Multiple Objects Tracking based on Auto-selection of Color Image Processing Methods

Kazuyuki Morioka Non-member (University of Tokyo, morioka@hlab.iis.u-tokyo.ac.jp)

Yoichi Kuroda Non-member (Tokyo University of Science, yoichi@hlab.iis.u-tokyo.ac.jp)

Joo-Ho Lee Member (Ritsumeikan University, leejooho@ieee.org)

Hideki Hashimoto Member (University of Tokyo, hashimoto@iis.u-tokyo.ac.jp)

Keywords: Image Sequence Processing, Object Tracking, Color Histogram, Clustering, Mean Shift

In order to achieve the applications, such as surveillance systems, ITS and intelligent environments, using vision sensors, it is important to extract and track the moving objects from image sequences. Some tracking systems often prepare the models of the objects in advance. It is difficult to adopt this model based tracking in the general environment where many kinds of objects exist. On the other hand, model based tracking is efficient for the tracking in the crowded environment. In this paper, we propose the color histogram based tracking system including the object model construction and object correspondence between frames by the fusion of the bottom-up process and top-down process.

Fig. 1 shows the two processes which configure whole tracking algorithm. Object finding process is used for the extraction of new objects and correspondence of the objects between frames. The information of the new objects are also obtained in the object finding process. This color model is defined based on the color histograms of the objects separated from background image. Feature vectors based on the color histogram of the object region are clustered in the feature space for recognition of the new objects.

Tracking process is used for the objects tracking in the complex environment in which object finding process doesn't work well. Initial color model is constructed by using the object ID and size of object region given by object finding process in tracking process. Object color model is compared with the color histogram of the object candidate regions in the next frame. Object candidate region that has largest similarity is decided as the object region in the next frame by Mean Shift algorithm. These two processes are selected automatically according to the tracking condition in the proposed

tracking method.

Fig. 2 and Fig. 3 show the experimental results of the object tracking. As shown in Fig. 2, two objects are tracked under the occlusion stably. Process selection result in the tracking of Fig. 2 is shown in Fig. 3. Selection of the process to Object 1 and Object 2 was performed like this figure.

Fig. 2. Tracking results

Fig. 3. Result of process selection

Fig. 1. Structure of tracking algorithm

色情報処理手法の自動選択による複数物体追跡

非会員 森岡 一幸* 非会員 黒田 陽一**
正員 李 周浩*** 正員 橋本 秀紀**

Multiple Objects Tracking based on Auto-selection of Color Image Processing Methods

Kazuyuki Morioka*, Non-member, Yoichi Kuroda**, Non-member,
Joo-Ho Lee***, Member, Hideki Hashimoto**, Member

In order to achieve the applications using vision sensors, it is important to extract and track the moving objects from image sequences. Some tracking systems often prepare the models of the objects in advance. It is difficult to adopt this model based tracking in the general environment where many kinds of objects exist. On the other hand, model based tracking is efficient for the tracking in the crowded environment. We propose the color histogram based tracking system including the object model construction and object correspondence between frames by the fusion of the bottom-up process and top-down process.

In this paper, two processes that configure the tracking system are described. Then, process selection method in the proposed tracking system is explained. Finally, experimental results of objects are shown and effect of the proposed tracking system is also explained.

キーワード：動画処理，物体追跡，色ヒストグラム，クラスタリング，Mean Shift

Keywords: Image Sequence Processing, Object Tracking, Color Histogram, Clustering, Mean Shift

1. はじめに

観測画像から移動物体を抽出し，追跡・位置同定を実現することは，視覚センサを用いたシステムにおいて最も重要な技術のひとつである。視覚センサを用いたシステムのアプリケーションとして，街や公的空間の監視⁽¹⁾やITSにおける車両の認識や追跡⁽²⁾，ロボット技術を拡張した空間知能化⁽³⁾などが挙げられる。特に，空間知能化においては，空間に大量に視覚センサを分散配置することで，人物やロボットなどの種類によらない移動物体の抽出と追跡を行ない，空間側の知識獲得に結びつけることが不可欠である。

移動物体の抽出・追跡と位置同定のために，対象物体の変形や隠れなどに比較的安定であるという特長を持つ，対

象物体の色情報の特徴量として，物体追跡を実現する研究例が数多く見受けられる⁽⁴⁾⁻⁽⁷⁾。物体の追跡のためには，特徴となる色情報の獲得方法がまず問題となる。中村らによる色情報のオンライン学習に基づくロボットの行動を実現した例⁽⁸⁾ではFuzzy ART ニューラルネットワークを使用した環境の変動にロバストな色情報の獲得方法が示されている。また，追跡対象色情報と背景情報を融合した物体抽出法⁽⁹⁾では，画像のピクセル座標と3次元YUV色情報を組み合わせた5次元空間での最近傍識別による色情報の学習が検討されている。物体の追跡のために次に問題となるのが，獲得した特徴量に基づいた追跡アルゴリズムである。物体領域の色情報を事前情報として保持し，追跡を行なう研究のうち代表的なものは，Comaniciuらによる研究⁽¹⁰⁾である。この研究では，カーネル関数に基づいて追跡対象領域のカラーヒストグラムを構成し，次フレームの候補領域をカルマンフィルタとMeanShift法によって推定している。Satoらの研究⁽¹¹⁾では，Comaniciuによる手法をベースとしながら，追跡対象物体の各部分毎にカラーヒストグラムを構成し，パーティクルフィルタを用いて次フレームの物体位置を推定することで，追跡性能を向上させている。

