

# 色ヒストグラム特徴とパン・チルト・ズームカメラを用いた 高速物体探索法——動的アクティブ探索法——

川西 隆仁<sup>†</sup>      村瀬 洋<sup>†</sup>

Quick Object Search Using Color Histograms and Pan-Tilt-Zoom Camera  
—— Dynamic Active Search ——

Takahito KAWANISHI<sup>†</sup> and Hiroshi MURASE<sup>†</sup>

あらまし 本論文では、3次元実環境中にある物体をパン・チルト・ズームカメラを用いて高速に探索する手法を提案する。これは画像から特定の物体を高速に探索する技術「アクティブ探索法」を拡張したものである。実環境中の物体は、照明条件や物体の向きやカメラパラメータの違いにより大きさや見え方が異なる。これに対処するには、様々な見え方に対応して参照画像を増やすことが有効であるが、単純に増やすだけでは探索時間が増大する。それに加えて、物体が遠方にある場合には段階的にパン・チルト・ズームを行う必要があり、これによりカメラ制御回数も膨大になる。提案手法では、まず増加した参照画像間のヒストグラムの類似性を用いて探索する時間を削減する。更に、本手法をカメラ制御の予測に動的に利用し、カメラ制御回数も削減する。実験の結果、室内環境においてカメラから5m以内の位置にある小物体を正しく探索するのに、従来のアクティブ探索法と単純なカメラ制御を用いた従来法に比較し、探索速度は4倍以上向上し、本手法の有効性を確認した。

キーワード 物体検出、パン・チルト・ズームカメラ、アクティブ探索法、カメラ制御

## 1. ま え が き

視覚をもつロボットや監視システムを実現するためには広い視野の中から目的の物体を高速・高精度に検出する機能が必要となる。本論文では、広い視野を取得できるパン・チルト・ズームカメラを用いて特徴照合により室内環境から物体を探索する技術について扱う。差分情報を使用していないために動きのない静止物体の検出も可能であり、差分情報を利用する手法との併用も可能である。

室内環境中の物体をカメラ制御して探索する際の問題点は、物体の向き、照明条件などの違いなどにより物体の見え方の変化が大きいため、パン・チルト・ズームを制御して撮影した非常に多数の画像を扱う必要があることである。

複数センサ [1]、パン・チルト・ズームカメラ [2]、全

方位センサ [3] などを用いて広い視野を取得する従来手法では、これらの問題を回避するために、あらかじめ用意した全周囲背景画像との差分や前フレームの画像との差分を利用して動きのある物体を検出 [4] ~ [6] している。しかし、このような差分を用いた手法では、環境中の静止物体を検出するなどの用途には適用できない。差分を使用しない手法には、ロボットの誘導のために室内環境中のランドマークを検出する方法 [7], [8] が提案されている。これらの方法は特定の人工的なマーカにしか適用できず、一般の物体を物体固有の特徴に基づいて検出することはできない。

一方、特徴照合による物体探索では、カメラで取得した画像上の局所領域に着目し、あらかじめ用意した物体の参照画像との間で照合する。代表的な手法にテンプレートマッチング法がある。しかし、物体までの距離や方向が未知であるため、大きさ・位置の異なる多数の局所領域との間で照合する必要がある上に、物体の向きや照明条件の変動を考慮に入れた場合には更に多くの参照画像を用意して照合する必要があり、探索時間が指数関数的に増大する。探索速度の改善のた

<sup>†</sup> 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所、厚木市

NTT Communication Science Laboratories, NTT Corporation, 3-1 Morinosato-Wakamiya, Atsugi-shi, 243-0198 Japan

めにいくつかの手法 [9] ~ [11] が提案されているが、カメラのパン・チルト・ズームの制御までも考慮するような広大な探索空間に対しては、照合コストが大きく十分な速度は得られていない。

一方、膨大な照合回数を大幅に削減する技術として、アクティブ探索法 [12] ~ [15] が提案されている。アクティブ探索法は特徴照合の際にヒストグラムを用い、ヒストグラムの性質を用いて精度を落とさないことを保証しながら探索に不要な照合を削減する手法であり、これにより極めて高速な物体探索が可能となる。更には、ヒストグラム照合を利用するため多少の形状変動に対しても頑健である。しかしながら、照明条件や物体の向きの違いにより特徴が大きく変化する場合には、単一の参照画像のヒストグラムでは正しく探索できない。この変化を吸収するには、様々な条件で物体を撮影した複数の参照画像のヒストグラムを用意し探索する必要がある。また、ヒストグラム特徴には、物体領域が小さくなるほど識別性能が低下する [16] 性質があるため、ズームを細かく制御する必要がある。このため、数多くのカメラの制御回数が必要になる。

