

# 最近傍識別による背景差分と色検出の統合 ——事例に基づく情報統合

加藤 丈和<sup>†</sup> 和田 俊和<sup>†</sup>

我々は現在までに、最近傍識別器のコンピュータビジョンへの応用として、画像の各画素の色ベクトルを識別することで、色ターゲットを検出する方法を提案した。本稿では、その拡張として、あらかじめ用意した背景画像の画素の色3次元(YUV)と、入力画像の対応する画素の色3次元(YUV)を合わせた6次元のベクトルを識別することにより、背景差分によって得られる効果と、色検出によって得られる効果を両方兼ね備えたターゲット検出システムを構築する。実験では、半透明物体と背景が変動する場合に対するターゲット検出実験を示し、その有効性を実証する。

## Integration between Background Subtraction and Color Detection Based on Nearest Neighbor Classifier: Instance Based Multimodal Information Integration

TAKEKAZU KATO<sup>†</sup> and TOSHIKAZU WADA<sup>†</sup>

We proposed a color-target detection that regarded each pixel of images as 3-d vector (YUV) and classified the pixel by nearest neighbor classifier. In this paper, we expand the color-target detection into target-detection integrating a color-detection and a background subtraction. It regards each pixel as 6-d vector that consists of 3-d vector in a pixel in input images and 3-d vector in a background image. Experimental results demonstrate the effectiveness of our method against semitransparent target and variable background.

### 1. はじめに

多くのコンピュータビジョンのアルゴリズムでは、様々な物理モデルや統計モデルを用いてシーンの解析を行う方法がとられてきた。このようなアプローチの欠点としては、モデルに関する事前知識を必要とする点、事前に与えたモデルで表現できない複雑なシーンや、例外的な現象に対処することが困難な点などがあげられる。このような問題に対し、我々は、物理モデルや統計モデルを用いず、事例によって事象を説明する Case-Based Reasoning (CBR)<sup>1)</sup> のアプローチに基づくコンピュータビジョンアルゴリズムの研究を行っている。

CBRの中核となりうる基本的な技術として、最近傍識別器<sup>4)</sup>があげられる。最近傍識別器は、トレーニングパターンをプロトタイプとして記憶し、入力パターンと最も近いプロトタイプが属するクラスによって対象を識別する手法であり、1) 十分なトレーニング

パターンが与えられれば、バイズ誤り確率の2倍以下の誤り確率を達成できる、2) 分布モデルなどの事前知識を必要としない、3) そのままで多クラスの識別が可能である、などの優れた性質を持っている。一方で、識別速度が遅い、メモリを大量に消費するなどの欠点も持っているため、いままで最近傍識別器がコンピュータビジョンの実用的なタスクに応用された例は少ない。

我々は、このような問題点を解決するために、最近傍識別器の高効率化、高速化<sup>5),8)</sup>を行うとともに、最近傍識別器のコンピュータビジョンへの応用<sup>6),7)</sup>についての研究を行ってきた。文献7)では、画像の各画素をその色ベクトルによって識別することで、精度の良い色ターゲット検出を実現した。また、文献6)では、画像の領域を二値画像として表現し、それを高速に識別する手法を提案した。

本稿では、画素単位の識別において、各画素の色ベクトルだけでなく、他の特徴量を組み合わせることによって、高機能なターゲット検出を実現する。具体的には、入力画像の画素の色と背景画像の画素の色の組合せによる、色検出と背景差分の効果を兼ね備えた

<sup>†</sup> 和歌山大学情報通信システム学科

Department of Computer and Communication Science,  
Wakayama University

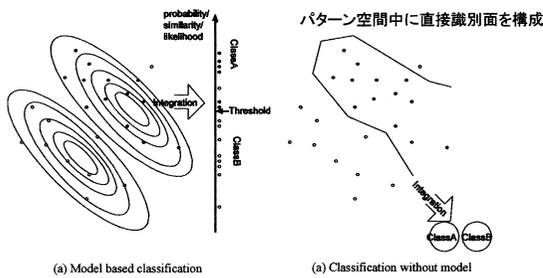


図 1 モデルに基づく識別とモデルを使わない識別  
Fig. 1 Model based and instance based classification.