これらの先行研究では，色情報による物体追跡の有効性が示されているものの，一般的に，物体の色情報獲得と物体追跡は別の問題として取り上げられている。自律的な空

* 東京理科大学 理工学部 電気電子情報工学科
〒278-8510 千葉県野田市山崎 2641

Tokyo University of Science

2641, Yamazaki, Noda, Chiba, 278-8510

** 東京大学生産技術研究所

〒153-8505 東京都目黒区駒場 4-6-1

Institute of Industrial Science, University of Tokyo

4-6-1, Komaba, Meguro-ku, Tokyo, 153-8505

*** 立命館大学 情報理工学部 情報コミュニケーション学科

〒525-8577 滋賀県草津市野路東 1-1-1

Ritsumeikan University

1-1-1, Noji Higashi, Kusatsu, Shiga, 525-8577

間認識機能の実現には、追跡対象物体を自動認識し、認識結果に基づいて追跡を行なう対象物体の検出と追跡に関する処理が含まれている必要がある。本研究では、画像中の物体の状態に応じて、物体の存在を確定させ、色情報を獲得し追跡するプロセスと、物体の存在の仮定に基づいて物体領域のカラートラッキングを行なうプロセスを自動的に選択することによって、物体に関する事前情報のない状態から追跡を開始できる複数物体追跡手法を提案する。次章以降、まず2章では提案する複数物体追跡アルゴリズムの概要について述べる。3章と4章では、物体追跡アルゴリズムを構成する各プロセスについて説明し、5章でプロセスの選択方法について述べる。6章では実験結果を示し、最後に結論として本研究のまとめと今後の課題について述べる。

2. 物体追跡アルゴリズム

本研究で用いる追跡のための物体特徴は、画像から抽出される物体領域の色情報である。本研究で提案する複数物体追跡アルゴリズムは、以下の図1に示す例のように、オブジェクト発見プロセスとトラッキングプロセスから構成され、オブジェクト発見プロセスをベースとしながら適宜プロセスの選択を行なうことで物体を追跡する。本アルゴリズムのコアとなる二つのプロセスでは以下の処理を行なう。

● オブジェクト発見プロセス

本研究では、固定カメラを用いたシステムを考えており、入力画像を背景画像と比較することで、ある程度は物体の抽出が可能である。このプロセスでは、背景画像情報に基づいて、新規物体と考えられる領域を検出し、カメラの観測領域内に存在する物体数と各物体のカラーモデルを獲得する。さらに、新規物体を発見するだけでなく、獲得したカラーモデルに基づいてフレーム間の物体の対応づけをすることで追跡も行なう。

● トラッキングプロセス

オブジェクト発見プロセスにより確定した物体の存在情報（物体数、物体位置）から、物体が存在すると想定される領域の色情報に基づいた物体領域追跡を実現する。

図1の例では、フレーム n からフレーム $n+5$ までの物体 m および $m+1$ の追跡を行なっている際の様子を示している。物体 m 、 $m+1$ はそれぞれフレーム n 以前にオブジェクト発見プロセスにより検出および追跡された状態であったが、追跡状態に応じて、物体 m に関してはフレーム $n+1$ から $n+2$ まで、物体 $m+1$ に関してはフレーム $n+2$ から $n+4$ まで、トラッキングプロセスが起動し追跡を継続するというプロセスの切替による追跡手法の概念を示している。オブジェクト発見プロセスでは、背景差分による物体領域抽出を前提としているため、物体に背景色と類似する色情報が含まれているときや、ノイズなどの影響で物体領域の検出が困難であるとき、また物体同士のオクルージョンにより各物体に属する領域を分割できないときなどの追跡は不可能である。そのような場合に、複雑な

図1 トラッキングアルゴリズムの構成

Fig. 1. Structure of tracking algorithm.

環境でも背景色の影響などを受けずに物体領域の追跡が可能なアルゴリズムを用いることによって、物体の追跡を継続する。

オブジェクト発見プロセスとトラッキングプロセスの融合により、トラッキングの状況に応じて適切なプロセスが選択され、オクルージョンなどの発生時にも安定したトラッキングを行なうことが可能である。この二つのプロセスの融合による物体追跡アルゴリズムは、得られる画像情報から未知の物体のモデルを自動生成し、追跡・位置認識を行なう、画像内の色情報からのボトムアップ的な処理と、あらかじめ物体の存在が仮定された状態からのトップダウン的な処理が統合され、互いに補完的に動作することで複数物体の追跡を実現する。また、各プロセスは物体毎に別々に動作するものであり、佐竹らの提案⁽¹²⁾に見られるような、注視点制御に基づく追跡といった要素も含んでいる。

3. オブジェクト発見プロセス

3.1 物体候補領域の抽出

固定カメラによるシステムでは、背景画像との比較による移動物体領域と背景画像を分離することで、対象物体の大まかな情報を獲得することができる。移動物体抽出の前提となる背景画像は文献(13)の手法に基づいて生成され、入力画像との差分により物体候補領域を抽出する。背景差分で得られた二値画像について膨張・収縮処理、ノイズの除去などの処理を加えてさらにラベリングを施し、各物体候補ごとの領域に分割する。一定量の画素数に至らない物体候補領域はここで除去する。以上の処理により、図2に示すような、物体候補領域が抽出される。

まず、図3に示すような形で、各物体候補領域ごとのカラーヒストグラムに基づいた、各物体領域の画像フレーム毎での物体候補領域を表す特徴ベクトルを定義する。ここでは、時刻 t における画像フレームで、 j 番目の物体候補領域を表す特徴ベクトルを p_j^t とする。このとき、物体を表す特徴ベクトル p_j^t の各成分 $(p_j^t)_u$ ($u = 1, \dots, m$) は、以下の式(1)のように計算される。

図 2 抽出された物体候補領域

Fig. 2. Extracted object candidates.

