

# 追跡対象と周辺領域の関係性を利用した協調的な物体追跡

山下 隆義<sup>†</sup> 藤吉 弘亘<sup>††</sup>

<sup>†</sup> オムロン株式会社

〒 525-0025 滋賀県草津市西草津 2 丁目 2 - 1

<sup>††</sup> 中部大学

〒 487-8501 愛知県春日井市松本町 1200

E-mail: <sup>†</sup>takayosi@omm.ncl.omron.co.jp, <sup>††</sup>hf@cs.chubu.ac.jp

**あらまし** 物体追跡は、追跡対象の形状変化や照明変化による見え変化、物体の移動やカメラ移動による動き変化、他の物体や追跡対象自身による遮蔽など様々な変化が生じる複雑な状況下でも対象を追跡し続けることが求められている。従来の物体追跡は、これらの問題に対して、追跡物体と背景領域を区別するのに有効な情報を捉えることに主眼を置いてきた。我々は視点を変え、追跡物体と背景領域を区別するだけでなく、背景領域から追跡対象の状況を理解して、その情報を活用することができる新たな物体追跡のモデルを提案する。提案モデルでは、追跡対象とその近傍領域との空間的な関係性と特徴の類似性をもとに、追跡対象に生じる誤追跡の発生確率を周辺状況として導く。この発生確率をもとに、追跡の振る舞いを変えることで、類似する物体や背景への誤追跡を防ぐことが可能となる。提案手法の有効性を示すために、提案する追跡モデルと関連する従来手法と追跡性能の比較を行った。その結果、提案手法は複雑背景や類似物体との遮蔽が生じる様々なシーンにおいて従来手法より追跡性能が向上していることを示すことができた。

**キーワード** 物体追跡, 周辺理解, 協調的, 周辺領域, 勾配法

## 1. はじめに

物体追跡は、監視セキュリティやヒューマンインタラクション、ビデオ編集等の用途において重要な技術の一つである。時系列において、追跡対象は時々刻々と形状や見えが変化すると共に、背景変化やカメラ移動などが同時に発生する。このような複雑な状況での追跡を目指してこれまで様々な手法が提案されている。物体追跡の課題として、追跡対象または背景領域の見え変化や動き変化によるものがある。見え変化は、追跡対象の形状変化や照明変化、背景変化などである。動き変化は、追跡対象の移動方向や速度の変化、カメラの移動、さらにはフレームレートの変化などがあげられる。見えの変化への対応には、色ヒストグラムを用いる手法 [4] [3] [5], 照明変化などが生じる場合には複数の色ヒストグラムを利用する手法 [2] [15], 単一色の場合には追跡対象と背景領域を比較しオンラインで特徴選択する手法 [8] [14] [16] などが提案されている。また、形状変化や照明変化に対しては、追跡対象を SIFT や SURF などの不変特徴量の集合と見なす手法もある [7]。

一方、動き変化に対しては、追跡対象の存在する領域を尤度をもとにサンプリングするパーティクルフィルタ [1] [12] や追跡対象の動きごとに最適なダイナミクスを導入した手法 [13] がある。さらに、低フレームレートでも追跡するために、複数の観測モデルを用いた手法 [9] も提案されている。

このように、見え変化や動き変化に対する解決手法は多数提案されている。これらの解決手段は、Collins [14] らの手法に代表されるように、追跡対象内もしくは追跡対象と背景領域を区別するのに有効な特徴を捉えることに主眼を置いているものが多い。そのため、追跡対象と類似する物体を背景に含むような場合や他の物体による遮蔽が生じるシーンにおいては、追跡することが困難である。特に、類似物による遮蔽が生じるようなシーンにおいて、物体を追跡するためには、追跡対象以外の情報を有効に活用することが重要である。Yang らは背景領域のうち、追跡対象とともに安定して存在する領域をデータマイニングにより抽出する手法を提案している [6]。この手法では、誤追跡が発生した場合に、継続的に存在する背景領域との位置関係から追跡対象の領域を推定する。また、Grabner らも同様に、継続的に現れる背景領域の不変特徴量と追跡対象の位置関係から追跡対象の領域を推定する手法を提案している [10]。この手法は、背景領域が動くとき追跡対象の領域を推定することが困難となる。このように、背景領域を活用する手法は提案されている一方、あくまで位置関係に基づく補完的な役割に留まっている。

そこで、我々は、これまでの追跡対象中心の手法から視点を変え、追跡対象と背景領域との関係性を捉えることに主眼を置く。すなわち、追跡対象とともに背景領域を捉え、それらの関係性から追跡対象周辺の状況を理解する。そして、その周辺状況に応じて追跡方法を動的に

切り替えることで、見えや動き変化だけでなく、類似物を含むシーンにおいても追跡対象を途切れなく追跡することが可能となる。ここで、「状況の理解」とは、「追跡対象が別の人物と話しながら歩いている」というような高次な理解から、「追跡対象に発生する誤追跡の可能性」のような特定事象の理解まで様々考えられる。我々は、追跡対象とその近傍領域の空間的な関係性と特徴の類似性をもとに誤追跡の発生確率を導出し、その誤追跡の発生確率を「状況の理解」として定義する。本稿では、このような追跡対象周辺の状況から得た情報をもとに追跡の観測モデルと遷移モデルを動的に選択可能なフレームワークを提案する。

## 2. 従来の物体追跡モデル

確率的なベイズ推論の観点において、式(1)のように、追跡対象の現時刻までの推定位置  $\bar{X}_t = \{x_1, \dots, x_t\}$  と観測値  $\bar{Z}_t = \{z_1, \dots, z_t\}$  から、事後確率が最大化となる位置  $\bar{x}_t^*$  を追跡対象の位置として推定することができる。

$$\bar{x}_t^* = \operatorname{argmax}_{\bar{X}_t} p(\bar{X}_t | \bar{Z}_t) \quad (1)$$