ターゲット検出手法を提案する．また、ハッシュ表を用いた識別結果のキャッシュによる、実時間ターゲット検出のための最近傍識別の高速化手法を提案する．

## 2. 事例に基づくターゲット検出と情報統合

### 2.1 事例に基づくターゲット検出

画像中からターゲットを検出する問題は、コンピュータビジョンの中でも重要な課題と1つとしてあげられる．ターゲット検出問題は、画像中の各画素を、その画素値に基づいて、ターゲットと非ターゲットに識別する問題と見なすことができる．

このような問題に対し、従来のコンピュータビジョンやパターン認識の多くの研究では、図 1 (a) に示すように、ターゲットを物理的、統計的なモデルによって表現し、モデルとの一致度合いを「らしさ」を表す尺度として、確率や尤度、類似度などによって評価する手法がとられてきた．しかし、このような手法では、対象のモデルを事前知識として与える必要があり、対象とするシーンが複雑に変化する場合など、事前に与えたモデルでは表現するのが困難であるような場合や、例外的な現象に対して精度が良くないという問題点があった．

このような問題点を解決するためには、図 1 (b) に示すように、対象をモデルによって表現することなく、トレーニングデータを用いて、直接パターン空間中に識別面を構成すればよい．我々は、このような考えに基づき、最近傍識別器を用いた色ターゲット検出手法<sup>7)</sup>を提案した．最近傍識別器は、与えられたトレーニングパターンに対して、最大マージン基準を与える区分的識別面を持ち、図 1 (b) に示すような識別面を容易に得ることが可能である．提案手法では、画像の各画素の色ベクトルを識別することによって頑健かつ柔軟なターゲット検出を実現した．

我々は、このようにモデルの推定を行うことなく、トレーニングデータそのものを使って対象を表現する

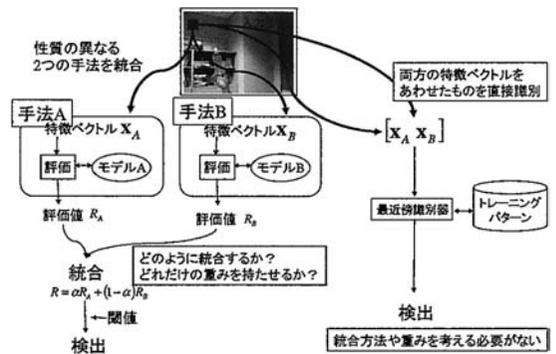


図 2 情報統合  
Fig. 2 Information integration.

アプローチを、事例に基づくコンピュータビジョンと呼んでいる．

### 2.2 事例に基づく情報統合

多くの要素が複雑な相互作用を持つ実世界に対して有効なコンピュータビジョンアルゴリズムの実現のためには、個々の要素を解析するアルゴリズムを検討するだけでは不十分であるという考えから、近年、多角的情報の統合<sup>9)</sup>のアプローチによる研究が数多くなされている．多角的情報の統合とは、複数の異なる性質を持つ特徴量や、複数のメディアを解析するモジュールを統合することによって、複雑な対象や事象を解析しようとするアプローチである．典型的な例としては、図 2 (a) に示すように、画像から複数の特徴量（特徴 A、特徴 B）を抽出し、これらを異なる手法（手法 A、手法 B）によって評価を行い、これらの評価値に重みを持たせて統合することによって、全体の評価値とするものである．

しかし、このような多角的情報の統合には、個々の手法によって得られた評価をどのように統合するか、どのような重みを持たせるかという問題が生じ、かえって問題を複雑化してしまうという問題点がある．これに対し、我々が提唱する事例に基づくコンピュータビジョンのアプローチでは、図 2 (b) に示すように、異なる性質を持つ特徴ベクトルを、個別に解析することなく、そのまま組み合わせさせて識別することによって、2つの手法を組み合わせたと同じ結果を得ることが可能である．このような特徴ベクトルを直接組み合わせさせて解析するアプローチを、我々は事例に基づく情報統合と呼ぶ．これによって、2つの手法をどのように統合するかを考える必要がなく、問題を単純化したままで解析することが可能となり、複雑な対象や事象の解析にも有効であると考えられる．

### 3. 最近傍識別器を用いたターゲット検出

本章では、事例に基づく情報統合のアプローチによって、すでに提案している最近傍識別器を用いた色ターゲット検出<sup>7)</sup>を拡張し、色検出と背景差分を統合する方法と色情報を統合する方法について述べる。

#### 3.1 色ターゲット検出

文献 7) では、画像の各画素を YUV の 3 次元ベクトルと見なし、3 次元パターン空間内での識別を行うことで色ターゲット検出を実現した。この手法では、ターゲットと非ターゲットのトレーニングデータをユーザがインタラクティブに指示することによって、カラーLED など、従来安定な検出が困難であったシーンに対して、安定な検出を実現した。

背景差分の問題は、あらかじめ与えた背景画像と入力画像との画素値の差を用いて背景領域かターゲット領域かを決定する問題ととらえることができ、色ターゲット検出の問題と同様に考えることができる。濃淡画像に対する背景差分では、単純に濃度の差分値に対する閾値の決定問題と見なすことができるが、カラー画像に対する背景差分では、色の違いをどのように評価するかということが重要な問題となる。

従来では、色を表すベクトル間にユークリッド距離などの距離を定義し、その距離に対して閾値を設定する方法や、色ベクトルの各要素に閾値を設定する方法、色ベクトル空間での差分値の分布に対して統計的なモデルをあてはめ、モデルとの類似度を評価する方法などが用いられてきた。このような手法では、どのような色空間を用いるか、また、色ベクトルの各要素をどのように評価するかが重要な問題となっていた。

