

追跡対象数の変動を確率的に考慮した擬似独立な複数物体追跡

宮野 博義[†] 石寺 永記[†]

[†] NEC 情報システムズ 〒 213-8511 神奈川県川崎市高津区北見方 2-6-1

E-mail: [†]h-miyano@cq.jp.nec.com, ^{††}ishide@bq.jp.nec.com

あらまし 複数物体の追跡技術は時系列画像処理の重要課題である。追跡対象の数が時々刻々と変化する場合の対応方法は主に二つある。一つは、追跡対象が取りうる数や場所を組み合わせで考慮しながら追跡する手法であり、比較的安定に追跡可能だが高速化には不向きである。もう一つは、物体が検出される毎に個別に追跡する手法であり、高速化には向くが誤検出の影響を受けやすく数の変動に不安定である。そこで、個別に追跡する方法を基に、数の変動に適切に対応することで精度と速度を両立させる手法を提案する。提案手法は、追跡対象毎にその追跡対象が実在している確率を設定し、物体検出結果を用いて確率値を逐次更新することで、追跡対象数の変化に対応する。実験で提案手法の優位性が確認できた。

キーワード 複数物体追跡、パーティクルフィルタ、最適割当問題

1. はじめに

物体追跡技術は、監視システムや機械による状況理解など、さまざまな用途における基本的な技術である。ある特定の物体を追跡する手法として、カルマンフィルタやパーティクルフィルタ [1] [2] といった追跡手法が提案され、幅広い用途に用いられている。

近年では、複数物体を同時に追跡する手法に関する研究が積極的に進められている。パーティクルフィルタを複数物体に拡張する方法としては、複数物体すべてに対する状態をまとめて 1 つの joint state と呼ばれる状態として表現し追跡するという手法が提案されている (文献 [3] [4] [5] [6] [7])。joint state を用いることで、複数物体を比較的安定に追跡することが可能であり、追跡対象数の変動にも対応可能である。しかし、本アプローチは複数の追跡対象の組み合わせを考慮するため高次元の状態空間となり、その結果精度を出すために膨大な数のパーティクルが必要になることから高速化には不向きであるという問題がある。

一方、複数物体をそれぞれ個々独立に、パーティクルフィルタで追跡するというアプローチも考えられ、比較的高速な処理を実現することができる (文献 [8] [9] [12] [13])。しかし、追跡対象数が時々刻々と変化する際に、追跡対象を新規に生成したり削除したりする処理が必要であるが、物体検出での誤検出の影響を受けやすく追跡対象数の変動には不安定になりがちである。

そこで、個別の追跡処理を基に、追跡対象数の変化に適切に対応することで、精度と速度を両立させる方式を提案する。提案手法は、追跡対象毎にその追跡対象が実在している確率を設定し、物体検出結果を用いて逐次確率値を更新することで、追跡対象数の変化に対応する。実験で提案手法の優位性が確認できた。

2. 関連研究

複数物体を個々独立に追跡するアプローチでは、個々の追跡処理の改善以外にも、1) 複数の追跡処理が同一物体を追跡してしまうという問題 (“hijacking” 問題 [8] と呼ばれる) にどう対応するか、2) 追跡対象数の変動にどう対応するか、といった 2 つの課題が存在する。

1) hijacking 問題への対応という課題に対しては、完全に個々の物体を独立に追跡するのではなく、追跡物体間の関係も見ながら追跡を行う処理が必要である。そのために、追跡対象を他の対象と識別できるよう識別学習を行うアプローチ (文献 [12])、事前に複数の追跡対象と複数の物体検出結果を排他的に対応付けを行う data association と呼ばれる処理を行うアプローチ (文献 [9] [13] [14])、追跡対象同士の重なりを評価する特徴量 (interaction 特徴 [8] と呼ばれる) を追加し、重なりに対してペナルティーをあたえることで hijacking 問題を回避するというアプローチ (文献 [8] [11]) が知られている。上記のいずれの処理も、追跡物体間の関係を見ているという意味では、物体毎に完全に独立な追跡処理では無く、その意味で、文献 [8] と同様、本稿でもこれらのアプローチを疑似独立な追跡と呼ぶことにする。

2) 追跡対象数の変動への対応という課題については、従来では大きく 2 つのアプローチが存在する。

2-a) 1 つ目の方法は、物体別に予め一定数のパーティクルを散布しておき、パーティクルが収束したら新規物体が存在すると判定して追跡処理を開始し、逆にパーティクルが発散したら物体が存在しないと判定する方法である (文献 [13] [19])。この方法では、パーティクルの収束を待つために追跡対象を登録するまでに時間がかかるため、長期にわたり画像中に存在するような追跡対象でないと適用が難しく、また、同時に複数の新規追跡対