本論文で提案する手法は、従来のアクティブ探索法の考え方を以下の 2 点において発展させたものである。

第 1 に、本手法では局所領域と複数の参照画像との照合を行うが、その際に、ある参照画像の照合結果を利用して他の参照画像の照合を適応的に削減する。ここでの照合の削減は、従来のアクティブ探索法と同様に原理的に総当たり照合と同じ解を保証している。照合が省略可能かどうかは、照合結果の類似値と参照画像間の類似値で決定され、参照画像間で類似している場合は大きく削減し、類似していない場合にはあまり削減しないと制御が自動的かつ厳密に行える。このため、精度を保つために対象に応じてパラメータを調整する必要はない。更に、今回のように特徴の変動が大きい場合に効率的なヒストグラム区画の構成法についても改良している。

第 2 に、画像上の局所領域と参照画像との照合結果を、カメラの制御回数の削減に利用する。照合には適さない小さい局所領域に関しても、先に述べた高速探索法を利用して物体の存在の可能性を予測し、その予測値をもとに優先的にズームインするカメラの方向を決定することでカメラの制御回数を削減する。

本論文では、2. で提案手法を、3. で探索精度と探索速度の評価実験とその結果を、4. でまとめを述べる。

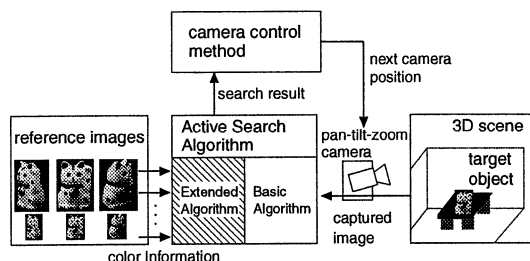


図 1 動的アクティブ探索法の処理の概要  
Fig.1 Dynamic Active Search procedure.

## 2. 動的アクティブ探索法

本手法は、画像から特定の物体を高速に探索するアクティブ探索法を 3 次元環境中の物体探索に適用できるように拡張したものである。具体的には室内環境などにパン・チルト・ズームカメラを設置し、パン・チルト・ズームのカメラパラメータを動的に制御しながら、あらかじめ登録された物体と類似した物体を特徴照合を用いて高速に探し出す。

### 2.1 手法の概要

図 1 に動的アクティブ探索法の処理の概要を示す。参照画像データベース (図 1 左) とアクティブ探索部 (図 1 中央下段) とカメラ制御部 (図 1 中央上段) からなる。以下、動的アクティブ探索法の処理の概要を説明する。

まず準備としての学習段階では、物体に対して様々な条件で撮影した画像を用意し、参照画像データベースに蓄積する。次に、データベースに登録された参照画像からアクティブ探索に適したヒストグラム区画を生成する。更に、カメラ制御のために、3 次元環境全体を見渡すパン・チルト・ズームのカメラパラメータ集合を求めらる。

探索段階では、まずカメラ制御部はカメラパラメータ集合の最初のパラメータをカメラに通知する。カメラは受け取ったカメラパラメータを用いて画像を撮影し、その画像をアクティブ探索部へ送る。アクティブ探索部では物体を画像中から高速に探索し、しきい値を超えた領域を探索結果としてカメラ制御部へ通知する。カメラ制御部では物体を検出した場合には物体をユーザに表示して探索を終了し、検出しなかった場合には探索結果に基づいてカメラを制御する順序を動的に切り換え最も優先度の高いカメラパラメータで撮影する。

## 2.2 複数参照画像の高速探索法

動的アクティブ探索法では、照明などの環境の変動に頑健にするために、探索処理自体も従来のアクティブ探索法の改良を行っている。具体的には、参照画像の色分布を前もって学習し、アクティブ探索を効率化するヒストグラム区画を生成する点と、探索段階において複数の参照画像間の関係を利用して照合の削減を行う点の改良である。

### 2.2.1 効率的なヒストグラム区画の設計

従来のアクティブ探索では、RGBの各軸を等間隔に分割することによって減色した各色の頻度を計数しヒストグラムを生成した。しかし、このような減色は照明の変動や、物体の向きの変化などに対して頑健でない。あらかじめ室内環境中の様々な位置、向きで撮影された物体の画像が撮影できる場合にはその物体の色を学習サンプルとしてベクトル量子化(VQ)することによって効率の良い減色を行うことができる。従来、アクティブ探索にVQを利用する方法として[14],[15]が提案されている。しかしこれらはVQを探索の前処理として行っており、本研究のようなリアルタイムで入力される大量の画像情報に対して行うことは想定されていない。このためVQに伴う距離計算などに大きな計算量がかかり速度に問題がある。そこで本研究[17]では、距離計算を省略するために符号化テーブルを用いる。符号化テーブルはRGBの各軸を32分割(32<sup>3</sup>区画)した各色の小区画がどのVQ符号に属するかを示すもので、定数時間で色とVQ符号とを対応づける。また参照画像に含まれる色に対してVQを行えるため、参照画像に含まれない色を表す特別な符号0を導入することによって、参照画像に含まれる色に効率良く符号を割り当てることができる。この方法では符号数が少なくなるため、ヒストグラム照合演算が高速になる。更に、物体の写らない領域を高速に枝刈りすることができる。手順の詳細を以下に述べる。