このような問題に対して我々のアプローチでは、単純に色ベクトル間の差分ベクトルの識別問題と考え、差分ベクトルを直接識別することで、色空間や色ベクトルの評価法に影響されにくい安定な背景差分を実現することができる。

たとえば、画素の色を YUV 色空間で表現する場合、入力画像のある画素の色を  $Y_f, U_f, V_f$ 、背景画像の対応する画素の色を  $Y_b, U_b, V_b$  としたとき、色ベクトル間の差  $c_{bg}$  を次式のように与える。

$$c_{bg} = [|Y_f - Y_b| |U_f - U_b| |V_f - V_b|]^T \quad (1)$$

このようにして得られた差分ベクトル  $c_{bg}$  を、3 次元パターン空間で最近傍識別によって識別することで、背景領域かターゲット領域かを決定することができる。

#### 3.2 色検出と背景差分の統合

色ターゲット検出には、背景にターゲットと似た色が存在する場合に検出が不安定になるという問題があ

る。そこで本研究では、2.2 節で述べた、事例に基づく情報統合のコンセプトに従い、入力画像の色ベクトルに、他の特徴ベクトルを組み合わせることで識別することによって、入力画像の色だけでは検出が困難な画素に対して安定な検出を実現する。

色検出以外によく用いられるターゲット検出手法として背景差分がある。背景差分は、あらかじめ用意した背景画像と入力画像を画素ごとに比較して、その違いによって背景画像にないターゲットを検出する手法である。つまり、画像全体に対して同じ色モデル（学習データ）と比較する色ターゲット検出に対して、背景差分は入力画像の画素ごとに異なるモデル（背景画像の対応画素）との比較を行う手法であるといえ換えることができ、色ターゲット検出とは性質の異なる手法であるといえる。

具体的にいえば、背景差分では、背景部分にターゲットと似た色が存在しても、ターゲットが同じ場所に重ならない限り影響されない。一方で、背景差分には、半透明物体のようにターゲット色が背景に影響されて変化する場や、物体の影のように逆にターゲットによって背景が変化する場などに安定な検出が困難であるという問題がある。本研究では、このように相補的な手法である、色ターゲット検出と背景差分を統合したターゲット検出手法を提案する。

このように異なる 2 つの検出手法を組み合わせる方法として、それぞれの結果の論理和や論理積をとる方法や、それぞれの手法の評価値を重み付けをして足し合わせる方法などがあるが、本研究では事例に基づく情報統合のコンセプトに従い、それぞれの手法が用いる特徴ベクトルをそのまま組み合わせるという手法を用いる。

背景差分で用いる特徴ベクトルは、入力画像の各画素の色ベクトルと、背景画像の対応する画素の色ベクトルとの差分である。これを色検出で用いる入力画像の各画素の色ベクトルと組み合わせることで識別を行うが、実際には、入力画像の色ベクトルと、背景差分の色ベクトルがあればその差分は計算可能であるので、これらを組み合わせた特徴ベクトルを構成し、識別を行う。

画素の色を YUV 色空間で表現し、入力画像のある画素の色を  $Y_f, U_f, V_f$ 、背景画像の対応する画素の色を  $Y_b, U_b, V_b$  としたとき、次式のような 6 次元ベクトル  $c_{6d}$  を構成する。

$$c_{6d} = [Y_f \ U_f \ V_f \ Y_b \ U_b \ V_b]^T \quad (2)$$

このようにして得られた  $c_{6d}$  を 6 次元パターン空間で最近傍識別によって識別することで、色検出と背景差分を統合したターゲット検出を実現する（図 3）。

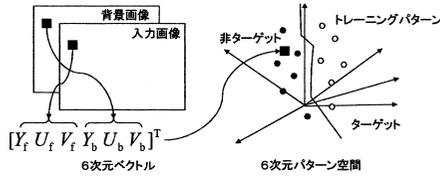


図 3 6次元ベクトルと6次元パターン空間

Fig. 3 Six dimensional vector and pattern space.

#### 4. ターゲット検出のための最近傍識別器の高速化

画像の各画素の識別によるターゲット検出では、画像中のすべての画素を識別する必要があるため、非常に高速な識別が必要となる。たとえば  $640 \times 480$  のサイズの画像からターゲット検出を行うためには、約 30 万回の識別を行う必要があり、これをビデオレート (30 fps) で動作させるためには、1 個のパターンに対して約  $0.1 \mu\text{s}$  以内に識別が終了しなければならない。また、結果を見ながらトレーニングデータを追加できるインタラクティブなシステムの構築のためには、学習時間も高速である必要がある。

