最近傍識別器を用いた色ターゲット検出

和田俊和†

カラーマーカの検出や肌色検出などの色ターゲット検出は，ターゲットと非ターゲットの色の識別問題と見なすことができ，ターゲットを構成する画素とそれ以外を識別する手法が研究されてきた．この種の研究では，ほとんどの場合，まず多次元の特徴ベクトルから確率や類似度など「ターゲットらしさ」を表す 1 次元の量を求め，その大小に基づく識別が行われてきた．本稿ではこういったパターンの持つ多様性を 1 つのスカラー値に縮約する「統合演算」を先に行っていたことが識別の精度や柔軟性を阻害してきたという考察に基づき，統合演算を必要としない最近傍識別器を用いた色ターゲット検出機構を提案する．このシステムは，ターゲットと非ターゲットの色を人間の教示からインタラクティブに学習し，ビデオレートでターゲット検出を行うものである．実験では，肌色とカラー LED の検出実験を示し，その有効性を実証する．

Color-target Detection Based on Nearest Neighbor Classifier

Toshikazu Wada†

Color target detection tasks can be regarded as a color classification problem. Most of the classification methods first integrate features into a single value representing the probability or similarity, and the classification is done by comparing or thresholding these values. This integration, i.e. mapping from multi-dimensional pattern space to 1-dimensional space removes the manifold of pattern distributions in the original space, and hence, the exceptional patterns can be missclassified. In this report, we propose a novel method for color target detection based on nearest neighbor classifier, which doesn't require the feature integration and can be trained by specifying target and nontarget color examples. Since the color classifier is implemented as LUT, our system can detect targets in video-rate. Experimental results demonstrate the effectiveness of our method.

1. はじめに

コンピュータビジョンにおける色の研究は，次の 2 つに大別される．

- 色恒常性の実現や，物体の材質の推定などを目的とした物理学/心理物理学的な研究
- 対象の検出・識別などを目的としたパターン認識的な研究

前者は物理量/心理物理量としての色そのものに関する研究であるが，カラーマーカの検出や肌色検出などの問題に，これらの手法を適用しても正しく動作しないことが多い．この原因としては，1）ターゲットが多色であり単一の物体色モデルと整合しない，2）正事例のみからターゲットのモデルを構成するため非ターゲットを誤検出してしまう，3）照明条件や物体の材質，カメラの特性に関する前提条件が成立しない，などのケースがあることがあげられる．

一方，後者の研究は色を物理量/心理物理量ではなく単にパターンと見なして，色に基づくターゲットの検出や識別問題などを取り扱う立場である．この立場の研究では

- 画像内にターゲット以外の未知の色を持つ物体が侵入することによる誤検出，
- 照明の変化などによって対象の色が変化することによる未検出，

などの問題は考慮されていないことが多い．しかし，上記の誤検出や未検出が起きた場合に，その色を再教示することができるため，このことが致命的な欠点になるケースは少ない．

後者の研究では「ターゲット色らしさ」を表す確率や尤度，類似度などの単一のスカラー値に縮約する統合演算を行い，そのスカラー値の大小により識別を行うケースがほとんどである．この場合，**図 1** (a) に示すように，多次元ベクトルからスカラー値への写像を行う際に，確率分布モデルや類似度などの人為的なモ

† 和歌山大学システム工学部
Faculty of Systems Engineering, Wakayama University

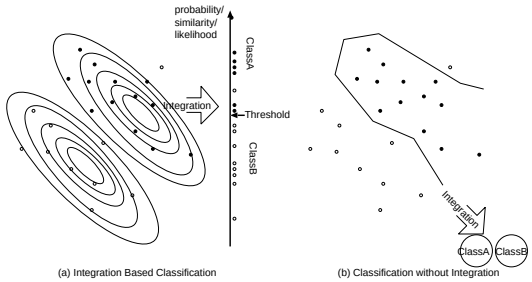

図 1 情報統合を行う識別と，統合を行わない識別
Fig. 1 Recognitions with integration and without integration.

デルを導入することになるが，実際のパターンの分布と合致したモデルを用いない限り，完全な 2 クラスの分離に使える精度の高いスカラー値が得られない．この結果，互いに区別しにくい例外的な色を誤識別するケースが発生しうる．この問題を回避するには同図 (b) のように識別を行う前に「らしさ」を表すスカラー値への縮約を行わずに，色空間内で単に識別面だけを設定するようにすればよい．

このように特定のモデルを用いずに識別面を求める問題は，パターン識別理論の中で取り扱われてきた．1967 年に Cover らが提案した最近傍識別器[1] は確率や尤度などの計算をともなわない識別器である☆．この識別器は，基本的に全トレーニングパターンとその所属クラスを記憶しておき，入力パターンと最も近いトレーニングパターンを見つけ，その所属クラスに入力パターンを分類するという単純なものである．

パターン間の距離を計測していることから，他の識別法と同様にスカラー値への縮約を行っていると思われがちであるが，この距離はターゲット/非ターゲットの各クラスへの帰属度とは無関係である．すなわち，他の識別法ではクラス（パターンの集合）と入力パターン（要素）との類似度あるいは相違度を求めているのに対し，最近傍識別器ではパターン（要素）間の比較を行っているだけであり，入力に最も近いパターン（最近傍パターン）と入力パターンとの距離の大小にかかわりなく，入力パターンはつねに最近傍パターンの所属するクラスに分類される．

最近傍識別器は優れた識別性能を持っており，特に大量のトレーニングパターンが与えられる問題では，**Bayes Error** の 2 倍未満の誤り確率が達成できる[1] という特長がある．一方，問題点としては，