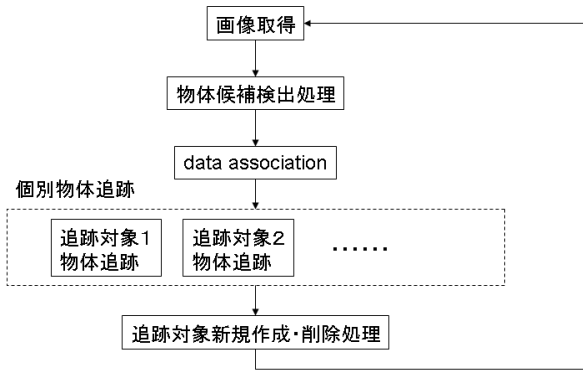


図 1 疑似独立な複数物体追跡の処理フロー

象が観測される時の対応方法も考えなければならず、追跡対象の数が次々刻々と変化するような状況下での複数物体追跡の性能向上は容易では無い。

2-b) もう 1 つの方法は、新たな追跡対象が物体検出処理によって確認できた時点で新規に物体追跡器を追加し、追跡対象が一定期間検出されなかったら削除する、というものである (文献 [9] [13])。具体的には、物体検出結果と現在の追跡対象との対応関係を求め、現在の追跡対象との対応が無い物体候補を新規登録する。追跡対象の数が次々刻々と変化するような状況下に関しては、2-a) の方法よりもこちらの方が対応が容易である。しかしこの手法では、ノイズのような誤った物体検出結果に対して追跡処理を継続するリスクが生じないようにすると、一時的に隠れなどで観測できなかった追跡対象を追跡できなくなってしまうという問題が特に重大になる。

従来からの個別に追跡する方法では、追跡対象数の変動に対応しながらもノイズの誤追跡軽減と、隠れへの追従性の確保を実現することは難しい。本稿ではこの問題に対して解決策を提示する。以降の節では、まず 3 節で、複数物体追跡の従来技術について説明する。そして 4 節で提案手法の内容について説明し、5 節で提案手法に関する評価実験の内容と結果について説明する。

3. 複数物体追跡の従来技術

3.1 概要

本節では複数物体追跡の従来技術について説明する。本稿で取り扱う複数物体追跡の処理フローは図 1 で示されるように、物体候補検出処理、data association、個別物体追跡、追跡対象新規生成・削除処理から構成される。

3.2 物体候補検出処理

物体候補検出では、追跡したい物体の候補を検出する処理を行う。例えば人物を追跡する場合には、HOG 特徴を用いた人物検出手法 (文献 [28]) などを用いれば良い。

3.3 data association

data association では、追跡対象が物体検出のどの結

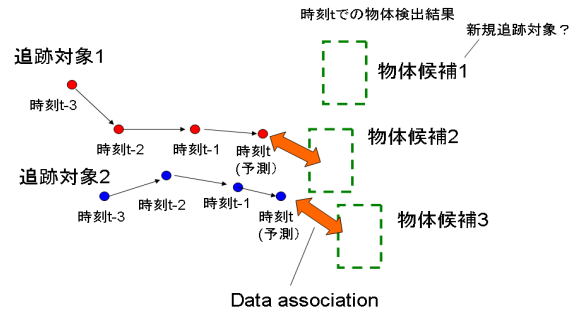


図 2 data association の例

果に対応するかの対応関係を決定する。例えば図 2 に、2 つの追跡対象と 3 つの物体候補との data association の例を挙げる。この場合、追跡対象 1 と物体候補 2 が対応づき、追跡対象 2 と物体候補 3 が対応づく。このように対応関係を決めれば、以降は個々独立にパーティクルフィルタ等の単一物体の追跡技術を用いて追跡を行うことができる。

data association は、追跡対象と物体候補をそれぞれノードとする 2 部グラフの割当問題とも見なすことができる。各追跡対象と各物体検出結果との間のコストを定義し、コストの総和が最も小さくなるような対応関係を求める最適割当問題の枠組みで解くことが良くなされている (文献 [9])。最適割当問題は、ハンガリアン法 [17] など各種最適化手法によって容易に解を算出できる。

また、最適割当となる 1 つの対応関係のみ利用するのではなく、全ての対応関係を考慮する手法も提案されている (文献 [13] [14])。1 つの対応関係を θ とし、観測 Z に対し対応関係が θ となる確率 $P(\theta|Z)$ を用いて、追跡対象 k が検出した物体候補 m と結び付く重み β_{km} を、以下のように算出する (文献 [13])。

$$\beta_{km} = \sum_{\theta \in \Theta_{km}} P(\theta|Z) \quad (1)$$

ただし、 Θ_{km} は追跡対象 k と物体候補 m が対応づく対応関係の集合を表す。 β を求めた後は、各物体候補を β_{km} の重みで利用し、個別の追跡処理を行う。

3.4 個別物体追跡

3.4.1 更新式

data association の後では、個別の物体追跡技術を適用する。本稿では個々の物体追跡技術としてパーティクルフィルタを採用する。以下では時刻 t における N 個のパーティクルを $\{x_t^{(i)}, w_t^{(i)}\} (i = 1, \dots, N)$ と表現する。ここで、それぞれの状態ベクトル x は $x = \{x, y, u, v\}$ とする。ただし、 (x, y) は追跡物体の位置を、 (u, v) は追跡物体の速度を表す。また、 w はパーティクルの重みを表す。追跡対象の時刻 t における位置は、 $x_t = \sum_i w_t^{(i)} x_t^{(i)}$ 、