図 3 カラーヒストグラムに基づく抽出物体の特徴ベクトル

Fig. 3. Feature vector of extracted object based on color histogram.

$$(p_j^t)_u = \frac{1}{n_j^t} \sum_{i=1}^{n_j^t} \delta[b(x_i) - u] \dots\dots\dots (1)$$

ここで、 $\{x_i\}_{i=1, \dots, n_j^t}$ は、物体候補領域のピクセル位置を示している。 n_j^t はこの物体候補領域内に含まれる総ピクセル数であり、領域の大きさで正規化されている。 u は色空間の分割の際に割り当てた色番号を示している。例えば、RGBの各色をそれぞれ8分割したヒストグラムの場合、 u は1から $m = 512$ の値をとることになる。 $b(x_i)$ は、物体領域内の座標 x_i のピクセル値を色空間の次元から1次元へと変換するときに割り当てられる色番号である。本研究では色空間として、2次元の正規化色空間 $r = \frac{R}{R+G+B}$, $g = \frac{G}{R+G+B}$ を用いている。すなわち、各物体候補領域は、カラーヒストグラムの各ビンの要素をベクトルとした特徴ベクトル p_j^t で表される。ただし、 δ はクロネッカーデルタ関数である。

3.2 物体候補領域色情報の特徴 カラーヒストグラムで表現された物体候補領域の特徴ベクトルの分布の特徴について考察する。図4は、特徴ベクトルが検出される画像の例、および特徴ベクトルを入力パターンとした際に一定時間内に得られる移動物体のパターンの集合を特徴空間にプロットしたものである。図3.2の画像では、カメラの観測領域内に3個の物体が存在している。図3.2は、そのような3個の物体が常に存在している数秒間の画像系列から、フレーム毎に抽出される物体候補領域の特徴ベクトルの全てを特徴空間に表したものである。ここでの特徴空間とは、特徴ベクトルの次元数と同一の次元を持つ多次元空間であるため、その分布状態を可視化するために全成分から2成分のみを選択してプロットしている。パターンの分布からわかるように、物体の数に対応したクラスターが得ら

図 4 入力パターンの分布

Fig. 4. Distribution of input patterns.

図 5 オンライン物体認識の概念

Fig. 5. Concept of on-line object recognition.

れる。クラスターの中心から離れるパターンは画像中に現れた雑音、もしくは物体同士の短時間のオクルージョンなどにより変動した特徴ベクトルを示している。つまり、ある一定時間内に得られる、物体を構成するカラーヒストグラムによる特徴ベクトルを、その特徴が類似しているグループにまとめ、毎フレーム入力されるパターンの分布に従っていくつかのクラスターに分類することで、カメラの観測領域内に存在する物体数が確定でき、また各クラスターの標本ベクトルを対応させることで、各物体のカラーモデルを獲得することができると考えられる。

このような処理では、教師なし学習型で入力データを受け取るとすぐにパラメータを修正するオンラインの分類アルゴリズムが必要である。また、クラス数は既知でないため、オンラインでの分類の際に既存のクラスターに属するのか新規クラスターであるかを判断する処理が必要とされる。

3.3 オンラインクラスタリングによる物体情報獲得 カメラ観測領域内に存在する物体数および各物体に対するカラーモデルを獲得するため、特徴空間において物体情報をオンラインでクラスタリングするオブジェクト発見プロセスにより、各物体毎のカラーモデルを獲得することを考える。概念図を図5に示す。時系列情報として各物体候補の特徴ベクトルが毎フレームごとに得られる。得られた特徴ベクトル群を特徴空間に投影し、特徴空間内で適切に分割することで、物体数とカラーモデルを確定させる。ここでは、観測領域内で物体が検出された場合、各物体に対応した一定のサイズ以上のクラスターが一定時間内に特徴空間内に表れるものと仮定する。

図6は、オブジェクト発見プロセスの処理の流れを示し

図 6 オブジェクト発見プロセス

Fig. 6. Object finding process.

ている。前述の通り、はじめに背景差分により物体と考えられる領域を切り出し、各領域に対してカラーヒストグラムに基づいて特徴ベクトルを構成する。一定時間内(図6の例では $t-s$ から t までの s フレーム)に、 a 個以上の特徴ベクトルで構成されたクラスタが生成されたときには、新規物体として認定する。認定された新規物体には物体 ID が付与され、オブジェクト発見プロセスおよびトラッキングプロセスを通じて使用される。それ以外の特徴ベクトルは、ノイズにより生じたもの、あるいはオクルージョンの発生などによる物体切り出しの失敗によるものとする。

一定時間内のクラスタ生成状況により物体の存在を認定するのは、次のような理由からである。第一に、前述した通り、背景差分により検出された領域は、物体が背景の色情報と同一色を含む場合や、ノイズなどの影響で、基本的に不安定である。そのため、物体カラーモデルの安定化のためには、各物体について、いくつかの特徴ベクトルからの情報が得られることが望ましいからである。第二には、移動などに伴う物体の色情報の変動のため、観測されるすべての特徴ベクトルからなるクラスタの形状は複雑化し、クラスタの標本ベクトルを物体カラーモデルとすることの妥当性は失われる。そのため、クラスタ生成のために有効な時間を制限することが必要である。

複数の物体が同時に発見される場合、特徴ベクトル群はオンラインクラスタリングにより、いくつかのカテゴリに分割され、各カテゴリが発見された物体に対応する特徴ベクトルの集合となる。新規に獲得された特徴ベクトルが、既存のクラスタに属するか、新たなクラスタとして扱うべきであるかは、クラスタの標本ベクトルとの距離に応じて決定される。最初に、式(2)により、時刻 t における、既存クラスタの参照ベクトルと現フレームで取得された物体の特徴ベクトルの類似度 S を計算することで、最近傍のクラスタを検出する。ここで、 r_k^t は時刻 t におけるオブジェクト k の参照ベクトルを示している。類似度は Bhattacharyya 係数により計算する。 p_j^t および r_k^t を用いた場合、Bhattacharyya 係数は式(2)で表される。Bhattacharyya 係数は 0 以上 1 以下の値をとり、両ベクトルが完全に一致する場合、最大値 1 をとる。

$$S(p_j^t, r_k^t) = \sum_u \sqrt{(p_j^t)_u (r_k^t)_u} \cdots \cdots \cdots (2)$$

ただし、 $(p_j^t)_u, (r_k^t)_u$ ($u = 1, \dots, m$) はそれぞれ特徴ベクトル p_j^t と標本ベクトル r_k^t の成分を表している。

