 実験環境

提案手法の有効性を検証するために、東京大学柏図書館の一階入口上部にカメラを設置して得られた映像に提案手法を適用する実験を行った。

### 3.2 実験方法

背景画像は M-estimation による適応的背景推定手法 [2] を用いて作成した。また実験で用いたパラメータは経験的に表 1 のように設定した。また、動画は画面内の人数が 1 人 → 2 人 → 1 人と変化するものを使用した。

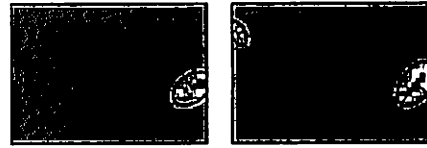


図 3: Increase of the number of classes

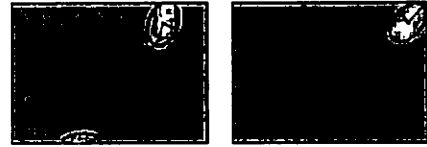


図 4: Decrease of the number of classes

### 3.3 実験結果

画面の左上から人が進入した場合 (Fig.3) と、画面の左下から人が画面外へ出た場合 (Fig.4) を示す。画面内の人数の変化に応じてクラス数が適切に増減していることが確認できた。

### 3.4 まとめと展望

本研究では単一カメラによる動画像中の複数人物の自動検出方法を提案した。検出を行うために、移動物体領域の分散の推定値と実測値を比較し複数人物の判定を行った。人数の増加に対応することで、画面内に新たに人物が入ってきた際に生じる尤度の低い点群を新たな人物領域として設定した。また人数の減少に対応するため、評価用にクラス数を 1 減らしたパラメータで計算することにより人数の減少に対応した。

今後の展望として多数の歩行者がいる動画像にも提案手法を適用し、有効性を確認する。また、非線形モデルを用いたオクルージョンに頑健な手法についても検討する。

## 参考文献

- [1] 西村 他: “隠れマルコフモデルによる歩行パターンからの目的地推定 -推定結果評価手法の検討-”, 日本機械学会, ROBOMEC07, Vol.7, No.2, pp. 2P1-C10(1)-2P1-C10(3), 2007.
- [2] 上端 他: “動画像における移動物体抽出のための相関分析による影領域判定に関する研究”, 計測自動制御学会, 第 7 回 SICE SI 部門講演会講演論文集, pp. 1348-1349, 2006.
- [3] 森下 他: “固定カメラ画像における人物領域の重心からの足元位置同定に関する研究”, 計測自動制御学会, 第 7 回 SICE SI 部門講演会講演論文集, pp. 1380-1381, 2006.
- [4] 広瀬 他: “混合分布推定による動画像からの複数人物の位置検出”, 第 8 回 計測自動制御学会 SI 部門, 3L2-2, 2007.
- [5] AP Dempster 他: “Maximum likelihood from incomplete data via the em algorithm”, Journal of Royal Statistical Society. Series B, Vol.39, No.1, pp.1-38, 1977
- [6] 森下 他: “動画像中の人数判定のための移動物体領域の分布パラメータ推定”, 第 8 回 計測自動制御学会 SI 部門, 3L2-1, 2007

表 1: Experimental settings

フレームレート	15 [fps]
サンプリング数 $n$	230
楕円領域に複数人いるか判断する閾値 $\theta_1$	2.7
$\beta_n$ の外れ値 $\gamma$ に用いる閾値 $\theta_2$	$1.0 \times 10^{-6}$
クラス追加に用いる外れ値の割合の閾値 $\theta_3$	0.1