$y_t = \sum_i w_t^{(i)} y_t^{(i)}$ として出力することができる。

パーティクルの状態や重みは、運動モデル $p(x_t^{(i)} | x_{t-1}^{(i)})$ や観測モデル $p(z_t | x_t^{(i)})$ を用いて更新される。本稿では、運動モデルについては、等速直線運動を仮定する。観測モデルとしては、文献 [9] [10] にあるような、

$$p(z|x) \propto p(z_D|x) \cdot p(z_C|x) \cdot p(z_A|x) \quad (2)$$

を採用する。ただし $p(z_D|x)$ は物体検出で得られた物体候補矩形から計算される尤度、 $p(z_C|x)$ は物体検出で得られる状態 x に対する物体らしさから計算される尤度、 $p(z_A|x)$ は色・形状などの特徴の類似性から計算される尤度である。具体的には、 $p(z_D|x)$ は、

$$p(z_D|x) \sim \exp(-((x - d_x)^2 + (y - d_y)^2)/2\sigma_D^2) \quad (3)$$

と設定する。ただし、 (x, y) は追跡対象の位置、 (d_x, d_y) が対応づく物体検出結果の物体候補の位置を表し、 σ_D は予め定めるパラメータである。 $p(z_C|x)$ は、物体検出で得られる状態 x に対応する物体らしさの信頼度 $f(x)$ を変換関数 g で尤度の尺度に変換した値 $p(z_C|x) \sim g(f(x))$ である。また、 $p(z_A|x)$ は、

$$p(z_A|x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-D_B^2/2\sigma_B^2) \quad (4)$$

とする。ここで、 D_B は Bhattacharyya 係数から求めた距離であり、

$$D_B^2 = 1 - \sum_{u=1}^U \sqrt{q^*(u)q(u)} \quad (5)$$

である（文献 [7] [8] [13]）。ただし、 q^* は追跡対象の物体モデルとして予め記録されている色ヒストグラムであり、 $q(u)$ は状態 x に対する色ヒストグラムである。また、 u はヒストグラムの各 bin をあらわす。

3.4.2 オンライン学習

尤度 $p(z_A|x)$ の計算時に用いる q^* は、追跡対象の変化やカメラアングルからの見え方の変化等によって時々刻々と変わっていく可能性がある。そこで、新しく観測された結果 q_t で以前に保持されているデータを、

$$q^* \leftarrow (1 - \alpha)q^* + \alpha q_t \quad (6)$$

のように更新する。ただし、 α は予め定めた定数である。また本稿では、追跡対象の大きさ情報も保持しておき、オンライン学習によって更新することにする。

3.4.3 interaction 特徴を用いた追跡

複数物体追跡においては、一般的に追跡対象が重なりすぎた状態の頻度が小さいことが期待できる。そこで、複数の追跡対象間の重なりを考慮する特徴量を計算し、本特徴量の値が小さいほど大きくなるように尤度を設定し、個別の追跡処理を行うという方法が知られている（文献 [8] [11]）。本稿でも、対象とする追跡対象の外接矩形の中で、対象とする追跡対象以外の外接矩形との共通部分が占める面積の比率を interaction 特徴として採用し、個別の追跡処理を行う。

表 1 追跡対象の分類

	実在する	実在しない・分からない
観測できる	A (確率 η)	×
観測できない	B (確率 ξ)	C (確率 $\zeta = 1 - \eta - \xi$)

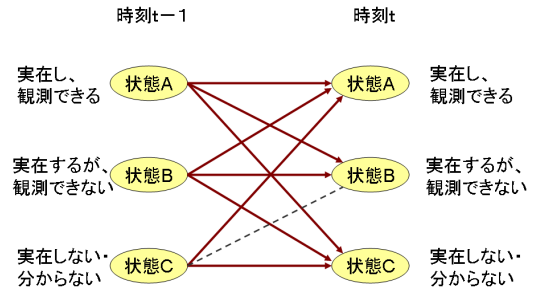


図 3 追跡対象の状態の遷移

3.5 追跡対象新規生成・削除処理

複数物体追跡では、新たな追跡対象が発生したり、追跡対象が画面外に消えるといった追跡対象数の変動に対応することが必要である。従来では、図2の物体候補1のように、既存追跡対象と対応づかない物体検出結果を新規追跡対象とし、物体検出結果と τ 回以上対応づかない追跡対象や画面外に出た追跡対象を削除するという処理が一般的に行われている（文献 [13] [9] [15] [24]）。 τ をある程度大きく設定することで、一時的に他の移動物体等による隠れで物体検出結果として得られないような追跡対象でも、削除することなく追跡が継続できる。

3.6 従来手法の課題

従来な追跡手法では、追跡対象数が時々刻々と変化する際に、ノイズなどの誤検出と追跡対象の一時的な隠れに同時に対応することが難しい。ノイズのような誤った物体検出結果に対して追跡処理を継続することを回避するには τ を小さくした方が良いが、一時的に隠れなどで観測できなかった追跡対象でも追跡を継続できるようにするためには、 τ を大きくした方が良いため、両者を同時に解決することが難しく、この問題を解決することが課題である。

4. 提案手法

4.1 確率に基づく追跡対象新規作成・削除処理

前節で示した課題解決のため、追跡対象毎にその追跡対象が実在している確率を設定し、物体検出結果を用いて逐次更新することで、追跡対象数の変化に対応する手法を提案する。

追跡対象の状態は、実在し観測できる状態 A、実在はするが観測できない状態 B、実在しないもしくは分からない状態 C の3種類に分類される（表1）。追跡対象がそれぞれの状態となっている確率を物体検出結果を用いて以下のように更新し、追跡対象の新規生成や削除の処理を実施する。

