

空間的及び時間的な属性情報を導入した特徴点ベースの物体追跡

山下 隆義[†] 藤吉 弘亘^{††}

Object Tracking based on Feature Points with Spatial and Temporal Attributes

Takayoshi YAMASHITA[†] and Hironobu FUJIYOSHI^{††}

あらまし 本論文では、複雑な背景下でも追跡対象に生じる見えの変化に対して頑健に追跡するために、所属属性と記憶属性の二つの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。所属属性は各特徴点が追跡対象の点か背景領域の点かを区別するためのものである。所属属性を用いることで、追跡対象の領域内に含まれる背景の特徴点を排除することができる。また、背景の特徴点が追跡対象に近づいていることを判断できるため、隠れが生じるか否かを事前に推定することが可能となる。記憶属性は出現頻度に応じて特徴点の保持期間を変えらるための属性である。出現頻度が高い特徴点を長期間保持することで、追跡対象に一時的な変化が生じても信頼のおける過去の特徴点を保持しているため、モデルの劣化を抑えることが可能となる。比較実験では、これらの属性情報を導入することで、従来手法で課題であった見えの変化時の追跡失敗を防ぐことが可能であることを示す。

キーワード 物体追跡, 遮へい, 所属属性, 記憶属性, 記憶モデル

1. ま え が き

物体追跡は、複雑な背景下で追跡対象の形状や照明の変化、部分的若しくは全体の隠れなど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して、変化に頑健な特徴量とともに、追跡対象のモデル生成及びその更新方法に着目した様々な手法が提案されている。追跡対象のモデルは、追跡対象の方形全体を捉える方法 [2]~[5] と局所領域の集合とみなす方法 [10], [13], [15] 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], [7], [9], [11], [14] 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、推定するための情報が少なくなり、類似する位置を推定することが困難となり、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中で追跡対象の位置を推定し続ける。追跡対象や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。このような状況でモデルを更新すると、背景の混入によりモデルの劣化が生じ、追跡が失敗しやすくなる。

本論文ではこれらの問題点を解決するモデルの生成方法及び更新方法を提案する^(注1)。モデルは特徴点の集合とみなし、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。各特徴点には、追跡対象か背景領域かまたは悪影響を及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ともに、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

[†] オムロン株式会社, 草津市
OMRON Corporation, 2-2-1 Nishikusatsu, Kusatsu-shi,
525-0025 Japan

^{††} 中部大学, 春日井市
Chubu University, 1200 Matsumoto-cho, Kasugai-shi, 487-
8501 Japan

(注1): 本論文は、画像の認識・理解シンポジウム (MIRU2012) の発表内容に基づいている [22].

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。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を長期間保持することで、追跡対象の変化が生じた場合でも過去の特徴点と対応付けでき、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。モデルの生成では、追跡対象領域内の色やこう配の分布をヒストグラム特徴として用いる方法 [2], [5], 部分空間を用いて追跡対象の方形全体を特徴空間に射影する方法 [3], [11] がある。このような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 [6], [7], [9]。検出器を用いる手法は、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できず、追跡が失敗しやすい。

追跡対象のモデルとして、不変特徴量や一般化ハフ変換を利用して、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 [10], [13], [20]。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。

一方、モデルの更新は、オンライン学習による逐次的な適応方法が提案されている。オンライン学習は、見えの変化などにより追跡対象方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。そのため、ラベル付きデータを併用して背景の情報が含まれてもモデルの劣化を防いでいる方法 [12] や時系列で学習サンプルを蓄積する方法 [14] が提案されている。