また、 c がその特徴ベクトルが属するべきクラスタを表すとすると、次の式(3)のように求められる。ただし、この時点での既存のクラスタ数を N とする。

$$c = \begin{cases} \arg \max_k S(p_j^t, r_k^t), & \text{for } S(p_j^t, r_k^t) > T \\ N + 1, & \text{otherwise} \end{cases} \cdots \cdots \cdots (3)$$

ここで、 T は特徴ベクトル間の類似度を評価する閾値であり、クラスタの数や大きさに影響を与える。 T は、いくつかの動画像を用いた実験の結果より、経験的に決定される。新規クラスタとして判定された特徴ベクトルは、次フレームの処理以降、新規クラスタの標本ベクトルとして用いられる。

各クラスタの標本ベクトルは式(4)によって更新される。更新されたベクトルは、時刻 $t+1$ における標本ベクトルとして使用される。ここで、 α は学習係数であり、 δ_{ck} は標本ベクトル r_k^t を更新するべきかどうかを判定する。 δ_{ck} はクロネッカーデルタ関数であり、オブジェクト k がクラスタ c に属するときのみ 1 となり、標本ベクトル r_k^t が更新される。更新された r_k^{t+1} は全成分の合計が 1 になるように正規化される。

$$r_k^{t+1} = r_k^t + \alpha \cdot \delta_{ck} \{p_j^t - r_k^t\} \cdots \cdots \cdots (4)$$

各物体毎の安定したカラーモデルがこのオブジェクト発見プロセスによって得られる。次フレーム以降も、同様の処理が繰り返され、すでに発見された物体に関しては、式(2)(3)によってフレーム間での対応づけが行なわれ追跡される。

4. トラッキングプロセス

4.1 物体領域カラートラッキング トラッキングプロセスは、オブジェクト発見プロセスにおいて雑音やオクルージョン等の影響で、物体 ID が付与された物体の対応付けに失敗するとき、物体の存在を仮定した上で、各物体ごとに起動されるプロセスである。オブジェクト発見プロセスとトラッキングプロセスは物体追跡というタスクにおいて互いに相補的であり、どちらか一方の処理に大きく依存するべきではなく、最小限の情報のやりとりだけでプロセスを切替えられることが望ましい。ここでは、トラッキングプロセスではオブジェクト発見プロセスで得られた物体情報のうち、物体 ID と物体の大きさを含めた位置情報をトラッキングプロセスの初期情報として使用する。本研究では色ベーストラッキングにおける代表的な位置同定手法である MeanShift 法⁽¹⁰⁾を改良した領域のカラートラッキング手法をトラッキングプロセスとして使用する。

図7 トラッキングプロセスのためのカラーモデル

Fig. 7. Color model for tracking process.

4.2 トラッキングのための物体領域モデル トラッキングのための物体カラーモデルは、トラッキングプロセス開始フレームにおける物体領域を囲む領域から、式(6)により色ヒストグラムとして生成する。まず、オブジェクト発見プロセスからトラッキングプロセスに状態が遷移したとき、物体位置とその大きさの情報に基づいて、図7に示すように、物体領域にカーネル関数 $k(x)$ を用いた重みづけを行ない、背景色による影響を抑制した物体領域の色特徴を定義する。本研究では式(5)によって表される Triangle Kernel 関数を利用する。また、 $b(x)$ はオブジェクト発見プロセスと同様に物体領域に含まれるピクセル値を色空間の次元から1次元に変換する関数であり、ここでは、フレーム t におけるオブジェクト k に対するモデル領域 Q_k^t ならびに、候補領域 P_k^t について説明している。

$$k(x) = \begin{cases} 1 - |x|, & \text{for } |x| < 1 \\ 0, & \text{otherwise} \end{cases} \quad \dots\dots\dots (5)$$

• Target Model: $Q_k^t = \{(q_k^t)_u\}_{u=1,\dots,m}$

$$(q_k^t)_u = C \sum_{i=1}^n k(\|x_i^*\|) \delta[b(x_i) - u] \quad \dots\dots\dots (6)$$

$$C = \left\{ \sum_{i=1}^n k(\|x_i^*\|^2) \right\}^{-1} \quad \dots\dots\dots (7)$$

• Target Candidate: $P_k^t = \{(p_k^t)_u\}_{u=1,\dots,m}$

$$(p_k^t)_u = C_h \sum_{i=1}^{n_h} k(\|x_i\|) \delta[b(x_i) - u] \quad \dots\dots\dots (8)$$

$$C_h = \left\{ \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \right\}^{-1} \quad \dots\dots\dots (9)$$

ここでも、オブジェクト発見プロセスと同様に、色空間として正規化 RGB カラーを用いる。 x_i^* は中心を0としたモデル領域内のピクセル座標を指す。なお、式(7)および式(9)は、それぞれ $\sum_{u=1}^m (q_k^t)_u = 1$, $\sum_{u=1}^m (p_k^t)_u = 1$ になるよう標準化するための関数である。式(8)(9)において、 n_h は候補領域内に含まれる総ピクセル数、 x_i は中心を y とした候補領域内のピクセル座標を示している。 h は $(y - x_i)$ の最大値である。この二つのヒストグラムを利用して物体の追跡・位置同定を行っていく。

4.3 MeanShift 法を利用したトラッキングプロセス ここでのトラッキングとは、モデルを利用して後続するフ

レームから得られる候補領域の色ヒストグラムが最も類似すると考えられる領域を探索する処理のことを指す。この二つのヒストグラムの類似度もオブジェクト発見プロセスと同様に Bhattacharyya 係数により計算する。モデル Q_k^t および候補 P_k^t の色ヒストグラムを用いた場合、Bhattacharyya 係数は式(10)で表される。

$$d(Q_k^t, P_k^t) = \sum_{u=1}^m \sqrt{(p_k^t)_u (q_k^t)_u} \quad \dots\dots\dots (10)$$