(1) data association で既存追跡対象と対応づかない物体検出結果を新規追跡対象として仮登録し、各状態の確率値に初期値をセットする。

(2) data association の結果に応じて各状態に属する確率を更新する。data association によって物体候補と対応づいた場合には状態 A に属する確率を高める。状態 A に属する確率が予め定めた閾値以上になった時に初めて追跡対象が実在すると判定する。

(3) data association の結果に応じて物体候補と対応づかない場合には状態 A に属する確率を下げ、状態 B と状態 C に属する確率を高める。状態 C に属する確率が予め定めた閾値以上になれば追跡対象を削除する。

各状態における確率を計算・更新するためには、各状態間の遷移を考え必要があるが、状態 C から状態 B への遷移は考えなくて良い(図3)。なぜなら、物体検出の情報だけでは、状態 B と状態 C の区別はつかず、追跡対象が実在することが分かって初めて状態 B であると判断できるからである。よって、状態 B になるためには一度状態 A を経由するという遷移上の制約が発生する。この制約により、例えばノイズのような短い時間でしか物体検出結果との対応付けができない追跡対象は、状態 A に十分遷移する前に状態 C に遷移するため削除が促進される。また、しばらく物体候補との対応付けが行われた追跡対象に対しては、十分状態 A に遷移できるため、その後物体候補としばらく対応付けができなくても、その追跡対象は状態 B に遷移可能となり、一時的な隠れと判断され削除されにくくなる。このように遷移に制約を加えて確率化することで、ノイズの追跡を抑えつつ、一時的な隠れへの追従性も確保できる。

以降では、具体的な確率の更新式を示す。今、時刻 t において追跡対象 k が状態 A にある確率を $\eta_{k,t}$ 、状態 B にある確率を $\xi_{k,t}$ 、状態 C にある確率を $\zeta_{k,t} = 1 - \eta_{k,t} - \xi_{k,t}$ とする。また、 P_D^m は物体候補 m が検出すべき物体である確率とする。追跡対象 k が観測可能な状態で存在するのであれば、物体検出処理で検出されることが期待できることから、 $\eta_{k,t}$ を、

$$\eta_{k,t} = (1 - \alpha) \cdot \eta_{k,t-1} + \alpha \eta_k \quad (7)$$

のように更新する。ただし、 $\eta_k = \sum_m \beta_{km} P_D^m$ である。

次に $\xi_{k,t}$ であるが、状態 A になる確率 $\eta_{k,t}$ が増える場合は、必ず状態 B である確率は減少する。この時、 B である確率と C である確率との比率は保持すると仮定すれば、

$$\xi_{k,t} = \xi_{k,t-1} - \frac{\xi_{k,t-1}}{1 - \eta_{k,t-1}} (\eta_{k,t} - \eta_{k,t-1}), \text{ if } \eta_{k,t} \geq \eta_{k,t-1}$$

という更新式が得られる。一方、状態 A である確率 $\eta_{k,t}$ が減る場合には、その分状態 B である可能性も増加するので、

$$\xi_{k,t} = \xi_{k,t-1} + \delta_k (\eta_{k,t-1} - \eta_{k,t}), \text{ if } \eta_{k,t} < \eta_{k,t-1}$$

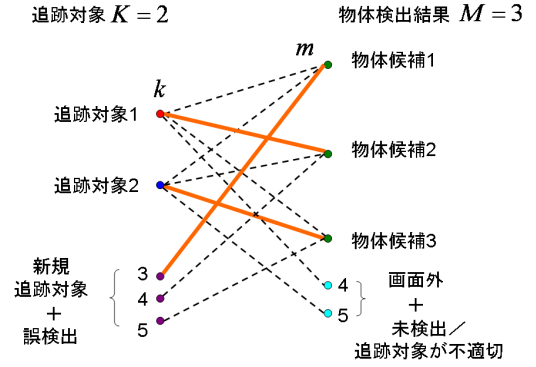


図4 物体数変動を考慮した data association

のように更新する。ただし、 δ_k は追跡対象 k が状態 C では無く B であろうと判断できる確率である。状態 B には、他の移動物体による隠れや背景固定物による隠れ、あるいは単純な検知漏れが考えられる。他の移動物体による隠れであれば、少なくとも追跡対象のそばの物体検出信頼度は高いことが期待できることから、今回は、 δ_k を、

$$\delta_k = (1 - \gamma) \sum_i w^{(i)} b(x^{(i)}) + \gamma p_o$$

と計算することにした。ただし、 $b(x^{(i)})$ は i 番目のパーティクルが表す矩形内の前景比率から後述の5. 記載の変換テーブルによって確率化したものである。本式の第1項は他の移動物体による隠れの可能性、第2項は背景固定物や検知漏れの可能性、 γ はその比率を示し、 p_o は背景固定物によりオクルージョンが生じる比率を表し、それぞれ、今回は固定値として $\gamma = 0.95$ 、 $p_o = 1$ と実験的に定めている。本稿では追跡対象の実在する確率という概念を導入することによる追跡対象数の変化への対応容易性を示すことが主目的であったため、パラメータを単純化したが、一時隠れの状況には場所依存性や他追跡対象との関係も重要な要素であり、今後改善の余地がある。

以上のように、提案手法は追跡対象が実在しない場合も考慮した追跡を行う。従って、data association の処理についても追跡対象が実在しない場合も想定したものに拡張する必要がある。それについては次節で説明する。