また、時系列でのマルコフ性を考慮することで、時刻  $t$  における事後確率は式(2)となる。

$$\begin{aligned} p(x_t | \bar{Z}_t) &= \frac{p(z_t | x_t, \bar{Z}_{t-1}) p(x_t | \bar{Z}_{t-1})}{p(z_t | \bar{Z}_{t-1})} \\ &= \frac{p(z_t | x_t) \int_{x_{t-1}} p(x_t | x_{t-1}) p(x_{t-1} | \bar{Z}_{t-1}) dx_{t-1}}{p(z_t | \bar{Z}_{t-1})} \end{aligned} \quad (2)$$

これより、時刻  $t$  における位置  $x_t$  の事後確率  $p(x_t | \bar{Z}_t)$  は、現時刻での観測モデル  $p(z_t | x_t)$  と過去の時刻の状態  $p(x_{t-1} | \bar{Z}_{t-1})$  および遷移モデル  $p(x_t | x_{t-1})$  から表現することができる。ここで、 $p(z_t | \bar{Z}_{t-1})$  は正規化係数である。観測モデルは、テンプレートやヒストグラムで表現された追跡対象のモデルと時刻  $t$  における位置  $x_t$  の特徴の類似性を観測する。遷移モデルは、追跡対象の過去の位置をもとに、等速運動やランダムウォークを仮定して各位置での時刻  $t$  における追跡対象の出現確率を算出する。すなわち、物体追跡は、時系列における追跡対象の予測位置から決定される探索範囲内において、観測モデルが最も類似する位置を追跡対象の位置として推定している。

追跡過程において、追跡対象の動きや見えの変化により、追跡対象の位置推定が困難なケースがある。Yangらは、時系列で安定して出現する背景領域を補助領域として捉え、追跡対象と補助領域の関係性を考慮した物体追跡をモデル化している。このモデルによる追跡対象の事後確率は、補助領域の事後確率と位置関係を加えた式(3)のようになる。

$$p(x_{t,0} | Z_{t,0}) \propto p_0(z_{t,0} | x_{t,0}) p_0(x_{t,0} | Z_{t-1,0}) \prod_k m_{k0}(x_{t,0}) \quad (3)$$

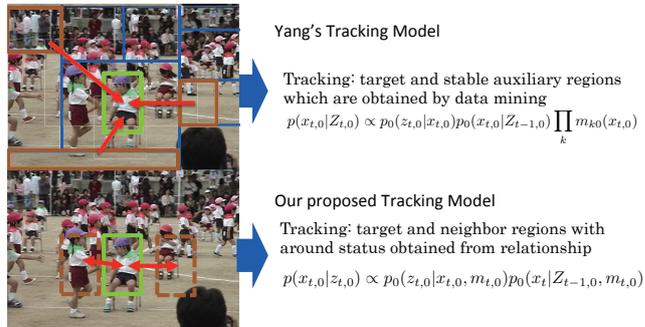


図1 従来手法と提案手法の追跡モデルの違い

ここで、 $x_{t,0}$  は時刻  $t$  における追跡対象の位置であり、 $k$  は補助領域 ( $k = 1, \dots, K$ ) である。また、 $m_{k0}(x_0)$  は補助領域  $k$  から追跡対象へ伝達された情報を示しており、以下のようになる。

$$m_{k0} = \int_{x_{t,k}} p_k(x_{t,k} | Z_{t,k}) \psi_{k0}(x_{t,k}, x_{t,0}) dx_k \quad (4)$$

$\psi_{k0}(x_{t,k}, x_{t,0})$  は、データマイニングにより得られた継続的に出現する補助領域と追跡対象の位置関係に対する確率である。補助領域は、追跡対象と同じ動きをしている領域 (例えば頭部追跡の場合は胴体など) が選ばれることが多い。

## 3. 周辺状況を理解する物体追跡モデル

### 3.1 提案する物体追跡モデル

Yangらの追跡モデルは、図1に示すように、背景の中で継続的に捉えることができる領域と追跡対象との位置関係から追跡対象の位置を推定する。このモデルは、あくまで背景の中で有効な補助領域との位置関係を追跡結果に反映させるのみに留まっており、追跡対象周辺に生じる状況を理解していないといえる。そのため、追跡対象の周辺に類似する対象が存在する場合や遮蔽が発生すると誤追跡が生じる。追跡対象が見えや動き変化に対応するだけでなく、類似領域が周辺に存在するような状況でも、ロバストに追跡するためには、追跡対象の周辺に生じる状況を理解することが重要である。このような視点をもとに、我々は図1に示すような追跡対象周辺の領域から追跡対象の状況を理解する追跡モデルを提案する。我々は、追跡対象周辺の視覚的情報を利用して導いた追跡対象に生じる誤追跡の発生確率を「状況の理解」として定義する。また、視覚的情報は、追跡対象と周辺領域の空間的な関係性と特徴の類似性から表現される。我々の提案手法は、このような視覚的情報から導出した誤追跡の発生確率をもとに追跡の振る舞いを動的に選択する追跡モデルとなる。提案する追跡対象の事後確率は次のようになる。

$$p(x_{t,0} | Z_{t,0}) \propto p_0(z_{t,0} | x_{t,0}, m_{t,0}) p_0(x_{t,0} | Z_{t-1,0}, m_{t,0}) \quad (5)$$