我々は、高速な最近傍識別アルゴリズムである K-D Decision tree (KDDT)<sup>5)</sup> を提案している。この手法は、既存の最近傍探索アルゴリズムの中で最も高速なアルゴリズムの 1 つである、Approximate Nearest Neighbor (ANN)<sup>2)</sup> や、識別結果に影響を与えない学習データをあらかじめ除去しておく Condensing<sup>3)</sup> の方法を用いるのに対して、最大で約 400 倍の高速化を実現した手法であるが、この手法でも 1 個のパターンに対して、3 次元で  $0.9 \mu\text{s}$ 、6 次元で  $14.1 \mu\text{s}$  の識別時間が必要であり十分な識別速度とはいえない。

また、KDDT は、次元数やサンプル数が増加すると学習に時間がかかり、6 次元では数十秒から数分の時間がかかるという問題点もある。

また、色ターゲット検出<sup>7)</sup> では、Look Up Table (LUT) を用いた高速化により、ビデオレートの色ターゲット検出を実現している。しかし、この手法では、色空間全体のクラス情報を LUT 上に記録するため、6 次元の識別に用いるには、各次元が 128 段階の場合  $256^6 = 4$  テラバイトのメモリを必要とし、現実的ではない。また、LUT の更新アルゴリズムより、パターン間の距離尺度が市街地距離に限定され、ユークリッド距離などの他の距離定義に適用することができないという問題点もある。

##### 4.1 キャッシュを用いた識別の高速化

本研究で扱う特徴ベクトルは画素値であり離散値を

とるという点と、実際の画像には、画素値の偏りがあり、同じ色が繰り返し出現することが多いという点に着目し、キャッシュを用いた高速化を行う。つまり、一度識別を行った識別結果をメモリ上に記憶しておき、次に同じ色が出現した場合には、記憶した結果を用いて識別を行う。

キャッシュアルゴリズムには、入力パターンをキーとしたハッシュ表を用いた。ハッシュ表はキーから計算されるハッシュ値と呼ばれる整数を使ってテーブルを検索する手法であり、ハッシュ値の衝突がなければ、入力パターンの種類やデータ数に依存せず  $O(1)$  の検索が可能であり、使用メモリもテーブルのサイズとテーブルの各エントリのサイズをあらかじめ決めておくことで、容易にコントロール可能であることから今回の目的に合致していると判断した。

ハッシュ表において、高速な検索を実現するには、衝突しにくいハッシュ値を使うことが重要である。そこで、入力パターンに対応するハッシュ値は、実際の画像では、似た色が連続して出現することが多いという画像の特性を考慮し、次式のように計算した。

$$\begin{aligned} \text{hash}(Y_f, U_f, V_f, Y_b, U_b, V_b) = & (Y_f \bmod \hat{Y}_f) * \hat{U}_f * \hat{V}_f * \hat{Y}_b * \hat{U}_b * \hat{V}_b \\ & + (U_f \bmod \hat{U}_f) * \hat{V}_f * \hat{Y}_b * \hat{U}_b * \hat{V}_b \\ & + (V_f \bmod \hat{V}_f) * \hat{Y}_b * \hat{U}_b * \hat{V}_b \\ & + (Y_b \bmod \hat{Y}_b) * \hat{U}_b * \hat{V}_b \\ & + (U_b \bmod \hat{U}_b) * \hat{V}_b \\ & + (V_b \bmod \hat{V}_b) \end{aligned} \quad (3)$$

ただし、 $Y_f, U_f, V_f$  は入力画像の画素の色、 $Y_b, U_b, V_b$  は背景画像の画素の色である。また、 $\hat{Y}_f, \hat{U}_f, \hat{V}_f, \hat{Y}_b, \hat{U}_b, \hat{V}_b$  は定数であり、 $\bmod$  は剰余演算を表す。

式 (3) は、入力ベクトルの各要素の剰余を組み合わせたものであり、似ているが異なる色に対して異なるハッシュ値が得られるようにしている。こうすることによって、似た色が多く出現する場合に、ハッシュ値の衝突を避けることができ高速な検索が可能になる。

また、この計算は、各画素値の法が 2 のべき乗のときには、各画素値に対する論理積、論理和、ビットシフトのビット演算のみで計算することが可能であり、高速に計算することができる。なお、同じハッシュ値を持つ結果に対しては、ハッシュ値の再計算は行わず、ハッシュ表の各エントリをリストで管理する。

図 4 にアルゴリズムの概要を示す。まず、入力画像の各画素に対応するパターンについてハッシュ表を検索する。ハッシュ表に同じパターンの結果が存在しなければ、最近傍識別によってターゲットのクラスを決

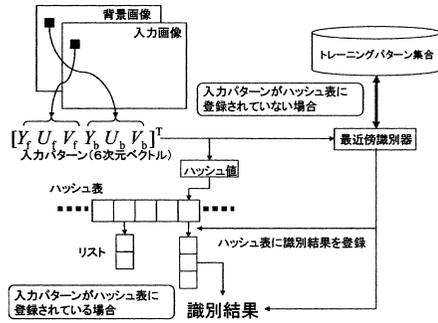


図 4 ハッシュ表による識別結果のキャッシュ

Fig. 4 An algorithm for caching classification results with a hash table.