次に、トラッキングの方法について説明する。本研究では、MeanShift 法とカルマンフィルタを用いたトラッキング手法を用いる。この手法は、対象物体の座標と速度ベクトルを含む、状態変数の予測、MeanShift 法による物体の位置推定、測定、更新を繰り返しながらモデルと類似した候補領域を探索するものである。MeanShift 法は図8に示すような、モデルと候補領域の間に発生する類似度の勾配を利用しながら、モデルとの類似度が最も高くなる頂点に向かってシフトを繰り返し、勾配が0となる部分を追跡対象物体の移動位置として確定する方法である。図8では、例として現在のフレームと20フレーム後の画像を比較したときの類似度の様子を示しているが、実際には毎フレーム同様の処理を行なう。物体が移動する場合、現在のフレームにおける物体候補領域の中心座標を $\hat{y}_1$ 、一つ前のフレームにおける物体領域の位置位置を $\hat{y}_0$ とすると、 $\hat{y}_1$ と $\hat{y}_0$ の間には、図8のように類似度空間上に勾配差が発生する。この勾配差を利用し、式(11)(12)から探索領域に色の類似度に応じた重みづけを行なう。 $\hat{y}_0$ からのシフト量が一定値以下まで減少するまで同一フレームにおける同一物体に対して数回の処理を繰り返すことで、追跡対象物体の位置 $\hat{y}_1$ を求める。 $\hat{y}_1$ は式(11)(12)により得られる。さらに、MeanShift 法による探索の初期値 $\hat{y}_0$ をカルマンフィルタによる予測位置とすることで、ある程度高速な移動物体への対応も可能である。

$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \quad \dots\dots\dots (11)$$

$$w_i = \sum_{u=1}^m \sqrt{\frac{(q_k^t)_u}{(p_k^t)_u}} \delta[b(x_i) - u] \quad \dots\dots\dots (12)$$

照明、日光など環境の影響による追跡対象物体の色変化への対策として、式(13)のように、前フレームにおけるモデル Q_k^{t-1} に、現在のフレームから得られる候補領域の特徴量 P_k^t の数%を加算していくモデル更新を行なう。急激なモデルの更新は物体本来の色情報を損失する原因となるため、 P_k^t の加算率を変更させることとする。トラッキングプロセスのみによる追跡では、オクルージョンなどの無い場合、同一物体に対する Bhattacharyya 係数の平均値 μ 、標準偏差 σ は時間の経過に応じて徐々に一定値に収束する様子が観測でき⁽¹⁴⁾、平均値と標準偏差を用いて、追跡時の同一物体における Bhattacharyya 係数の分布を表現するこ

図 8 MeanShift の概念

Fig. 8. Concept of MeanShift.

とができる。本研究では、平均値 μ 、標準偏差 σ を利用して、Bhattacharyya 係数の分布を式 (13) に示す正規分布により近似する。確率密度関数 $g(d)$ により現フレームでの物体位置での色情報のモデルとの類似度の妥当性を評価し、式 (14) に示すように、 $\beta\%(1\% \sim 5\%)$ を P_k^t の加算率としてモデルを更新する。

$$Q_k^t = (1 - \beta g(d))Q_k^{t-1} + \beta g(d)P_k^t \dots\dots\dots (13)$$

$$g(d) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d-\mu)^2}{2\sigma^2}\right) \dots\dots\dots (14)$$

また、物体のサイズ変化に対応しつつトラッキングを継続させるためには、物体領域のサイズを状況に応じて変更させる必要がある。1 フレーム前における領域のサイズを基準とした幾つかのサイズを用意し、これら領域から得られる特徴量とモデルの比較の結果、最も類似度が高い領域のサイズを選択しサイズ変化を行っている。

5. プロセス選択方法

5.1 両プロセスの特徴 前章までに述べた二つのプロセス、オブジェクト発見プロセスおよびトラッキングプロセスは、それぞれ長所と短所があり、ひとつのプロセスのみで複数物体の安定した追跡を実現することは困難である。そこで、そのプロセスの融合と効果的な切替の仕組みが必要である。

これまで述べたように、オブジェクト発見プロセスは新規物体を発見する機能があり、追跡アルゴリズムの基盤となるプロセスである。特徴としては、物体位置を安定して観測することが可能であるが、基本的に背景画像から物体ごとの抽出が困難な、複雑な環境ではうまく働かない。一方、トラッキングプロセスは、オブジェクト発見プロセスで発見された物体の存在情報に依存するものの、複雑な環境でもカラートラッキングが継続できるよう処理される。しかし、物体位置検出が不安定であることから、長時間の追跡では物体領域以外の色情報を含んだ状態でカラーモデルを更新することが多くなり、物体の本来の色情報が徐々に失われていき、追跡誤差が蓄積していく傾向がある。この両プロセスの特徴を生かしたプロセスの選択方法を検討する必要がある。

図 9 二つのプロセスの融合

Fig. 9. Fusion of two processes.

両プロセスの特徴を考慮すると、以下のような場合にオブジェクト発見プロセスとトラッキングプロセスの切替が行なわれることが妥当であると考えられる。

・オブジェクト発見プロセス トラッキングプロセス

オブジェクト発見プロセスにおいて、特徴空間内の物体を表すクラスタから外れた特徴ベクトルとして観測され、フレーム間の物体の対応づけに失敗したとき。この場合、物体が存在していることは確定しているので、前フレームでの物体の位置と大きさに基づき、スムーズにトラッキングプロセスへ移行できる。

・トラッキングプロセス オブジェクト発見プロセス

トラッキングプロセス実行中に、並行して動作しているオブジェクト発見プロセスで新規物体に相当するクラスタが生成され、両プロセスで物体位置と大きさが類似するとき。この場合、二つのプロセスでの処理結果を比較したうえでのプロセス選択が必要となる。