- トレーニングパターンが追加されていくと，大量のメモリを消費する，
- 識別速度が遅い，
- あるクラスのトレーニングパターンが他クラスの分布の中に入り込んでいる場合，誤識別率が増加する，

などがあげられる．

最近傍識別器が提案された当時は，計算機の記憶容量が少なかったことや，実用的に用いることができる最近傍探索手法がなかったため，最近傍識別器が実際的な問題に用いられるケースは少なく，主にパターン認識理論の研究対象としてしか扱われてこなかった．しかし，現在では，識別能力をできるだけ維持して記憶容量を削減する Condensing ☆☆の研究も進み，実用的な最近傍探索アルゴリズムの登場[6]，計算機の高速化，メモリの大容量化などを背景として，この強力な識別手法が現実の問題に適用できる条件が整いつつある．

また，モデルを用いてクラス概念を帰納する学習を行わず，トレーニングパターン（事例）そのものを用いてクラス概念を記述し入力パターンを分類する「事例ベース学習」[7]（CBL: Case Based Learning あるいは IBL: Instance Based Learning）でも，最近傍識別器が主な手法の 1 つとして採用され，研究が続けられている．この事例ベース学習という枠組みが強調される理由は，特定のモデルを採用するための前提条件が成り立たない問題への対処や，特定のモデルを採用することによって生じるパターン分布の推定に生じる癖を排除することにある．

本稿では，まず色ターゲット検出の問題に最近傍識別器を用いることを提案する．ここで述べる色ターゲット検出問題とはある画像中で検出したいターゲットを色によって指定して検出することを指しており，背景やターゲットの色が変化するなどの状況は特に想定していない．次に，最近傍識別器を用いて試作した色ターゲット検出システムについて述べる．このシステムは，

- 画像上の領域を指定することにより，ターゲットと非ターゲットの色をユーザが教示し，
- 教示色から最近傍識別を行いターゲット画素の検

☆ k-最近傍識別器と呼ばれる手法は，入力に対して最も近接する k 個のパターンを探し，それに基づいて識別を行う手法であり，本稿で述べる手法はその特殊ケースすなわち，1-最近傍識別器と見なすことができる．

☆☆ Condensing と類似した概念として Editing がある．これは，トレーニングパターンを削減することにより，識別率を向上させることを指すが，歴史的に Condensing と Editing という 2 つの用語が混同されてきた．本稿では，正確さを期すためにこれらを区別して用いる．

出を行う，
という操作を繰り返すことができるため，教示を行うことにより既知の色の変化であれば背景とターゲットの色の変化にも対処できる．

このシステムでは，入力されうるすべての色を最近傍識別器で識別した結果を記録した Look Up Table（LUT）を学習時に生成し，この LUT を参照して識別を行うため，きわめて高速なターゲット検出が行える．また，トレーニングパターンは LUT 内部に記憶されるため，記憶容量が増えることもない．このように，色ターゲット検出の問題に LUT を用いることによって，最近傍識別器の問題点のうち，速度と記憶容量の問題をうまく回避することができる．

この方法では，LUT 内の各色に関して最近傍探索を行った結果のクラスラベルを逐一書き込んでおかなければならないが，この計算は各色を最近傍識別するのと同じだけの時間を要する．本研究では，インタラクティブシステムとして対話性を失わないようにするための高速な LUT 更新アルゴリズムも提案する．

実験では，肌色とカラー LED の検出問題について，EM アルゴリズムを用いて推定した混合正規分布から LUT を生成した場合と，本手法を用いた場合について比較し，速度および検出精度の両面で本手法がこれらに比べて優れた性質を持つことを確認する．

2. 色ターゲット検出問題

物理量としての色に関する研究は，光の反射や透過による色の変化を解析し，物体固有の色を求めようとする研究がほとんどである．特に，反射光を body reflection と surface reflection の 2 つの和から成るものとしてとらえる「二色性反射モデル（Dichromatic Reflection Model）[8]」は，よく知られており，このモデルをベースとした対象検出手法も提案されている[9]．この研究では，**図 2** に示すように，ボディ反射色 C_B と表面反射色 C_S の 2 つの色ベクトルが張る色空間内の平面（Dichromatic Plane）によって対象の色を表現している．この方法では，Dichromatic Plane のうちで C_S 成分が正になる範囲と，RGB Cube との交差部分を対象の色としてモデル化しているが，

- Skewed T と呼ばれる部分以外もモデルに含まれてしまうこと，
- Dichromatic Plane に厚みを持たせなければならないが，その基準が不明確であること，
- 黒や白といった無彩色部分をモデルに含めることができない，

などの問題がある．また，対象物体が完全拡散面に近いとき C_B に相当する色が観測されるが，その色分布を色空間内の楕円で近似する方法[10] もある．これに関しても同様の限界がある．

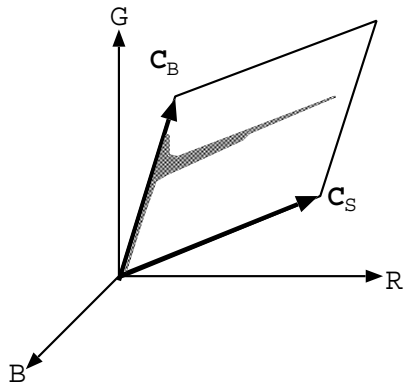

図 2 2 色性反射モデルにおける Skewed T と Dichromatic Plane
Fig. 2 Skewed T and Dichromatic Plane in dichromatic reflection model.