これまでの追跡対象のモデル生成及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 [8] は、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している。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前提となっている。Grabner ら [15] も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いるが、類似物による隠れなどの隠れたかどうか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 [19] 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 [21] は、追跡対象とその周辺の背景領域を同時に追跡するとともに、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、モデルの劣化及び隠れに対する追跡失敗を解決しようと試みているが、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。本論文では、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_i^* を求める。時刻 t の特徴点 $F_{t,i} = \{x_i, \mathbf{v}_i\}$ は位置 x_i と特徴ベクトル \mathbf{v}_i をもち、集合 FP の特徴点 $F_{m,j} = \{y_j, \mathbf{v}_j, \bar{x}_j\}$ は位置 y_j と特徴ベクトル \mathbf{v}_j とともに、追跡対象の中心からの変位 \bar{x}_j をもっている。対応付けは、式 (1) のように、 $F_{t,i}$ の特徴ベクトル \mathbf{v}_i と最も類似する特徴ベクトルをもつ特徴点 $F_{m,j}$ 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。

$$\rho(F_{t,i}) = \max_{F_{m,j} \in FP} D(\mathbf{v}_i, \mathbf{v}_j) \quad (1)$$

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ゆが度を $p(x'|x_{c,i})$ とし、式 (2)

のように位置 x' のゆう度を求める。ゆう度 $p(x')$ が最も高い位置を追跡対象の中心位置 x_t^* としている。

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

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、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が、追跡対象が遮へいされる場合や背景を含むような場合は、正しく推定できない。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 の二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、追跡に悪影響を及ぼす特徴点を示す Discard-Feature (DF) の三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) 及び刷込み記憶を示す ImprintingMemory (FM) の四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、遮へい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しつつ、モデルの劣化を防ぐ。

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 及び DF 属性の部分集合 FS_d の和となる。追跡対象の位置に対するゆう度は、所属属性が TF と BF の特徴点から求める。 α 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を決める。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

図 1 に提案手法の流れを示す。まず、各時刻において特徴量抽出を行い、現時刻の特徴点 $F_{t,i}$ と蓄積している特徴点 $F_{m,j}$ との対応付けを行う。対応した特徴点の所属属性及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。蓄積している各特徴点は、位置関係や対応付けの回数をもとに所属属性及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 [18] により抽出し、特徴ベクトルは、Grabner ら [15] 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 30×30 ピクセルの注目領域を設定し、16 方向のこう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 3×3 のブロックに分割し、各ブロック内について 8 方向のこう配ヒストグラムを算出する。特徴量は、各こう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。

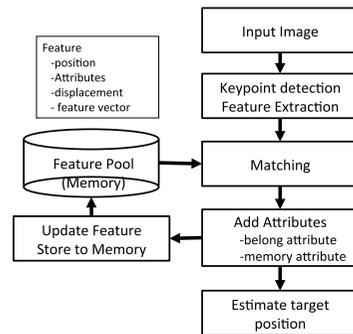


図 1 提案する物体追跡の流れ
Fig. 1 Tracking flow of proposed method.

4.2 特徴点の対応付け

現時刻で抽出した特徴点 $F_{t,i}$ と蓄積している特徴点 $F_{m,j}$ の対応付けを行う。それぞれの特徴ベクトルを v_i 及び 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。蓄積し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、追跡対象の方形内に存在する場合は TF 属性、追跡対象の方形外の場合は BF 属性を付与する。また、追跡対象の方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、追跡対象の方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

隠れが生じている状態では、追跡対象を表す TF 属性の特徴点数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、背景から追跡対象領域に混入した特徴点を示す 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 α 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。ここで、背景領域から中心位置を推定する際、追跡対象と背景領域の各特徴点は、相対的な位置関係が大きく変化しないことを前提としている。

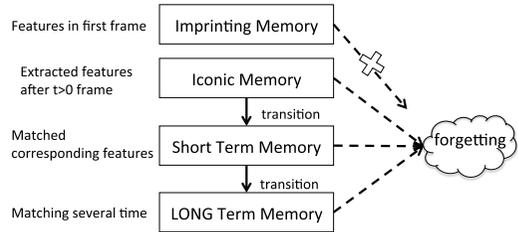


図2 各記憶属性の遷移
Fig. 2 Translation of memory attributes.

