

制約付き EM アルゴリズムによる対象個数推定

瀬藤 英隆[†] 加藤 文和[†] 和田 俊和[†]

[†] 和歌山大学 システム工学研究科

〒 640-8510 和歌山県 和歌山市 柴谷 930

E-mail : hidetaka@vrl.sys.wakayama-u.ac.jp, t.kato@ieee.org, twada@vrl.sys.wakayama-u.ac.jp

あらまし

本稿では、個数が未知の対象に対して、EM アルゴリズムに基づきその個数を推定しながら画像中の対象を検出・追跡する方法について述べる。EM アルゴリズムは、観測データに対して複数の分布モデルを当てはめる手法である。背景差分により画像から検出した変化画素を観測データとし、各検出対象を分布モデルと見なすことで複数対象の検出・追跡に適用することができる。さらに、MDL を用いて分布モデルの複雑さとデータへの当てはまりの良さを評価することにより、個数推定を同時に行う方法も提案されている。しかし、単純な EM アルゴリズムと MDL の組み合わせでは、EM アルゴリズムの初期値依存性や対象の近接、立ち位置の違いによる対象の見かけの大きさが変化する等の影響を受け、個数推定が正しく行えないケースがある。そこで、本研究では各対象領域の大きさや形状は画像中の位置と強い相関があることに着目し、正規分布の中心座標に対する共分散行列パラメータの関係を 5 次元空間内の超平面として学習し、この平面内で分布のパラメータを推定することで、個数推定を安定化することができる事を実験を通じて示す。

キーワード EM アルゴリズム, MDL, 分布の形状制約

Object Counting based on Constrained EM-Algorithm.

Hidetaka Seto[†], Takekazu Kato[†], Toshikazu Wada[†]

[†]Graduate School of Systems Engineering, Wakayama University,
930 Sakaedani, Wakayama, 640-8510, Japan

E-mail : [†] hidetaka@vrl.sys.wakayama-u.ac.jp, t.kato@ieee.org, twada@vrl.sys.wakayama-u.ac.jp

Abstract This paper presents a method for detecting and tracking unknown number of targets in image sequences based on EM algorithm. EM algorithm can fit mixture of distributions to observed data. Applying this algorithm to the foreground pixels obtained by background subtraction, we can estimate the assumed number of object positions, and shape parameters. Also, we can estimate the number of objects based on MDL criterion, where the valid number of objects provides the minimum DL. A simple combination of EM algorithm and MDL criterion, however, will produce incorrect object number. This is because the constraint on shape and size of each object is too loose and EM algorithm may estimate inconsistent fittings. We propose a robust method for counting the number of objects by introducing a constraint between the object position and shape parameters, i.e., covariance matrix. This constraint can be represented by a hyper plane in 5-D space spanned by positions (x,y) and covariance parameters ($\sigma_x^2, \sigma_{xy}, \sigma_y^2$). By confining the distribution parameters on this plane in EM algorithm, we can robustly estimate the number of objects.

Keywords EM algorithm, MDL, form Constrained

1 はじめに

カメラ映像から個数が未知の対象の検出・追跡を行なうことが出来れば、デパートなどで入店あるいは出店する人数の時間毎の傾向や、授業における出席人数など様々な分野に应用可能である。このような複数対象の検出・追跡を行なう際、対象の個数の推定が重要な問題になる。本研究では、個数が未知である対象を検出・追跡するために、その個数推定を行なう事を目的とする。

本研究では、画像上で変化している画素を検出し、それらを対象毎にクラスタリングすることによって、対象の検出、追跡を行なう。変化画素のクラスタリング手法としては正規分布の混合分布を推定する EM アルゴリズムを用いる。

EM アルゴリズムを用いた対象検出では、画像中で変化した画素を観測データとして、それらに混合正規分布を当てはめることによって検出を行なう。この手法では対象検出と追跡を同時に行なう

ために、追跡に失敗しても自動的に再検出が行えるという利点がある。EM アルゴリズムは分布モデルの個数が既知という制限があるが、MDL と組み合わせることで個数の推定を同時に行うことができる [1]。この手法では、分布モデルの記述長と残差の記述長の和を表す情報量を最小化することにより対象の個数を推定する。