また、二つのプロセスによる追跡の持つ意味は図 9 に示すような図で表される。物体同士のすれ違いなどの際に物体領域がしばらくの間抽出できない場合、物体を表す色特徴が変動している可能性がある。その際にオブジェクト発見プロセスでの追跡が不可能である時間の前後に表れるクラスタを結びつける処理が、トラッキングプロセスで行なわれる。

5.2 プロセス選択方法 プロセスの選択で大きな問題となるのは、トラッキングプロセスからオブジェクト発見プロセスへの切替である。図 10 は、オブジェクト発見プロセスおよびトラッキングプロセスによる処理の関係を表したものである。同一フレームにおいて、両プロセスから互いに重なる追跡ウィンドウが得られたとき、その位置関係やサイズにより適するプロセスを選択する。なお、追跡ウィンドウとして、オブジェクト発見プロセスでは背景から分離され物体 ID が存在する物体を完全に含む最小の矩形領域、トラッキングプロセスでは追跡アルゴリズムで推定された物体領域を用いる。

図 10 に示すように、トラッキングプロセスにおける追跡ウィンドウとオブジェクト発見プロセスによる新規物体ウィンドウを比較したとき、ウィンドウ同士の関係は以下のようなになる。

図 10 両プロセスの関係

Fig. 10. Relationship between two processes.

図 11 プロセスの選択方法

Fig. 11. Selection of the process.

- ウィンドウ同士の重なりが全くないとき
- 各プロセスの出力が一对一に対応するとき
- どちらかのプロセスのウィンドウに、もう一方のプロセスでの複数の追跡ウィンドウが重なるとき

ウィンドウ同士の重なりがないときは、各プロセスの出力ウィンドウがトラッキングアルゴリズム全体の出力となる。また、各プロセスのウィンドウが一对一に対応するときには、オブジェクト発見プロセスによるウィンドウを全体の出力とする。

問題となるのは、図 11 に示すような、複数の物体が画像内で分離または結合する場合である。図 11 において、実線はオブジェクト発見プロセスによる追跡、点線はトラッキングプロセスによる追跡を示している。図 11 (a) は、Object1 として認識、追跡されている物体が分離した場合である。オブジェクト発見プロセスでは前フレームとの対応付けに失敗し Object1 に対するトラッキングプロセスが実行されるが、実際には Object2, Object3 が新規発見され追跡されることが望ましい。また、図 11 (b) は Object1 と Object2 がオブジェクト発見プロセスにより認識、追跡されている場合である。両物体の接近によりオブジェクト発見プロセスでの追跡は失敗し、各物体に対してトラッキングプロセスが起動する。その際、物体の接近した状態が継続することで、結合した二つの物体を新たなひとつの物体として誤

認識してしまう可能性があるが、それを棄却しトラッキングプロセスによる追跡のみを有効とする必要がある。

ここでは、両プロセスによる追跡結果が出力として得られる場合、互いの出力位置とウィンドウの比較・対応付けを改めて行うことで、この問題に対処する。まず、オブジェクト発見プロセスによる追跡ウィンドウを基準とし、画像内で重なる部分のあるトラッキングプロセスで追跡中の物体をプロセス切替の候補物体として選択する。例えば、図 11 (a) では、Object1 の分離後に新規発見された Object2 および Object3 が、トラッキングプロセスにより追跡継続中の Object1 の追跡ウィンドウと重なるため、Object1 がプロセス切替候補物体となる。図 11 (b) では、トラッキングプロセスにより追跡される Object1 と Object2 が切替候補物体となる。プロセス切替候補物体の追跡ウィンドウが、対応するオブジェクト発見プロセスによる追跡ウィンドウと画像内で重なる面積を評価することで、プロセスの切替を決定する。ここでは、トラッキングプロセスによる追跡ウィンドウがオブジェクト発見プロセスによるウィンドウ内部に一定量以上含まれる場合（80%以上）にはトラッキングプロセスによる処理結果を全体の出力とする。それ以外の場合にはトラッキングプロセスによる追跡結果に対応する物体は存在しないものとし、トラッキングプロセスを終了させる。図 11 (a) の例では、Object1 の追跡は終了し、新たに Object2 と Object3 の追跡がオブジェクト発見プロセスから開始される。図 11 (b) では、物体の結合した状態が新規物体として認定されても、トラッキングプロセスによる追跡が優先される。

6. 実験

6.1 物体追跡実験結果 アルゴリズムの有効性を示すために、動画像を用いた追跡実験の結果を示す。実験では、15 fps の 320 × 240 の RGB カラー画像を 140 フレーム使用した。オブジェクト発見プロセスとトラッキングプロセスにより二つの物体が追跡されている様子、併せて各物体の特徴空間におけるクラスタの変動の様子を図示している。

図 12 トラッキング結果

Fig. 12. Tracking results.

図 13 プロセス選択の結果

Fig. 13. Result of process selection.

図 12 では、物体追跡の様子を示している。図 12 (a) および図 12 (b) は、それぞれ物体 1 および物体 2 が新規物体として発見されたフレームである。この実験では、図 12 (c) から図 12 (f) までは、両物体ともトラッキングプロセスでの処理に移行して追跡されている。図 13 は、プロセスの選択の様子を示している。図 13 に示すように、物体 1 はオクルージョンの生じたフレーム 74 からフレーム 99 までトラッキングプロセスにより追跡されたが、その後オブジェクト発見プロセスにより追跡された。一方、物体 2 でも同様にオクルージョン発生時にトラッキングプロセスが開始したが、フレーム 100 以降の抽出物体の色特徴がオクルージョン前の色特徴と異なっており、同一物体であるとの判定はできなかったため、トラッキングプロセスによる処理を継続し、オブジェクト発見プロセスにおいてフレーム 125 までに発見された新規物体とトラッキングプロセスの結果を比較することで、プロセス間の対応づけを行ない、その後はオブジェクト発見プロセスによる追跡が行なわれた。