4.2 data association の拡張

以降では、時刻 t における追跡対象数を K 、物体検出処理で検出された物体候補数を M とする。追跡物体数の増減も考慮するために、図4のような追跡物体数の増減も考慮したダミーノードも追加する。図4を見れば分かるように、 $K + M$ 個ずつのノードの2部グラフに対して data association を考えることになる。 $k > K$ のノードは物体検出結果で得られる物体候補が既存の追跡対象では無い事象を表す。具体的には、新規追跡対象が誤検出を表している。また、 $m > M$ のノードは追跡対象が物体検出結果で得られていない事象を表す。すなわち、画面外に消えたか、未検出、もしくはノイズ等を誤って

追跡していることを表している。このような生成・削除を反映するダミーのノードを追加するという考え方自体は、文献 [27] 等、従来手法でも見られる考え方である。ただし、従来手法では追跡対象が必ず存在するという前提で data association を実施するが、本稿では追跡対象自体が実在しない可能性も含んでおり、追跡対象が実在しない可能性も考慮したコスト関数の設計を行っている点が新規である。

まず、 β の計算に必要となる $P(\theta|Z)$ の計算方法について説明する。 $P(\theta|Z)$ は、文献 [13] の考えを拡張して、次のように計算する。

$$P(\theta|Z) \propto \prod_{e_{km} \in \theta} p(e_{km}|Z) \quad (8)$$

ただし、 e_{km} は対応関係 θ の中で追跡対象 k と物体候補 m を結び付けるエッジを表す。また、

$$p(e_{km}|Z) = \begin{cases} ((1 - P_D^m) + \epsilon P_D^m) s_o, & k > K \\ (\Delta_k + Q) s_o, & m > M \\ (\eta_k + \xi_k) P_D^m (1 - Q) s_{km}, & \text{Otherwise} \end{cases} \quad (9)$$

である。ここで、 P_D^m は物体候補 m が検出すべき物体である確率である。また、 Q は追跡対象が未検出または実在しない確率である。さらに、 s_{km} は物体候補 m の追跡対象 k に対する尤度である。また、 $\eta_k + \xi_k$ は前節で述べた、追跡対象 k が実在すると推定できる確率であり、追跡対象が実在しているものを物体候補との対応付けにおいて優先している。

式 (9) で $k > K$ の場合は、物体候補が追跡対象と対応づかない場合であり、誤検出となる可能性 $1 - P_D^m$ と新規の追跡対象となる可能性 ϵP_D^m を考慮している。ただし、 ϵ は新規追跡対象らしさを表す補正項であり、本稿では 3.4.3 節で述べた interaction 特徴によって求められる尤度を採用する。すなわち、他の追跡対象との重なりが小さい物体候補は誤検出ではなく新規人物らしいと判定されやすくなる。

式 (9) で $m > M$ の場合は、追跡対象が物体候補と対応づかない場合であり、追跡対象が画面外に出てしまった可能性 Δ_k が、未検出 / 追跡対象が不適切の可能性 Q を考慮している。 Δ_k は、追跡対象 k に対する画面外のパーティクルの重み付き平均として求める。

式 (9) で Otherwise の場合は、追跡対象が物体候補と対応づける場合である。対応づいた場合、追跡対象と物体候補との近さを表現する尤度 s_{km} を考える。 s_{km} としては、式 (2) に述べた尤度を採用する。

なお、追跡対象と物体候補が対応づかない場合 ($m > M$ または $k > K$ の場合) には一方が観測されていないことから近さを表現する尤度の計算ができない。そこでこの場合は固定値 s_o で代用することにする。 s_o は様々な追跡対象と物体検出結果の組み合わせで計算される尤度の平均値とする。これは、予め他の学習用

DB を用意して算出し、利用すれば良い。

data association の結果、既存追跡対象と対応づいていない物体候補 m は新規追跡対象の候補となる。 m が新規追跡対象として登録される確率 P_{new} は、式 (9) $k > K$ の場合における第 1 項と第 2 項の比に対応し、

$$P_{new} = \frac{\epsilon P_D^m}{(1 - P_D^m) + \epsilon P_D^m} \beta_{km}, \quad k > K \quad (10)$$

となる。 P_{new} の大きさに応じて追跡対象が実在する確率がセットされる。一方、物体候補と対応づいていない追跡対象 k は削除される追跡対象の候補となる。追跡対象 k が画面外で削除すべきと判定する確率 P_{out} は、式 (9) $m > M$ の場合における第 1 項と第 2 項の比に対応し、

$$P_{out} = \frac{\Delta_k}{\Delta_k + Q} \beta_{km}, \quad m > M \quad (11)$$

となる。 P_{out} が予め定めた閾値以上であれば画面外で削除すべきと判定し、対応する追跡対象の削除ができる。

4.3 オンライン学習の制御

複数物体追跡の際には、例えば複数物体がすれ違う状況のように、対応関係がはっきりしない場合にオンライン学習は悪影響が生じる場合もある。そこで、本稿では、 $\beta_{km} > th$ の時に追跡対象 k に対するモデルを物体候補 m の情報を用いてオンライン学習する、という設定にした。これによって、 β_{km} が十分大きな値となった場合にのみ、追跡対象 k と対応づいている物体候補は m であると判断し、物体 m の結果から追跡対象 k のモデルをオンライン学習する、という制御を行うことができる。本稿では、 $th = 0.95$ と固定値を設定した。