しかし、この手法では、パラメータに制約のない正規分布モデルを用いているため、複数の対象が近づいた場合などに、EM アルゴリズムによって対象の形状としてあり得ない形状や大きさの分布を推定し、その結果個数も正しく推定できないというケースが多数存在するという問題がある。

そこで、本研究では追跡対象が決まっている場合、その観測データの分布形状はその位置との相関が強いことに着目する。具体的には対象の位置に応じた観測データの分布形状をあらかじめ学習しておき、EM アルゴリズムで推定する分布の形状に制約を与えることで、クラスタリングと MDL

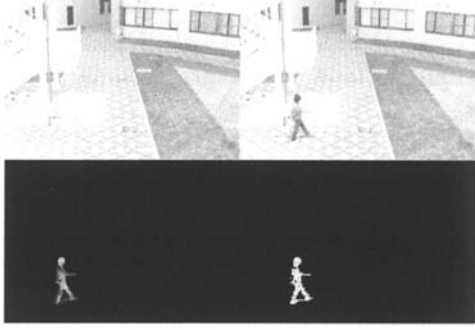


図 1: (左上) 背景画像, (右上) 入力画像, (左下) 背景と入力画像との差分, (右下) サンプリングにより獲得した座標データ

による個数推定を安定化する手法を提案する。

以下第 2 章では準備として EM アルゴリズムと MDL を用いた複数対象の検出と個数推定の方法について述べ、第 3 章では提案手法として位置に応じた分布形状の学習とそれに基づく分布形状の制約による個数推定の安定化について述べる。第 4 章に実験結果と考察を述べ第 5 章でまとめる。

2 従来研究

2.1 EM アルゴリズムによる複数対象検出

EM アルゴリズムによる複数対象の検出では、まず入力画像から背景差分によって変化画素を検出し、それらを観測データとして対象毎にクラスタリングすることにより、対象検出を実現する。本研究では、図 1 に示すような、背景差分の結果を EM アルゴリズムの観測データとして用いる。また、前フレームのクラスタリング結果を次フレームの EM におけるクラスタリングの初期値とすることで推定する分布をフレーム間で対応付ける事ができ、追跡を行なうこともできる。

EM アルゴリズムは、複数のパラメトリックな分布から発生した観測データに対して、各分布のパラメータと各データがどの分布から発生したかを最尤推定によって求める数値解法であり、条件付き期待値を計算する E(Expectation) ステップと、その期待値を最大化するモデルパラメータを推定する M(Maximization) ステップを繰り返し、局所最適解を得る方法である。

N 個の観測データを $\mathcal{Y} = \{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(N)}\}$ としたとき、観測データが K 個の正規分布からなる混合正規分布に従うとすると、 k 番目の分布は次

式のように表される。

$$P(\mathbf{y}|\Theta_k) = \frac{1}{(2\pi)^{\frac{M}{2}}} |R_k|^{-\frac{1}{2}} \exp(E) \quad (1)$$

$$E = -\frac{1}{2} (\mathbf{y} - \mu_k)^T R_k^{-1} (\mathbf{y} - \mu_k) \quad (2)$$

ただし、 μ_k は分布モデル k の重心、 R_k は分布モデル k の分散共分散行列、 M はデータの次元数である。

この場合の EM アルゴリズムの流れは以下のようになる。

初期化:

分布のパラメータの初期値を $\Theta_k^{(0)}, \pi_k^{(0)} = \frac{1}{K}$ を与え、ループカウンタ l を 1 とする。ここで、対象追跡の場合には初期値 $\Theta_k^{(0)}, \pi_k^{(0)} = \frac{1}{K}$ は、前フレームの推定結果とする。

E ステップ:

観測データ $\mathbf{y}^{(n)}$ が分布モデル k に従う確率 (分布モデル k の事後分布) を計算する。

$$m_k^{(n,l)} = \frac{p(\Theta_k^{l-1}|\mathbf{y}^{(n)})}{\sum_{i=1}^K \frac{P(\mathbf{y}^{(n)}|\Theta_i^{(l-1)})\pi_i^{(l-1)}}{P(\mathbf{y}^{(n)}|\Theta_i^{(l-1)})\pi_i^{(l-1)}}} \quad (3)$$

