色情報と深度情報を併用した自己遮蔽に頑健な物体追跡手法

今井 順一*・柏木 雄平**・木辻 亮***

Robust Object Tracking against Self-occlusion Using Color and Depth Information

Jun-ichi IMAI*, Yuhei KASHIWAGI** and Ryo KITSUJI***

Visual object tracking techniques are widely required by many vision applications. The color-based particle filter is known as one of useful methods for robust object tracking. However, the conventional color-based particle filter has a problem that it is not robust against self-occlusion. Self-occlusion occurs when a part of a target object is hidden by itself from a camera. When the target object moves or rotates, a part of the target disappears because the self-occlusion occurs and other part appears because the self-occlusion is resolved. The conventional color-based particle filter often fails to follow such a change of the target's appearance due to self-occlusion during the tracking process. In this paper, we propose a novel method for robust object tracking against the self-occlusion. The proposed method is based on the color-based particle filter, and it also uses depth information obtained by an RGB-D camera. When the self-occlusion occurs and the target's appearance changes, the proposed method extracts a region for the target object in the input image by the graph cuts based on depth information. However, this process often includes unnecessary regions, especially when some objects are close to the target. Then, the proposed method distinguishes the region for the target from unnecessary ones by investigating expanse of colors around the target. Therefore, the target model is correctly updated and the robust tracking is achieved. In order to verify the effectiveness of the proposed method, we carried out an experiment to compare the proposed method with the conventional one. Experimental results show that the proposed method works well.

Key Words: object tracking, self-occlusion, color-based particle filter, RGB-D camera, graph cut

1. はじめに

移動物体の視覚的追跡は画像に基づくセンシングにおいて広く応用される重要な技術である。たとえば、人物の侵入検出を利用したセキュリティシステム¹⁾や視覚追跡を利用した移動ロボットの人物追従制御^{2)~4)}、ヒューマン・ロボット・インタラクションにおけるロボットビジョン⁵⁾など、さまざまな場面で利用されている。

物体追跡の代表的な手法として、色分布による物体モデルを利用したパーティクルフィルタによる追跡手法⁶⁾がある。これには、

- 物体の変形や画像平面上での回転、部分的な遮蔽に頑健であること

- 事前知識を必要とせず、その時点で保持する情報のみから物体モデルを作成できること

などの利点がある。特に後者の利点は状況に応じて追跡対象を切り替える必要がある応用において重要な条件となる。たとえば、人間とロボットとのインタラクションにおいて注目すべき物体が対話内容に応じて移り変わる場面はしばしば見られる。このとき、ロボットは注目物体のモデルを必要に応じて新規に作成しながら追跡を行なうことが求められる。また、Imaiらは追跡対象を自律的に切り替えることで他物体による遮蔽に頑健な物体追跡手法を提案している⁷⁾。この手法では色分布モデルによるパーティクルフィルタが利用されており、事前知識を必要とせずにモデル作成が可能な特徴が生かされている。

しかし、色分布による物体モデルは追跡開始時にカメラから見える外見をもとに作成されるため、自己遮蔽によって外見の色の配分が大きく変化し得る物体に対して適切に対応できない問題がある。自己遮蔽とは物体の一部が自分自身によりカメラから見えなくなる現象を指す。追跡中、物体が撮影される角度が変化し自己遮蔽が発生することによってそれまで見ていた領域が隠れたり、逆に自己遮蔽の解消により隠れていた領域が現れることが頻繁に起こり得る。このとき、追跡対象の認識に必要な色分布も変化するため、安定な追跡

* 千葉工業大学情報科学部 習志野市津田沼 2-17-1

** (株) エム・ソフト 東京都台東区東上野 2-18-10

*** (株) カイカ 東京都目黒区大橋 1-5-1

* Faculty of Information and Computer Science, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino

** M.SOFT Co., Ltd., 2-18-10 Higashiueno, Taito-ku, Tokyo

*** CAICA Inc., 1-5-1 Ohashi, Meguro-ku, Tokyo

(Received April 6, 2018)

(Revised September 24, 2018)

のためには物体モデルを更新する必要がある。しかし、更新時に追跡対象以外の不必要な色情報をモデルに取り込んでしまうとそこから連鎖的に追跡対象が広がり、最終的に追跡が破綻する。したがって、どこまでが追跡対象と同一物体であるかを的確に抽出し、モデルを更新する必要がある。

そこで本論文では、自己遮蔽に対して頑健な物体追跡を実現する手法を提案する。追跡に用いる状態推定アルゴリズムとしてのパーティクルフィルタの枠組みは保持したまま、その観測モデル、すなわち追跡対象物体のモデルおよびその重み評価を改善することでこれを実現する。具体的には、追跡中に深度情報によって追跡対象と同一物体の領域を抽出し、色分布による物体モデルをオンラインで更新する仕組みを導入する。画像情報に加えて深度情報取得のため、提案手法では RGB-D カメラを利用する。追跡対象物体の深度変化は連続的であることを仮定し、深度情報に基づく Graph Cut^{8),9)}によりモデルに追加すべき領域を抽出する。その際、追跡対象と接する他物体が存在するとその領域を追跡物体領域から分離できず、誤ってモデルに追加してしまう問題が生じる可能性がある。そこで提案手法では、追跡対象物体の見かけ上の大きさが急激に変化しないことを仮定することでこの問題を解決する。本論文では評価実験を行ない、提案手法の有効性を確認する。

2. 関連研究

物体追跡において、追跡中の外見変化に対する頑健性の向上は重要な課題の 1 つであり、多くの研究が行なわれてきた。

アプローチの方法はいくつか考えられる。その 1 つとして、追跡対象を背景領域から分離するためのモデルを学習により得る方法が考えられる。たとえば、追跡対象と背景を識別する検出器を学習していく方法^{10),11)}や、外見変化に不変な特徴を学習する方法^{12),13)}、対象の見えの変化をモデルとして獲得する方法^{14),15)}などが提案されている。