本実験結果より、複数対象を発見し追跡している様子が伺える。また、トラッキングプロセスに移行した後の追跡の様子とプロセスの切替えによる追跡の様子も示されている。なお、本実験における各パラメータは、クラスタリングの際の有効フレーム数と必要となる特徴ベクトル数 $s = 50$, $a = 25$, 類似度に関する閾値を $T = 0.8$, 学習係数 $\alpha = 0.1$ とした。また、トラッキングプロセスにおけるモデル Q_k^t 生成のための加算率は $\beta = 2\%$ とした。各パラメータは、あらかじめ取得したいいくつかの動画像を使用して追跡のシミュレーションを行なうことにより決定した。これらのパラメータはトラッキング性能や対象物体の色情報の変動への色モデルの追従性能を決定するものであり、本実験で用いた値付近では、その値にわずかな変動があっても追跡性能が著しく劣化することはなく、大まかなパラメータ設定でも十分動作する。

また、物体 1 と物体 2 の特徴ベクトルの変動の様子を図 14 および図 15 に示す。ここでは、多次元の特徴ベクトルを、可視化のために主成分分析して 3 次元で表している。特徴ベクトルとして、正規化 RGB カラーによる 2 次元の色空間をそれぞれ 32 分割した、計 1024 次元の特徴ベクトルで各物体領域の色情報を表したが、例として挙げた実験による追跡では、3 次元程度で色の特性の変動を失わずに表現することが可能であった。物体 2 に関しては、トラッキングプロセスが起動する前後で物体の特徴ベクトルは大きく変動し、異なるクラスタとして表されるが、プロセス間の対応付けの結果、両クラスタが同一物体であると認定され、追跡が継続された。

6.2 考察 ここでは、実験結果に関係する、本手法の処理時間、および追跡対象物体の個数、大きさや得られる色情報などについての考察を行なう。

前述したとおり、オブジェクト発見プロセスは背景差分

図 14 物体 1 の色情報

Fig. 14. Color information of object 1.

図 15 物体 2 の色情報

Fig. 15. Color information of object 2.

をベースとしているため、基本的に処理時間は大きくなく、トラッキングプロセスにおけるフレームごとの処理時間よりも小さい。一方、トラッキングプロセスにおいて主に使用される MeanShift 法による追跡は、計算コストが少ないことが知られている⁽¹⁰⁾。トラッキングプロセスにおける人物追跡数および、毎フレームでの MeanShift の繰り返しの回数が最も処理時間の増大に影響を与えと考えられる。実験結果で示した動画像において、トラッキングプロセス実行時の毎フレームでの MeanShift の繰り返し回数は、各物体に対して平均して 4 回程度である。文献 (10) においても、複数物体追跡の際、同様に平均 4 回程度の繰り返しを行ない、一般的な計算機を用いて 30 fps 以上の高速処理を実現しているとの報告もなされている。本手法は、複数の手順を含んでいるが、どちらのプロセスも実時間処理が可能であり、また追跡状態に応じてプロセスを適宜切り替えることで、処理時間を急激に増大させることなく、文献 (10) による手法を物体の自動検出から安定した追跡まで拡張しているものと位置付けられる。

また、実験結果では、二つの人物の追跡結果を示したが、オブジェクト発見プロセスおよびトラッキングプロセス双

方において、3 つ以上の複数物体への対応が可能であることは確認されており^{(15) (16)}、ソフトウェアを拡張することによって、融合した一連のトラッキングにおいても 3 つ以上の複数物体への対応は可能となる。さらに、オブジェクト発見プロセスおよびトラッキングプロセスそれぞれにおいて、オブジェクト領域の色情報を適宜更新しており、本実験結果で示すように、周囲の照明環境や物体の色構成に急激な変動が無いかぎりに対応可能である。ただし、色情報をベースとした手法であるため、影などの影響や得られる物体領域が小さいときなど、十分な色情報 (彩度, 明度) が獲得できない場合、また背景色と物体色が広範囲にわたって類似する場合などへの対処は困難である。

7. 結 論

本論文では、カメラ画像による空間認識のために必要な基本処理である複数物体トラッキングについての研究として、背景画像と抽出された物体領域の色情報のクラスタリングによるボトムアップな物体発見と追跡および物体の存在を仮定した状態からのトップダウンな物体領域の追跡の両者の組み合わせにより、動画像上で物体を追跡するアルゴリズムを検討した。さらに、実験結果により、その有効性を示した。

今後の課題としては、各プロセスについて、照明の変動などに対する頑健性を定量的に評価することが必要である。本実験では互いに異なる色で構成される物体の追跡の様子を示したが、全く同一の色構成からなる物体の場合には、うまく働かないことが予想される。そのような場合に対処するため、オブジェクト発見プロセスでの新規物体の認識の際には、色情報による分類に加えて位置情報を利用した学習を行なうなどの処理が必要であろう。トラッキングプロセスにおいても同一色の物体に対処できるよう改善することが必要である。また、カメラ単体だけではなく、他のカメラとの協調により広域でのシームレスな追跡のため、本研究での追跡アルゴリズムを利用して、オブジェクト発見プロセスにより得られた色情報に基づいたグローバルモデルを自動構築し、異なるカメラ間での対応づけのための検討を行なう。将来的には大量のカメラを配置した分散視覚による複数物体の自動追跡システムを構築する予定である。(平成 16 年 11 月 25 日受付, 平成 17 年 6 月 28 日再受付)

文 献

- (1) R.T. Collins, A.J. Lipton, H. Fujiyoshi, and T. Kanade: "Algorithms for Cooperative Multisensor Surveillance", *Proceedings of The IEEE*, Vol.89, No.10, pp.1456-1477 (2001-10)
- (2) S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi: "Traffic Monitoring and Accident Detection at Intersections", *IEEE TRANSACTION on ITS*, Vol.1 No.2, pp.108-118 (2000-6)
- (3) J-H. and Lee H. Hashimoto: "Intelligent Space - concept and contents", *Advanced Robotics*, Vol.16, No.3, pp.265-280 (2002-8)
- (4) Y. Wu and T.S. Huang: "Color Tracking by Transductive Learning", *Proc. of IEEE Conf. on CVPR*, Vol.I, pp.133-138

