

# 属性情報を導入した隠れに頑健な特徴点ベース物体追跡

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 2-2-1

<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、複雑な背景下で追跡対象の形状変化や照明変化、隠れ等の様々な要因により生じる見えの変化に対して頑健に追跡することが求められている。見えの変化に対応するためには、変化に頑健な特徴量とともに、追跡対象のモデル生成およびその更新方法が重要である。本稿では、モデル生成及び更新時に生じる背景の混入による誤追跡を抑制するために、属性情報を導入した特徴点ベースの追跡手法を提案する。提案手法では2種類の属性情報を導入する。1つ目は所属属性であり、追跡対象または背景領域の特徴点であるかを区別するだけでなく、背景から追跡対象内に混入した特徴点かどうかを区別する。2つ目は記憶属性であり、人の記憶モデルにもとづき、特徴点を一時記憶、短期記憶、長期記憶、さらに刷り込み記憶の4つに分類し、それぞれ異なる保持期間で追跡対象のモデルに含まれる。これらの属性情報を導入することで、従来手法が問題とした隠れに対する頑健性が低いこと、およびモデルの劣化を解決し、追跡失敗を防ぐことが可能となる。

**キーワード** 物体追跡, 遮蔽, 所属属性, 記憶属性, 記憶モデル

## 1. はじめに

物体追跡は、コンピュータビジョンの重要なトピックの一つであり、複雑な背景下で追跡対象の形状変化、照明変化、部分的もしくは全体の隠れなどの様々な要因で生じる見えの変化に対して、頑健に追跡することが求められている。このような見えの変化に対応するためには、変化に頑健な特徴量とともに、追跡対象のモデル生成およびその更新方法が重要であり、様々な手法が提案されている。追跡対象のモデルは、追跡対象の矩形全体を捉える方法 [2] [3] [4] [5] と局所領域の集合と見なす方法 [11] [14] [16] がある。また、モデルの更新方法は、オンライン学習によりモデルを逐次更新する方法 [6] [8] [10] [12] [15] が提案されている。従来手法は、見えの変化に対して一定の成果をあげているが、以下のような問題点がある。

(1) 隠れに対する頑健性が低い。

追跡対象のモデルは、一般的に追跡対象の特徴を用いて生成される。その際、背景との区別の有効な特徴量を用いて生成される場合もある。そして、生成されたモデルと類似する位置を探索し、追跡対象の位置を推定する。しかしながら、隠れが生じた場合、類似する位置を推定することができず、追跡に失敗する。

(2) 追跡対象のモデル劣化による追跡失敗。

物体追跡は、初期に与えられた追跡対象のモデルをもとに、時々刻々と変化する環境のなかで追跡対象の位置を推定し続ける。追跡対象または環境の変化に対応するためには、モデルの更新が重要である。しかしながら、常に追跡対象の位置を正確に推定することは困難であり、

不正確な状況でもモデルは更新される。このような不正確な状況でモデルを更新すると、背景の混入によるモデルの劣化が生じ、追跡が失敗しやすくなる。

本稿ではこれらの問題点を解決するモデルの生成方法および更新方法を提案する。モデルは追跡対象だけでなく、背景領域のモデルも同時に生成する。モデルを特徴点の集合と見なし、各特徴点に追跡対象か背景領域かまたは悪影響をおよぼすかを表す所属属性を付与する。各特徴点に所属属性を付与することで、追跡対象の領域内に含まれる背景の特徴点を排除でき、モデルの劣化を防ぐと共に、背景の特徴点が追跡対象に近づいているのかを判断することができ、隠れが生じるか否かを事前に理解することが可能となる。

モデル更新は、各特徴点の出現頻度に応じて保持する期間を決める更新モデルを導入する。人は物事を記憶する場合、繰り返し出現する情報は長期間保持し、一時的に必要な情報は短期間のみ保持している。このような人の記憶モデルを参考に、特徴点の記憶期間を決める記憶属性を各特徴点に付与する。出現頻度が高い特徴点は長期間保持することで、追跡対象の変化が生じても過去の特徴点を記憶することができ、モデルの劣化を抑えることが可能となる。我々は特徴点ベースの物体追跡において、各特徴点に2種類の属性情報を付与し、それを活用することで隠れや追跡対象の見えの変化に頑健な物体追跡手法を提案する。

## 2. 関連研究

追跡対象の見えの変化に対する頑健性を向上させるために、様々なモデルの生成や更新方法が提案されている。

Isard [2] や Comaniciu ら [5] は、追跡対象領域内の色や勾配の分布をヒストグラム特徴としてモデルを生成している。Black [3] や Ross ら [12] は、部分空間を用いて追跡対象の矩形全体を特徴空間に射影して追跡対象のモデルを生成している。追跡対象の領域全体に着目した場合、類似する物体を誤追跡しやすい。

また、見えの変化に対応するために、物体検出を応用して追跡対象と背景との区別に有効な特徴量を用いてモデルを表現する方法がある。Collins ら [6] は、追跡対象と背景を区別できる色特徴量を選択している。Avidan ら [10] は、色特徴と勾配ヒストグラムを利用して、追跡対象と背景を区別する複数の検出器を結合して追跡している。Grabner ら [8] は、追跡対象全体を複数の局所特徴で表現し、ブースティングにより局所特徴を選択することで見えの変化に対応している。このような検出器をベースとする手法は、追跡対象領域内に背景が混入すると、正しいモデルを生成することができず、追跡が失敗しやすい。