4.4 近似計算による高速化

本手法の実現にあたっては、 β の算出が必要であり、そのためには全ての対応関係を考慮する必要がある。しかし対応関係の数は、追跡対象の数や物体候補の数が増えると指数関数的に増加するため、近似計算による高速化が必要である (文献 [22] [23])。ただし、今回は追跡対象の生成・削除、あるいはオンライン学習に β を活用することが主目的であるため、より簡易に、対応関係の最適解 $\hat{\theta} = \arg \max P(\theta|Z)$ を計算し、

$$\beta_{km} = \begin{cases} \frac{P(\hat{\theta}|Z)}{P(\hat{\theta}|Z) + P(\hat{\theta}_{-km}|Z)} & \text{if } e_{km} \in \hat{\theta} \\ 0 & \text{Otherwise} \end{cases} \quad (12)$$

として近似する。ただし、 $\hat{\theta}_{-km}$ は追跡対象 k と物体候補 m とを結びエッジ e_{km} を除去した時の最適解である。評価関数 S を、

$$S(\theta) \equiv -\log P(\theta|Z) \propto \sum_{e_{km} \in \theta} (-\log p(e_{km}|Z)) \quad (13)$$

と定義すれば、最適な対応関係を求める問題はコストを $-\log p(e_{km}|Z)$ とする最適割当問題となるので、ハンガリアン法によって多項式時間で求められる。

5. 実験

5.1 実験環境

評価には、PETS2006 [20] の Dataset S1 カメラ 3 のデータを用いた。今回選んだ DB は屋内の駅のホームを監視する映像で、映像内での人物の出入りが多いデータである。従って、追跡対象数の変化への適応性を評価するのに適切であると考えられる。データは全部で 3021 フレームあり、その中で人物がとらえられているフレームが 2538 フレーム存在している。なお評価を行うために、予め本データに対し手動で全フレームに対する正解付けを行ってある。

人物検出手法については、今回の評価データが屋内環境で人物同士の重なりが大きいことから、文献 [28] にあるような人物全体を探索する手法ではなく、遮蔽に対応しやすい頭部検出結果を活用して人物検出を行うアプローチをとった。ただし、カメラから遠く人物が小さく映る領域では頭部検出での人物検出は困難であることから、背景差分法 [18] による移動体検知も併用して人物検出を行った。

頭部検出器については、INRIA データ [26] にある 2416 枚の人体から頭部領域を手動で抽出し、抽出されたデータに摂動を加えることで 20,000 個の学習データを作成し、文献 [25] に示される検出器学習手法を用いて作成した。また、人物領域は、対象人物の身長が平均身長であると仮定して、DB に付随するキャリブレーション情報を用いて算出した。得られた人物検出結果に対し、人である確率を、背景差分法で得られる人物領域内の前景画素の比率を特徴量として算出する。算出方法については、別途独自に採出した屋内 DB から変換テーブルを学習した。

実験評価にあたっての尺度について、以下に説明する。まず、人物検出の性能を評価するために、適合率 (precision)、再現率 (recall) 及び適合率と再現率の調和平均である F1 値を用いる。これらの値は高ければ高いほど良い性能であることを示している。また、追跡性能の評価には、人物検出精度だけでなく、同一の人物が同一の ID で追跡し続けることができるか、といった評価も必要である。ただ、評価尺度の選定には注意が必要である。例えば、単純に人物が現れてから同一 ID で追跡し続けられたフレーム数を評価尺度とすると、たまたま追跡開始直後に追跡失敗が生じた場合、以降の追跡性能が良くても評価値が大きく下がるので、適切な評価にはなりにくい。適切な評価尺度については文献 [21] などで議論されているので、今回はそれに従い、追跡対象の ID が途中で変化する数と、人物検出も含めて追跡についての全体的な性能を評価する尺度である MOTA を採用する。ID 変化数は、小さければ小さいほど良い追跡性能であることを示す。また、MOTA は、

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \quad (14)$$

と定義される量であり、大きければ大きいほど良い性能であることを示す。ただし、 m_t 、 fp_t 、 mme_t 、及び g_t はそれぞれ時刻 t における未検出の数、誤検出の数、ID 変化数、及び真の追跡対象数である。

5.2 比較手法

比較に用いる従来手法としては、3 節に示した手法 (従来手法 1) を採用する。また、文献 [9] では、さらに細かい工夫として、既存追跡対象との重なりが無いものに限定して登録したり、直前の物体検出結果との重なりが大きい物体候補のみを採用したりすることで、誤追跡を削減する工夫も見られている。そこで、従来手法 1 の条件に追加して、既存追跡対象との重なりが無いものに限定して登録する手法 (従来手法 2) と、さらに従来手法 2 の条件に追加して、直前の物体検出結果との重なり比率が閾値 (実験では 0.5) 以上の物体候補に限定して登録する手法 (従来手法 3) も評価の比較対象とすることにした。上記従来手法には、3 節で述べたようなパラメータ τ が存在しているが、それらは $\{1, 2, 4, 10, 20, 40, 60, 80\}$ の中から、評価 DB に対し最も F1 値が大きいものと ID 変化数が最も小さいものの 2 通りを選択し、評価実験を行った。