一方、追跡対象と背景との分離をなるべく単純なモデルを用いて実現し、その単純さによって頑健性を確保する手法も考えられる。たとえば対象の輪郭¹⁶⁾や色分布（ヒストグラム）^{6),17),18)}に基づく方法などが提案されている。

第 1 章でも述べたように、後者の手法に属する色分布モデルによる追跡はその単純さから物体の変形や部分遮蔽などに頑健性をもち、モデルを作成しやすいといったメリットももつ。このメリットを生かすため、本研究で提案する手法でも色分布に基づいた物体モデルを採用する。しかし、そのままでは自己遮蔽に対応できないため、外見の変化に応じて新しい情報をモデルに取り込んでいく必要がある。その際、誤って追跡対象以外の情報を取り込むことは追跡の破綻につながるため、モデル更新の際には追跡対象と背景領域を正確に分離する必要がある。その意味で、本論文の提案手法はモデル更新の際に前者の手法群の考え方を適用しているといえる。

ただし、前者の手法群のように追跡の過程で詳細な識別モデルを学習することはせず、その場での分離を可能にする目

的で、本論文では画像情報に加えて深度情報を利用し、深度に基づいた Graph Cut により追跡対象と背景領域の分離を行なう点に特徴がある。

3. 従来手法

本章では、まずモデルとして色分布を用いたパーティクルフィルタによる物体追跡手法について概説し、その後その問題点について述べる。

3.1 パーティクルフィルタ

画像処理における物体追跡など、直接観測できない隠れ状態の時系列推定問題に対し、各時刻での隠れ状態の分布をサンプルと重みで表現することで時系列ベイズ推定を実現する手法としてパーティクルフィルタ^{19),20)}が知られている。

$\mathbf{s}$, $\mathbf{z}$ をそれぞれ対象の推定状態および観測結果とする。事後分布 $p(\mathbf{s}|\mathbf{z})$ はサンプルと重みのペアの集合

$$S = \{(\mathbf{s}_i, \tilde{w}(\mathbf{s}_i)) | i = 1, 2, \dots, N\} \quad (1)$$

で近似される。各サンプル（パーティクル） $\mathbf{s}_i$ は対象の状態に対する 1 つの仮説を表わし、 $\tilde{w}(\mathbf{s}_i)$ はその妥当性を表わす重みである（ $\sum_{i=1}^N \tilde{w}(\mathbf{s}_i) = 1$ ）。

各パーティクルはダイナミクスモデルに従って新しい状態へと遷移し、観測結果をもとに重み付けされる。その後、次時刻のサンプル集合 S の生成のため、 N 個のパーティクルが復元抽出される（リサンプリング）。パーティクル $\mathbf{s}_i$ が選択される確率は $\tilde{w}(\mathbf{s}_i) \propto p(\mathbf{z}|\mathbf{s}_i)$ である。

パーティクルフィルタは同時に複数の状態仮説を保持するため、頑健な推定が可能となる。

3.2 色分布モデル

Nummiaro らはパーティクルフィルタにおける追跡対象のモデルとして色分布を用いた物体追跡手法を提案した⁶⁾。本論文で提案する手法はそれを改良して利用する。

色分布は HSV 表色系における HS 平面上のヒストグラムとして表現される。色相 (H)、彩度 (S) の値はそれぞれ 16 段階に離散化され、色分布のヒストグラムは $16 \times 16 = 256$ 個のビンをもつ。照明条件の影響を低減するため明度 (V) は利用しない。

パーティクル $\mathbf{s}(t)$ は時刻 t における画像内の楕円領域のパラメータを表わす。

$$\mathbf{s}(t) = (c_x(t), c_y(t), a_x(t), a_y(t), c_x(t-1), c_y(t-1), a_x(t-1), a_y(t-1))^T \quad (2)$$

ここで、 (c_x, c_y) は楕円の中心座標、 a_x, a_y はそれぞれ楕円の x 軸方向、 y 軸方向の半径、 T は転置を表わす。 $\mathbf{s}$ が定める楕円領域 $\mathcal{E}$ 内の色分布 $p(u; \mathcal{E})$ ($u = 1, 2, \dots, 256$) を求め、それをもとにパーティクル $\mathbf{s}$ の重みを評価する。なお、本論文では $p(\cdot; \mathcal{E})$ を必要に応じて $p(\cdot; \mathbf{s})$ と書く。

3.3 ダイナミクスモデル

各パーティクルの次時刻の状態はつぎのダイナミクスモデ

ルによって予測される.

$$\mathbf{s}(t) = A\mathbf{s}(t-1) + \mathbf{v} \quad (3)$$

ここで, A は状態遷移の決定論的な成分を表わす行列, $\mathbf{v}$ は多次元確率変数を表わす. 本論文では A としてつぎのような行列を用いる. これは追跡対象が等速度のダイナミクスをもつことを表わす.

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (4)$$

また, 確率変数 $\mathbf{v} = (v_1, v_2, v_3, v_4, 0, 0, 0, 0)^T$ の前半 4 成分はそれぞれ独立なコーシー分布に従う.

$$p(v_k) = \frac{1}{\pi} \cdot \frac{\tau}{v_k^2 + \tau^2} \quad (k = 1, \dots, 4) \quad (5)$$

ここで τ はコーシー分布のスケールパラメータである.