追跡対象のモデルとして、追跡対象全体を捉えずに局所領域の集合として扱う方法がある。He ら [14] は、追跡対象を不変特徴量の一つである SURF の集合として表現し、追跡対象の矩形内に背景が含まれても追跡が失敗しないようにしている。Godec ら [21] は、一般化ハフ変換により、各特徴点と追跡対象の中心位置との関係を投票して推定することで背景情報の混入に対応している。Yang ら [11] は、追跡対象と背景を区別できる局所領域の集まりで表現し、各局所領域の中心位置との位置関係により背景の混入に対応している。これらの手法は背景の混入による追跡失敗を抑制しているものの、背景に対象と類似した領域があると領域全体をモデル化する場合と同様に追跡が失敗する。Fan ら [18] は背景の混入を防ぐために、セグメンテーションを同時に行う方法を提案している。この方法は、形状の変化や背景の混入による追跡失敗を抑制することができるが、全体の隠れに対応できず、計算コストが大きい。

一方、モデルの更新は、近年オンライン学習による逐次的な適応方法が盛んに提案されている。オンライン学習は、見えの変化などにより追跡対象矩形内に背景が混入した場合、背景の情報を含めてモデルを更新するため、モデルが劣化し追跡が失敗しやすい。Grabner [13] は、Semi-online Boosting により、ラベル付きデータとラベルなしデータを併用することで背景の情報が含まれてもモデルの劣化を防いでいる。Babenko ら [15] は、モデル更新の過程で、学習サンプルを蓄積しておき、その蓄積された学習サンプルと背景を区別するようなモデルに更新している。追跡対象の蓄積サンプル内に1つ以上の正しいサンプルがあれば、追跡失敗を防ぐことができる。

これまでの追跡対象のモデル生成および更新は追跡対象のみに着目、または追跡対象と背景との区別に主眼がおかれている。Yang ら [9] は、追跡対象を補助する領域

を見つけ、追跡対象と補助領域が協調して追跡する手法を提案している。追跡対象と同時に頻繁に出現する背景の領域を補助領域として捉え、追跡対象に隠れが生じた場合に、補助領域から位置を推定することができる。しかしながら、補助領域は、追跡対象と同じような動きをしていることが求められている。Grabner ら [16] も背景に存在する特徴点から追跡対象までの位置関係を利用して、追跡対象の位置を推定している。これらは、追跡対象に隠れが生じた際に、背景領域の情報を利用して追跡対象の位置を推定しており、類似物による隠れなどの隠れたかどうかを判断しにくい場合は、うまく推定することができない。Dinh ら [20] は、追跡対象と類似する物体が背景に含まれるか否かをあらかじめ求めておくことで類似物への誤追跡を防いでいる。山下ら [22] は、追跡対象とその周辺の背景領域を同時に追跡すると共に、追跡対象のモデルを背景の追跡に利用して類似物体が存在するか否かを誤追跡の発生確率として求めている。これらは、隠れに対する頑健性を向上させているが、追跡対象のモデルに背景が混入した場合に追跡が失敗する。

このように、追跡対象のモデル生成や更新方法について様々な手法が提案され、モデルの劣化および隠れに対する追跡失敗を解決しようと試みている。しかしながら、モデルの劣化と隠れを同時に解決することはできていない。本稿では、特徴点に属性情報を導入することで、モデル生成と更新時に背景情報の混入を防ぐと同時に隠れに対する頑健性を向上させる追跡モデルを提案する。

### 3. 属性を付与した物体追跡モデル

#### 3.1 従来の特徴点ベースの物体追跡

特徴点ベースの物体追跡は、前時刻までに蓄積された特徴点の集合  $FP = \{F_{m,1}, \dots, F_{m,j}\}$  と時刻  $t$  で検出された特徴点の集合  $F_t = \{F_{t,1}, \dots, F_{t,i}\}$  の対応付けを行い、対応関係から追跡対象の位置  $x_t^*$  を求める。現時刻  $t$  の特徴点  $F_{t,i} = \{x_i, v_i\}$  は位置  $x_i$  と特徴ベクトル  $v_i$  を持ち、集合  $FP$  に蓄積された特徴点  $F_{m,j} = \{y_j, v_j, \bar{x}_j\}$  は位置  $y_j$  と特徴ベクトル  $v_j$  とともに、追跡対象の中心からの変位  $\bar{x}_j$  を持っている。対応付けは、式 (1) のように、 $F_{t,i}$  の特徴ベクトル  $v_i$  と最も類似する特徴ベクトルを持つ特徴点  $F_{m,j}$  を内積  $D$  などにより  $FP$  から探し、スコア  $\rho(F_{t,i})$  を計算する。

$$\rho(F_{t,i}) = \max_{F_{m,j} \in FP} D(v_i, v_j) \quad (1)$$

時刻  $t$  の特徴点のうち対応付けができた特徴点の集合を  $FS$  とし、対応付けができた特徴点ペアから追跡対象の中心位置  $x_{c,i}$  を  $x_{c,i} = x_i + \bar{x}_j$  のように推定する。推定位置  $x_{c,i}$  を中心とした位置関係の尤度を  $p(x'|x_{c,i})$  とし、式 (2) のように位置  $x'$  の尤度を求める。尤度  $p(x')$  が最も高い位置を追跡対象の中心位置  $x_t^*$  としている。