4.5 記憶属性について

我々はスクワアの記憶分類 [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) の三つからなる。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、時間とともに忘却される。SM の記憶のうち、緻密化リハーサルと呼ばれる繰り返し処理をすることで 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。各記憶属性の遷移について、図 2 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基づき、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。

4.7 追跡対象の位置推定

追跡対象の位置は、従来手法と同様に、各特徴点による推定位置を統合して決定する。各特徴点からの推定位置 $x_{c,i}$ を中心に、位置関係に関するう度 $p(x'|x_{c,i})$ 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 $x_{c,i}$ 及びう度 $p(x'|x_{c,i})$ 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 ρ 及び所属属性に関する重み α から $p(x')$ を求める。そして、う度 $p(x')$ が最大となる位置 x' を推定位置 x^* として求める。

また、追跡対象の方形の大きさは、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 及び前時刻の変位 \bar{x}_j から算出した方形の倍率 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 属性情報に基づいた特徴点の更新

各特徴点について、対応付けの結果及び蓄積している各特徴点の記憶属性に応じて、特徴点の更新処理及び忘却処理を行う。まず、追跡対象の方形領域内に含まれる場合は所属属性を TF 属性、方形領域外の場合は BF 属性とする。追跡対象の方形領域内に存在する特徴点が BF 属性をもつ特徴点と対応付けされた場合は DF 属性とする。

次に、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 属性に変更する。また、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、対応付けができ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。

蓄積する各特徴点 $F_{m,j}$ は、対応付けできた現時刻の特徴点 $F_{t,i}$ から位置を $y_j = x_i$ 、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 $(x^* - x_i)$ 及び更新係数 β から式 (8) のように算出する。

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

蓄積している特徴点 $F_{m,j}$ 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、未対応回数が T_i を超えると記憶から忘却する。また、SM 及び LM の場合も同様に未対応回数がそれぞれ 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、対応付けができた場合にカウントをリセットする。

4.9 提案手法の前提条件

提案手法は、特徴点の属性をもとに周辺状況や追跡対象位置の推定に用いる特徴点を決める。属性情報は、初期フレームにおける特徴点の属性をもとに更新していくため、初期フレームの追跡対象内には背景の特徴点がないことを前提としている。また、隠れが生じている間は背景の特徴点から追跡対象の位置を推定しているため、その間は追跡対象周辺の背景が大きく変化しないことも前提としている。

5. 比較実験

5.1 実験概要

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、方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及び、属性情報の有無による追跡性能の比較実験を行う。方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 [14] 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 [13] と背景を補助的に利用する Grabner らの手法 [15] 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し、全フレームでの平均誤差を比較する。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

一つ目の比較実験には、Babenko らの公開評価データベースを用いる [14]。撮影環境及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。全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及び高さが正解情報として入力されている。また、解像度は 320×240 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で不自然な動きも含まれている。

二つ目の比較実験には、六つのホームビデオを用いる。各シーケンスは、人の見えの変化が伴うシーン (HumanPose1, 320 フレーム), (HumanPose2, 700 フレーム), (HumanPose3, 450 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 320×240 ピクセルの RGB カラー画像である。

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、[14] は特徴量に Haar-like 特徴, [13] は特徴点に SURF を用いる。また、[15] 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 β 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。

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 1 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、提案手法は五つのシーケンスで最も誤差が小さく、残りの三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 3 に代表的なシーケンスにおける提案手法が推定した中心位置と追跡対象方形を示す。追跡対象方形

は、TF 属性の特徴点から推定した場合は青色、遮へい時などに 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 3(a) の David Indoor や図 3(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 3(b) の Occluded Face は、人を追跡対象とし、顔の遮へいが生じる映像である。顔に遮へい物が重なった場合、追跡対象の方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。このような遮へいが生じても遮へい物を誤追跡せず、追跡対象を正しく追跡できている。図 3(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。遮へい時は 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、遮へい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。Sylvester や Tiger, Coke Can のような追跡対象が小さいシーケンスの場合、提案手法を含む特徴点ベースの手法は特徴点数が少ないため、追跡対象の位置を正しく推定できない場合が多くなる。一方、

表 1 公開データベースでの中心位置の誤差 [pixel]
Table 1 Error of center position in public DB.