3.4 観測モデル

2 つの色分布 p, q の距離 d_b を Bhattacharyya 係数²¹⁾ b に基づいてつぎのように定義する.

$$d_b(p, q) = \sqrt{1 - b(p, q)} \quad (6)$$

$$b(p, q) = \sum_{u=1}^{256} \sqrt{p(u)q(u)} \quad (7)$$

各パーティクル $\mathbf{s}_i$ は, 対応する色分布 $p(\mathbf{s}_i)$ と追跡対象モデルの色分布 q との距離 d_b に基づく重み関数 w によって評価される.

$$w(\mathbf{s}_i) \equiv w(p(\mathbf{s}_i)) = \exp\left(-\frac{d_b(p(\mathbf{s}_i), q)^2}{2\sigma^2}\right) \quad (8)$$

w は $[0, 1]$ の値をもつ関数である. 重み $w(\mathbf{s}_i)$ はリサンプリング時に正規化される.

$$\tilde{w}(\mathbf{s}_i) = \frac{w(\mathbf{s}_i)}{\sum_{k=1}^N w(\mathbf{s}_k)} \quad (9)$$

色分布モデルを用いたパーティクルフィルタによる物体追跡の例を Fig. 1 に示す. この例では Fig. 1 (a) に示すボールが追跡対象である. 図中の黒い楕円は個々のパーティクル $\mathbf{s}_i$ の状態仮説を, 灰色の楕円は全パーティクルの平均状態 $\bar{\mathbf{s}}$, すなわちシステムとしての追跡対象の予測位置を表わす.

$$\bar{\mathbf{s}} = \sum_{i=1}^N \tilde{w}(\mathbf{s}_i) \cdot \mathbf{s}_i \quad (10)$$

追跡対象の位置にパーティクルが収束し追跡できているようすを Fig. 1 (b), (c) に示す. 一方, 画像中に追跡対象が存在せず, パーティクルが収束せずに散らばっているようすを Fig. 1 (d)

Fig. 1 Examples of object tracking using color-based particle filter

Fig. 2 An example of self-occlusion

に示す.

3.5 問題点

第 1 章で述べたように, 色分布を用いたパーティクルフィルタによる物体追跡手法には自己遮蔽に対応できない問題点がある.

面によって色が異なるキューブの例を Fig. 2 に示す. Fig. 2 (a) の状態で色分布モデルを作成し, 追跡を開始したと仮定する. このときパーティクルフィルタは「黄色 (前面) と青色 (上面) が半々程度になる領域」を探して追跡を行なう. ここでカメラもしくは自身の移動により物体が 45° 回転すると (Fig. 2 (b)), これまで見えていた黄色い面の面積が小さくなり, 逆に自己遮蔽によって隠れていた赤色の側面が現れる. しかし, パーティクルフィルタにとっての追跡対象は引き続き黄色と青色の面のみであり, 赤色の面は追跡対象の一部とは認識されない. さらに回転すると (Fig. 2 (c)) 黄色い面は自己遮蔽によって隠され, 代わりに赤色の面が大きく見えるようになる. このとき, パーティクルフィルタが追跡対象と認識できるのは青色の面のみであり, 黄色の成分が含まれないため追跡精度は急激に悪化する. ここでさらに青色の面が隠れ赤色の面のみが見える状態になると, 追跡対象であるキューブが見えているにもかかわらずパーティクルフィルタはそれを見失う.

この問題を解決する方法の 1 つとして, 初期状態で見えない部分も含め追跡対象のもつ色をすべて事前に調査し, 色分布モデルに反映することが考えられる. しかし, これでは「事前知識なしでモデルを作成可能」という利点を失う. そこで

考えられるのが、初期状態では隠れていた部分の色情報を追跡中に随時モデルに取り込んでいく手法である。しかし、対象以外の領域の色情報を誤って取り込んだ場合、そこから追跡対象が連鎖的に広がり、追跡の破綻につながる。したがって、どこまでが追跡対象の一部であるかを正確に認識し、その領域の情報のみをモデルに取り込むことが必要となる。本論文ではこれらを実現する手法を提案する。

4. 提案手法

本章では、色情報に加えてカメラからの深度情報を併用することで自己遮蔽に頑健な物体追跡を実現する手法を提案する。提案手法では深度情報を取得する必要から RGB-D カメラを利用する。

4.1 仮定

本論文ではつぎの 3 点を仮定する。

- (1) 画像上の追跡対象物体の領域における色分布は、連続する 2 フレーム間で大きく変化しない。
- (2) 画像上の追跡対象物体の領域の大きさは、連続する 2 フレーム間で大きく変化しない。
- (3) 同一フレームにおける追跡対象物体の領域内の深度変化は連続的である。

仮定 1 は自己遮蔽の発生や解消によって生じる追跡対象の外見変化が 1 時刻で見失うほど急激ではないことを、仮定 2 はカメラへの接近や離反、自己遮蔽の発生や解消によって生じる追跡対象の見かけ上の大きさが急激に変化しないことを仮定している。これらはいずれも、外見変化を捉えるのに十分なフレームレートを確保できることを仮定している。

仮定 3 は、同一物体であれば物理的に連結しているため、観測される深度も連続的に変化することを仮定している。形状および撮影角度によっては同一物体の一部でも深度が連続しない領域が存在することも考えられるが、本論文においてはそれらを同一物体と（その時点では）認識しないこととする。追跡処理の過程でそれらの領域の深度が連続した状態で観測されればその時点で同一物体と認識される。

4.2 提案手法の概要

まず、提案手法の概要について述べる。提案手法は主につぎの 3 点からなる。

4.2.1 複数の色分布による物体モデル

Fig. 2 のような自己遮蔽を考える場合、物体のモデルとして追跡開始時に作成した色分布だけでは不十分である。そこで提案手法では単独の色分布ではなく複数の色分布の集合を追跡対象のモデルとし、追跡の過程で見かけが変わった追跡対象は新しい色分布として後から追加していくことで物体モデルを適応的にする。