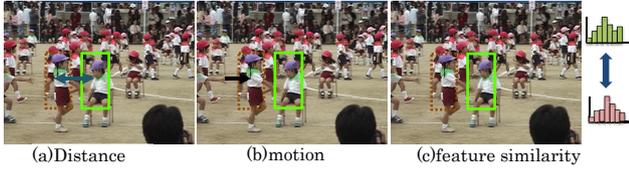


図2 追跡対象と周辺領域の関係

ここで、 $p_0(z_{t,0}|x_{t,0}, m_{t,0})$  は、時刻  $t$  における追跡対象の観測モデル、 $p_0(x_{t,0}|Z_{t-1,0}, m_{t,0})$  は、時刻  $t$  における追跡対象  $x_t$  に対する遷移モデルである。追跡対象の観測モデルおよび遷移モデルは、時刻  $t$  における周辺状況を表す  $m_{t,0}$  を条件とすることで、周辺状況を考慮した事後確率を得ることができる。また、周辺状況  $m_{t,0}$  は、全ての周辺領域の状況であり、以下のように定義する。

$$\begin{aligned}
m_{t,0}(x_{t,0}) &= \prod_k p_k(x_{t,k}|Z_{t,k})\psi'_{k0}(x_{t,k}, x_{t,0}) \\
&\quad - \prod_k p_k(x_{t,k}|Z_{t,0})(1 - \psi'_{k0}(x_{t,k}, x_{t,0})) \\
&= \prod_k p_k(z_{t,k}|x_{t,k})\phi_k(x_{t,k})\psi'_{k0}(x_{t,k}, x_{t,0}) \\
&\quad - \prod_k p_k(z_{t,0}|x_{t,k})\phi_k(x_{t,k})(1 - \psi'_{k0}(x_{t,k}, x_{t,0}))
\end{aligned} \tag{6}$$

ここで、 $\phi_k(x_{t,k})$  は遷移モデルを表す。1つの周辺領域に対して、2つのモデルを適用する。1つは第1項の周辺領域自身の特徴を用いた追跡であり、2つ目は第2項の追跡対象の特徴による追跡である。第1項は周辺領域自身を追跡し、第2項は追跡対象と類似する領域が追跡対象周辺に存在するかどうかを得ることができる。Yangらは、追跡対象と周辺領域の関係  $\psi_{k0}(x_{t,k}, x_{t,0})$  を位置関係に基づいて算出している。一方、本モデルの  $\psi'_{k0}(x_{t,k}, x_{t,0})$  は、追跡対象と周辺領域との空間的な関係性と特徴の類似性をもとに算出する。 $\psi'_{k0}(x_{t,k}, x_{t,0})$  が大きいと  $m_{t,0}$  は正となり、第1項で求める周辺領域自身の追跡結果が重要となる。すなわち、追跡対象の周辺に類似する物体が存在しないことを表している。一方、 $\psi'_{k0}(x_{t,k}, x_{t,0})$  が小さいと  $m_{t,0}$  は負となり、第2項の追跡対象の特徴による観測モデルが重要となる。これは、追跡対象の周辺に類似する物体が存在する確率が高いことを意味している。このように、周辺状況を捉える関数  $\psi'_{k0}(x_{t,k}, x_{t,0})$  から追跡領域の周辺状況を理解することができる。

### 3.2 周辺状況の理解

周辺状況を理解するためには、特徴空間において、追跡対象と類似する領域を捉えることが重要である。また、類似する領域との位置関係だけでなく、移動方向も重要な情報である。追跡対象と周辺領域との空間的な関係性と特徴の類似性を考慮することで、周辺状況を類似物体

への誤追跡の発生確率として理解することが可能となる。これらの関係を図2に示す。これより、周辺状況を理解するために重要な要素は、追跡対象と周辺領域との位置関係  $S_d$ (図2(a))、移動方向の違い  $S_m$ (図2(b))、特徴の類似性  $S_a$ (図2(c))である。これらの要素から追跡対象に対する誤追跡の発生確率を式(7)のように表すことができる。この発生確率を周辺状況を示す値とする。

$$\psi'_{k0}(x_{t,k}, x_{t,0}) = \exp[-S_d(x_{t,k}, x_{t,0})S_m(x_{t,k}, x_{t,0})S_a(x_{t,k}, x_{t,0})] \tag{7}$$

また、位置関係  $S_d$ 、移動方向  $S_m$ 、類似性  $S_a$  はそれぞれ式(8)、式(9)、式(10)のようになる。

$$S_d = \begin{cases} 1 & (\|x_{t,k} - x_{t,0}\|_2) < T_d \\ 0 & otherwise \end{cases} \tag{8}$$

$$S_m = \begin{cases} 1 & \frac{x_{t,k} - x_{t-1,k}}{\|x_{t,k} - x_{t-1,k}\|} \cdot \frac{x_{t,0} - x_{t-1,0}}{\|x_{t,0} - x_{t-1,0}\|} < 0 \\ 0 & otherwise \end{cases} \tag{9}$$

$$S_a = D[x_{t,0}, x_{t,k}] = Q(x_{t,k}) \ln \frac{Q(x_{t,k})}{P(x_{t,0})} \tag{10}$$

式(10)の  $Q$  は周辺領域の特徴ベクトル、 $P$  は追跡対象の特徴ベクトルであり、類似性  $S_a$  は KL divergence により求める。

$\psi'_{k0}(x_{t,k}, x_{t,0})$  は追跡対象と周辺領域の位置関係が一定より近く、移動方向が近づく方向の場合、類似性  $S_a$  が高いと小さくなる。このとき、周辺状況  $m_{t,0}$  は第2項が大きな役割を占め、負の値をとる。すなわち、誤追跡の発生確率が高くなる状況として理解される。

一方、それ以外の場合、 $m_{t,0}$  は第1項が大きな役割を占め、正の値となる。この場合は、誤追跡の発生確率が低い状況として理解される。最終的に、追跡対象の遷移モデルおよび観測モデルは、伝達される値  $m_{t,0}$  の値を条件とした条件付き確率として算出される。

周辺領域は、領域数  $k$  とその位置、および式(8)のしきい値  $T_d$  により理解する範囲を決定することができる。しきい値  $T_d$  を大きくし、領域数  $k$  を増やすことによって画像全体を理解することも可能である。本稿では、誤追跡の発生確率を導出するために、周辺領域を追跡対象の近傍に配置している。

## 4. 提案手法による物体追跡

### 4.1 周辺状況を考慮した物体追跡のフレームワーク

我々が提案する周辺状況を考慮した物体追跡モデルを用いた物体追跡の実現方法について述べる。提案モデルを実現するフレームワークを図3に示す。時刻  $t=0$  において、追跡対象および周辺領域のテンプレートを生成する。また、遷移モデルの初期化を行う。基準となる遷移モデルは等速運動と仮定する。それらをもとに初期フレームにおける周辺状況を理解する。時刻  $t=1$  では、