$$p(x') = \sum_{F_{t,i} \in FS} p(x'|x_{c,i})\rho(F_{t,i}) \quad (2)$$

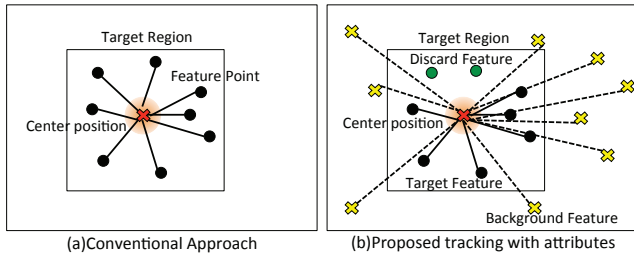


図 1 追跡モデルの比較

### 3.2 属性情報を導入した物体追跡モデル

従来の特徴点ベースの物体追跡は、検出された特徴点のうちの特徴点を用いて追跡対象の位置を推定するかが重要となる。従来手法は、図 1(a) に示すように、追跡対象の領域に含まれる特徴点をもとに、追跡対象の位置を推定している。しかしながら、追跡対象が遮蔽される場合や背景を含むような場合は、正しく追跡できない。そこで我々は、周辺状況を理解する所属属性  $A_s$  と特徴点の重要性を示す記憶属性  $A_m$  の 2 つの属性情報を各特徴点に与えることにより、このような問題を解決する。所属属性  $A_s$  は、追跡対象の特徴点を示す TargetFeature(TF)、背景の特徴点を示す BackgroundFeature(BF) とともに、追跡に悪影響をおよぼす特徴点を示す DiscardFeature(DF) の 3 つの属性を定義する。また、記憶属性  $A_m$  は、人の記憶モデルを参考に一時記憶を示す IconicMemory(IM)、短期記憶を示す ShortMemory(SM)、長期記憶を示す LongMemory(LM) および刷り込み記憶を示す ImprintingMemory(FM) の 4 つの属性を定義する。

このような属性情報を利用することにより、検出された特徴点は  $F_{t,i} = \{x_i, v_i, A_{s,i}, A_{m,i}\}$ 、蓄積される特徴点は  $F_{m,j} = \{x_j, v_j, \bar{x}_j, A_{s,j}, A_{m,j}\}$  となる。追跡対象の位置は、図 1(b) に示すように、遮蔽など追跡対象の状況に合わせて TF 属性と BF 属性の特徴点を用いて推定する。また、DF 属性から追跡対象に隠れが生じるか否かを判断することができる。一方、記憶属性は、追跡対象のモデルを更新する際に、各特徴点の出現頻度に応じて記憶する保持期間を設定することで、オンライン更新による過適合や隠れによる追跡失敗を防ぎ、追跡対象の見えの変化に対する追従性の向上に寄与する。

従来の特徴点による物体追跡の式 (2) は、属性情報を考慮することで、式 (3) のように定義することができる。

$$p(x') = \alpha \sum_{F_{t,i} \in FS_o} w(A_{m,i}) p(x' | x_{c,i}) \rho(F_{t,i}) + (1 - \alpha) \sum_{F_{t,i} \in FS_b} w(A_{m,i}) p(x' | x_{c,i}) \rho(F_{t,i}) \quad (3)$$

特徴点の集合  $FS$  は  $FS = FS_o \cup FS_b \cup FS_d$  のように、TF 属性の部分集合  $FS_o$ 、BF 属性の部分集合  $FS_b$  および DF 属性の部分集合  $FS_d$  の 3 つの部分集合の和

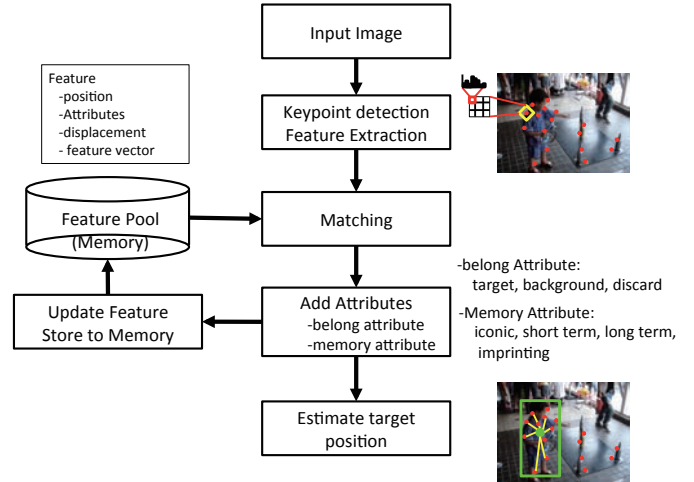


図 2 提案する物体追跡の流れ

となる。提案手法では、追跡対象の位置に対する尤度を所属属性が TF と BF の特徴点から求める。 $\alpha$  は所属属性に対する重みであり、追跡対象または背景のいずれの推定位置を重視するか決める値となる。また、 $w(A_{m,i})$  は、記憶属性に対する重みであり、現時刻においてどの記憶を重視するか決める値である。それぞれの重みは、逐次更新する。

### 4. 提案手法の流れ

我々が提案する物体追跡の流れを図 2 に示す。まず、各時刻において特徴点の検出および特徴量抽出を行い、現時刻の特徴点  $F_{t,i}$  と記憶されている特徴点  $F_{m,j}$  との対応付けを行う。対応した特徴点の所属属性および記憶属性を現時刻の特徴点に付与し、追跡対象の位置  $x_t^*$  を推定する。記憶されている各特徴点は、位置関係や対応付けの回数をもとに所属属性および記憶属性を更新し、その属性情報から変位や特徴ベクトルを更新する。また、記憶の保持期間をもとに不要な特徴点を集合  $FP$  から削除し、現時刻で検出された特徴点を集合  $FP$  に追加する。

#### 4.1 特徴量抽出

特徴点は、FAST Detector [19] により抽出し、特徴ベクトルは、Grabner ら [16] が用いている SIFT ベースの簡易な特徴記述を用いる。抽出した各特徴点について、特徴点を中心に 30x30 ピクセルの注目領域を設定し、16 方向の勾配ヒストグラムを算出する。ヒストグラムの値が最大となる方向を主方向とし、注目領域を回転させる。次に、注目領域を 3x3 のブロックに分割し、各ブロック内について 8 方向の勾配ヒストグラムを算出する。特徴量は、各勾配ヒストグラムを線形結合した 72 次元の特徴ベクトル  $v$  となる。

#### 4.2 特徴点の対応づけ

現時刻で抽出した特徴点  $F_{t,i}$  と記憶されている特徴点  $F_{m,j}$  の対応付けを行う。それぞれの特徴ベクトルを  $v_i$

および  $v_j$  として、対応付けは式 (4) のように正規化相互相関を用いて行う。

$$\rho(F_{t,i}) = \max_{F_{m,j} \in FP} \frac{\sum_{n=1}^N v_i^n v_j^n}{\sqrt{\sum_{n=1}^N v_i^{n2}} \sqrt{\sum_{n=1}^N v_j^{n2}}} \quad (4)$$