4.2.2 追跡対象物体領域の抽出

自己遮蔽により隠れていた領域が現れ、その外見に応じた色分布を新たに追加するためには、どこまでが追跡対象に相当する領域であるかを認識する必要がある。その領域を抽出するために、提案手法では仮定 3「追跡対象物体の深度変化

Fig. 3 Processing flow of proposed method

は連続的である」に基づき、Graph Cut による深度情報に基づいたセグメンテーションを利用してモデルに追加すべき領域を抽出する。

4.2.3 除外領域の抽出

深度情報のみに基づいて候補領域を抽出した場合、追跡対象に接触する他物体がその領域に含まれる問題がある。3.5 節でも述べたように、本来不要なこれらの領域の色情報をモデルに取り込むことは追跡が破綻する原因となる。この問題を解決するために、提案手法では追跡対象物体の周囲に広がる色情報に基づいて候補領域から除外すべき領域を検出する。

提案手法の処理の流れを Fig. 3 に示す。以下、各処理の詳細について順に述べる。

4.3 複数の色分布モデルによる追跡

まず、従来手法と同様に色分布モデルを用いたパーティクルフィルタによって物体追跡を行なう。このとき、4.2.1 項で述べたように、提案手法では複数の色分布の集合 C を追跡対象のモデルとする。重み関数 $w(\cdot|C) \in [0, 1]$ はつぎのように定義される。

$$w(s|C) = \max_{1 \leq k \leq |C|} \left\{ \exp \left(-\frac{d_b(p(s), q_k)^2}{2\sigma^2} \right) \right\} \quad (11)$$

ここで q_k ($k = 1, 2, \dots, |C|$) は C に含まれる k 番目の色分布を表わす。 q_1 は常に現時刻での外見に相当する色分布を表わすものとし、毎時刻更新する。 q_k ($k \neq 1$) は、追跡対象の外見が変化した結果、 C に含まれる既存の色分布だけでは十分な精度で追跡できなくなった場合に追加される。この処理については 4.6 節で述べる。

得られた全パーティクルの平均状態 $\bar{s}$ をもとに、現在のモデルで追跡対象と認識できる領域 $\mathcal{R}_{obj}$ を抽出できる。

$$\mathcal{R}_{obj} = \mathcal{E}(\bar{s}) \cap \mathcal{R}_{col} \quad (12)$$

$$\mathcal{R}_{col} = \{u | u \in \mathcal{I}, \max_{p_k \in C} \{p_k(h(u))\} > p_{th}\} \quad (13)$$

ここで、 $\mathcal{E}(\bar{s})$ は $\bar{s}$ が表わす楕円領域、 $\mathcal{R}_{col}$ は色分布集合 C に含まれるいずれかの色分布において尤度が p_{th} 以上の色をもつ画素の集合、 $h(u)$ ($= 1, 2, \dots, 256$) は画素 u がもつ色

の色分布ヒストグラムにおけるビンを表わす。

4.4 追跡対象物体領域の抽出

4.2.2 項で述べたように、自己遮蔽により隠れていた領域が現れ、その外見に応じた色分布を新たに追加するためには、どこまで追跡対象の一部であることを認識する必要がある。

4.1 節の仮定 3 から、4.3 節で抽出した領域 $\mathcal{R}_{obj}$ と画像上で連結し、かつ深度が連続的に変化する領域は追跡対象物体の一部であると考えられる。その領域を抽出するために、提案手法では Graph Cut による深度情報に基づいたセグメンテーションを利用する。

4.4.1 Graph Cut の概要

まず、Graph Cut^{8),9)}の概要について述べる。

$L = \{L_1, L_2, \dots, L_u, \dots, L_{|\mathcal{I}|}\}$ を画像 $\mathcal{I}$ におけるラベル分けの状態とする。ここで u は画像 $\mathcal{I}$ の画素を表わす。 L_u は画素 u のラベルとし、追跡対象物体領域を表わす obj 、またはそれ以外の背景領域を表わす bkg の 2 値を取るとする。

このとき、ラベル分け L に対するエネルギー $E(L)$ をつぎのように定義する。

$$E(L) = R(L) + B(L) \quad (14)$$

ここで $R(L)$, $B(L)$ はそれぞれ領域およびその境界の特徴を表わす項であり、つぎのように定義する。

$$R(L) = \sum_{u \in \mathcal{I}} R_u(L_u) \quad (15)$$

$$B(L) = \sum_{\{u,v\} \in \mathcal{P}} B_{\{u,v\}} \cdot (1 - \delta(L_u, L_v)) \quad (16)$$

ここで、 $\mathcal{P}$ は画像 $\mathcal{I}$ において隣接する画素の非順序対 $\{u, v\}$ の集合、 δ はクロネッカーのデルタ関数を表わす。また、 $R_u(L_u)$ は画素 u のラベルが L_u であるときのコスト、 $B_{\{u,v\}}$ は隣接する画素対 $\{u, v\}$ を繋ぐエッジのコストを表わし、具体的には求めたいラベル分けの性質に応じて定義される。

Graph Cut はエネルギー (14) を最小にするラベル分けを探索する。最適化手法として min-cut/max-flow アルゴリズム²²⁾を利用する。画像 $\mathcal{I}$ の全画素に対応するノード、および 2 つの終端ノード $\mathcal{S}$, $\mathcal{T}$ をもつグラフ $\mathcal{G}_d$ を作成する。本論文では、 $\mathcal{T}$ とのエッジをもつ画素が追跡対象物体 obj に、 $\mathcal{S}$ とのエッジをもつ画素が背景 bkg に相当する。隣接する 2 つのノードを繋ぐエッジを n-link、各ノードと $\mathcal{S}$ または $\mathcal{T}$ とをつなぐノードを t-link とよぶ。