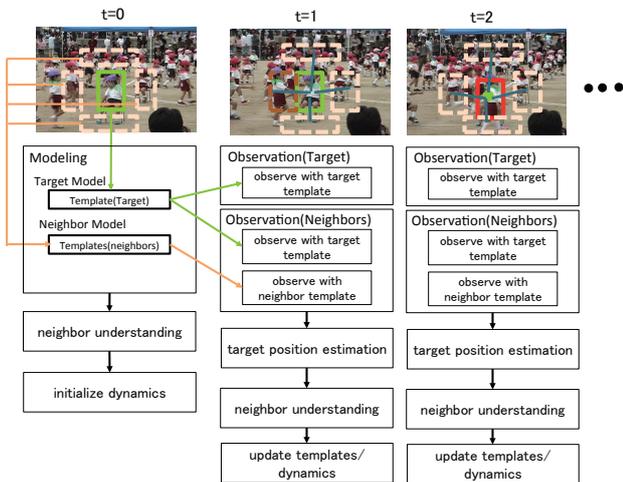


図3 提案手法のフレームワーク

追跡対象と周辺領域をそれぞれ追跡する。追跡対象は、遷移モデルをもとに決められた範囲内で追跡対象の探索を行い、追跡対象位置を推定する。周辺領域も同様に、遷移モデルをもとに、一定範囲内で探索する。各周辺領域に対して、周辺領域自身のテンプレートと追跡対象のテンプレートの両方を用いて追跡処理を行う。すなわち、各周辺領域に対して、2つの追跡結果を得ることができる。次に、推定された結果から追跡対象と周辺領域との関係性を求め、周辺状況を理解する。周辺状況から得た誤追跡の発生確率をもとに、最終的な追跡位置を追跡対象と周辺領域の追跡結果から推定する。また、周辺状況に従い、観測モデルと遷移モデルを更新する。更新された観測モデルおよび遷移モデルをもとに、次フレーム以降は、追跡対象の追跡を行う。

本フレームワークにおける大きなポイントは、追跡対象と周辺領域の関係性をもとに追跡対象周辺の状況を理解し、それに応じた観測モデルと遷移モデルにより最終的な追跡物体の位置を決定できる点である。追跡対象周辺の状況を理解することで、誤追跡が生じる確率の高い領域が存在するかを判断することが可能となる。誤追跡が生じる可能性が高い場合、周辺領域の観測モデルから得られた位置をもとに追跡対象の位置を推定し、それ以外は追跡対象自身の追跡結果から位置を決定する。また、遷移モデルは、周辺状況により、追跡対象をもとにした遷移モデルを利用するか、周辺領域から得られた遷移モデルを利用するかを動的に選択している。ここでは、本フレームワークで用いる観測モデルと遷移モデルおよび周辺状況を考慮した各モデルの選択方法について述べる。

## 4.2 観測モデル

人体や手等のような姿勢変化や形状変化により見えが異なる非剛体の物体を追跡するために、特徴量として色分布を適用する。色分布はヒストグラムで表現し、色の変化に対するロバスト性のあるHSV色空間を用いる。また、白色や黒色の物体を追跡する場合、照明変化が生じ

るとHSV色空間は、色相の値が不安定となる。そこで、照明変動に対するロバスト性を高めるために、色相と彩度および彩度と明度の2つの共起ヒストグラムを結合したジョイントヒストグラムを用いる。

ヒストグラムをベースとした特徴量を用いて、追跡対象のテンプレートと候補領域との類似度を観測する。Mean Shiftなどのカーネルベースの追跡では、勾配法を用いることで、類似度をもとに局所解までの移動距離を得ることができる。候補領域の中心位置を $y_k$ とし、候補領域における位置 $x_i$ に対する追跡対象との類似度を式(11)のように重み付き平均として算出する。これにより、中心座標 $y_k$ からの移動距離 $\hat{x}_k$ を算出することができる。距離 $\hat{x}_k$ が収束するまで繰り返し演算を行い、追跡対象の位置 $x_k^*$ を推定する。

$$\hat{x}_k = \frac{\sum_{i=1}^N (x_i - y_k) w(x_i) K(x_i, y_k)}{\sum_{i=1}^N w(x_i) K(x_i, y_k)} \quad (11)$$

$$x_k^* = x_k + \hat{x}_k \quad (12)$$

ここで、 $w(x_i)$ は重み、 $K(x_i)$ はカーネル関数を表している。重み $w(x_i)$ は、追跡対象のテンプレートと候補領域におけるヒストグラムの類似度から算出される。類似度の算出は、Nguyenら[5]と同様にKL divergenceを用いる。KL divergenceは、非負な値をとり、テンプレートと候補領域の特徴量が一致した場合ゼロとなる特長がある。また、バチャリア係数を用いる場合に比べて、追跡精度が良い。KL divergenceを用いることにより、探索領域の位置 $x_i$ における重み $w_i$ は、

$$w(x_i) = \sum_{u=1}^M \ln\left(\frac{r_u}{1-r_u} \sum_{j=1}^M \frac{s_j}{1-s_j}\right) \delta[b(x_i) - u] \quad (13)$$

$$r_j = \frac{1}{\sum_{x_i \in \chi_M} K(x_i, y_k)} K(x_i, y_k) \delta[b(x_i) - j] \quad (14)$$

$$s_j = \frac{1}{\sum_{x_i \in \chi_D} K(x_i, y_k)} K(x_i, y_k) \delta[b(x_i) - j] \quad (15)$$