一方，このように物理量としての色の特性を考慮しない識別的手法についても研究が行われてきた．SCARF[11] では，道路に関する色モデルを生成するために，入力画像中の道路とそれ以外の領域の色を混合正規分布モデルを用いて表現し，誤識別が最小となるようにモデルパラメータを更新して，道路面の色モデルを生成するという手法が用いられている．この場合，ターゲットと非ターゲットの両クラスのトレーニングパターンを用いているものの，混合正規分布を用いているために，

- ターゲットと非ターゲットのわずかな色の違いが識別しにくい，
- 正規分布の個数を経験に基づいて決定しなければならない，

などの問題が生じる．また，Fuzzy ART モデルを用いて直方体状の色モデルを更新する手法[12] もある．色モデルの更新が行える点は優れているが，これもまた SCARF 同様に色検出精度の点で問題がある．

3. 最近傍識別器 vs SVM

色ターゲット検出問題を識別問題と見なした場合の特性は以下のようになる．

- ターゲット色の色空間内での分布形状は未知であり，特定の確率分布モデルを用いてターゲット色の分布を表現することはできない．
- 一般にターゲット色の数は多く，数万から数十万色のトレーニングパターンが一度に与えられる

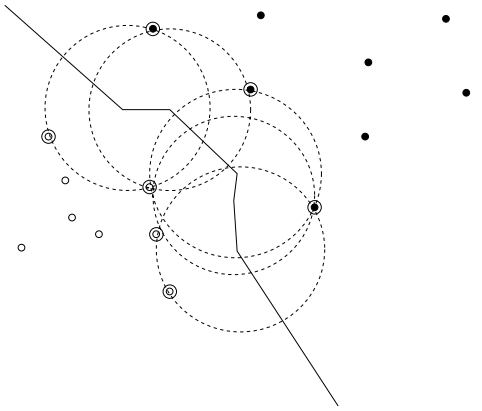

図 3 最近傍識別器の識別面と Voronoi Condensing によって残されるパターン
Fig. 3 Classification boundary of NN classifier and patterns obtained by Voronoi Condensing.

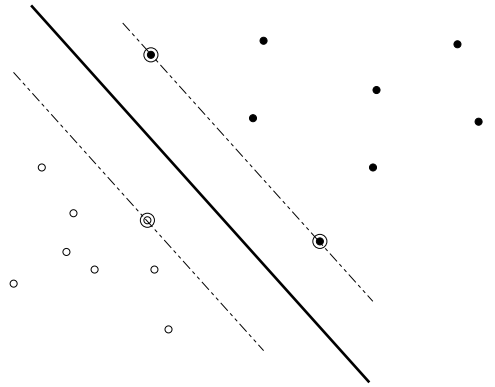

図 4 Linear SVM における識別面と Support Vector
Fig. 4 Classification boundary of Linear SVM and Support Vectors.

ケースもある．

- 色空間の次元数は 3 次元であり，高次元空間への拡張が行える必要性はない．

このような特性を念頭に置き，色ターゲット検出を行うのに適した識別器を選定しなければならない．ここでは，近年，高い識別能力を持つ 2 クラスの識別器として脚光を浴びている Support Vector Machine（SVM）[13] を取り上げ，色ターゲット検出問題への適合性を比較する．

3.1 最近傍識別器と SVM の類似性

最近傍識別器は，入力パターンに対して最近傍のトレーニングパターンを探索し，そのトレーニングパターンが属するクラスにパターンを分類するという単純な識別器である．この識別規則は**図 3** に示すように，異なるクラスに属するトレーニングパターンから見て等距離の位置に区分的線形識別面を設定しているのと等価なものである．

このように異なるクラスに属するパターンから識別面までの最短距離が最大になるように識別面を決める基準を「最大マージン基準」と呼ぶ．この識別器は，識別のために多くのパターンを記憶しなければならないという欠点を持っているが，以下に述べる Condensing [2),3)] と呼ばれる方法で記憶するパターン数を削減することができる．

図 3 の識別面を規定しているトレーニングパターンは，図中◎で示したものだけである．これらの点は，全トレーニングパターンを母点としてパターン空間を Voronoi 分割した際に，互いに異なるクラスの Voronoi 領域が接するとき，それらの領域の母点となっている．これらの点だけを残して，他のパターンを削除してしまう方法を Voronoi Condensing という．Voronoi Condensing によって最近傍識別器の識別性能が低下することはなく，Condensing の中では，これが最も安全なアルゴリズムである．これ以上にデータを削除しつつ，できるかぎり識別性能を維持する Gabriel Condensing，さらにパターン数を削減する相対近傍 Condensing などの手法も提案されている．

一方，SVM は基本的に**図 4** に示すように，最大マージン基準によって線形識別面を設定する手法である．図中◎で示した識別面から最短の距離にあるトレーニングパターンは Support Vector と呼ばれる．この識別面は，Support Vector の線形結合として表現することができる．

3.2 識 別 面

SVM と，最近傍識別器は次の 2 点できわめてよく似た性質を持っているといえる．

- 最大マージン基準によって識別面を設定していること．
- トレーニングパターンの一部によって識別面が規定されること．

ところが，SVM は，線形識別面を維持したまま kernel 関数[15] によって空間を歪める[16] という進化をとげた結果，

- 元の空間での最大マージン基準が成立しないため，ノイズなどによるパターンの変動に対する識別の頑健性が保証できない，
- 線形識別面に固執した結果，基本的には 2 クラスの識別問題しか扱えない，

という 2 つの欠点を持つことになった．SVM はさらに，線形識別不可能な場合に計算を安定化させる Soft Margin の導入[14] という改良を加えて現在の姿になり，現在も学習計算の高速化や省メモリ化など様々な