ここで、 $n$  は特徴ベクトルの要素位置である。記憶されている全ての特徴点との相関係数をスコアとして算出し、スコア  $\rho(F_{t,i})$  がしきい値  $TH_m$  より大きい特徴点ペアの特徴点  $F_{t,i}$  を集合  $FS$  に含める。

### 4.3 所属属性について

所属属性  $A_s$  は、追跡対象の矩形内に存在する場合、TF 属性を付与し、追跡対象の矩形外の場合は、BF 属性を付与する。また、追跡対象の矩形内への背景からの不要な点の混入を防ぐために、DF 属性を導入する。DF 属性は、追跡対象の矩形領域内に存在する特徴点  $F_{t,i}$  が BF 属性をもつ特徴点  $F_{m,j}$  と対応付けされた場合、付与される。DF の所属属性を導入することで、追跡対象領域内にある不要な特徴点を削除でき、更新時のモデルの劣化による追跡失敗を防ぐことが可能となる。また、DF 属性の特徴点をもとに、周辺状況を理解し、隠れが生じるかどうかを判断することができる。

### 4.4 所属属性による周辺状況の理解

DF 属性をもつ特徴点は、背景領域から追跡対象領域に混入した特徴点であり、この特徴点が多いと追跡対象に隠れが生じた状態を示している。隠れが生じている状態では、追跡対象を表す TF 属性の特徴点が少なくなっており、これらの点から追跡対象の位置を推定すると、誤追跡が生じる可能性が高くなる。このような場合は、背景領域を表す BF 属性の特徴点を利用して追跡対象の位置を推定する方がよい。そこで、DF 属性の特徴点数と TF 属性の特徴点数に応じてどちらの推定結果を重視するかを重みとして決める。式 (3) の重み  $\alpha$  を周辺状況を理解する値とし、以下のように求める。

$$\alpha = \begin{cases} 1 - \frac{(|FS_d|)^2}{(|FS_o|)^2 + 1} & |FS_o| \geq |FS_d| \\ 0 & otherwise \end{cases} \quad (5)$$

DF 属性の特徴点数が増えるに従い、背景領域からの推定位置を重視し、TF 属性の特徴点数より多くなると、背景領域から追跡対象の中心位置を推定する重みとなっている。

### 4.5 記憶属性について

我々はスクワイアの記憶分類 [1] をもとに、各特徴点を分類し、記憶属性を付与する。スクワイアの記憶分類では、人の記憶は感覚記憶 (IM)、短期記憶 (SM)、長期記憶 (LM) の 3 つからなる。IM は、一時的な記憶であり、人の場合は 1 秒から 2 秒程度である。SM は、20 秒程度保持される記憶であり、時間と共に忘却される。SM の記

表 1 記憶属性の分類

記憶属性	属性名	記憶時間
感覚記憶	Iconic Memory	1-2 sec
短期記憶	Short Term Memory	< 20 sec
長期記憶	Long Term Memory	> 20 sec
刷り込み記憶	Imprinting Memory	—

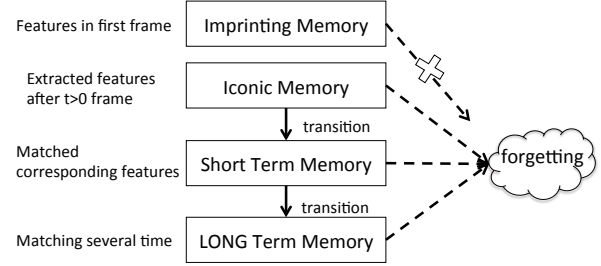


図 3 各記憶属性の遷移

憶のうち、緻密化リハーサルと呼ばれる繰り返し処理を行うことで SM から LM に移行される。我々は、このような人の記憶特性とともに、動物がもつ刷り込み (FM) を記憶属性として導入する。FM は、特定の物事が短時間で長期的な記憶として保持される学習機能である。このような 4 つの記憶属性および記憶時間の範囲をそれぞれ表 1 に示す。また、各記憶属性の遷移について、図 3 に示す。FM 属性は、最初のフレームで抽出された追跡対象の特徴点に付与され、忘却せずに恒久的に記憶しておく。また、IM 属性は各時刻で検出した特徴点を一時的に記憶する。IM 属性や SM 属性は対応付けの回数にもとづき、SM 属性や LM 属性に遷移し、特徴点は各記憶属性に応じた保持期間記憶される。このように、人の記憶モデルのように、対応付けの状況に応じた記憶属性を付与し、記憶することで、対応付けの精度を向上させることが可能である。

### 4.6 各記憶属性の重要性

対応付けできた特徴点の集合  $FS$  から各記憶属性の重要性を重み  $w(A_{m,i})$  として求める。人の記憶は、記憶属性により保持期間が異なるだけでなく、直前に正しく判断した記憶属性の情報ほど重きを置く。そこで、この特性を導入し、多くの対応付けができた記憶属性ほど大きな重みを与えるようにする。IM 属性の特徴点の部分集合を  $FS_I$ 、SM 属性の特徴点の部分集合を  $FS_S$ 、LM 属性の特徴点の部分集合を  $FS_L$ 、FM 属性の特徴点の部分集合を  $FS_F$  とした場合、各記憶属性の重みは、式 (6) のようになる。

$$w(A_{m,i}) = \begin{cases} \frac{|FS_I|}{|FS|} & (A_{m,i} = IM) \\ \frac{|FS_S|}{|FS|} & (A_{m,i} = SM) \\ \frac{|FS_L|}{|FS|} & (A_{m,i} = LM) \\ \frac{|FS_F|}{|FS|} & (A_{m,i} = FM) \end{cases} \quad (6)$$