- (2000-6)
- (5) P. Perez, C. Hue, J. Vermaak, and M. Gangnet: "Color-Based Probabilistic Tracking", Proc. European Conf. on Computer Vision, pp.661-675 (2002-5)
 - (6) Y. Raja, S.J. McKenna, and S. Gong: "Colour Model Selection and Adaptation in Dynamic Scenes", Proc. European Conf. on Computer Vision, pp.460-475 (1998-6)
 - (7) T.-L. Liu and H.-T. Chen: "Real-Time Tracking Using Trust-Region Methods", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.26, No.3, pp.397-402 (2004-3)
 - (8) T. Nakamura and T. Ogasawara: "Toward a Visual Tracking System with On-Line Visual Learning Capability", Journal of Robotics Society of Japan, Vol.18, No.7, pp.1047-1054 (2000-10) (in Japanese)
中村恭之・小笠原司:「オンライン学習機能を持つ視覚追跡システムの提案」, 日本ロボット学誌, 18, 7, pp.1047-1054 (2000-10)
 - (9) N. Ukita: "Data Learning for Color Target Detection with Nearest Neighbor Classifier in the Position-Color Space", IPSJ SIG Technical Reports CVIM-145, pp.85-92 (2004-9) (in Japanese)
浮田宗伯:「位置・色調統合空間におけるターゲット色検出のためのサンプルデータ学習法」, 情報処理学会 研究報告 CVIM-145, pp.85-92 (2004-9)
 - (10) D. Comaniciu, V. Ramesh, and P. Meer: "Kernel-Based Object Tracking", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.25, No.5 pp.564-577 (2003-5)
 - (11) Y. Satoh, T. Okatani, and K. Deguchi: "A Color-based probabilistic Tracking by Using Graphical Models", Proc. of 2004 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp.631-636 (2004-9)
 - (12) J. Satake and T. Shakanaga: "Multiple Human Tracking on Image Sequence under Hierarchical Attention Control", The Transactions of the IEICE D-II, Vol.J86-D-II, No.8, pp.1212-1221 (2003-8) (in Japanese)
佐竹純二・尺長 健:「階層的注視点制御による動画像上での複数人物追跡」, 信学論誌 D-II, J86-D-II, 8, pp.1212-1221 (2003-8)
 - (13) S. Takatsuka, K. Morioka, N. Ando, O. Ono, and H. Hashimoto: "Robust Background Update Method for Moving Object Extraction Under Varying Environment", JSME Annual Conference on Robotics and Mechatronics 2003, p.1P1-2F-E1 (2003-5) (in Japanese)
高塚壮介・森岡一幸・安藤慶昭・小野 治・橋本秀紀:「移動物体分離のための環境変化に対してロバストな背景更新法」, 日本機械学会 ロボティクス・メカトロニクス講演会 2003, p.1P1-2F-E1 (2003-5)
 - (14) K. Morioka, Y. Kuroda, S. Takatsuka, and H. Hashimoto: "Seamless Object Tracking in Distributed Vision Sensor Network", SICE Annual Conference 2004, pp.1031-1036 (2004-8)
 - (15) K. Morioka and H. Hashimoto: "Intelligent Environment Based on Distributed Vision Sensors with On-Line Learning of Color Model", SICE System Integration Division Annual Conference 2003, pp.311-312 (2003-12) (in Japanese)
森岡一幸・橋本秀紀:「カラーモデルのオンライン学習機能を持つ分散視覚センサを用いた空間知能化」, 計測自動制御学会 システムインテグレーション部門 講演会 SI2003, pp.311-312 (2003-12)
 - (16) Y. Kuroda, K. Morioka, J.-H. Lee, H. Yashima, and H. Hashimoto: "Mobile Objects Tracking Using Modified Hybrid Algorithm", JSME Annual Conference on Robotics and Mechatronics 2004, p.2P1-L1-40 (2004-6) (in Japanese)
黒田陽一・森岡一幸・李 周浩・八嶋弘幸・橋本秀紀:「新ハイブリッドアルゴリズムによる複数物体のトラッキング」, 日本機械学会 ロボティクス・メカトロニクス講演会 2004, p.2P1-L1-40 (2004-6), IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.25, No.5 (2003-5)

森 岡 一 幸 (非会員) 1977 年。2000 年 3 月東京大学工学部

電気工学科卒業。2005 年 3 月同大学大学院工学系研究科電気工学専攻博士課程修了。同年 4 月東京理科大学理工学部電気電子情報工学科助手, 現在に至る。博士 (工学)。移動ロボット, ロボットビジョンの研究に従事。IEEE, 日本ロボット学会会員

黒 田 陽 一 (非会員) 1980 年生。2003 年 4 月東京大学生産技術研究所研究実習生。2004 年 3 月東京理科大学工学部経営工学科卒業。在学中, 画像処理, ロボットビジョンの研究に従事。

李 周 浩 (正員) 1970 年生。1999 年 3 月東京大学大学院工学系研究科電子情報工学専攻博士課程修了。同年 4 月同大学生産技術研究所研究員。2003 年 4 月東京理科大学工学部経営工学科助手。2004 年 4 月立命館大学情報理工学部情報コミュニケーション学科助教授, 現在に至る。博士 (工学)。ロボティクス, 空間知能化の研究に従事。IEEE, 日本ロボット学会, 日本機械学会会員。

橋 本 秀 紀 (正員) 1957 年生。1987 年 3 月東京大学大学院工学系研究科電気工学専攻博士課程修了。同年 4 月同大学生産技術研究所講師。1990 年同大学助教授, 現在に至る。1989-90 年 MIT 客員研究員。工学博士。知的制御システム, 制御工学, ロボティクス及び ITS の研究に従事。IEEE, 日本ロボット学会, 計測自動制御学会会員。