4.4.2 領域抽出への適用

グラフ $\mathcal{G}_d$ の各エッジに対し、Table 1 に示すコストを設定する。ここで、 $\mathcal{R}_{obj}$, $\mathcal{R}_{bkg}$ は追跡対象物体 obj および背景 bkg のシードとなる領域、 $\mathcal{R}_{udf}$ はラベルが未定義の領域を表わす。 $\mathcal{R}_{obj}$ は 4.3 節で得たものを利用する。残りの 2 つの領域はそれぞれつぎのように定義する。

$$\mathcal{R}_{bkg} = \mathcal{I} \setminus \mathcal{E}'(\bar{s}) \quad (17)$$

$$\mathcal{R}_{udf} = \mathcal{I} \setminus (\mathcal{R}_{obj} \cup \mathcal{R}_{bkg}) \quad (18)$$

Table 1 Costs for edges in graph $\mathcal{G}_d$

Edge		Cost	for
n-link	$\{u, v\}$	$B_{\{u,v\}}$	$\{u, v\} \in \mathcal{P}$
t-link	$\{u, \mathcal{S}\}$	K	$u \in \mathcal{R}_{obj}$
	$\{u, \mathcal{S}\}$	$R_u(bkg)$	$u \in \mathcal{R}_{udf}$
	$\{u, \mathcal{S}\}$	0	$u \in \mathcal{R}_{bkg}$
t-link	$\{u, \mathcal{T}\}$	0	$u \in \mathcal{R}_{obj}$
	$\{u, \mathcal{T}\}$	$R_u(obj)$	$u \in \mathcal{R}_{udf}$
	$\{u, \mathcal{T}\}$	K	$u \in \mathcal{R}_{bkg}$

Fig. 4 Definition of regions $\mathcal{R}_{obj}$, $\mathcal{R}_{bkg}$ and $\mathcal{R}_{udf}$

ここで、 $\mathcal{E}'(\bar{s})$ は楕円領域 $\mathcal{E}(\bar{s})$ と中心座標が同じで x 軸方向と y 軸方向の半径をそれぞれ ϵ ずつ拡大した楕円領域を表わす。 $\mathcal{R}_{obj}$, $\mathcal{R}_{bkg}$ および $\mathcal{R}_{udf}$ の関係を Fig. 4 に示す。図の白い領域が $\mathcal{R}_{obj}$ 、黒い領域が $\mathcal{R}_{bkg}$ 、灰色の領域が $\mathcal{R}_{udf}$ を表わす。

n-link に与えるコスト ((16) 式の $B_{\{u,v\}}$) をつぎのように定義する。

$$B_{\{u,v\}} = \exp \left(-\frac{|d_u - d_v|^2}{2\sigma_{B_d}^2} \right) \quad (19)$$

ここで、 d_u , d_v は画素 u , v の深度、 σ_{B_d} は深度のばらつきを表わすパラメータである。また、t-link に与えるコスト ((15) 式の R_u) をつぎのように定義する。

$$R_u(obj) = -\log p_d(u) \quad (20)$$

$$R_u(bkg) = -\log(1 - p_d(u)) \quad (21)$$

$$K = 1 + \max_{u \in \mathcal{I}} \left\{ \sum_{v: \{u,v\} \in \mathcal{P}} B_{\{u,v\}} \right\} \quad (22)$$

ここで p_d は領域 $\mathcal{R}_{obj}$ 内の深度分布に基づく尤度関数であり、つぎのように定義する。

$$p_d(u) = \exp \left(-\frac{|d_u - \bar{d}|^2}{2\sigma_d^2} \right) \quad (23)$$

ここで $\bar{d}$ および σ_d は領域 $\mathcal{R}_{obj}$ 内の深度の平均および標準偏差を表わす。

以上のように設定したグラフ $\mathcal{G}_d$ に min-cut/max-flow アルゴリズムを適用し、未定義領域 $\mathcal{R}_{udf}$ の各ピクセルに対するラベルを決定することで、追跡対象と同一物体の画像領域の候補 $\hat{\mathcal{R}}$ を抽出する。

4.5 除外領域の抽出

4.5.1 除外処理の概要

前節の処理では深度情報のみに基づいて候補領域 $\hat{\mathcal{R}}$ を抽出

$$R_u(obj) = -\log p_{obj}(h(u)) \quad (28)$$

$$R_u(bkg) = -\log p_{bkg}(h(u)) \quad (29)$$

$$K = 1 + \max_{u \in \mathcal{I}_s} \left\{ \sum_{v: \{u,v\} \in \mathcal{P}} B_{\{u,v\}} \right\} \quad (30)$$

ここで, p_{obj} , p_{bkg} はそれぞれ領域 $\mathcal{R}_{obj}$, $\mathcal{R}_{bkg}$ における色分布, $\Delta_{col}(u, v)$ は画素 u , v のもつ色の HS 平面上での距離, σ_{B_c} は HS 平面上でのばらつきを表わすパラメータである.

以上のように設定したグラフ $\mathcal{G}_c$ に min-cut/max-flow アルゴリズムを適用し, 除外領域 $\overline{\mathcal{R}}$ を抽出する. 最後に, 除外領域 $\overline{\mathcal{R}}$ を追跡候補領域 $\widehat{\mathcal{R}}$ から除外することで, 追跡対象物体のみからなる領域 $\mathcal{R}$ を抽出する.

$$\mathcal{R} = \widehat{\mathcal{R}} \setminus \overline{\mathcal{R}} \quad (31)$$

抽出された除外領域 $\overline{\mathcal{R}}$ の例を Fig. 7(e) に, 追跡対象物体領域 $\mathcal{R}$ の例を Fig. 7(f) に示す. Fig. 7(f) から, 人物の指やマットが除外された領域を抽出できていることがわかる.