各記憶属性の重みは、各時刻で逐次更新する。

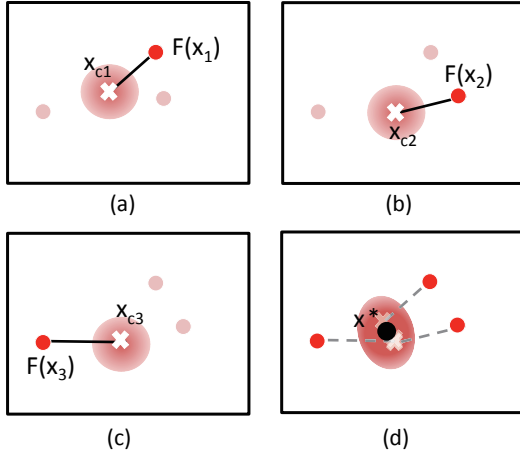


図4 各特徴点からの位置の推定およびその統合処理

#### 4.7 追跡対象の位置推定

図4に示すように、従来手法と同様に、各特徴点ごとに中心位置を推定する。各特徴点からの推定位置  $x_{c,i}$  を中心に、位置関係に関する尤度  $p(x'|x_{c,i})$  を正規分布として仮定する。集合  $FS$  に含まれる全ての特徴点について、推定位置  $x_{c,i}$  および尤度  $p(x'|x_{c,i})$  を算出し、式(3)のように、記憶属性の重み  $w$  と対応付けのスコア  $\rho$  および所属属性に関する重み  $\alpha$  から  $p(x')$  を求める。そして、尤度  $p(x')$  が最大となる位置  $x'$  を推定位置  $x^*$  として求める。

また、追跡対象の矩形の大きさは、式(7)のように、現時刻の中心位置  $x^*$  と集合  $FS_o$  に含まれる特徴点の位置  $x_i$  および前時刻の変位  $\bar{x}_j$  から算出した矩形の倍率  $S$  をもとに大きさを更新する。

$$S = \frac{1}{|FS_o|} \sum_{F_{t,i} \in FS_o} \frac{\|x^* - x_i\|}{\|\bar{x}_j\|} \quad (7)$$

#### 4.8 属性情報にもとづいた特徴点の更新

各特徴点について、対応付けの結果および記憶されている各特徴点の記憶属性に応じて、特徴点の更新処理および忘却処理を行う。まず、追跡対象の矩形領域内に含まれる場合は、所属属性をTF属性とし、矩形領域外の場合はBF属性とする。また、追跡対象の矩形領域内に存在する特徴点がある特徴点と対応付けされた場合、DF属性とする。次に、対応付けが出来なかった現時刻の特徴点  $F_{t,i} \notin FS$  は、記憶属性をIM属性とする。対応付けが出来た特徴点  $F_{t,i} \in FS$  について、特徴点からの推定位置  $x'$  と最終的な推定位置  $x^*$  の差がしきい値  $T_d$  より大きな場合、不要な点と判断し削除する。一方、推定位置に大きなズレがない特徴点は、対応する特徴点  $F_{m,j}$  の位置を  $y_j = x_i$ 、特徴ベクトルを  $v_j = v_i$  として更新する。また、変位は、前時刻の変位  $\bar{x}_j$  と現時刻の変位  $(x^* - x_i)$  および更新係数  $\beta$  から式(8)のように算出する。

$$\bar{x}_j = \beta(x^* - x_i) + (1 - \beta)\bar{x}_j \quad (8)$$

対応付けができた特徴点  $F_{t,i}$  について、対応する特徴点  $F_{m,j}$  の記憶属性がIMの場合、記憶属性をSM属性に変更し、SM属性の場合は、対応付けの回数をカウントし、 $T_m$  を超えるとLM属性に変更する。

一方、記憶されている特徴点  $F_{m,j}$  で対応付けがされなかったものは、未対応回数をカウントする。記憶属性がIM属性の場合、未対応のフレームが  $T_i$  を超えると記憶から忘却する。また、SMおよびLMの場合も同様に未対応のフレームがそれぞれ  $T_s$ ,  $T_l$  を超えると記憶から忘却する。対応付け回数は、累積回数であり、対応付けが断続的であっても累積する。一方、未対応回数は、対応付けが出来た場合にカウントをリセットする。

## 5. 比較実験

### 5.1 実験概要

本章では、提案手法と他の追跡手法の追跡性能の比較を行う。提案手法は追跡対象を特徴点の集合として捉え、属性情報を利用して追跡対象とともに背景の情報も考慮している。そこで、矩形全体を捉える手法と特徴点の集合として捉える手法との比較および、属性情報の有無による追跡性能の比較実験を行う。矩形全体を捉える手法は、近年優れた追跡性能が報告されている検出器ベースの追跡手法である Multiple Instance Learning による追跡手法 [15] を用いる。特徴点の集合として捉える手法は、特徴点ベースの SURF トラッキング [14] と背景を補助的に利用する Grabner らの手法 [16] を用いる。評価方法は、追跡対象の中心位置について正解位置との誤差を計測する。また、各手法の初期位置は同じ中心座標と幅、高さを与える。

### 5.2 評価データベース

1つ目の比較実験には、Babenko らが公開している評価データベースを用いる [15]。撮影環境および被写体の種類が異なる8種類のシーケンスから構成されている。すべてのシーケンスには5フレーム間隔で追跡対象の中心位置と幅および高さが正解情報として入力されている。また、解像度は320x240ピクセルのグレースケール画像であり、シーケンスの長さは異なっている。Babenko の評価データベースは、追跡性能を確認するために物体の動きや見えの変化が多様であるが、あくまで性能を測るため不自然な動きも含まれている。