改良が行われ続けている．

これに比べて，最近傍識別器は，

- 元の空間内で，最大マージン基準の意味で過不足ない区分的線形識別面が作られる，
- 識別面は閉曲面であっても開曲面であっても，連結な曲面の個数がいくらあってもかまわず，不連続なパターン分布に対しても誤識別を起こさない識別面が作られる，
- そのままで多クラスの識別が行える，

という特長がある．しかし，トレーニングパターンに過剰にフィットした識別面が得られるケースがあるため，あるクラスのトレーニングパターンが他クラスの分布の中に入り込んでいるようなケースでは，誤識別率が増加するという問題がある．この問題の詳細については，3.5 節で述べる．

3.3 学習計算のコスト

学習に要する計算コストに関して見ると，SVM では n 個のトレーニングパターンからその要素である Support Vector を抽出するために，Kuhn-Tucker 条件を前提とした Lagrange の未定係数法により凸二次最適化問題を構成し，それを解いているが，その過程で n^2 のオーダのメモリと，計算時間を消費している．これと比べて，最近傍識別器では，

- Condensing を行わない場合には，トレーニングパターンを追加するだけで，学習を終えるため，オーダ n のメモリしか消費せず，学習のために特別な計算は要らない，
- Condensing を行う場合にも，オーダ n のメモリしか消費せず，パターン空間の次元数を D とした場合の計算時間は，Voronoi Condensing で $n^{[D/2]}$ オーダ，Gabriel Condensing の場合，Dn^3 オーダのアルゴリズムの存在が知られている[3)☆]．

このように，少なくともトレーニングパターン集合から識別に必要な要素を抽出する処理に必要以上のメモリを消費しない点では，最近傍識別器の方が優れている．

3.4 識別計算コスト

識別計算に要するコストは，識別機構をどのように実装するかによって変化するが，一般的にいえば他の手法と比べて最近傍識別器の識別時間は遅いといえる．しかし，ANN[6)] などのアルゴリズムを用いれば，次元数が極端に高い場合を除いてある程度実用に耐える性能を持たせることができる．また，ディジタル化された色のように，すべてのパターンを計算機メモリ上に保持できる場合には，後述するように Look Up Table（LUT）を用いて高速な識別が行えるため，この問題を回避することができる．

3.5 識別性能

Vapnik は，「教師信号と識別器の出力の差」を測る**連続損失基準**に基づいて計算した期待リスクの上限を示し，SVM を用いてこの上限値を最小化する Structual Risk Minimization が実現できることを論じている[13)]．しかし，実際的な意味での誤識別は，**0-1 損失基準**，すなわち「うまく識別ができるか否か」という基準で測られるべきであるという見方もできる．

クラスを ω_i $(i = 1, \cdots, M)$，クラス ω に属するパターン x の生起確率密度関数を $p(x|\omega)$，クラス ω に含まれるパターンが生起する確率を $P(\omega)$ で表したとき，Bayes 決定則に基づく 0-1 損失基準での誤り確率の最小値，すなわち **Bayes Error** は，次式で表される．

$$P_e = 1 - \int_{-\infty}^{\infty} \max_i^M [p(x|\omega_i)P(\omega_i)]dx \qquad (1)$$

注意すべきことは，式 (1) において，$p(x|\omega_i)$ は真の確率分布を表しており，$p(x|\omega_i)$ が有限の台を持ち，かつ，その台が互いに素になる場合，つまり「**異なるクラス間でパターン分布が分離可能な場合**」には，$P_e = 0$ となることである．最近傍識別器の場合には，トレーニングパターン数を無限大に近づけたときの漸近的な誤り確率を P_{NNerr} とすると，次式が成り立つ[1)]．

$$P_e \leq P_{NNerr} \leq P_e \times \left(2 - M\frac{P_e}{M-1}\right) \qquad (2)$$

式 (2) は，トレーニングパターン数を無限大に近づけたとき，最近傍識別器の誤り確率が式 (1) の P_e の 2 倍を超えない範囲に漸近的に収束することを表している．これは当初，あるクラスの確率分布についてのみ成立すると考えられていたが，その後，どのような確率分布の場合に関しても成立することが証明されている[17)]．これは，Bayes Error が 0 にならないケース，すなわちあるクラスのトレーニングパターンが他クラスの分布の中に入り込んでいるようなケースでは，誤識別率が増加するという最近傍識別器の問題も，十分なトレーニングパターンが与えられていれば理想的な場合の 2 倍未満の誤識別率の範囲内におさえられることを意味している．

このことから，特に，

- 正確な確率密度分布が推定できるだけの十分な事

☆ この計算オーダは，全トレーニングサンプルについて Voronoi 分割や Gabriel グラフを求める際の計算に基づくものであり，識別境界付近のデータについての探索を行う効率的な Condensing アルゴリズムが提案されている[4),5)]．

前知識が与えられておらず，
- 十分な数のトレーニングパターンが利用できる，

という色ターゲット検出の問題に最近傍識別器は適しているといえる．

4. LUT による最近傍識別の高速化

人間と識別器が教示と識別を繰り返すインタラクティブなシステムでは，多くのトレーニングパターンを集めやすいという好条件がある一方で，インタラクションを阻害しないように学習時間を短くしなければならないという要請もあり，最近傍識別器のアプリケーションとして適している．特に色識別の問題では，
- LUT を用いることによって識別計算時間を短縮することができる，
- 全パターンを記憶することができるため Condensing を行う必要がない，