4.6 物体モデルの更新

前節までの処理で抽出された追跡対象領域 $\mathcal{R}$ における色分布を作成し, 色分布集合 C の q_1 として置き換える. q_1 は常に追跡対象の最新の外見に対応する色分布とし, 毎時刻更新する.

つぎに, q_1 と C に含まれるほかのすべての色分布 q_k ($k \neq 1$) との距離 (6) d_b を求める. もしすべての色分布との距離が事前に設定した閾値 d_{th} を超えている場合, C は追跡対象の最新の外見を表現できていないことになる. そこで, q_1 を $q_{|C|+1}$ として色分布集合 C に追加する. q_1 と $q_{|C|+1}$ は同じものであり, 一時的に重複するが, q_1 は次時刻に更新される.

物体モデルの更新後, 時刻を更新し次フレームに対する追跡処理を開始する.

以上の一連の処理により, 自己遮蔽が発生する状況においても頑健な物体追跡が可能となる.

5. 評価実験

提案手法の有効性を検証するため, 評価実験を行なう.

5.1 実験設定

以降のすべての実験において, RGB-D カメラとして Microsoft 社製 Kinect for XBOX360 を用いる. また, 深度情報取得および画像処理にはそれぞれ OpenNI, OpenCV ライブラリを利用する. 使用した PC の OS は Windows 7 Professional SP1, CPU は Intel 社製 Core i5-4590 (3.30 GHz), メモリは 4GB であり, プログラミング環境として Visual Studio 2013 (C++言語) を用いた.

追跡対象物体には面によって色が異なるキューブ (Fig. 2) および紙袋 (Fig. 8) を用いる.

処理を行なう画像サイズは 320×240 画素とする. そのほかのパラメータは, パーティクル数 $N = 200$, コーシー分布 (5) のスケールパラメータ $\tau = 3$, 重み関数 (11) のパラメータ

Fig. 8 Paper bag as tracking target

Fig. 9 Experimental result for moving camera

タ $\sigma = 0.25$, 領域 $\mathcal{R}_{col}$ の定義 (18) における閾値 $p_{th} = 0.01$, $\mathcal{G}_d$ における n-link のコスト (19) のパラメータ $\sigma_{B_d} = 5$, 深度分布の尤度関数 (23) のパラメータ $\sigma_d = 0.25$, 楕円領域および探索領域の拡張幅 $\epsilon = 15$, $\mathcal{G}_c$ における n-link のコスト (27) のパラメータ $\sigma_{B_c} = 5$, 物体モデル更新の閾値 $d_{th} = \exp(-1) \approx 0.368$ と設定する. これらの値は予備実験により経験的に定めた.

5.2 実験 1

まず, 提案手法により自己遮蔽に頑健な物体追跡が実現できることを確認する実験を行なう.

背後に段ボール箱を置き, 床面にマットを敷いた環境で, カメラの移動により自己遮蔽が発生・解消する条件での追跡を行なった. 提案手法による追跡結果を Fig. 9 に示す. Fig. 9 の左側の図がカメラ画像を表わし, 図中の楕円が追跡対象を表わす. また, 右側の図は抽出した追跡物体領域を表わす. 追跡開始時はキューブの赤色の面を追跡していたにもかかわらず (Fig. 9 上段), カメラの移動とともに現れた青色の面も追跡対象に取り入れ (中段), 逆に赤色の面が自己遮蔽により見えなくなっても安定して追跡を継続できている (下段) ことがわかる.

同様の実験を, 木目の机の上に置かれた紙袋を移動させ, 固

Fig. 10 Experimental result for fixed camera

Fig. 11 Experimental result for target held by hand

定したカメラから追跡する条件でも行なった。提案手法による追跡結果を Fig. 10 に示す。Fig. 9 と同様、追跡開始時に見えていた面が自己遮蔽により隠れたにもかかわらず (Fig. 10 下段)、安定して追跡を継続できている。

つぎに、Fig. 9 と同様の環境において人間の手でキューブを把持した状態での追跡を行なった。提案手法による追跡結果を Fig. 11 に示す。Fig. 11 (a) が原画像、(b) が深度のみで抽出した追跡対象候補領域 $\hat{\mathcal{R}}$ を表す。Fig. 11 (b) に把持する人物の指や床面のマットの一部が誤って含まれていることがわかる。Fig. 11 (c) は除外色をもつ除外領域 $\bar{\mathcal{R}}$ 、(d) は除外処理後の追跡対象物体領域 $\mathcal{R}$ を表す。Fig. 11 から、提案手法により追跡対象候補領域から周辺に広がる手やマット

Fig. 12 Comparison between conventional method and proposed method

の色を除外色として認識し、適切に除外できていることがわかる。

最後に、移動するカメラから机の上に置いたキューブを追跡する条件で、自己遮蔽に対する処理を行なわない従来手法と提案手法との比較を行なった。追跡結果を Fig. 12 に示す。Fig. 12 (a) が従来手法、(b) が提案手法による結果を表す。図中の黒い楕円は個々のパーティクルによる状態仮説を表す。Fig. 12 から、従来手法が自己遮蔽の発生・解消により生じる追跡対象の外見変化に対応できず見失っているのに対し、提案手法では適切に追跡を継続できていることがわかる。

以上の結果から、提案手法により自己遮蔽に頑健な物体追跡が実現できていることを確認した。

5.3 実験 2

つぎに、4.5 節で述べた除外処理の有効性を確認する実験を行なう。

実験環境を Fig. 13 に示す。実験 1 と同様、背後に段ボール箱を置き、床面にマットを敷いた環境で、Fig. 13 (a) では手袋をした人物の手でキューブを把持し、(b) ではキューブを床面のマットだけでなく背後の段ボール箱にも接触させた状態での追跡を行なった。4.5 節の除外処理を行なわない場合と行なう場合とで比較を行なう。