定し、その結果をハッシュ表のエントリのリストの先頭に挿入する。また、そのパターンの識別結果がハッシュ表に存在する場合は、それを識別結果として用いる。このとき、検索したパターンをリストの先頭に移動することで、同じパターンを連続して検索したときに高速に検索可能にする。

## 5. 実験結果

以下の 3 種類の検出アルゴリズムを用いて文献 7) と同様のターゲット検出システムを構築した。

- (1) 入力画像の YUV を 3 次元ベクトルとして識別 (色検出)
- (2) 入力画像と背景画像の YUV それぞれの差分値を 3 次元ベクトルとして識別 (背景差分)
- (3) 入力画像の YUV 背景画像の YUV を合わせた 6 次元ベクトルとして識別 (統合手法)

本システムではユーザが画像上の領域を指定して、ターゲット 1 からターゲット 255、および非ターゲットのクラスを対応付けることで、インタラクティブに学習を行うことができる。

最近傍識別アルゴリズムには Arya らが提案した Approximate Nearest Neighbor (ANN<sup>2)</sup>) を用いた。ANN では許容誤差を設定できるが、ここでは許容誤差は 0 (誤差なし) とした。実験では、キャッシュを用いる提案手法と、キャッシュを用いない通常の ANN で識別を行った場合とで検出速度の比較を行った。

また、手法 (1), (2) では Y, U, V それぞれを 128 段階に量子化した YUV 色空間を用い、手法 (3) では 64 段階に量子化した YUV 色空間を用いた。手法 (3) の場合に 64 段階の量子化を行った理由は、最近傍識別ではパターン空間が広がると、より多くのトレーニングパターンが必要になり、インタラクティブにトレーニングパターンを指示する場合に時間がかか

るためである。なお、式 (3) のハッシュ値の計算に用いる定数  $\hat{Y}_f, \hat{U}_f, \hat{V}_f, \hat{Y}_b, \hat{U}_b, \hat{V}_b$  はすべて 32 とした。

使用機材は、計算機: Intel Xeon 2.4GHz Dual CPU, カメラ: SONY DFW VL500 を用い、入力画像は  $640 \times 480$  の YUV カラー画像を用いた。

### 5.1 半透明物体の検出

まず、半透明物体の検出結果を図 5 に示す。この実験では、半透明物体であるペットボトルをターゲットとし、これを画面全体をまんべんなく移動させながら撮影した 300 フレームの画像を使って、できるだけペットボトルだけが検出されるようにトレーニングを行い、トレーニングに用いていない画像シーケンスに対して検出を行った。

色検出 (b) では、ターゲットの部分で背景色が透けて見えているため、背景の色と似た色となるため、背景との切り分けができていないことが分かる。また、キャップの白い部分についても、背景に非常に明るい部分があったため、切り分けが困難となっている。背景差分 (c) では、キャップや手などの不透明な部分に関しては、検出ができていますが、半透明な部分では、背景色との違いが小さいため、検出が困難であることが分かる。

これに対して統合手法 (d) では、ほぼ完全に背景色が透けている部分で若干の抜けはあるものの、安定に検出できていることが分かる。また、背景差分では、手の部分もターゲットとして検出しているが、統合手法では手の部分は非ターゲットとして切り分けができていくことが分かる。これは、統合手法は色検出の機能も兼ね備えているため、ターゲットとして検出したい部分だけをターゲットとしてトレーニングし、それ以外を非ターゲットとしてトレーニングすることで、このような画像の変化があるが、非ターゲットであるような領域を切り分けることができる。

### 5.2 カラーディスプレイを背景とした検出

次に CRT の前にターゲットを置いて検出した場合の結果を図 6 に示す。この場合はターゲットは青色の本とし、これを薄い青色を表示したディスプレイの前で本の傾きを変えながら撮影した 300 フレームの画像を用いて、半透明物体の場合と同様に本だけが検出されるようにトレーニングを行い、トレーニングに用いなかったシーケンスに対して検出を行ったときの結果である。