Video Clip	MIL [14]	ST [13]	SP [15]	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

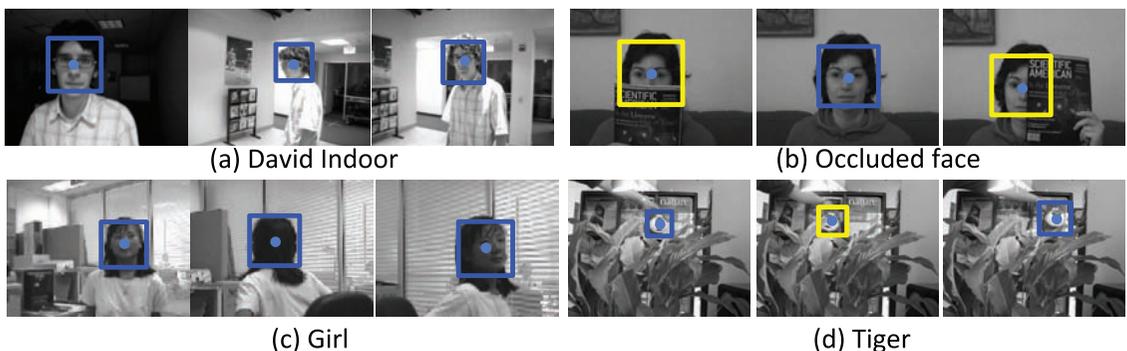


図 3 公開 DB の各シーケンスにおける提案手法の追跡結果例
Fig. 3 Tracking result of proposed method in public DB.

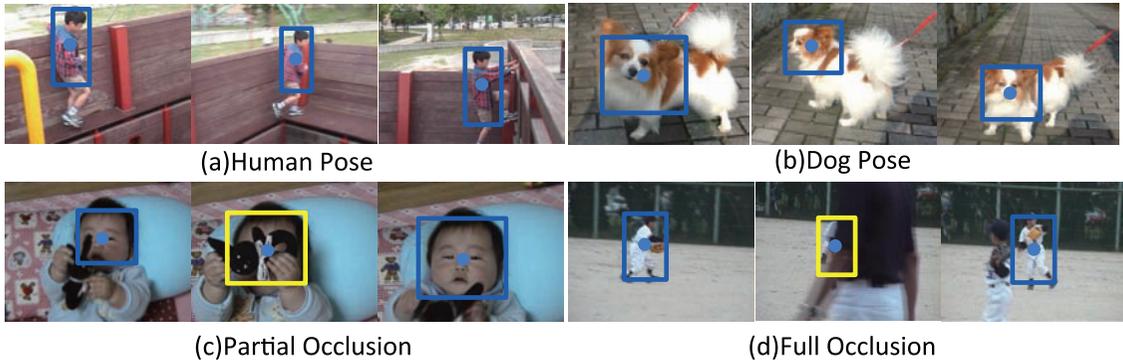


図 4 ホームビデオでの各シーケンスにおける提案手法の追跡結果例
Fig. 4 Tracking result of proposed method in Home video.

表 2 ホームビデオでの中心位置の誤差 [pixel]
Table 2 Error of center position in home video.

Video Clip	MIL [14]	ST [13]	SP [15]	Ours
HumanPose1	18	34	30	15
HumanPose2	48	59	43	21
HumanPose3	16	23	24	11
DogPose	24	42	38	16
PartialOcclusion	40	30	21	9
FullOcclusion	33	45	35	17
average	29	38	31	14

MIL のように対象全体を捉える手法は、追跡対象が小さい場合でも追跡対象を検出できるため、追跡対象の位置を正しく推定できる場合が多く、誤差が小さくなっている。

5.5 ホームビデオでの性能比較

表 2 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方形を図 4 に示す。追跡対象方形は、TF 属性の特徴点から推定した場合は青色、遮へい時などに 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 4(a) (b) のように人や犬の姿勢変化が伴う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 4(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分かる。