従来手法、提案手法ともに、個別追跡は同じであり、パーティクルの数は $N = 100$ とする。従来手法の data association については、新規人物や画面外のノード以外については、提案手法と共通のコスト関数を用いる。また、オンライン学習のパラメータについては、 $\alpha = 0.1$ とする。

提案手法に関して、実験的にパラメータ $Q = 0.2$ と定めた。また、 η と ξ を用いた生成・削除の制御については以下を行う。追跡対象仮登録後、 $\eta > th_1$ となった時に初めて、追跡対象が実在するという判定を行う。また消滅については、実在しない確率が十分高ければ削除する。ただし実在するが観測できない状態が長すぎる場合に、追跡対象を全て覚えておくのでは処理量が膨大となる可能性もあるため、追跡対象の η が小さいものについては、追跡を打ち切ることとする。具体的な判定式は、 $\zeta > th_2$ 、または $\zeta > th_3$ かつ $1 - \eta > th_2$ である。パラメータ $th_1 = 0.9$ 、 $th_2 = 0.95$ 、 $th_3 = 0.8$ については実験的に定めた。また、 β を計算し活用することの有効性を示すために、 $\beta = 1$ に固定した場合の評価も合わせて実施することにした。

5.3 実験結果

従来手法と提案手法とを比較した結果を表 2 に示す。なお、従来手法 1 は $\tau = 10$ の時に F1 値が最も大きくなり、 $\tau = 40$ の時に最も ID 変化数が小さくなった。同様に、従来手法 2 は $\tau = 20$ と $\tau = 80$ が選ばれた。従来手

表 2 従来手法と提案手法の比較

	Prec.	Rec.	F1	MOTA	ID 変化
人物検出	0.587	0.620	0.603	-	-
従来 1($\tau = 10$)	0.710	0.652	0.680	0.394	24 回
従来 1($\tau = 40$)	0.757	0.564	0.677	0.347	6 回
従来 2($\tau = 40$)	0.722	0.661	0.690	0.417	8 回
従来 2($\tau = 80$)	0.664	0.707	0.685	0.400	6 回
従来 3($\tau = 40$)	0.761	0.637	0.694	0.445	8 回
提案 ($\beta = 1$)	0.750	0.678	0.712	0.461	0 回
提案 (β 推定)	0.785	0.683	0.731	0.505	0 回

法 3 は、どちらも $\tau = 40$ の時となった。

表 2 を見ると分かるように、従来手法のほぼすべての手法と比較して、提案手法は、再現率と適合率の双方ともに良い値となった。また、従来手法と比較して、提案手法は ID 変化数を大幅に抑制することができた。このことから、提案手法は、誤った物体検出結果に対する追跡を抑制しながら、一時的に観測できない状態での追跡継続が出来ていると考えられる。また、提案手法において β を用いることの有用性も確認できた。

次に、実験結果に関する具体的な例を、図 5 に示す。これらの図において、上から順に物体検出結果、従来手法 2 の追跡結果 ($\tau = 2$)、従来手法 2 の追跡結果 ($\tau = 80$)、提案手法の追跡結果を表す。追跡結果において、同じ色で示した人物は同一人物として追跡処理が行われていることを示している。また、太色で示された人物はオンライン学習がなされている状態であることを示している。提案手法では左から 2 枚目の絵でグレーで表示されている人物矩形があるが、これは実在すると判定される前の仮登録の状態を表す。図 5 を見て分かるように人物検出結果は、人物が密集して移動していることから未検出や誤検出が発生する。これに対し従来手法では、 τ が小さい場合には左から 3 枚目から 4 枚目への推移をみれば分かるように、誤検出による誤まった追跡対象を早期に削除できているものの、同時に正しい人物追跡結果も削除されてしまっている。一方、 τ が大きい場合には左から 2 枚目より右への推移を見れば分かるように、検出結果による誤った追跡対象を長期にわたり継続し、さらに正しい人物検出結果が新規人物として登録されず正しい人物の正しい追跡もできないという悪影響も引き起こしている。これに対し、提案手法では、誤った物体検出結果に対する追跡を抑制しながら同時に正しい人物追跡が維持できていることが確認できる。

ただ、従来手法 / 提案手法ともに、追跡対象の増減に関して調整すべきパラメータは少ないとは言えない。学習データを用いたパラメータの推定については今後の課題である。

6. おわりに

本稿では、追跡対象を個別に追跡する方法を基に、追

跡対象数の変化に対応しながらも、誤った物体検出結果に対する追跡を抑制しながら、一時的に観測できない状態でも追跡を継続できるようにすることで、精度と速度を両立した手法を提案した。提案手法では、追跡対象毎にその追跡対象が実在している確率と、一時的に観測できなくなっている確率を設定し、物体検出結果を用いて逐次確率値を更新することで、追跡対象数の変化に対応する。

実験の結果、従来手法のほぼすべての手法と比較して、提案手法は、再現率と適合率の双方ともに良い値を示していること、さらに ID 変化数を大幅に抑制されているという結果となり、誤った物体検出結果に対する追跡を抑制しながら、一時的に観測できない状態での追跡継続が出来ていることを確認できた。