となる。ここで、 $u$ はおよび $j$ 、 $b(x_i)$ は $M$ 個のビンで構成されるヒストグラムのビン番号を表している。 $r_j$ は追跡対象のテンプレートにおけるビン $j$ の値をもとに算出した値であり、 $s_j$ は探索領域のテンプレートにおけるビン $j$ の値をもとに算出した値である。また、追跡対象の領域を $\chi_M$ 、探索領域を $\chi_D$ としている。 $\delta[\cdot]$ はクロネッカー関数であり、 $b(x_i) - j = 0$ のとき1、それ以外の場合は0となる。カーネル関数 $K(x_i, y_k)$ には、式(16)のようなEpanechnikovカーネルを用いる。

$$K(\cdot) = \begin{cases} 1 - g(x_i, y_k)^2 & \text{if } \|g(x_i, y_k)\| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

$$g(x_i, y_k)^2 = \frac{4\|y_k - x_i\|^2}{h^2} \quad (17)$$

$h$  はカーネルのバンド幅である。このような勾配法を用いて物体領域 ( $k = 0$ ) および周辺領域 ( $k = 1, \dots, K$ ) の各対象の位置  $x_k^*$  を探索する。各対象に対する探索領域の中心位置  $y_k$  は遷移モデルに基づいて決める。また、重み関数  $w(x_i)$  は、周辺状況に応じて、追跡対象のテンプレートまたは周辺領域のテンプレートのどちらを適用するかを決定する。

### 4.3 周辺状況を考慮した追跡対象の位置推定

式 (12) による追跡対象および周辺領域の探索と、式 (7) から周辺状況を理解することができる。この状況をもとに時刻  $t$  における最終的な追跡対象の位置  $x_{t,0}$  を推定する。周辺状況を理解する値  $m_{t,0}$  が負の場合、追跡対象周辺に類似する対象が存在するため、誤追跡が生じる確率が高くなる。一方、 $m_{t,0}$  が正の場合は、誤追跡の確率が低い。そこで、 $m_{t,0}$  の値から追跡対象と周辺領域のいずれの観測モデルで観測できた結果を用いるかを選択する。 $m_{t,0}$  が正の場合、式 (12) から得た追跡対象の位置  $x_0^*$  をもとに、位置  $x_{t,0}$  を式 (18) のようにする。

$$x_{t,0} = x_0^* \quad (18)$$

一方、 $m_0$  が負の場合は、追跡対象と類似性がしきい値  $T_a$  より低い周辺領域  $D_k$  に対する追跡結果の重み付き平均として、式 (19) のように求める。

$$x_{t,0} = \frac{1}{C} \sum_{k \in K; D_k < T_a} p(x_k | z_k) x_{k,0}^* \quad (19)$$

ここで、 $C$  は類似性の低い領域の個数、 $x_{k,0}^*$  は周辺領域  $k$  から推定される追跡対象の位置である。

### 4.4 周辺状況を考慮した遷移モデルの選択

遷移モデルにより、過去の動きから推定した時刻  $t+1$  における探索領域の中心位置  $y_{t+1}$  を決める。位置推定には、周辺状況をもとに追跡対象の動きと周辺領域の動きを選択する。さらに、本研究では、周辺状況をもとにした時間軸の状況も考慮する。すなわち、周辺状況から誤追跡の生じる確率が高いとされた時刻が続いた場合、追跡対象が移動する範囲が広がる可能性が高い。そこで、周辺領域から推定するフレームが続いた場合、観測する範囲を広げるような遷移モデルを導入する。このような空間的および時間的な周辺状況を考慮した遷移モデル  $\phi(x_{t+1} | x_t, m_0)$  は、位置  $y_{t+1}$  と領域  $r_{t+1}$  から表すことができる。 $y_{t+1}$  および  $r_{t+1}$  を以下のように求める。

$$y_{t+1} = \frac{1}{T} \sum_{i=t-T}^T y_i \quad (20)$$

$$r_{t+1} = \begin{cases} 1.2r_t & \text{if } \alpha_t < 0 \\ r_t & \text{otherwise} \end{cases} \quad (21)$$

$\alpha_t$  は時刻  $t$  における周辺状況を表す式 (7) の  $\psi'_{k,0}(x_k, x_0)$

である。これより、時刻  $t+1$  での位置は、周辺状況に応じて、追跡対象の位置または周辺領域の位置をもとに決めることができる。

### 4.5 周辺領域の更新

周辺領域は周辺領域の状況理解および追跡対象の位置推定のために、重要な役割を果たしている。各周辺領域  $x_k$  に対して、周辺領域自身の特徴による事後確率  $p(x_k | z_k)$  と追跡対象の特徴による事後確率  $p(x_k | z_0)$  の2つの追跡を行う。常に有効な周辺領域を捉えるために、周辺領域の状況に応じて周辺領域の生成と消滅を行う。周辺領域  $x_k$  において、周辺領域自身の特徴による事後確率  $p(x_k | z_k)$  が高い場合、追跡が失敗するまで継続して行う。事後確率  $p(x_k | z_k)$  が低下して追跡が失敗した場合は、新たな周辺領域  $x_{k+1}$  を生成する。また、追跡対象との距離が離れすぎた場合も、その周辺領域を消滅させて新規に生成する。このようなケースは、追跡対象付近に類似した対象が存在していない場合であり、追跡対象または背景が動いているシーンである。

一方、周辺領域  $x_k$  において、追跡対象による事後確率  $p(x_k | z_0)$  が高い場合、すなわち周辺領域  $x_k$  が追跡対象と類似する場合は、現時刻以降で追跡対象が誤追跡する確率が高くなる。そのため、周辺領域において、追跡対象の特徴による追跡を継続して行う。また、追跡対象による事後確率が低い場合、すなわち周辺領域  $x_k$  がに類似する領域でない場合は、その周辺領域を消滅させ、新規に生成する。このように、周辺領域の状況に合わせて周辺領域の生成と消滅を行う。