など最近傍識別器の欠点が表面化しないことからも都合がよい．

識別器の学習は，色空間を量子化した LUT の各要素に識別結果となるクラスラベルを書き込むことによって実現される．具体的には，入力されうる色を最近傍識別した場合のクラスラベルを，対応する LUT の要素に書き込む．このようにしておけば，入力された色に対応する LUT 要素の読み出しだけで識別することができ，識別速度の問題も回避することができる．

LUT の更新は，LUT の各エントリにトレーニングパターンからの距離とクラスラベルを書き込んでいくことによってなされる．すでに他のトレーニングパターンによって書き込みがなされている場合は，書き込もうとする距離よりも大きな距離が書き込まれていれば上書きをし，そうでなければ書き込みをしないという方法である．この LUT の更新法は，色空間内で各トレーニングパターンの拡張を等速度で行ったとき，異なるクラスの境界がぶつかったところが識別面になることから「**色競合（Color Competition）**」と呼ぶことにする．

市街地距離（6 近傍距離）を用いた場合の色競合アルゴリズムをアルゴリズム 1（**図 5**）に示す．このアルゴリズムで使用されている変数および関数について以下に説明する．

- Queue，LUT の各要素は Color，Distance，Class の 3 変数を構成要素とする型である．
  - Color は Y，U，V の 3 つの値を持つレコード型変数である．
  - Distance は教示色からの距離を表すスカラー変数である．
  - Class は教示色の所属するクラスを表す変数である．
- Queue は待ち行列であり，EnQueue（データの追加）と DeQueue（データの読み出し）ができる．
- LUT[ ] は Color と同じ型を添え字として持つ配列である．
- CityNeighborEntry(Y,U,V) は (Y,U,V) の色から見た 6 近傍の各色を集合として返す関数である．
- foreach X in SET という構文は，SET の全要素についてその値を X にセットして後続する実行部を反復することを意味する．
- InRange(y,u,v) は，(y,u,v) という値が色のレンジを超えていなければ真を，そうでなければ偽を返す論理関数である．

```
clear LUT;
foreach (Y,U,V,Class) in TrainingSample
begin
    Entry.Color:=(Y,U,V);
    Entry.Distance:=0;
    Entry.Class:=Class;
    Enqueue(Queue,Entry);
end
while (not Empty(Queue))
begin
   Entry:=Dequeue(Queue);
   if LUT[Entry.Color].Distance>Entry.Distance then
      begin
        LUT[Entry.Color]:=Entry;
        Entry.Distance:=Entry.Distance+1;
        foreach (y,u,v) in CityNeighborEntry(Y,U,V)
        begin
          if InRange(y,u,v) then
          begin
            Entry.Color=(y,u,v);
            Enqueue(Queue,Entry)
          end
        end
      end
end
```

図 5 アルゴリズム 1：色競合による LUT 更新
Fig. 5 Algorithm 1: LUT updating algorithm for Color Competition.

このアルゴリズムは横型探索を用いて，各エントリへの上書きが最小限にとどめられるように設計されているため，ほぼ LUT の要素数に比例した時間で学習が終了する．さらに，色の追加や削除については，局所的なラベル書き換えで済むため，より短時間で実行できる．

4.1 インタラクティブ実時間色識別システム

ここでは以上の議論を基にして作成したインタラクティブな色ターゲット検出システムについて紹介する．このシステムは，以下に示す機能を持っている．

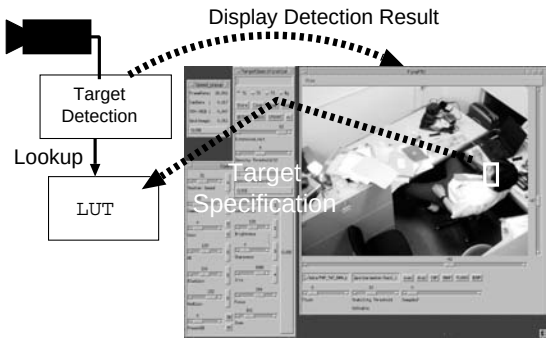

図 6 色ターゲット検出システムの概要
Fig. 6 Overview of Color Target Detection System.

- ターゲット色 3 クラスおよび背景色 1 クラスを教示することができる．
- VGA サイズの画像に対して，ターゲット画素の検出，密度解析，重心計算をビデオレートで行うことができる．
- 色を教示すると瞬時に LUT が更新され，ユーザはその LUT を用いて検出を行った結果を画面上で確認しながら，未検出・誤検出部分を再教示できる．
- このシステムは，過去に起きた未検出・誤検出についても再教示が行えるように画像蓄積用のリングバッファが付けられており，過去の画像を再生し未検出・誤検出が起きている部分で再教示が行える．

図 6 にシステムの構成と操作画面を示す．このシステムでは，画像上でユーザが指定した領域にターゲット 1〜3 または背景のクラスを対応付けることを繰り返し，LUT へのラベル付け，すなわち色競合による学習を行う．色競合の後，検出結果を即座に表示し，ユーザはこれを見ながら誤検出，未検出の部分について修正を加えていくことができる．このようにインタラクティブな教示と検出の機能を使うことによって，照明変化によって生じる色の変化や非ターゲットの誤検出などの問題にある程度対処できるようになる．

色ターゲット検出の際に参照される LUT のサイズは小さいことが望ましい．このため，LUT の各要素は**図 7** に示すように 8-bit で表し，上位 2-bit をクラスの種類を表すタグとして使い，下位 6-bit をトレーニングパターンからの距離として使う．この LUT を参照した色の識別は，各色に対応する LUT 要素のタグ bit を参照して行うことができる．