M ステップ:

事後分布の重み付けによる各分布の生起確率 $\pi_k^{(l)}$ およびそれらのパラメータ $\Theta_k^{(l)} = (\mu_k^{(l)}, R_k^{(l)})$ のアップデートを行なう。

$$\pi_k^{(l)} = \frac{1}{N} \sum_{n=1}^N m_k^{(n,l)} \quad (4)$$

$$\mu_k^{(l)} = \frac{1}{N\pi_k^{(l)}} \sum_{n=1}^N m_k^{(n,l)} \mathbf{y}^{(n)} \quad (5)$$

$$R_k^{(l)} = \frac{1}{N\pi_k^{(l)}} \sum_{n=1}^N m_k^{(n,l)} \kappa^{(n)} \kappa^{(n)T} \quad (6)$$

$$\kappa^{(n)} = \mathbf{y}^{(n)} - \mu_k^{(l)}$$

終了判定:

分布パラメータ $\pi_k^{(l)}, \Theta_k^{(l)}$ が前のイタレーションの推定結果 $\pi_k^{(l-1)}, \Theta_k^{(l-1)}$ から $\|\Theta_k^{(l)} - \Theta_k^{(l-1)}\| < \epsilon$ を満たせば終了。それ以外は $l := l + 1$ として E ステップに戻る。なお ϵ は式 (10) で計算される閾値である。

この結果得られるパラメータを用いて、全体の分布は次式で表される。

$$P_w(\mathbf{y}|\theta) = \sum_{k=1}^K \pi_k P(\mathbf{y}|\Theta_k) \quad (7)$$

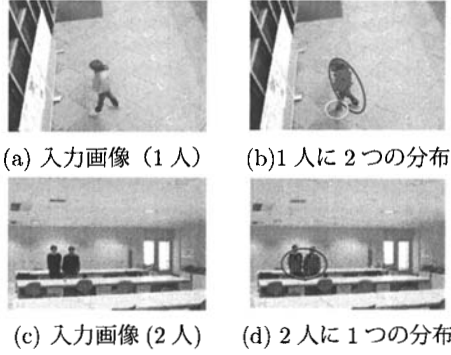


図 2: 従来手法の問題点

2.2 MDL による個数推定

EM アルゴリズムは、分布の個数が既知であるという制限があるが、分布の個数を情報量規準 (DL (Description Length)) を用いて推定する手法が提案されている [1]。この手法では、分布モデルの個数 K の値を変えて EM アルゴリズムを実行し、その都度 DL を計算し、DL がもっとも小さくなったとき (MDL - Minimum Description Length) の個数を推定結果とする手法である。DL が最小になるのはデータに対する当てはまりがもっともよく、かつモデルが単純 (分布モデルの個数が少ない) 場合である。DL の計算式は次式のように与えられる。

$$MDL(K, \Theta) = -\sum_{n=1}^N \log(\sum_{k=1}^K p(\Theta_k | y^{(n)}) \pi_k) + \alpha \quad (8)$$

$$\alpha = \frac{1}{2} (K(1+M + \frac{(M+1)M}{2}) - 1) \log(NM) \quad (9)$$

ただし、 α は分布モデルの個数に対するペナルティ項である。

EM アルゴリズムでの反復処理における M ステップにてパラメータの更新を行うと $MDL(K, \Theta)$ の値は減少する。前回の値からの減少量が以下の閾値 ε よりも小さくなったときに EM アルゴリズムの反復処理を打ち切り、クラスタリング結果とそのときの $MDL(K, \Theta)$ の値を記録する。

$$\varepsilon = \frac{1}{100} (1 + M + \frac{(M+1)M}{2}) \log(NM) \quad (10)$$

クラスタリング数 K を変えて再び EM アルゴリズムを実行し、 $MDL(K, \Theta)$ の値が最小になるときの K とクラスタリング結果を出力結果として採用する。