2つ目の比較実験には、個人がビデオカメラで撮影した実際のホームビデオを用いる。評価するシーケンスは4つあり、人の見えの変化が伴うシーン (HumanPose, 320フレーム)、動物の見えの変化とともにカメラの手ぶれが生じるシーン (DogPose, 400フレーム)、人が他の物体により部分的または全体が隠れるシーン (PartialOcclusion, 530フレーム, FullOcclusion, 200フレーム) である。それぞれの解像度は320x240ピクセルのRGBカラー画像である。

表 2 公開データベースでの中心位置の誤差 [pixel]

Video Clip	MIL [15]	ST [14]	SP [16]	Ours
David Indoor	23	44	32	21
Sylvester	11	18	21	12
Occluded Face	20	36	29	18
Occluded Face2	43	45	58	18
Girl	32	39	38	24
Tiger	15	34	31	18
Tiger2	17	28	24	17
Coke Can	21	34	28	25
average	22	34	32	19

### 5.3 パラメータ

比較手法は、各文献で述べられているように、[15] は特徴量に Haar-like 特徴、[14] は特徴点に SURF を用いる。また、[16] は、Harris オペレータにより抽出した特徴点と SIFT ベースの特徴量を用いる。提案手法の各パラメータは、予備実験により対応付けのしきい値  $TH_m$  を 0.7、更新係数  $\beta$  を 0.7、対応付けの回数  $T_m$  を 5 回、各記憶属性の保持期間である未対応フレーム数は、 $T_i$  を 5 フレーム、 $T_s$  を 30 フレーム、 $T_l$  を 100 フレームとしている。一方、しきい値  $T_d$  は、変位  $\bar{x}_j$  の 0.3 倍とする。

### 5.4 公開データベースでの性能比較

Babenko の評価データベースにおける正解位置からの誤差を表 2 に示す。各シーケンスで、最も誤差が小さい手法を赤字で示している。これより、提案手法は 5 つのシーケンスで最も誤差が小さく、残りの 3 つのシーケンスでも 2 番目の誤差である。また、全シーケンスの平均では提案手法が最も誤差が小さく、安定した追跡性能を得ることができている。

図 5 に各シーケンスにおいて、提案手法が推定した中心位置と追跡対象矩形を示す。追跡対象矩形は、TF 属性の特徴点から推定した場合は青色、遮蔽時などに BF 属性の特徴点から推定した場合は黄色で表している。図 5(a) の David Indoor や図 5(c) の Girl のように顔向き変化や照明変化などにより見えが変化するシーンでも正しく追跡することができている。図 5(b) の Occluded Face は、人を追跡対象とし、顔の遮蔽が生じる映像である。顔に遮蔽物が重なった場合、追跡対象の矩形は黄色になっており、BF 属性の特徴点により追跡対象の位置を推定している。このような遮蔽が生じても遮蔽物を誤追跡せず、追跡対象を正しく追跡できている。図 5(d) の Tiger は動物のぬいぐるみを追跡対象とし、人が操作して隠れや動きを生じさせている。遮蔽時は BF 属性の特徴点により追跡対象の位置を推定しているが、遮蔽時に大きく動いているため、BF 属性の特徴点からの推定する位置がずれており、従来手法より若干誤差が増加している。

表 3 ホームビデオでの中心位置の誤差 [pixel]

Video Clip	MIL [15]	ST [14]	SP [16]	Ours
HumanPose	18	34	30	15
DogPose	24	42	38	16
PartialOcclusion	40	30	21	9
FullOcclusion	33	45	35	17
average	29	38	31	14

### 5.5 ホームビデオでの性能比較

表 3 にホームビデオの各評価シーケンスにおける正解位置との誤差を示す。提案手法は、比較手法よりも各シーケンスにおいて誤差が小さく最も性能がよい。各シーケンスにおける提案手法で推定した中心位置と追跡矩形を図 6 に示す。追跡対象矩形は、TF 属性の特徴点から推定した場合は青色、遮蔽時などに BF 属性の特徴点から推定した場合は黄色で表している。図 6(a)(b) のように人や犬の姿勢変化がともなる場合でも追跡対象を正しく追跡することができている。また、図 6(c)(d) のように部分的または全身の隠れが生じるシーンでは、BF 属性の特徴点から追跡対象の位置を推定しており、追跡対象の位置を正しく推定することができていることがわかる。

## 6. 考 察

### 6.1 所属属性の効果

隠れが生じる前後において、特徴点に付与された所属属性の変化について調査する。図 7 に各フレームにおける特徴点を示す。TF 属性の特徴点を緑色、BF 属性の特徴点を黄色、DF 属性の特徴点を赤色、追跡対象の中心位置を青色で表している。また、DF 属性の特徴点が TF 属性の特徴点より多い場合は矩形を黄色、それ以外を青色で描画している。図 7(a) から図 7(e) は、公開データベースの occluded face での特徴点の分布である。図 7(b) で本が追跡対象矩形内に含まれており、背景から混入した特徴点として認識され、DF 属性が付与されている。図 7(c) において、DF 属性の特徴点が TF 属性の特徴点の数を上回ったため、BF 属性の特徴点から追跡対象の位置が推定されている。図 7(d)(e) では、本が追跡対象領域から外れたため、BF 属性が付与されている。また、TF 属性の特徴点が多くなり、TF 属性の特徴点から追跡対象の位置が推定されている。

図 7(f) から図 7(j) はホームビデオの Partial Occlusion の映像での特徴点の分布である。図 7(f) の画面左側にある黒いぬいぐるみが図 7(g) のように追跡対象の矩形内に含まれると、ぬいぐるみの特徴点は BF 属性から DF 属性になる。図 7(g) では、DF 属性の特徴点より TF 属性の特徴点の方が多く含まれるため、TF 属性の特徴点により追跡対象の位置が推定されている。一方、図 7(h) では、DF 属性の特徴点の方が TF 属性の特徴点より多くなっており、追跡対象の位置は BF 属性の特徴点から