5.6 追従性の評価

表 1 及び表 2 では、提案手法が比較手法と比較して安定した追跡ができていることを示した。しかしながら、正解位置からの誤差だけでは、追跡対象からず

れているだけなのか、異なる領域を追跡しているかの判断がつきにくい。そこで、提案手法により誤差が小さくなった三つのシーケンスにおいて、どのような追従性の改善が見られたのかを比較する。図 5(a) Occluded Face2 において、MIL (灰色の実線) は、頭を傾けるとは類似する追跡対象を正しく捉えることができずに背景を誤検出し、それ以降、正しい位置を追跡できていない。SP (黒色の破線) は傾いた場合でも追跡対象付近を追跡できているが、追跡対象の一部が隠れると追跡位置がずれてしまい、それ以降、正しい位置を追跡できていない。一方、提案手法 (黒色の実線) は傾いた場合でも SP と同様に追跡対象付近を追跡できている。また、隠れた場合は、属性情報により追跡位置がずれることがなく、それ以降も正しく追跡できている。

図 5(b) HumanPose2 では、MIL は姿勢が変化すると検出対象を見失い、追跡が失敗している。SP は、姿勢が変化すると背景の特徴点が含まれていくことで追跡が失敗している。一方、提案手法は、属性情報により追跡対象領域内の背景の特徴点を用いずに追跡するため、追跡対象を正しく追跡できている。

図 5(c) PartialOcclusion では、顔を隠す人形を SP は誤って追跡している。MIL も同様に人形を含む顔領域を記憶しているため、人形を誤って追跡している。一方、提案手法は人形付近の特徴量を BF 属性にすることができているため、正しく顔を追跡できている。

このように、提案手法は、従来手法の追跡が失敗するようなシーンにおいても追跡対象を正しく追跡することができている。

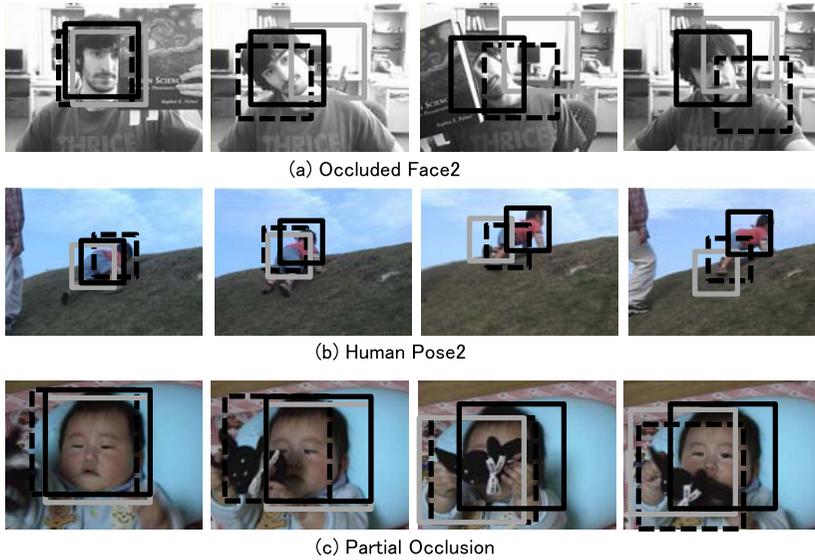


図 5 各手法の追跡結果例。黒色の実線は提案手法，黒色の破線は SP [15]，灰色の実線は MIL [14]

Fig. 5 Tracking result of proposed and conventional methods. Black rectangle is proposed method, Black rectangle with wave line is SP [15] and gray rectangle is MIL [14].