一時隠れの判定方法の高精度化や、学習データを用いた各種パラメータ推定の手法について検討することが今後の課題である。

文 献

- [1] B.Ristic, "Beyond the Kalman Filter : Particle Filters for Tracking Applications," Artech House, 2004.
- [2] M.Isard and A.Blake, "CONDENSATION - Conditional Density Propagation for Visual Tracking," International Journal of Computer Vision, vol.29, pp.5-28, 1998.
- [3] T. Yu and Y. Wu, "Collaborative tracking of multiple targets," Computer Vision and Pattern Recognition, 2004.
- [4] T. Yu and Y. Wu, "Decentralized multiple target tracking using netted collaborative autonomous trackers," Computer Vision and Pattern Recognition, 2005.
- [5] Z.Khan, T.Balch, and F.Dellaert, "MCMC-based particle filtering for tracking a variable number of interacting targets," Pattern Analysis and Machine Intelligence, vol.27, pp.1805-1819, 2005.
- [6] S.Särkkä, A.Vehtari and J.Lampinen, "Rao-blackwellized particle filter for multiple target tracking," Information Fusion,8,1,pp.2-15,2007.
- [7] K. Nummiaro, E. Koller-Meier and L.V.Gool, "A Color-based Particle Filter for Joint Detection and Tracking of Multiple Objects," IEEE International Conference on Acoustics, Speech, and Signal Processing, pp.217-220, 2005.
- [8] R.Hess and A.Fern, "Discriminatively Trained Particle Filters for Complex Multi-Object Tracking," In Proc. IEEE Conf. on Computer Vision and Pattern Recognition(CVPR), pp. 240-247, 2009.
- [9] M.D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier and L.V.Gool, "Robust Tracking-by-Detection using a Detector Confidence Particle Filter," IEEE 12th International Conference on Computer Vision, pp.1515-1522, 2009.
- [10] S. Stalder, H. Grabner and L.V.Gool, "Cascaded Confidence Filtering for Improved Tracking-by-Detection," Proceedings European Conference on Computer Vision (ECCV), 2010.
- [11] A.Beaugendre, H.Miyano, E.Ishidera and S.Goto, "Human Tracking System for Automatic Video Surveillance with Particle Filters," APCCAS, 2010.
- [12] K. Okuma, A. Taleghani, N. D. Freitas, J. J. Little,



図 5 追跡動作の例、追跡結果で太線／細線はオンライン学習の有無を表す
(上から順に人物検出、従来手法 2 ($\tau = 2$)、従来手法 2 ($\tau = 80$)、提案手法)

- and D. G. Lowe, "A boosted particle filter: multitarget detection and tracking," ECCV, 2004.
- [13] M.Jaward, L.Mihaylova, N.Canagarajah and D.Bull, "Multiple Object Tracking Using Particle Filters," Aerospace Conference, 2006.
- [14] K.Bai, "Particle Filter Tracking with Mean Shift and Joint Probability Data Association," International Conference on Image Analysis and Signal Processing (IASP), pp.607-612, 2010.
- [15] M. Yang, Fengjun Lv, Wei Xu, Yihong Gong, "Detection Driven Adaptive Multi-cue Integration for Multiple Human Tracking," IEEE 12th International Conference on Computer Vision, pp.1554-1561, 2009.
- [16] A.Elghamall, D.Harwood, L.Davis, "Nonparametric Model for Background Subtraction," 6th European Conference on Computer Vision(ECCV), vol.2, pp. 751-767, 2000.
- [17] H.Kuhn, "The hungarian method for the assignment problem," Naval Research Logistics Quarterly, vol.2, 83-97, 1955.
- [18] H.Ikeda and E.Ishidera, "輝度情報の時間的性質を用いた背景のモデル化とそれに基づく物体検出," FIT 2006, vol.3, pp.175-176, 2006.
- [19] 村北卓也, 石黒浩, "Sampled-Volume-based Data Association Filter による遮蔽ロバストな実時間複数人物追跡", 電子情報通信学会論文誌. D, 情報・システム J90-D(10), pp.2868-2883, 2007.
- [20] Ninth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. <http://www.cvg.reading.ac.uk/PETS2006/index.html>, 2006.
- [21] K.Bernardin and R.Stiefelhagen, "Evaluating multiple object tracking performance: The CLEAR MOT metrics," EURASIP Journal on Image and Video Processing, 2008.
- [22] S.Oh, and S.Sastry, "An Efficient Algorithm for Tracking Multiple Maneuvering Targets," 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference, pp. 4010-4015, 2005.
- [23] E. Aoki and K.Kienitz, "Suboptimal JPDA for Tracking in the Presence of Clutter and Missed Detections," 12th International Conference on Information Fusion, pp.818-825, 2009.
- [24] L.M.Fuentes and S.A.Velastin, "People Tracking in surveillance applications," Image and Vision Computing, pp.1165-1171, 2006.
- [25] 細井 利憲, 佐藤 敦, "一般化学習ベクトル量子化に基づく顔検出," 電子情報通信学会技術研究報告, vol. 102, No. 651, pp. 47-52, 2003.
- [26] INRIA Person Dataset, <http://pascal.inrialpes.fr/data/human/>
- [27] 西行健太, 波部育, 木戸出正繼, "局所的な特徴と大局的な特徴の統合による蓄積映像からの人物移動軌跡の獲得," MIRU2010,IS2-61, pp.1190-1197, 2010.
- [28] N.Dalal, B.Triggs, "Histograms of Oriented Gradients for Human Detection," Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.886-893, 2005.