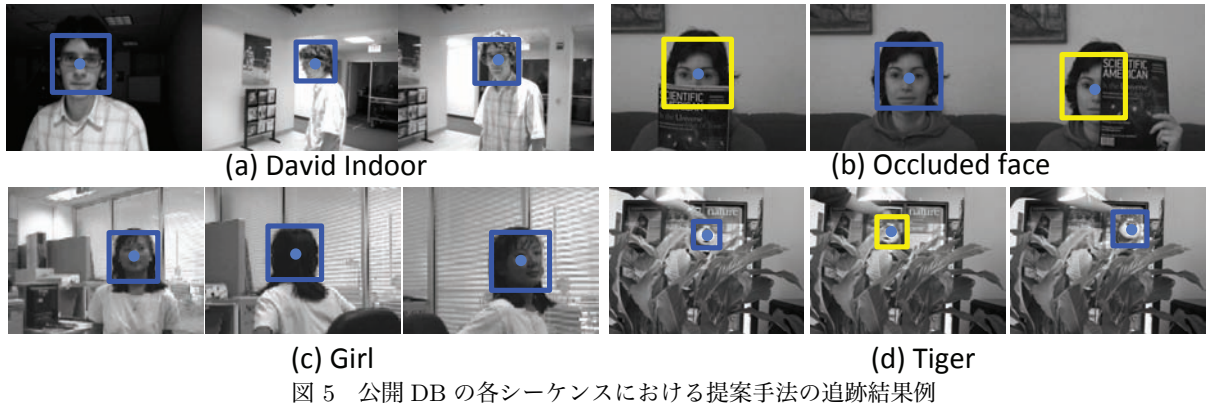


図 5 公開 DB の各シーケンスにおける提案手法の追跡結果例

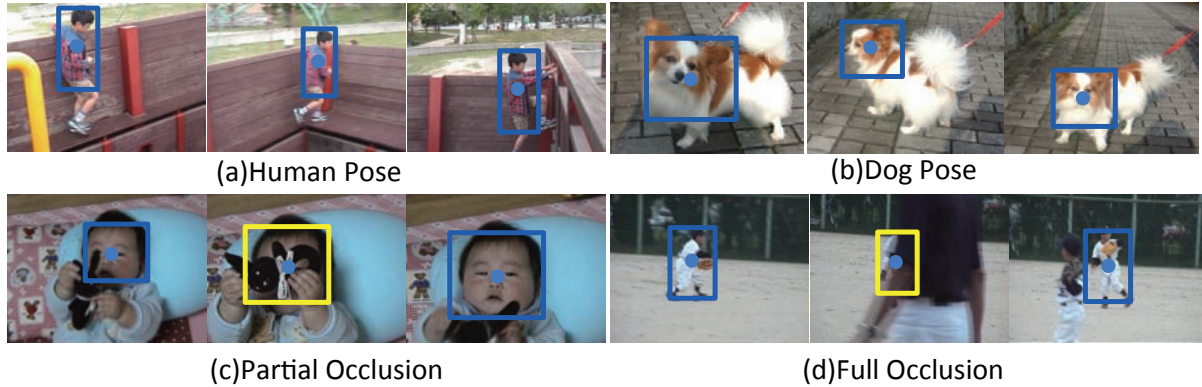


図 6 ホームビデオでの各シーケンスにおける提案手法の追跡結果例

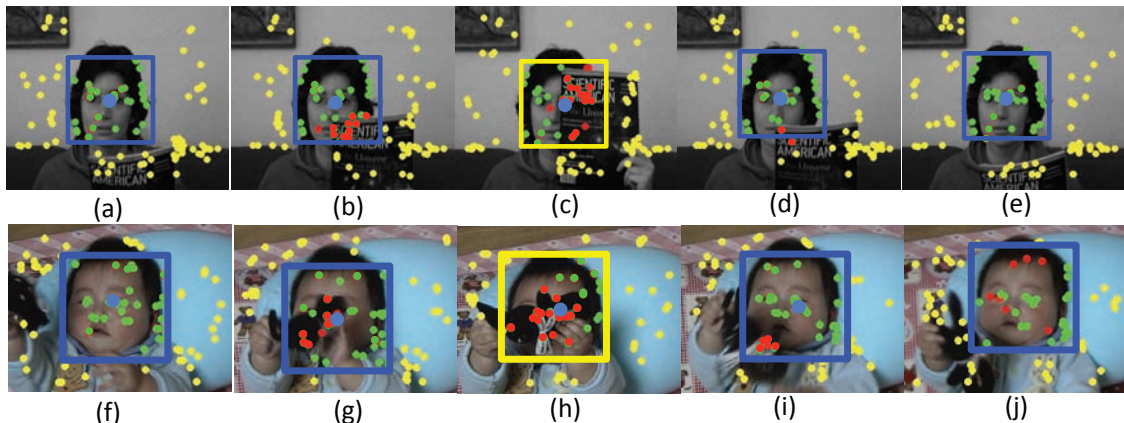


図 7 隠れが生じるシーンでの所属属性ごとの特徴点。緑色は TF 属性，黄色は BF 属性，赤色は DF 属性，青色は中心位置。

推定されている。これらの2つのシーケンスより、特徴点に所属属性を付与することで追跡対象に含まれる不要な背景の特徴点を排除することができ、追跡の失敗を防ぐことができている。

## 6.2 記憶属性の効果

図 8 に各フレームにおける記憶属性の変化を示す。IM 属性の特徴点を黄色，SM 属性をオレンジ色，LM 属性を赤色，FM 属性を緑色で表している。図 8(a) や (d) では、初期の追跡対象と類似していることから FM 属性の特徴点が多い。(b)(c) のように、顔の位置移動や微小な形状変化などで見えが変化するフレームでは、IM 属性または SM 属性の短期間保持される特徴点が多数対応付