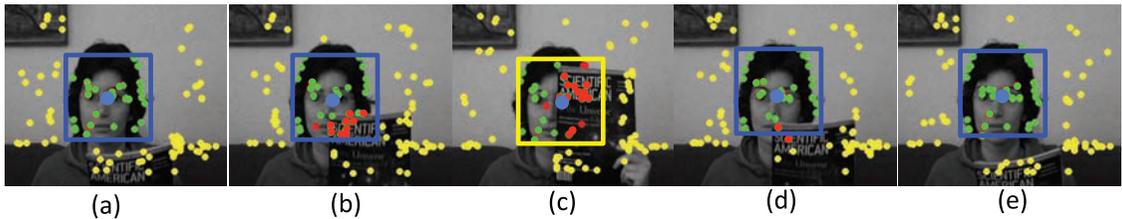


図 6 隠れが生じるシーンでの所属属性ごとの特徴点。緑色は TF 属性，黄色は BF 属性，赤色は DF 属性，青色は中心位置

Fig. 6 Each feature point with belonging attributes in occlusion scene. green is TF, yellow is BF, red is DF and blue is center position.

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 6 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方形を黄色，それ以外を青色で描画している．図 6(a) から図 6(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 6(b) で本が追跡対象方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 6(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 6(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．これより，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る．

6.2 記憶属性の効果

図 7 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

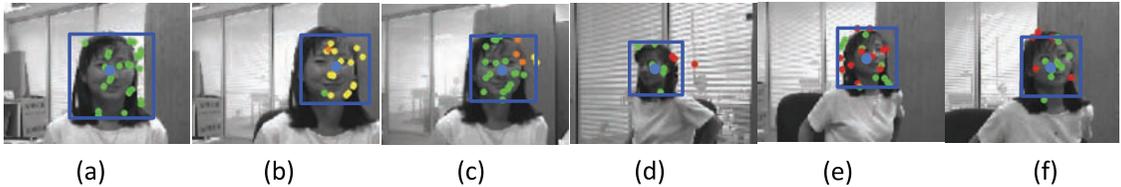


図 7 見え変化が生じるシーンでの記憶属性ごとの特徴点. 緑色は FM 属性, 黄色は IM 属性, オレンジ色は SM 属性, 赤色は LM 属性, 青色は中心位置
 Fig.7 Each feature point with memory attributes in various pose scene. green is FM, yellow is IM, orange is SM, red is LM and blue is center position.

属性を赤色, FM 属性を緑色で表している. 図 7(a) や (c) 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) 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 7 の (d) から (f) のように顔向きによる見えが変化し, 再度元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できており, 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しており, 四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.

6.3 隠れが生じた際の移動量による追跡性能

隠れが生じている際の移動量と追従性の関係について評価する. 評価データとして, 人物が映る画像に擬似的な隠れを発生させた後, 人物の位置をずらした画像である. また, 背景が単色なシーンと煩雑なシーンの 2 通りを用意している. 移動量は隠れの前後で追跡対象の幅の 10% から 100% まで変化させている. 初期フレームと隠れフレーム及び追跡限界前後のフレームにおける評価結果を図 8 に示す. 単色な背景の場合, 追跡対象の移動量が 30% を超えると追跡結果が大きくずれている. 一方, 煩雑な背景の場合は移動量が 70% まで追跡対象を捉えることができている. このとき, 煩雑な背景の場合の特徴点数は 43 点, 単色な場合は 3 点である. これより, 背景に十分な特徴点があると, 単色な背景のシーンよりも移動量が大きくても追跡することができる.

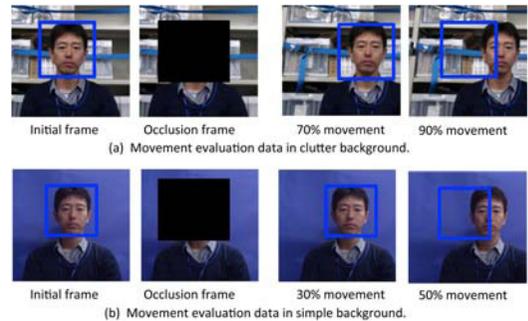


図 8 移動量評価結果
 Fig.8 evaluation result of movement.