Fig. 13 (a) の条件に対する結果を Fig. 14 に示す。Fig. 14 (a) および (b) は除外処理を行なわない場合の結果を表す。キューブを把持する手袋が追跡対象として含まれていることがわかる。一方、Fig. 14 (c)~(e) は除外処理を行なった場合の結果を表す。手袋、段ボール箱、マットの色の広がりを検出することでそれらを除外し、本来の追跡対象

Fig. 13 Experimental environments

(a) Tracking w/o exclusion (b) Target w/o exclusion

(c) Tracking with exclusion #1

(d) Exclusion region #1

(e) Target with exclusion #1

(f) Tracking with exclusion #2

(g) Exclusion region #2

(h) Target with exclusion #2

Fig. 14 Experimental results for Fig. 13 (a)

のみの領域を抽出できていることがわかる。また、ここからキューブを回転し隠された面が画面上に現れた場合の結果を Fig. 14 (f)~(h) に示す。追跡対象以外の物体を除外しつつ、新

Fig. 15 Experimental results for Fig. 13 (b)

Fig. 16 Experimental environment

たに現れた追跡対象の一部を領域へ含めることができています。

つぎに、Fig. 13 (b) の条件に対する結果を **Fig. 15** に示す。Fig. 15 (a) および (b) は除外処理を行わない場合の結果を表わす。キューブと接触し深度変化が連続になったことで段ボール箱が追跡対象の一部として抽出されている。一方、Fig. 15 (c)~(e) は除外処理を行なった場合の結果を表わす。Fig. 14 の場合と同様、除外色の検出により本来の追跡対象のみの領域を抽出できていることがわかる。

最後に、背景が平面でない環境で同様の実験を行なう。実験環境を **Fig. 16** に示す。紙袋を机の上に置き、ぬいぐるみを前方から接触させる環境とした。

Fig. 16 の条件に対する結果を **Fig. 17** に示す。Fig. 17 (b) は除外処理を行わない場合の結果を表わす。紙袋と接触し深度変化が連続になったぬいぐるみ、机の面および人物の手が追跡対象の一部として抽出されている。一方、Fig. 17 (c), (d) は除外処理を行なった場合の結果を表わす。これまでの実験結果と同様、除外色の検出によりぬいぐるみ、机の面および人物の手が取り除かれ、本来の追跡対象のみの領域を抽出できていることがわかる。

以上の結果から、除外色の検出および除外領域の抽出が有効に機能していることを確認した。

Fig. 17 Experimental results for Fig. 16

Fig. 18 Processing time and number of color distribution models

5.4 実験 3

最後に、提案手法の処理にかかる時間、および追跡中の色分布集合 C に含まれる色分布モデル数の変化を確認した。

Fig. 9 と同様の実験を再現したときの処理時間およびモデル数の変化のグラフを Fig. 18 に示す。追跡中の処理時間は平均 156.7 ms, 最大で 190.0 ms であった。フレームレートは平均 6.38 fps, 最低で 5.26 fps となる。本章で示した実験結果において、処理時間の遅さから 1 フレームで追跡対象を見失った例はなく、第 4 章冒頭で述べた仮定 1 は十分に満たされていたと考えられる。

また、追跡中に色分布モデル数 $|C|$ は最大 15 個まで増加した。これは、外見の変化に応じてそれを表現するのに必要な色分布を取り込んでいった結果である。Fig. 18 から、色分布モデル数と処理時間が必ずしも連動しているわけではないが、モデル数が増えるほど処理時間も増える傾向が見て取れる。

6. 考 察

前章の実験結果から、提案手法によって自己遮蔽に対する頑健な追跡を実現できることを確認した。特に、追跡対象に接触する物体が存在するときにその物体を追跡対象領域から除外できることを確認した。たとえば人物の手のように頻繁に現れる物体が事前にわかっている場合は事前に色を記憶し

除外対象とすることも可能であるが、一般には状況に合わせて除外できることが望ましい。Fig. 14 に示したとおり、提案手法では手袋をした手のように通常と異なる色であっても除外を実現できており、汎用性が高いと考えられる。また、壁面や床面に接触する状況でも頑健な追跡を実現できており、幅広い環境で特定の追跡対象物体を認識し続けることを可能にしている。

提案手法は、従来の色分布モデルによる追跡の利点である

- 物体の変形や画像平面上での回転、部分的な遮蔽に頑健であること
- 事前知識を必要とせず、その時点で保持する情報のみから物体モデルを作成できること

などの利点を維持したまま自己遮蔽に対する頑健性を実現した。したがって、第 1 章で述べたようなヒューマン・ロボット・インタラクションにおけるロボットビジョンへの応用や、Imai らの手法⁷⁾など事前知識を必要とせずにモデル作成が可能であることを前提とした追跡手法へも適用可能である。

一方で課題も残されている。たとえば、提案手法では外見変化に追従するため同時に複数の色分布を保持していた。しかし、Fig. 18 に示したように、外見変化が進み多数のモデルを保持すると処理に必要な時間が増大する傾向があり、追跡が長期間に渡る場合には問題になる可能性もある。これを解決するためには、距離の近い色分布同士を統合するなど効率的に整理・圧縮し、少ないモデル数で精度を維持するような表現方法を検討する必要がある。

また、追跡対象とそれ以外の物体を分離する方法をより頑健にすることも今後の課題として挙げられる。提案手法では除外領域を色の広がり調べることによって検出した。しかし、追跡物体と接触する他物体が同系色だった場合、提案手法では除外領域として認識できず、誤って追跡対象モデルに取り込んでしまう可能性がある。この問題を解決するためには、たとえば提案手法に加えて物体表面のテクスチャ情報を表現し利用する方法などが考えられる。