けされている。一方、図 8 の (e) から (h) のように顔向きによる見えが変化する、再度もとの姿勢に戻るような場合は、LM 属性や FM 属性の長期間保持されている特徴点が多数対応付けされている。IM 属性や SM 属性は、微小な見えの変化に逐次対応することが出来ており、LM 属性や FM 属性は大きな見えの変化や過去に現れた見えに対応することができている。このように、記憶属性はモデルの逐次更新により生じるモデルの劣化を防ぐことができており、4つの記憶属性を付与することで、見えの変化に対して頑健に追跡することが可能である。

## 7. まとめ

本稿では、特徴点ベースの物体追跡において、特徴点

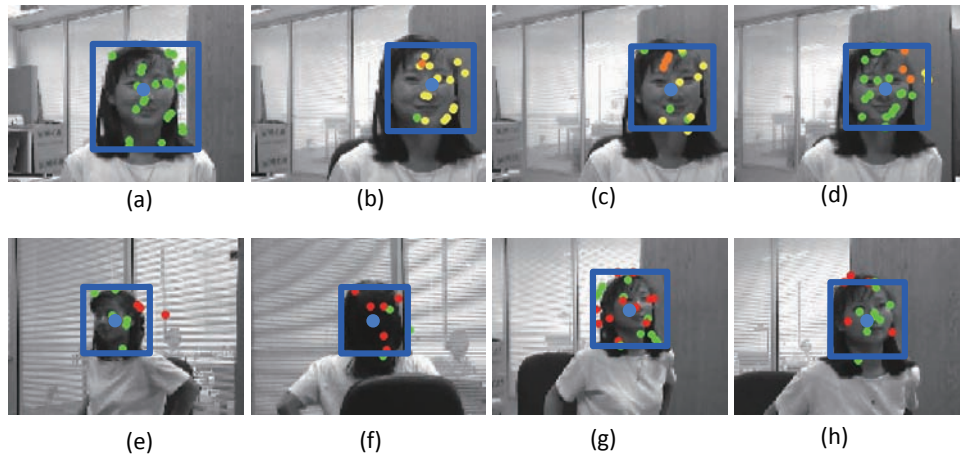


図 8 見え変化が生じるシーンでの記憶属性ごとの特徴点。緑色は FM 属性，黄色は IM 属性，オレンジ色は SM 属性，赤色は LM 属性，青色は中心位置。

に付与された属性情報を利用した物体追跡手法を提案した。属性情報には，周辺を理解することが可能な所属性と記憶モデルにもとづく記憶属性の 2 つを導入した。これにより，従来の手法が課題としていた追跡対象のモデルの劣化および隠れに対する頑健性を向上することできた。今後は，特徴点ベースの手法だけでなく，局所領域にもとづく手法など他の手法へ展開していく。

## 文 献

- [1] L.R. Squire, "Memory and Brain", Oxford University Press, 1987.
- [2] M. Isard and A. Blake, "Contour tracking by stochastic propagation of conditional density", European Conference on Computer Vision, pp.343-356, 1996.
- [3] M. J. Black and A. D. Jepson, "Eigentracking: Robust matching and tracking of articulated objects using a view-based representation", International Journal of Computer Vision, Vol.26, No.1, pp.63-84, 1998.
- [4] S. Birchfield, "Elliptical head tracking using intensity gradients and color histograms", IEEE Conference on Computer Vision and Pattern Recognition, pp.232-237, 1998.
- [5] 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.
- [6] 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.
- [7] A. Adam, E. Rivlin, I. Shimshoni, "Robust Fragments-based Tracking using the Integral Histogram", IEEE Conference on Computer Vision and Pattern Recognition, pp.798-805, 2006.
- [8] H. Grabner and H. Bischof, "On-line Boosting and Vision", IEEE Conference on Computer Vision and Pattern Recognition, pp.260-267, 2006.
- [9] 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.
- [10] S. Avidan, "Ensemble Tracking", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 29. No.2, pp.261-271, 2007.
- [11] M. Yang, J. Yuan, Y. Wu, "Spatial Selection for Attentional Visual Tracking", IEEE Conference on Computer Vision and Pattern Recognition, pp.1-8, 2007.
- [12] D. A. Ross, J. Lim, R-S. Lin, M-H. Yang, "Incremental Learning for Robust Visual Tracking", International Journal on Computer Vision Vol.77 No.3, pp.125-141, 2008.
- [13] H. Grabner, C. Leistner, H. Bischof, "Semi-supervised On-line Boosting for Robust Tracking", European Conference on Computer Vision, 2008.
- [14] W. He, T. Yamashita, H. Lu, S. Lao, "SURF Tracking", IEEE International Conference on Computer Vision, pp.1586-1592, 2009.
- [15] B. Babenko, H-M Yang, S. Belongie, "Robust Object Tracking with Online Multiple Instance Learning", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.33, No.8, pp.983-990, 2010.
- [16] 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.
- [17] J. Fan, Y. Wu, S. Dai, "Discriminative spatial attention for robust tracking", European Conference on Computer Vision, 2010.
- [18] J.Fan, X. Shen, Y. Wu "Closed-Loop Adaptation for Robust Tracking", European Conference on Computer Vision, 2010.
- [19] E. Rosten, R. Porter, T. Drummond, "Faster and better: a machine learning approach to corner detection", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.32, No.1, pp.105-119, 2010.
- [20] T. B. Dinh, N. Vo, G. Medioni, "Context Tracker: Exploring Supporters and Distracters in Unconstrained Environments", IEEE Conference on Computer Vision and Pattern Recognition, pp.1177-1184, 2011.
- [21] M. Godec, P. M. Roth, H. Bischof, "Hough-based Tracking of Non-Rigid Objects", IEEE International Conference on Computer Vision, pp.81-88, 2011.
- [22] 山下隆義, 藤吉弘亘, "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", 画像の認識・理解シンポジウム (MIRU2011), pp.56-63, 2011.