色検出 (b) の結果では、影になって暗い部分で検出が失敗していることが分かる。これは、ディスプレイの枠や後ろの電源を切ったディスプレイの色が暗

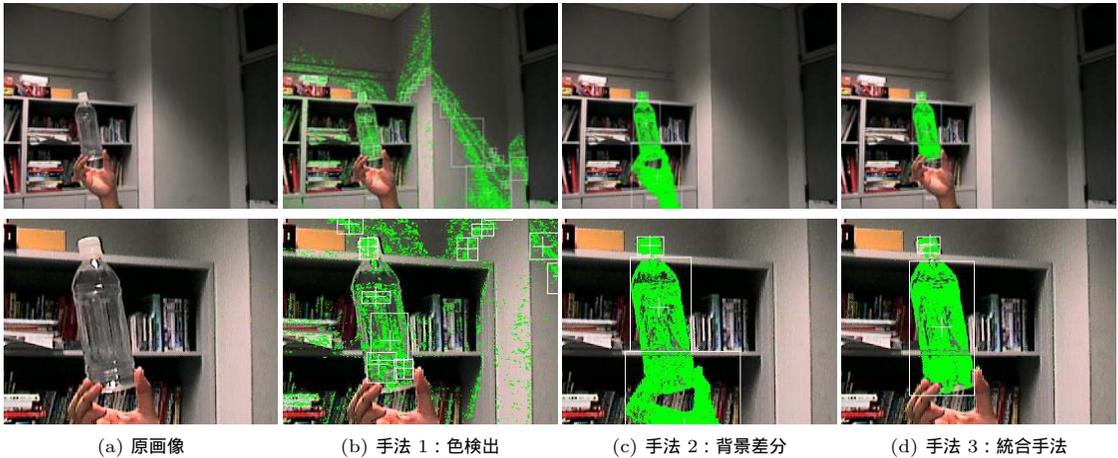


図 5 半透明物体の検出結果 (上段: 全画面表示, 下段: 拡大表示)

Fig. 5 Detection results of a transparent object (upper: full-screen, lower: enlarged image).

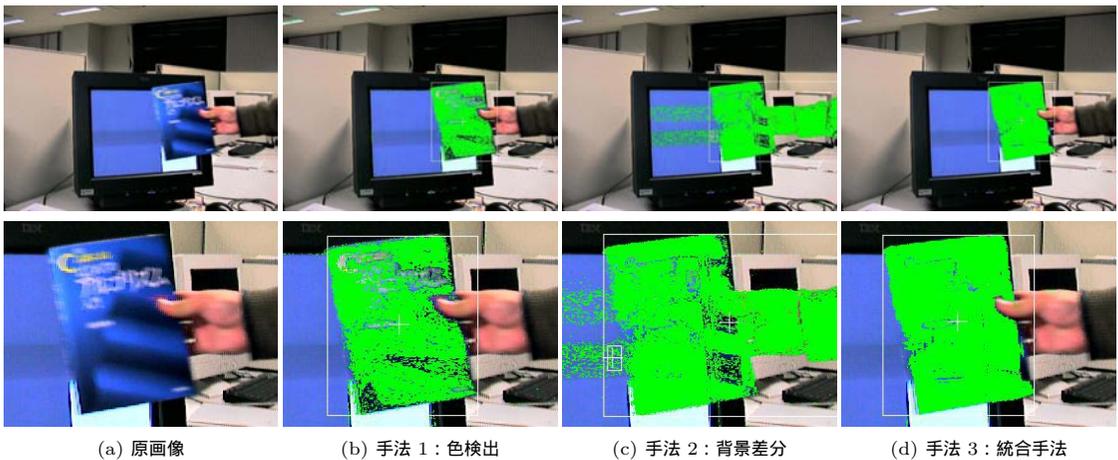


図 6 背景にディスプレイがあるシーンの検出結果 (上段: 全画面表示, 下段: 拡大表示)

Fig. 6 Detection results of an object before a display (upper: full-screen, lower: enlarged image).

いため、ターゲット中の暗い部分で検出に失敗したためである。また、背景差分 (c) の結果では、ターゲット領域はほぼ検出できているが、ディスプレイの部分に誤検出が存在していることが分かる。これは、ディスプレイの同期周波数とカメラのシャッタースピードとの関係によって、フリッカが生じディスプレイ上に帯状に暗い部分ができたためである。これらの結果に対し統合手法 (d) では、ほぼ完全にターゲットの部分だけを検出できていることが分かる。

### 5.3 検出速度の評価

次に、検出速度についての評価を行った結果を図 7 に示す。これは、図 6 の結果を含むシーケンスの検出を行ったときの結果であり、1 フレームごとのキャッシュのヒット率を (a) に、1 フレームあたりの検出時

間を (b) に、キャッシュサイズを (c) に示す。

キャッシュのヒット率 (a) の結果より、フレームの初めの部分では、キャッシュのヒット率が低く、その後ヒット率が上昇し、100 フレームほどでほぼ 1 に近くなっていることが分かる。また 100 フレームから 200 フレームのあたりでヒット率が低下しているが、これはこのフレームで入力画像に大きな変化があったためである。また、シーケンス全体を通してヒット率が最悪でも 0.99 程度であることから、同じ画素値を持つ画素が多く出現していることが分かり、キャッシュが非常に有効に働いていることが分かる。