システム構築に使用した機材は，計算機：Pentium4 Xeon 1.7 GHz Dual CPU（キャッシュ：L1 2 M-byte，L2 1 M-byte，L3 512 K-byte），カメラ：SONY DFW VL500 であり，計算機の OS は Linux kernel 2.4.18 である．このカメラは，YUV のディジタル画像出力を持ち，画素値のノイズのため Y，U，V の最下位ビットを無視しても検出精度にはほとんど影響を与えない．また高速な検出を行うには LUT がキャッシュメモリに収まるようにする必要があるため，128×128×128 のサイズ（2 M-byte）に量子化した YUV 色空間を LUT として用いている．入力画像のサイズは 640×480 であり，LUT の更新にはアルゴリズム 5 を用いている．この方式では，γ 補正や色空間に関する制限がまったくないため，すべてのカメラパラメータを自由に設定できる．

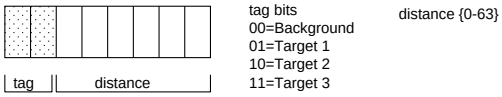

図 7 LUT 要素のデータ型
Fig. 7 Data structure of LUT element.

ユーザが領域を指定した際に行われる LUT の更新には約 0.5 s を要し，各ビデオフレームに対して施される処理は 22〜24 ms を要する．この内訳は以下のとおりであり，ビデオレートの処理速度が維持できる範囲内で，まだ 11〜9 ms の処理時間の余裕がある．

色検出，密度解析，重心計算，YUV→RGB 変換： 6〜7 ms
画面表示： 10〜12 ms
画像入力： 5 ms 以下

なお，このシステムでは同じ色に関して異なるクラスが教示された場合，その色は最後に教示されたクラスに帰属するものとして扱われる．

5. 実　　験

実験では，手法は以下の 3 手法についてターゲット/非ターゲットともに同じ学習パターンを与えて検出を行った結果について述べる☆．

EM-2： ターゲット/非ターゲットの各クラスに EM アルゴリズムを用いてそれぞれ 2 個の正規分布をあてはめ，その結果に基づいて LUT を作成する方法

EM-5： ターゲット/非ターゲットの各クラスに EM アルゴリズムを用いてそれぞれ 5 個の正規分布を

☆ SVM を用いた学習も試みたが，肌色検出では背景色の数が多く，学習に時間がかかりすぎるため省略した．また LED の抽出では単純な 3 クラスの識別問題であるため SVM は使用できなかった．

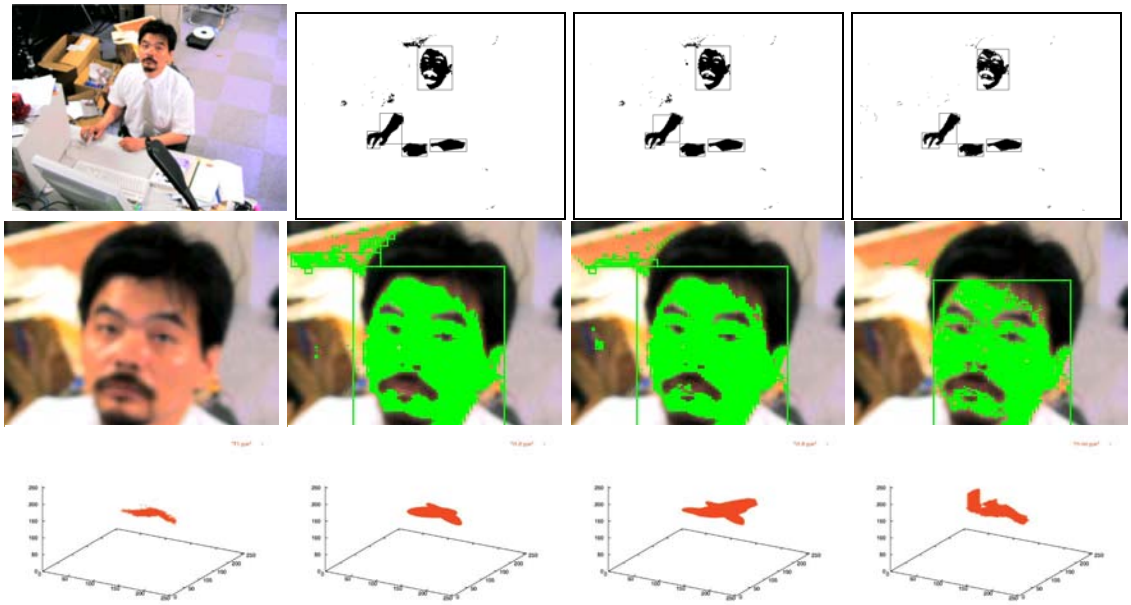

図 8 肌色領域抽出結果（上段左から：原画像，EM-2，EM-5，NN）（中段：上段の拡大表示）（下段左から：ターゲット色の教示色，EM-2，EM-5，NN で検出に使われた LUT）
Fig. 8 Skin color detection results (Top, from left to right: Original Image, Detection Result using EM-2, EM-5, NN) (Middle: Magnified image) (Bottom, from left to right: Training color, LUT used for EM-2, EM-5, NN).