3 提案手法

3.1 基本アイデア

EM アルゴリズムと MDL により人数が未知の複数人物を検出、追跡する場合、1 人の人物に対応

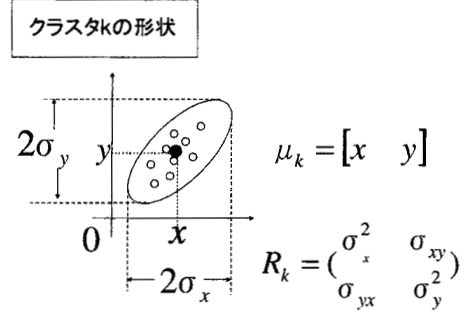


図 3: パラメータと分布形状との関係

する観測データに対して 1 つの分布を当てはめて推定する必要がある。しかし MDL による個数推定では、パラメータに制約のない正規分布モデルを用いているため、正しく人数を推定できない場合が多く存在する。例えば、図 2 に示すように、一人の人物に対して複数の分布を当てはめてしまう場合 (b) や、近接した複数の人物に対して 1 つの分布を当てはめてしまう場合 (d) がある。

監視カメラなどで撮影した人物の位置と分布形状すなわち対象の形と大きさには強い相関がある。そこで本研究では、あらかじめ追跡対象の分布形状を学習しておき、それを用いて EM アルゴリズムで推定する分布形状に対して制約をかけることで、上に述べたような誤った推定を避け、MDL による個数推定を安定化させる方法を提案する。分布形状に制約をかけて EM による推定を行なうことで、正しくない個数のクラスタリング結果のデータに対する当てはまりが悪くなり、情報量が増加する。その結果その個数は個数推定の選択肢から除外されやすくなり、最適な個数を推定しやすくなると考えられる。また、実際のシーンでは、画像中の位置によって奥行きの違いがありその分布形状や大きさも変化する。そこで、画像中の場所に応じた分布形状の制約を与えるため、対象の位置と分布形状との関係を学習によって求める。

3.2 学習に基づく制約付き EM アルゴリズム

EM アルゴリズムで推定する分布の形状を制限する。そのために、分布のパラメータの分散共分散 R_k について注目する。本研究では画像上の座標 (x, y) の推定を行うため、EM アルゴリズムにおける分散共分散の形式は 2 行 2 列になっており、そのパラメータと分布形状の関係は図 3 のようになっている。

対角成分 σ_x^2 , σ_y^2 の平方根 σ_x , σ_y は図 3 で表すようにそれぞれ直線上の x 方向と y 方向の標準偏

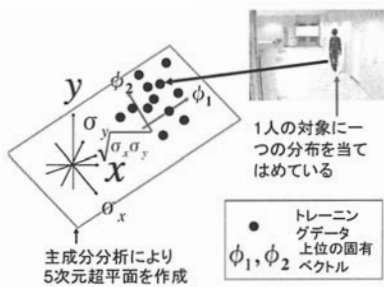


図 4: 超平面の作成

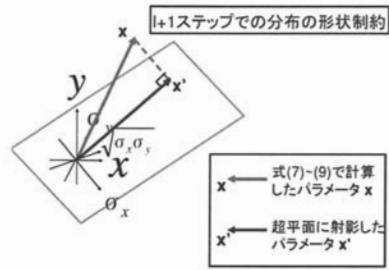


図 5: 超平面への射影

差を表し、分布を x 軸, y 軸に写像した時の幅になり, σ_{xy} の平方根 $\sqrt{\sigma_{xy}}$ は分布の傾きを表している。

3.2.1 主成分分析を用いた学習

1人の追跡対象に対して1つの分布推定したときの分布の重心 x, y と形状パラメータの分散共分散 $\sigma_x, \sqrt{\sigma_{xy}}, \sigma_y$ を5次元のベクトルとして、主成分分析によってこれらの関係を求め、固有値の高い固有ベクトルで張られる図4のような超平面を求める。