また、検出時間 (b) の結果より、ヒット率が高い部分では高速であり、ヒット率が低くなると検出に時間がかかっていることが分かる。この結果より、キャッ

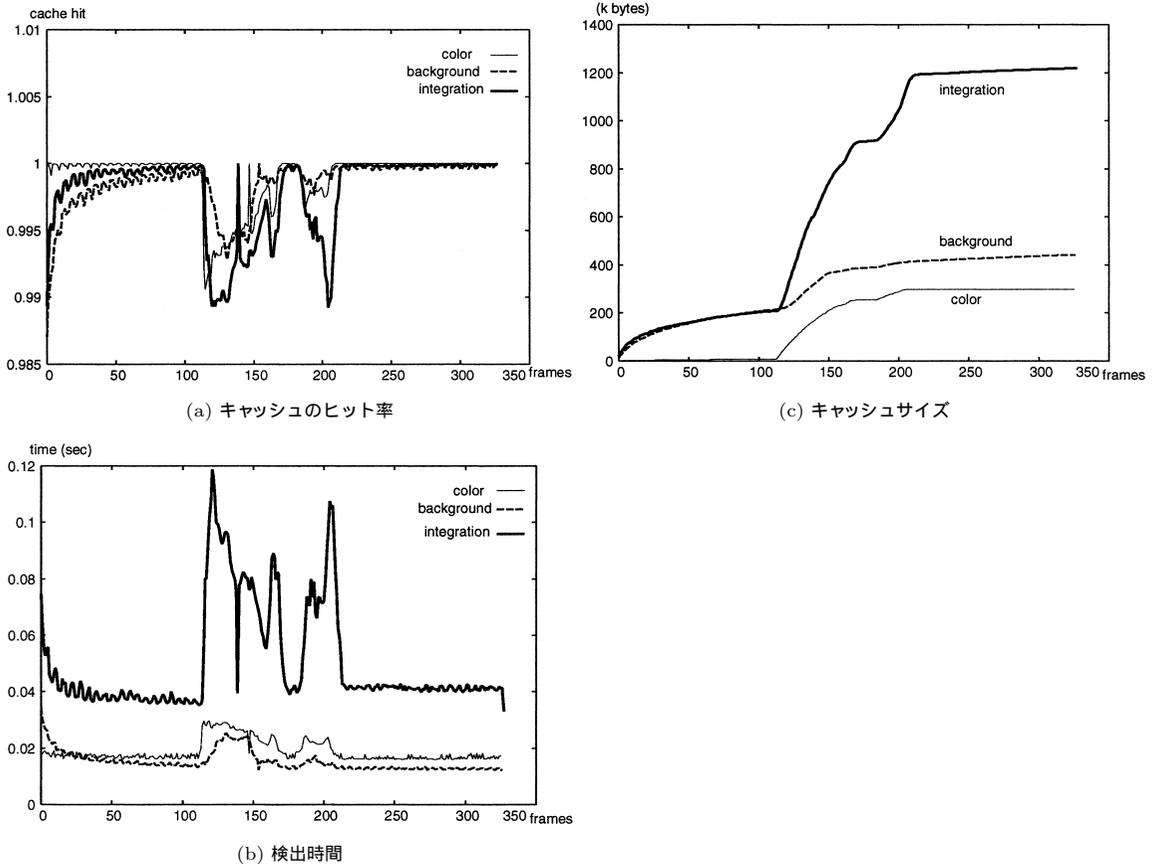


図 7 検出時間とキャッシュの状況  
 Fig. 7 Detection time and cache size.

ッシュが有効に働いている部分では高速な検出が実現できているといえる。また、6次元ベクトルを用いる統合手法では、特にその差が顕著であり、キャッシュによる高速化の効果が高いことが分かる。

なお、キャッシュを用いない通常のANNで識別を行った場合の処理時間は、1フレームあたりの平均で、3次元ベクトルを用いる場合で0.78秒、6次元ベクトルを用いる場合で1.23秒であり、10倍から30倍の高速化が実現できていることが確認できた。

次に(c)のキャッシュサイズについて、これはハッシュ表に登録されているエントリの数と、1エントリに必要なバイト数の積を求めたものである。この結果より、6次元ベクトルを用いる場合でも、最終的に1.2MB程度のメモリ使用量であり、メモリの使用効率も良好であることが分かる。

## 6. まとめ

本研究では、最近傍識別器を用いた高機能なターゲット検出手法として、文献7)で提案した色検出を

拡張し、背景情報を統合した手法を提案した。また、6次元の特徴ベクトルを用いた最近傍識別を高速に実行するために、ハッシュ表によるキャッシュを用いた高速化手法を提案した。

実験では、色検出と背景情報を統合することにより、半透明物体や背景が変動する場合など、単純な色ターゲット検出や、背景差分を個別に適用した場合には検出が困難なシーンでのターゲット検出を実現できていることを示した。また、検出速度に関しては6次元ベクトルを用いる場合でも1フレームあたり平均で約0.05秒、最悪でも0.12秒という、キャッシュを用いない場合の10倍から30倍の高速化を実現し、ビデオレートには至らないものの、実時間アプリケーションに適用するのに十分な検出速度が得られていることを確認した。