7. おわりに

本論文では、色分布モデルに基づくパーティクルフィルタを利用した物体追跡手法を拡張し、色情報に加えて深度情報を併用することで自己遮蔽に頑健な物体追跡を実現する手法を提案した。その際、追跡対象として含めるべきではない領域を色の広がりに基づいて除外し、追跡の頑健性を維持する手法を導入した。評価実験の結果、提案手法が有効に機能し、自己遮蔽に対して頑健な物体追跡を実現できていることを確認した。

今後の課題として、第 6 章で述べたような色分布モデルの整理・圧縮による処理時間の改善や、テクスチャ情報を利用した頑健性の向上などに取り組む予定である。

参 考 文 献

- 1) 神尾 崇：画像認識を応用した高度映像セキュリティシステ

- ム, Panasonic Technical Journal, **54-4**, 13/17 (2009)
- 2) 渡邊 睦: 人物自動追従・状態自動認識に関するビジョン技術, 日本ロボット学会誌, **27-6**, 626/629 (2009)
 - 3) 佐竹, 三浦: ステレオビジョンを用いた移動ロボットの人物追従制御, 日本ロボット学会誌, **28-9**, 1091/1099 (2010)
 - 4) 小山, 但馬, 田中, 廣瀬, 鋤柄, 藤井: ヒトを活動的にする移動支援ロボット—IDを発信する赤外LEDタグと全方位カメラを用いた利用者の追跡—, 計測自動制御学会論文集, **51-1**, 49/56 (2015)
 - 5) 尾関, 柏木, 井上, 岡: 粒子フィルタを用いたスポットライト群メタファに基づく視覚的注意モデル, 電子情報通信学会論文誌 D, **J95-D-11**, 1909/1918 (2012)
 - 6) K. Nummiaro, E. Koller-Meier and L.V. Gool: An adaptive color-based particle filter, Image and Vision Computing, **21-1**, 99/110 (2003)
 - 7) J. Imai and M. Kaneko: Visual tracking in occlusion environments by autonomous switching of targets, IE-ICE Transactions on Information and Systems, **E91-D-1**, 86/95 (2008)
 - 8) Y.Y. Boykov and M.-P. Jolly: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images, Proceedings of Eighth IEEE International Conference on Computer Vision, **1**, 105/112 (2001)
 - 9) 石川 博: グラフカット, 情報処理学会研究報告, 2007-CVIM-158, 193/204 (2007)
 - 10) H. Grabner and H. Bischof: On-line boosting and vision, Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 260/267 (2006)
 - 11) 山下, 勞, 藤吉: Online real boosting による人物追跡, 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, **1-1**, 73/82 (2008)
 - 12) I. Matthews, T. Ishikawa and S. Baker: The template update problem, IEEE Transactions on Pattern Analysis and Machine Intelligence, **26-6**, 810/815 (2004)
 - 13) A.D. Jepson, D.J. Fleet and T.F. El-Maraghi: Robust on-line appearance models for visual tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence, **25-10**, 1296/1311 (2003)
 - 14) 村瀬, ナイヤー: 2次元照合による3次元物体認識—パラメトリック固有空間法—, 電子情報通信学会論文誌 D-II, **J77-D-II-11**, 2179/2187 (1994)
 - 15) 米川, 川本, 井宮, 杉本: 見えの生成モデルに基づく物体追跡, 情報処理学会研究報告, 2010-CVIM-172, 1/7 (2010)
 - 16) M. Isard and A. Blake: Contour tracking by stochastic propagation of conditional density, European Conference on Computer Vision, 343/356 (1996)
 - 17) D. Comaniciu, V. Ramesh and P. Meer: Kernel-based object tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence, **25-5**, 564/577 (2003)
 - 18) Q.A. Nguyen, A. Robles-Kelly and C. Shen: Kernel-based tracking from a probabilistic viewpoint, Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, 1/8 (2007)
 - 19) M. Isard and A. Blake: CONDENSATION — Conditional density propagation for visual tracking, International Journal on Computer Vision, **29-1**, 5/28 (1998)
 - 20) A. Doucet, N. de Freitas and N. Gordon (Eds.): *Sequential Monte Carlo Methods in Practice*, Springer (2001)
 - 21) F. Aherne, N. Thacker and P. Rockett: The Bhattacharyya metric as an absolute similarity measure for frequency coded data, Kybernetika, **32-4**, 1/7 (1997)
 - 22) Y. Boykov and V. Kolmogorov: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, IEEE Transactions on Pattern Analysis and

Machine Intelligence, **26-9**, 1124/1137 (2004)

[著者紹介]

今井 順一 (正会員)

1997年北海道大学工学部情報工学科卒業。2003年北海道大学大学院工学研究科システム情報工学専攻博士後期課程修了。2003年室蘭工業大学サテライト・ベンチャー・ビジネス・ラボラトリー中核的研究機関研究員。2005年電気通信大学大学院電気通信学研究科助手。2007年同助教。2010年千葉工業大学情報科学部准教授。2016年同教授。現在に至る。電子情報通信学会、情報処理学会、日本感性工学会、映像情報メディア学会各会員。人間と知能システムとのインタラクション、コンピュータビジョン、感性情報処理などの研究に従事。博士(工学)。

柏木 雄平

2016年千葉工業大学情報科学部情報工学科卒業。現在、(株)エム・ソフトに勤務。在学中、物体追跡をはじめとする画像処理の研究に従事。

木辻 亮

2015年千葉工業大学情報科学部情報工学科卒業。現在、(株)カイカに勤務。在学中、物体追跡をはじめとする画像処理の研究に従事。