## 5. 評価実験

提案手法の有効性を示すために、提案手法と Yang らの手法 [6]、および周辺領域を利用しない場合 [5] と Mean-Shift を比較する。Yang らの手法は、追跡対象に楕円モデル、補助領域に色ヒストグラムを特徴量として用い、Mean Shift 法により追跡する。評価には、追跡対象と類似色の物体との重なりや追跡対象の移動、カメラの移動などが複合的に生じる一般的なホームビデオを用いる。各評価データの詳細は各比較実験で述べる。また、観測モデルとして用いる HSV の色ヒストグラムは、色相と彩度と明度のビン数をそれぞれ 16 としている。提案手法で用いる周辺領域の個数は 4 つとし、追跡対象の上下左右に配置する。周辺状況を理解するための各パラメータは、追跡対象のサイズを  $S_0$  として、 $T_d$  は  $S_0$  の 2 倍、 $T_m$  は  $S_0$  の 0.5 倍とする。追跡対象の初期位置は全ての手法で共通の位置とする。

### 5.1 類似物体が遮蔽するシーンでの性能

追跡対象と類似する色の服装をした人物が追跡対象を遮蔽して横切るような 150 フレームから構成されるシーンでの性能を比較する。提案手法および、従来手法に対

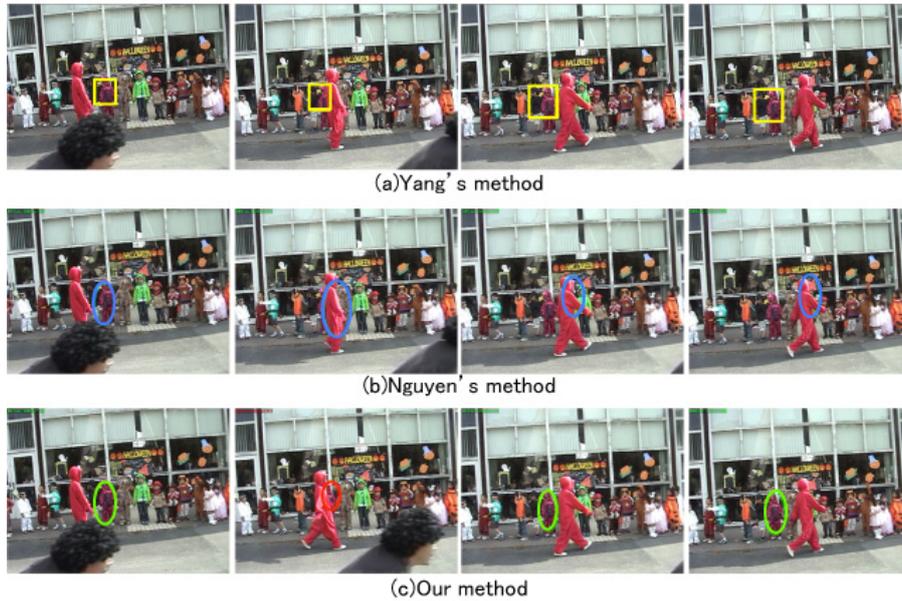


図 5 類似物体が遮蔽するシーンでの追跡結果例。各行の右から 60 フレーム目, 70 フレーム目, 80 フレーム目, 100 フレーム目の結果。

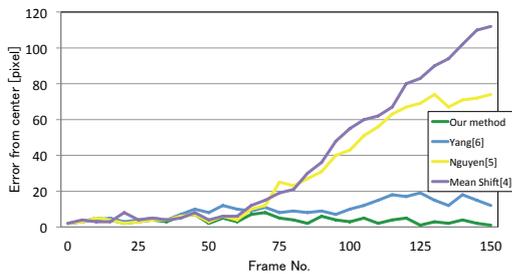


図 4 類似物体が遮蔽するシーンでの誤差。

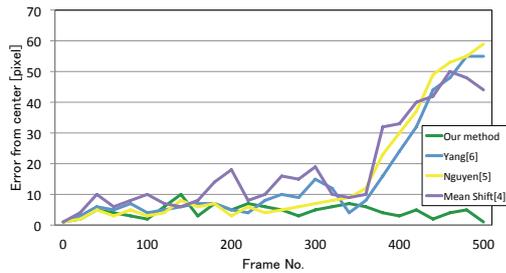


図 6 複雑背景下で交差が連続するシーンでの誤差。

する各フレームにおける中心位置からの誤差を図 4 に示す。また、代表的なフレームでの各手法の追跡結果を図 5 に示す。70 フレーム付近において、追跡対象の前を類似する人物が横切っている。このフレーム以降、周辺領域を用いない手法は横切る人物を誤追跡し、中心位置からの誤差が大きくなっている。一方、提案手法および Yang らの手法は、遮蔽が生じた際、周辺領域から追跡対象の位置を推定するため、80 フレーム付近で横切った後も正しく追跡することができている。ここで、提案手法の追跡結果が赤い楕円の場合、周辺状況から誤追跡の発生確率が高い状態を示している。

## 5.2 複雑背景下で交差が連続するシーンでの性能

複雑な背景状況において追跡対象の移動と遮蔽により

交差が連続して生じるシーンでの追跡性能を比較する。評価データは、運動会で追跡人物と類似する服装の人物が背景に存在している。また、追跡対象が類似する人物の前を横切り、椅子に座った後、その前を類似する服装の人物が通過する。評価データのフレーム数は、500 フレームである。提案手法および、従来手法に対する各フレームにおける中心位置からの誤差を図 6、代表的なフレームでの追跡結果を図 7 に示す。200 フレーム付近における追跡対象が類似する人物を横切るシーンでは、提案手法および従来手法とも乗り移ることなく追跡することができている。しかしながら、350 フレーム付近の追跡対象の前を類似する人物が横切るシーンにおいて、周辺領域を用いない手法だけでなく、Yang らの手法も横切る人物を誤追跡する。一方、提案手法はこのようなシーンでも追跡対象を正しく追跡することができている。Yang らの手法は追跡対象付近に類似する物体が存在しているかどうかを判断できないため、遮蔽が生じた際に対象物体または類似物のいずれを追跡しているのか判断ができない。また、背景領域の物体が動き、5.1 節のように安定的な補助領域がないため誤追跡が生じる。提案手法は、周辺状況から誤追跡の発生確率を導いて、誤追跡の確率が高い場合に周辺領域の中で追跡対象と類似しない領域をもとに位置を推定するため、遮蔽が生じた際も追跡対象の位置を正しく捉えることができている。