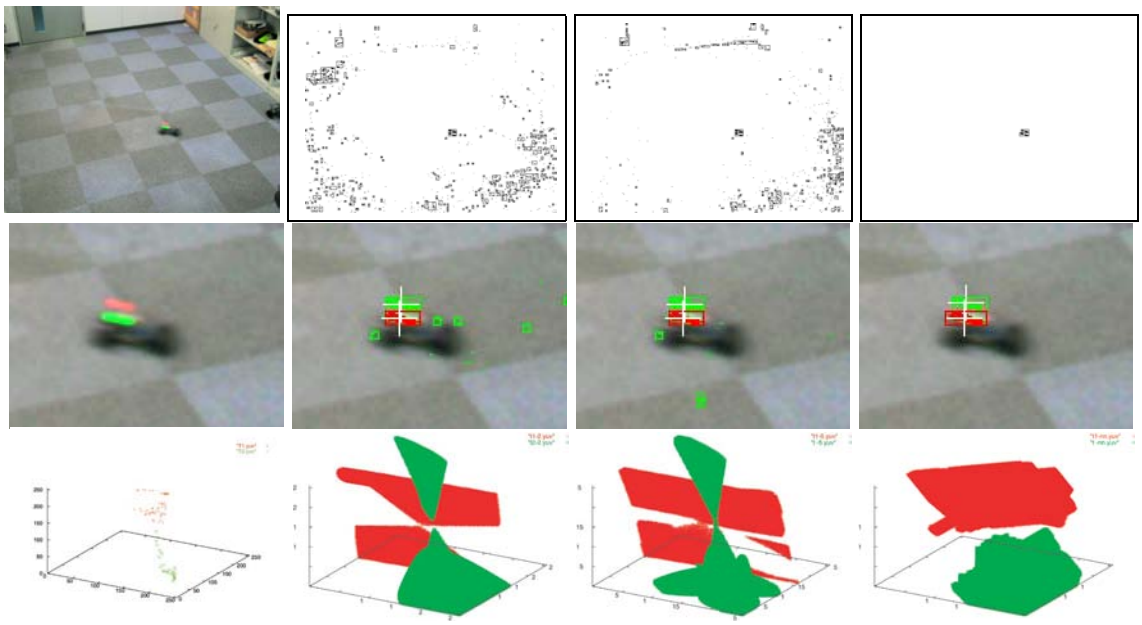

図 9 2 色の LED 検出結果（上段左から：原画像，EM-2，EM-5，NN）（中段：上段の拡大表示）（下段左から：ターゲット色の教示色，EM-2，EM-5，NN で検出に使われた LUT）
Fig. 9 LED color detection results (Top, from left to right: Original Image, Detection Result using EM-2, EM-5, NN) (Middle: Magnified image) (Bottom, from left to right: Training color, LUT used for EM-2, EM-5, NN).

あてはめ，その結果に基づいて LUT を作成する方法

NN： 色競合で LUT を作成する方法（提案手法）

なお，EM-2，EM-5 では各クラスの事前生起確率をすべて等しいものとし，確率密度の大小に基づいて LUT の各エントリの所属クラスを決定している．

5.1 肌色領域検出

蛍光灯による照明があたる屋内環境のもとで撮影した画像中の人物の肌色を検出した結果を**図 8** に示す．トレーニングパターン数は，肌色領域 1469 色，背景色 10002 色である．検出結果画像内の各矩形枠内の対象画素密度は 40%以上である．EM-2，EM-5，NN での LUT 生成時間はそれぞれ 22.1 s，40.7 s，0.5 s であり，NN（提案手法）が圧倒的に高速である．このことから，NN が教示と検出結果の提示を繰り返すインタラタクティブな使用に最も適しているといえる．図の上段，中段に示した検出結果を見ると，EM-2，EM-5 は肌色領域の抜けは少ないが，人物左側，頭部付近には木製やダンボールの箱など肌色と類似した部分での誤検出が目立つ．これは，EM アルゴリズムを用いた方法ではターゲット色の分布が汎化されすぎてしまったためこのような誤検出が起きているものと考えられる．これらに比べて本手法（NN）は肌色領域での抜けは多少あるものの，類似色の領域での誤検出はかなり抑制されている．下段の LUT を見ると，EM-2，EM-5 の LUT では混合正規分布を用いていることによる癖が現れた肌色分布となっていることが分かる．これに対して，NN の LUT は癖があるものの背景色の抑制が効いているため，汎化の方向が背景色のない方向に制限されている．

5.2 カラー LED 検出

屋内の蛍光灯による照明下で RC カーに取りつけた赤と緑の LED をそれぞれ別のクラスに属するパターンとして検出した結果を**図 9** に示す．この場合，LED 自体が発光するため 2 色性反射モデルなどの反射モデルを用いた手法は利用できない．各矩形枠内の対象画素密度は 10%以上，連結した矩形内の対象画素数が 3 以上の場合に，重心位置を ‘+’ で描いている．この場合，画像中から色をサンプルできる領域がきわめて狭いため，何度も領域を指定しない限り，十分な数のトレーニングパターンは得られない．この実験では，下段手法 1 の LUT に示すように，赤緑ともトレーニングパターンの数は非常に少なく，その色のばらつきも大きい．具体的な教示色の数は赤 LED，緑 LED，背景色，それぞれ 59 色，57 色，3199 色となっている．EM-2，EM-5，NN での LUT 生成時間はそれぞれ 13.2 s，33.4 s，0.5 s であり，やはり NN が最も高速であった．検出結果を見ると，EM-2，EM-5 ともに LED は検出できているが，背景部分での誤検出がかなり多い．これに対して NN では誤検出もまったくなく，LED の検出にも成功していることが確認できる．下段の LUT を見ると，EM-2，EM-5 の LUT では混合正規分布を用いていることと，背景の教示パターンが十分でないため，特定方向への汎化が起きており，元のトレーニングパターンにはない紡錘型の分布が推定されていることが確認できる．NN の場合は，汎化は起きているものの 2 つの LED の色分布はそれぞれまとまった領域になっている．