本手法の問題点としては、キャッシュのヒット率がほぼ100%に近い場合でも、文献7)のLUTを用いた手法に比べると検出に時間がかかっていることがあげられる。これは、キャッシュのメモリ使用量は少ない

ものの、識別結果の登録時に動的にエンタリ用のメモリの確保を行っているために、メモリの使用領域が広範囲に分散し、CPUのメモリキャッシュのヒット率が下がっているためであると予想される。このような問題に対しては、キャッシュエンタリに用いるメモリ領域をあらかじめ連続領域として確保しておき、その領域からエンタリ用の領域を確保することによって解決できると考えられる。

また、6次元ベクトルを用いた場合に、キャッシュを外れたときの速度低下が顕著であり、最近傍識別自体の高速化が必要である。我々は、最近傍識別器の高速化アルゴリズムとしてKDDT<sup>5)</sup>を提案している。しかし、このアルゴリズムは識別は高速であるが、学習に6次元では数十秒から数分かかり、また新たなトレーニングパターンが与えられると、そのつど再学習が必要であることから、本稿で述べるようなインタラクティブなシステムには適していない。

また、提案手法では、色検出の機能を持ち合わせているため、背景が多少変動しても、ターゲットを検出可能であるが、背景が大きく変動するような場合には、背景差分の影響が大きくなり、検出に失敗する可能性がある。このような場合では、あらかじめ背景画像を複数枚撮影しておき、それらをトレーニングデータに追加しておくことによって対応可能であると考えられる。

今後の課題としては、このような背景が大きく変動する場合への対応や、KDDTに対して学習時間の高速化、トレーニングの追加に対するインクリメンタルな学習の実現などの改善を行うことにより、KDDTを本手法に適用することで、さらなる高速化を実現することがあげられる。

謝辞 本研究の一部は、文部科学省科学研究費補助金若手研究(B)16700143の補助を受けている。

### 参 考 文 献

- 1) Aha, D.W.: Case Based Learning Algorithms, *Case based Reasoning Workshop*, pp.147-158 (1991).
- 2) Arya, S., Mount, D.M., Netanyahu, N., Silverman, R. and Wu, A.Y.: An optimal algorithm for approximate nearest neighbor searching in fixed dimensions, *5th ACM-SIAM Symposium Discrete Algorithms*, pp.573-582 (1994).
- 3) Bhattacharya, K., Poulsen, R. and Toussaint, G.: Application of Proximity Graphs to Editing Nearest Neighbor Decision Rule, *International Symposium on Information Theory* (1981).

- 4) Cover, B.T. and Hart, P.: Nearest neighbor classification, *IEEE Trans. Information Theory*, Vol.IT-13, No.1, pp.21-27 (1967).
- 5) Shibata, T., Kato, T. and Wada, T.: K-d Decision Tree: An Accelerated and Memory Efficient Nearest Neighbor Classifier, *ICDM'03*, pp.641-644 (2003).
- 6) 武本浩二, 加藤文和, 和田俊和: 画像の4分木表現に対する最近傍識別, 信学技報 PRMU, Vol.103, No.390, pp.1-6 (2003).
- 7) 和田俊和: 最近傍識別器を用いた色ターゲット検出 —「らしさ」に基づかない識別とコンピュータビジョンへの応用, 情報処理学会研究報告 CVIM134-3, No.134, pp.17-24 (2002).
- 8) 和田俊和, 加藤文和: 近接性グラフに基づく効率的 Condensing の理論, 信学技報 PRMU, Vol.103, No.96, pp.13-18 (2003).
- 9) 松山隆司: 画像理解の統合のための多角的情報の統合, 第19回画像工学コンファレンス, pp.97-102 (1988).

(平成16年3月29日受付)

(平成16年9月11日採録)

(担当編集委員 中原 智治)



加藤 文和

1974年生。1997年岡山大学工学部情報工学科卒業。2001年同大学院博士課程修了。2001年4月から2002年12月まで産業技術総合研究所。2003年1月より和歌山大学に勤務。コンピュータビジョン、パターン認識等の研究に従事。博士(工学)。電子情報通信学会、IEEE各会員。



和田 俊和 (正会員)

1987年東京工業大学大学院修士課程修了。1990年同大学院博士課程修了。同年岡山大学工学部助手。1994年同大学院自然科学研究科助手。1995年同大学工学部講師。1997年京都大学大学院工学研究科助教授。1998年同大学院情報科学研究科助教授。2002年から和歌山大学システム工学部教授。工学博士。画像理解、パターン認識の研究に従事。1995年David Marr賞、1997年情報処理学会山下記念研究賞、1999年電子情報通信学会論文賞各受賞。人工知能学会、電子情報通信学会各会員。