## 5.3 追跡対象の動きと遮蔽を伴うシーンの性能

複雑な背景下での追跡対象の移動と類似する人物との遮蔽が生じるようなシーンでの性能を比較する。追跡対象は右へ移動し、類似する人物が右から左へ動きながら追跡対象と重なるような 300 フレームからなる評価データである。このようなシーンにおける代表的なフレームでの追跡結果を図 8 に示す。170 フレーム付近で追跡対

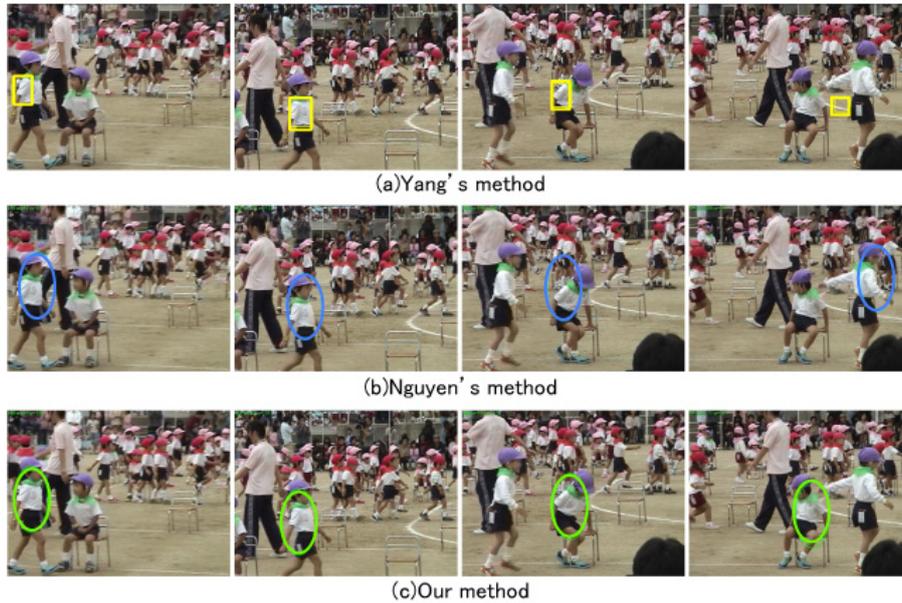


図 7 複雑背景下で類似物体の交差が連続するシーンでの追跡結果例. 各行の右から 200 フレーム目, 300 フレーム目, 350 フレーム目, 370 フレーム目の結果.

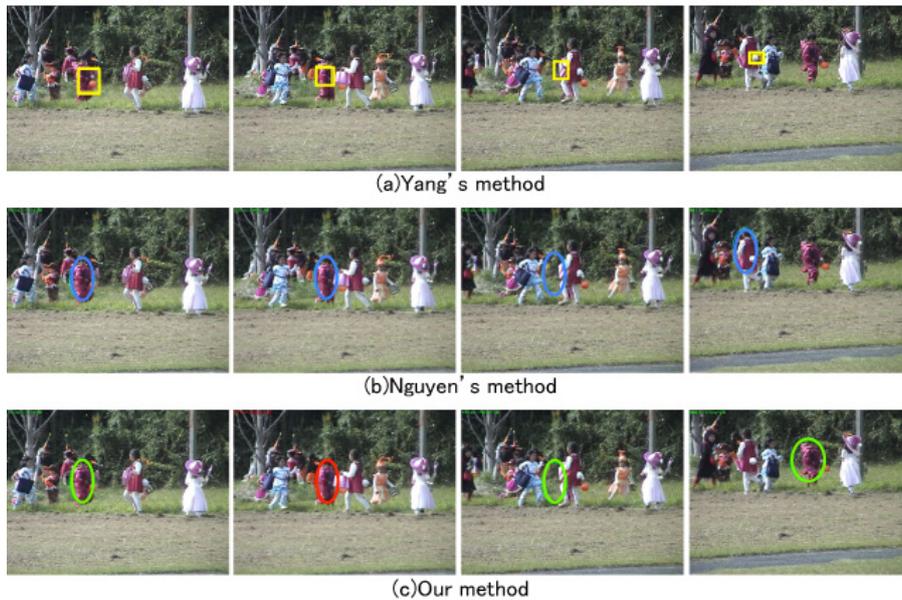


図 8 追跡対象の動きと類似物との遮蔽を伴うシーンでの追跡結果例. 各行の右から 160 フレーム目, 170 フレーム目, 180 フレーム目, 190 フレーム目の結果.

象と類似する物体との交差が発生し、180 フレーム付近で追跡対象が再度現れる。提案手法は、このような動きを伴う交差による重なりが発生した場合でも追跡対象の位置を正しく追跡することができている。一方、その他の従来手法は、交差する際に、別の人物を誤追跡している。このような、追跡対象が移動する場合においても提案手法は、類似する人物へ誤追跡することなく追跡対象を正しく追跡することができている。

#### 5.4 周辺領域数による追跡性能の比較

周辺領域の個数を変えた場合の追跡結果を比較する。評価には、5.2 節に用いた運動会での評価データを用いる。周辺領域は 0 個から 10 個まで 2 つずつ増やしてい

き、追跡対象を中心に円状に等間隔に基準位置を決める。各位置において、乱数ノイズを加えて、初期位置を位置をずらしている。評価方法には、追跡対象の外接円内に追跡対象の中心位置が含まれている場合、追跡成功とし、全フレーム数に対する追跡成功数を追跡成功率として算出する。追跡結果およびその際の処理速度を表 1 に示す。また、処理速度あたりの追跡性能をコストパフォーマンスを表す性能係数として示す。処理速度の計測は、Core2 Duo 2.4GHz のパソコンでの速度である。表 1 より、周辺領域を増やすと追跡性能成功率は向上するが 6 個の場合を境に低下する。これは、誤追跡の発生確率が高いフレームが増え、周辺領域からの推定位置と実際の位置関係がずれる場合が増加するためである。一方、追跡成功