6. ま と め

本稿では，色ターゲット検出の問題に対して最近傍識別器を用いる手法と，教示パターンが与えられた際に高速に LUT の内容を更新するアルゴリズムを提案した．実験では，混合正規分布モデルで各色の分布を表現する場合との比較を行い，提案手法の有効性を示した．さらに，この手法に基づいて，ビデオレートで色識別と，画素密度に基づくセグメンテーション，重心計算を行うシステムを構築した．このシステムは，下記のように幅広い用途で利用可能である．

- 人物の肌色検出を行えば，ジェスチャや顔認識，監視システムなどの用途に用いることができる．
- スタジオなどでは，クロマキー，ルマキーなどよりも一般的なバックグラウンドやフォアグラウンドの除去が実現できる．
- LED やカラーマーカの検出に用いれば，移動ロボットの位置同定や，モーションキャプチャシステムにも応用できる．
- その他，画像検索，画像編集などにおいて対象の検出や切り出しなどにも利用可能である．

また，最近傍識別器は，背景差分などに利用することも可能である．これまでの背景差分では，背景色の色と入力画像の色から何らかの操作を行うことによって，「背景らしさ」を表すスカラー値を算出し，これを閾値処理することによって対象のシルエットを求めていた．しかし，入力と背景の各色成分の差を表す 3 次元のパターン空間，あるいは入力と背景の色を組にした 6 次元のパターン空間で最近傍識別器による識別を行えば，より高精度な対象検出が行えるはずである．後者の場合は，色ターゲット検出と背景差分の両方の特色を備えた，より一般的なターゲット検出機構となるはずである．このように，最近傍識別器を識別エンジンとして用いれば，これまでの画像解析手法がより

強化される可能性が十分にある．

今後は，最近傍識別器の理論的側面の強化と，それを用いた背景差分，色競合に基づくターゲット検出の応用や色の変化に対する追従性について検討を行う予定である．

謝辞 本研究の一部は科学研究費基盤（A）（2）12308016 の支援を受けている．本研究のきっかけの 1 つは，UBC 計算機科学コース Vladimir Kravtchenko 氏の Dichromatic Reflection Model に基づく実時間対象検出の研究である．彼の研究指導を一緒に行った James Jim Little 教授との議論も研究の参考になった．ロボット学会誌編集委員会からは Physics Based Vision について執筆する機会をいただき，これが非 Physics Based Vision の研究を行うきっかけとなった．実装においては，九州大学吉本氏による色変換ルーチンを参考にした．

参 考 文 献

1) Cover, T.M. and Hart, P.E.: Nearest neighbor pattern classification, *IEEE Trans. Inf. Theory*, Vol.IT-13, No.1, pp.21–27 (1967).
2) Hart, P.E.: The condensed nearest-neighbor rule, *IEEE Trans. Inf. Theory*, Vol.IT-4, No.5, pp.515–516 (1968).
3) Bhattacharya, B.K., Poulsen, R.S. and Toussaint, G.T.: Application of Proximity Graphs to Editing Nearest Neighbor Decision Rule, *International Symposium on Information Theory*, Santa Monica (1981).
4) 和田，加藤：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).
5) 加藤，和田：近接性グラフに基づく効率的 Condensing のアルゴリズムと評価，信学技報 PRMU，Vol.103, No.96, pp.19–24 (2003).
6) Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R. and Wu, A.Y.: An optimal algorithm for approximate nearest neighbor searching, *J. ACM*, Vol.45, pp.891–923 (1998).
7) Aha, D.W.: Case Based Learning Algorithms, *Proc. Case based Reasoning Workshop*, pp.147–158 (1991).
8) Klinker, G.J., Shafer, S.A. and Kanade, T.: The Measurements of Highlights in Color Images, *Int. J. Comp. Vision*, Vol.2, No.1, pp.7–32 (1988).
9) Kravtchenko, V. and Little, J.J.: Efficient Color Object Segmentation Using the Dichromatic Reflection Model, *Proc. IEEE PACRIM' 99*, pp.90–94 (1999).
10) Rasmussen, C., Toyama, K. and Hager, G.D.: Tracking Objects By Color Alone, Technical Report TR1114, Dept. of Computer Science Yale University (June 1996).
11) Crisman, J. and Thorpe, C.: SCARF: A Color Vision System that Tracks Roads and Intersections, *IEEE Trans. Robotics and Automation*, Vol.9, No.1, pp.49–58 (1993).
12) Nakamura, T. and Ogasawara. T.: On-Line Visual Learning Method for Color Image Segmentation and Object Trackin, *Proc. IROS'99*, pp.222–228 (1999).
13) Vapnik, V.: *The Nature of Statistical Learning Theory*, Springer-Verlag, New York (1995).
14) Cortes, C. and Vapnik, V.: Support Vector Networks, *Machine Learning*, Vol.20, pp.273–297 (1995).
15) Aizerman, M.A., Braverman, E.M. and Rozoner, L.I.: Theoretical foundations of the potential function method in pattern recognition learning, *Automation and Remote Control*, Vol.25, pp.821–837 (1964).
16) Boser, B.E., Guyon, I.M. and Vapnik, V.: A trainig algorithm for optimal margin classifiers, *ACM Fifth Annual Workshop on Computational Lerning Theory*, Pittsburgh (1992).
17) Stone, C.J.: Consistent nonparametric regression, *Annals of Statistics*, Vol.5, pp.595–645 (1977).

(平成 15 年 4 月 3 日受付)
(平成 15 年 9 月 9 日採録)

（担当編集委員 日浦 慎作）

和田 俊和（正会員）

1990 年東京工業大学大学院博士課程修了．同年岡山大学工学部助手．1997 年京都大学大学院工学研究科助教授．2002 年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．