図 6: 屋内実験

3.2.2 超平面への射影を用いた形状制約

EM アルゴリズム内の M ステップにおけるパラメータのアップデートを行なう毎に、上記の主成分分析を用いて求めた超平面への射影を行なうことで分布形状への制約をかける。具体的には2.1章のEM アルゴリズムにおける M ステップ内の式(5), 式(6)で計算した形状パラメータ μ_k, R_k の各要素を5次元のベクトル $\mathbf{x} = \{x, y, \sigma_x, \sqrt{\sigma_{xy}}, \sigma_y\}$ として、式(11)を用いて図5のように学習で求めた超平面への射影を行なう。なお $\bar{\mathbf{x}}$ は \mathbf{x} の平均を, ϕ_i は対応する固有値の高い順に並べられた固有ベクトルを表す。

$$\mathbf{x}' = \sum_{i=1}^2 \phi_i \phi_i^T (\mathbf{x} - \bar{\mathbf{x}}) + \bar{\mathbf{x}} \quad (11)$$

この射影後のパラメータ \mathbf{x}' が収束するまでEM アルゴリズムは反復計算を行なう。

4 実験結果

実験では様々なシーンに対して、従来の分布形状に対して制約を行わないEM アルゴリズムを用いた場合と提案手法を適用した場合とを比較し、提案手法の有効性を確認する。本研究では、"1~(前

フレームの人数+1)"の数を候補として個数推定を行なった。ただし前フレームの人数が1人の場合のみ候補となる数は"1, 2, 3"とした。これは候補となる数が"1, 2"のみだと局所解に陥る可能性があるためである。

4.1 屋内での実験結果

まず図6に示すような屋内にて、奥から2人と手前から1人の合計3人の人物が登場するという状況において、従来手法と提案手法を比較実験した。図7, 図8はそれぞれ従来手法, 提案手法での個数推定の様子を表し、画像の右上の数字が推定された個数を表す。この場面では画像内に3人の人物が登場している。従来手法では奥にいる2人に対して一つの分布を当てはめ、かつ手前にいる1人に対して3つもの分布を当てはめてしまい個数推定を誤っている事が確認できる。一方、提案手法では奥の2人に対して小さな分布を1つずつ当てはめ、手前の大きく写っている人物には大きな分布を当てはめており、個数推定も正しく行なう事が出来ている。図9, 図10に個数推定の結果を示す。図7, 図8はこのグラフの200フレームにあたる。また、個数推定に失敗したフレーム数の割合は表1のようになり、従来手法と比べて提案手法



図 7: 従来手法での実験結果の例 (屋内)

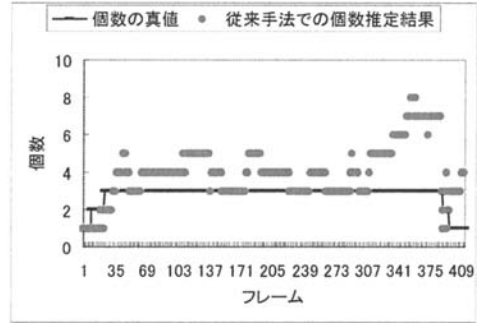


図 9: 従来手法での個数推定結果 (屋内)



図 8: 提案手法での実験結果の例 (屋内)

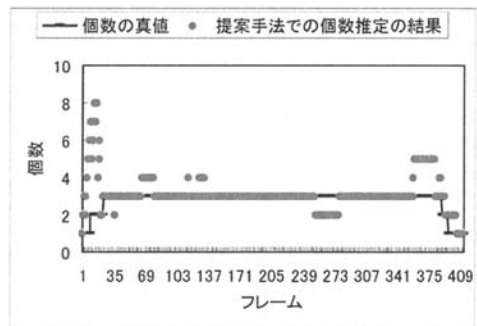


図 10: 提案手法での個数推定結果 (屋内)

は誤り率が約 46 ポイント改善していることが確認できる。

4.2 屋外での実験結果

次に図 11 に示すような高い位置から広く見下ろした視点で様々な位置から最大 8 人の人物が登場するという状況において従来手法と提案手法を比較した。図 12, 図 13 はそれぞれ従来手法, 提案手法での個数推定の様子を表し, 画像の右上の数字が推定された個数を表す。この場面では画像内に 8 人の人物が登場している。従来手法では狭い範囲に固まった複数の人物に対して一つの大きな分布を当てはめてしまい, 個数推定を誤っている。一方, 提案手法では一人に対して一つの分布を適切に当てはめ, 正しい個数を推定することが出来て

	誤り/全体 (フレーム数)	誤り率 (%)
従来手法	296/412	71.84
提案手法	106/412	25.73

表 1: 人数推定の誤り率 (屋内)