7. む す び

本論文では, 周辺を理解することが可能な所属属性と記憶モデルに基づく記憶属性の二つを特徴点に付与する物体追跡手法を提案した. 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及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, 局所領域に基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," Proc. European Conf. 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," Int. J. Comput. Vis., vol.26, no.1, pp.63-84, 1998.
- [4] S. Birchfield, "Elliptical head tracking using intensity gradients and color histograms," Proc. IEEE Conf. on Computer Vision, pp.232-237, 1998.
- [5] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol.25, no.5, pp.564-577, 2003.

- [6] R.T. Collins and Y. Liu, "Online selection of discriminative tracking features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.27, no.10, pp.1631-1643, 2005.
- [7] H. Grabner and H. Bischof, "On-line boosting and vision," *Proc. IEEE Conf. on Computer Vision*, pp.260-267, 2006.
- [8] M. Yang, Y. Wu, and S. Lao, "Intelligent collaborative tracking by mining auxiliary objects," *Proc. IEEE Conf. on Computer Vision*, pp.697-704, 2006.
- [9] S. Avidan, "Ensemble tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.29, no.2, pp.261-271, 2007.
- [10] M. Yang, J. Yuan, and Y. Wu, "Spatial selection for attentional visual tracking," *Proc. IEEE Conf. on Computer Vision*, pp.1-8, 2007.
- [11] D.A. Ross, J. Lim, R.-S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking," *Int. J. Comput. Vis.*, vol.77, no.3, pp.125-141, 2008.
- [12] H. Grabner, C. Leistner, and H. Bischof, "Semi-supervised on-line boosting for robust tracking," *Proc. European Conf. on Computer Vision*, pp.234-247, 2008.
- [13] W. He, T. Yamashita, H. Lu, and S. Lao, "SURF tracking," *Proc. IEEE Intl. Conf. on Computer Vision*, pp.1586-1592, 2009.
- [14] B. Babenko, H.-M. Yang, and S. Belongie, "Robust object tracking with online multiple instance learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.33, no.8, pp.983-990, 2010.
- [15] H. Grabner, J. Matas, and L. Van Gool, "Tracking the invisible: Learning where the object might be," *Proc. IEEE Conf. on Computer Vision*, pp.1-8, 2010.
- [16] J. Fan, Y. Wu, and S. Dai, "Discriminative spatial attention for robust tracking," *Proc. European Conf. on Computer Vision*, 2010.
- [17] J. Fan, X. Shen, and Y. Wu, "Closed-loop adaptation for robust tracking," *Proc. European Conf. on Computer Vision*, pp.411-424, 2010.
- [18] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.32, no.1, pp.105-119, 2010.
- [19] T.B. Dinh, N. Vo, and G. Medioni, "Context tracker: Exploring supporters and distracters in unconstrained environments," *Proc. IEEE Conf. on Computer Vision*, pp.1177-1184, 2011.
- [20] M. Godec, P.M. Roth, and H. Bischof, "Hough-based tracking of non-rigid objects," *Proc. IEEE Intl. Conf. on Computer Vision*, pp.81-88, 2011.
- [21] 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.
- [22] 山下隆義, 藤吉弘亘, "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," 第15回画像の認識・理解シンポ

ジウム (MIRU2012), OS13-04, 8p, 2012.

(平成 24 年 10 月 10 日受付, 25 年 2 月 8 日再受付)



山下 隆義 (正員)

2002 奈良先端科学技術大学院大学情報科学研究科博士前期課程了。同年オムロン(株)入社。2011 中部大学大学院博士後期課程了。物体検出, 物体追跡などパターン認識に関する研究に従事。工博。IEEE-CS 会員, 情報処理学会各会員。



藤吉 弘亘 (正員)

1997 中部大学大学院博士後期課程了。1997~2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow。2000 年中部大学講師を経て 2004 より同大准教授。2006 米カーネギーメロン大学ロボット工学研究所客員研究員。2010 中部大学教授。工博。計算機視覚, 動画像処理, パターン認識・理解の研究に従事。2005 年度ロボカップ研究賞, 2009 年度情報処理学会論文誌コンピュータビジョンとイメージメディア優秀論文賞, 2009 年度山下記念研究賞, 情報処理学会, 電気学会, IEEE 各会員。