表 1 周辺領域数による追跡性能の比較

周辺領域数	追跡成功率	処理速度	性能係数
0	72.2%	1.1ms	65.3
2	85.1%	1.16ms	73.3
4	89.1%	1.20ms	74.2
6	89.8%	1.23ms	72.6
8	87.8%	1.29ms	68.0
10	82.8%	1.34ms	61.8

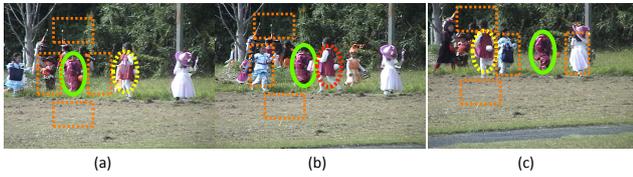


図 9 追跡対象と周辺領域の関係

率は、周辺領域を 6 個とした場合が最も良いが、処理速度あたりの性能は 4 個の場合の方がよい。これより、周辺領域を 4 個にした場合が最もコストパフォーマンスが高いといえる。

## 5.5 考察

図 9 に追跡対象と類似する人物が追跡対象の前を横切るフレームにおける追跡領域と周辺領域の追跡結果を示す。各図の緑の楕円は追跡結果であり、点線の矩形は周辺領域自身の観測モデルによる周囲領域の追跡結果、また点線の楕円は追跡対象の観測モデルを用いた周囲領域の追跡結果である。点線の楕円について、黄色は追跡対象と類似する領域が一定距離以上離れている場合、赤色は近い場合を示している。これより、追跡対象周辺において、類似する領域が一定距離以上離れている場合 (a)、周辺に類似する領域がなく、誤追跡の発生確率は小さくなっている。一方、追跡対象と類似する領域が一定距離以内に存在する場合 (b)、誤追跡の発生確率は高くなっている。また、追跡対象と類似する領域が再度離れると (c)、誤追跡の発生確率は小さくなる。このように、追跡対象周辺の状況を理解することで、類似領域が存在する場合でも途切れなく追跡することが可能である。

## 6. まとめ

本稿では、物体追跡の新たな手法として、追跡対象だけでなく周辺領域の状況を理解することが可能な追跡モデルを提案した。提案手法は、追跡対象の周辺領域における状況を理解して、その状況に応じて、追跡の振り舞いを変えることで、類似する物体や背景への誤追跡を防ぐことが可能となる。評価実験では、提案手法の有効性を示すために、類似する物体や背景が存在するシーンを対象に比較実験を行った。この評価実験を通して、提案手法が類似物体が近くにいたり重なりが生じるようなより困難なシーンでも追跡することが可能であることを示すことができた。

提案する追跡モデルは、特定の観測モデルや遷移モデ

ルによらず、勾配法以外への応用も可能である。今後は、Particle Filter などの他の手法への展開も検討する。

## 文献

- [1] M. Isard and A. Blake, “CONDENSATION—Conditional Density Propagation for Visual Tracking”, *International Journal of Computer Vision*, Vol.29, No.1, pp.5–28, 1998.
- [2] K. Nummiaro, E. Koller-Meier, L. Van Gool, “An adaptive color-based particle filter”, *Image and Vision Computing*, Vol.21, No.1, pp.99–110, 2003.
- [3] C. Yang, R. Duraiswami, D.Larry, “Efficient mean-shift tracking via a new similarity measure”, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.176–183, 2005.
- [4] D. Comaniciu, V. Ramesh, P. Meer, “Kernel-based object tracking”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol.25, No.5, pp.564–577, 2003.
- [5] Q. Nguyen, A. Robles-Kelly, C. Shen, “Kernel-based Tracking from a Probabilistic Viewpoint”, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–8, 2007.
- [6] M. Yang, Y. Wu, S. Lao, “Intelligent Collaborative Tracking by Mining Auxiliary Objects”, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.697–704, 2006.
- [7] W. He, T. Yamashita, H. Lu, S. Lao, “SURF Tracking”, *IEEE International Conference on Computer Vision*, pp.1586–1592, 2009.
- [8] Y. Wu, G. Hua, T. Yu, “Switching Observation Models for Contour Tracking in Clutter”, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.295–302, 2003.
- [9] Y. Li, H. Ai, T. Yamashita, S. Lao, M. Kawade, “Tracking in Low Frame Rate Video: A Cascade Particle Filter with Discriminative Observers of Different Life Spans”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.30, No.10, pp.1728–1740, 2008.
- [10] H. Grabner, J. Matas, L. Van Gool, “Tracking the invisible: Learning where the object might be”, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–8, 2010.
- [11] B. Babenko, H-M Yang, S. Belongie, “Visual Tracking with Online Multiple Instance Learning”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.983–990, 2010.
- [12] K. Okuma, A. Taleghani, N. Freitas, J. Little, D. Lowe, “A boosted particle filter: Multitarget detection and tracking”, *European Conference on Computer Vision*, pp.28–39, 2004.
- [13] K. Ishiguro, T. Yamada and N. Ueda, Simultaneous Clustering and Tracking Unknown Number of Objects, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–8, 2008.
- [14] R. T. Collins, Y.Liu, “Online selection of discriminative tracking features”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 27, No.10, pp.1631–1643, 2005.
- [15] 小関亮介, 藤吉弘亘, 秋田時彦, 柿並俊明, “協調的な複数 Mean-Shift トラックによる後方車両追跡”, *画像の認識・理解シンポジウム (MIRU2005)*, pp.419–426, 2005.
- [16] 山下隆義, 勞世紅, 川出雅人, “Online Real Boosting による人物追跡”, *情報処理学会論文誌 CVIM*, Vol.1, No.1, pp. 73–82, 2008.