いる。図 14, 図 15 に個数推定の結果を示す。図 12, 図 13 はこのグラフの 370 フレームにあたる。個数推定に失敗したフレーム数の割合は表 2 のようになる。提案手法は従来手法に比べて誤り率が約 17 ポイント改善していることが確認できる。また, 従来手法では画像上の人が狭い範囲に固まると実際の人数よりも少ない数の分布を当てはめて



図 11: 屋外実験

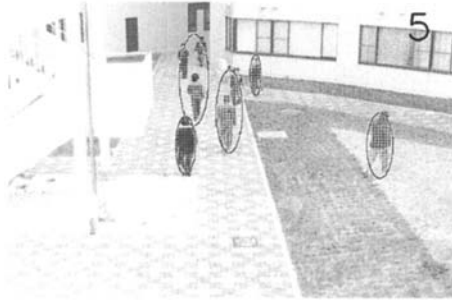


図 12: 従来手法での実験結果の例 (屋外)

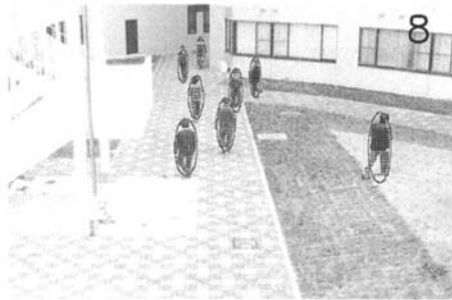


図 13: 提案手法での実験結果の例 (屋外)

しまうが、提案手法ではその問題を改善することが出来るということも確認した。

5 おわりに

本研究では、人数が未知の環境における複数人物の検出、追跡と人数推定を行なうことを目的とし、画像中の対象の形状や大きさに着目して、分布形状に制約を与えた制約付き EM アルゴリズムを提案した。トレーニングデータから対象位置と観測データの分布形状の関係を学習し、その関係を用いて分布の形状を制限しながら推定を行うことで、EM アルゴリズムと MDL の組み合わせによる人数推定を安定化する手法を提案した。この手法の有効性を確認するために、様々な視点の映像を用いて従来手法との比較実験を行ない、提案手法の有効性を確認した。

	誤り/全体 (フレーム数)	誤り率 (%)
従来手法	358/543	65.93
提案手法	263/543	48.43

表 2: 人数推定の誤り率 (屋外)

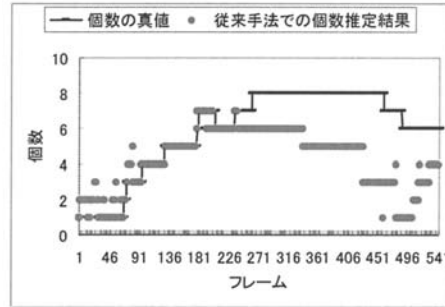


図 14: 従来手法による個数推定結果 (屋外)

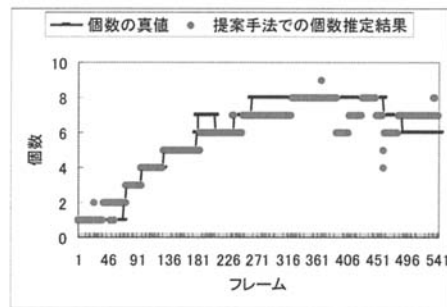


図 15: 提案手法での個数推定結果 (屋外)

今後の課題としては、EM アルゴリズムの特性でもある "初期値依存性" や "解の収束に対する保証がない" といった問題点を解決し、複数人物が近づいても各対象人物に対応する分布モデルが入れ替わらないように正しく追跡できるようにする方法について検討する。

参考文献

- [1] E.Render and H. Walker, "Mixture Densities, Maximum Likelihood and the EM Algorithm," SIAM Review, vol. 26, no. 2, April 1984.
- [2] Chris Stauffer, W.E.L. Grimson: "Adaptive background mixture models for real-time tracking", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99), vol.2 pp. 2246
- [3] M. Isard and A. Blake: "Condensation - conditional density propagation for visual tracking", International Journal of Computer Vision, 29, 1, pp. 5-28, 1998.