(1) 全体の画素の平均色、分散を求める。この平均色を代表色 $C_1$ とし、全画素を $C_1$ に帰属させる。

(2) 分散最大の代表色 $C_i$ を選ぶ。 $C_i$ の分散 $\alpha_i < \sigma$ であれば、ステップ(4)へ飛ぶ。

(3)  $C_i$ を二つに分割し、LBGアルゴリズム[18]~[20]により最適な代表色を決定する。全画素と全代表色の距離を計算し各画素を最も近い代表色に帰属させる。代表色に帰属された画素の色の平均色と分散 $\alpha$ を計算する。ステップ(2)へ戻る。

(4) 色空間の全色 $P_i$ に対して最も距離の近い代

表色 $C_j$ を求める。 $|P_i - C_j| < n\alpha_j$ であれば $P_i$ に $C_j$ の符号を割り当てる。さもなければ、 $P_i$ に符号0を割り当てる。

$n$ は学習に用いた画素のうち符号0に含まれる画素の割合が一定以下になるように経験的に2程度に決めた。 $n$ が少なすぎる場合には、符号0の割合が増加し、多すぎる場合には0以外の符号に物体の色に類似しない色が含まれる割合が増加する。どちらの場合も誤検出が増加する。一方、 $\sigma$ は、小さい場合には照明や向きの変動に弱くなり、大きい場合には誤検出が多くなる。このため物体の色分布に応じて決定する。

### 2.2.2 アクティブ探索法の基本

アクティブ探索法[12]は、入力画像中で位置や大きさの変化した物体を検出できるように、参照画像と、入力画像の局所領域との照合を行う手法であり、色ヒストグラム特徴の性質を利用した局所領域の枝刈りにより精度を保証したまま照合回数を大幅に削減するという特徴がある。

アクティブ探索は任意の形状の参照画像に適用できるが、簡単のため、方形の参照画像で基本アルゴリズム<sup>(注1)</sup>を説明する。まず、入力画像を符号化した画像に対し、参照画像と同じ大きさの探索窓をかけ、ヒストグラム特徴を作る。続いて、ヒストグラム同士の類似値(similarity)を計算する。ヒストグラム特徴 $H_M$ と $H_A$ の重なり率 $S_{AM}$ は、次のように定義される。

$$S_{AM} = \frac{1}{|M|} \sum_{i=1}^I \min(H_{M_i}, H_{A_i}) \quad (1)$$

ここで、 $H_M$ と $H_A$ は、それぞれ参照画像 $M$ と入力画像中の局所領域 $A$ に対するヒストグラムであり、 $H_{M_i}, H_{A_i}$ はそれぞれの $i$ 番目の符号をもつ画素数である。また $|M|$ は参照画像の画素数(参照画像の特徴の総数)であり、 $I$ は符号の種類である。 $S_{AM}$ が探索しきい値 $\theta$ 以上となる時、物体として検出する。

局所領域の枝刈り法を図2を用いて説明する。大きさの異なる二つの参照画像を $M, N$ とする( $|M| > |N|$ )。 $M$ と $N$ の色の変化はない、すなわち、 $N$ のヒストグラムは $M$ のヒストグラムの定数倍であり、 $N$ の色はすべて $M$ に含まれるとする(異なる色が含まれる場合への一般化は2.2.3(a)で扱う)。ここで、 $|A| = |M|, |B| = |N|$ を満たし重複領域をもつ局所領

(注1): ただし、探索窓の大きさによってヒストグラムの分布に変化が生じる場合を扱うので[12]と異なりヒストグラムを大きさと正規化しない。このため類似値は可換でない。 $S_{AM} \neq S_{MA}, S_{AM} \cdot |M| = S_{MA} \cdot |A|$

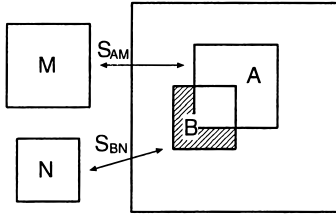


図 2 重なりを利用した照合のスキップ  
Fig. 2 Pruning using overlapping.

域  $A, B$  を考える．参照画像  $M$  と局所領域  $A$  との類似値  $S_{AM}$  と，参照画像  $N$  と局所領域  $B$  との類似値  $S_{BN}$  との間には式 (2) の不等式が成立する．

$$|N| \cdot S_{BN} < |M| \cdot S_{AM} + n \quad (2)$$

ここで， $|M| \cdot S_{AM}$  は  $A$  と  $M$  との間で色の同じ画素の組が何組あるかを意味する．これを同色画素数と呼ぶことにする．この式は「物体と  $B$  間の同色画素数 ( $|N| \cdot S_{BN}$ ) は物体と  $A$  間の同色画素数 ( $|M| \cdot S_{AM}$ ) が  $A$  と  $B$  の共通領域に局在し，かつ， $A$  に含まれない  $B$  の領域 (図 2 の斜線部分で画素数  $n$  とする) の画素すべてが物体と  $B$  の同色画素になった場合より多くなることはない」ことを意味する．すなわち， $S_{AM}$  が既知となったとき， $|N| \cdot S_{BN}$  の上限は式 (2) の右辺になる．この上限値が  $|N| \cdot \theta$  より小さければ  $N$  と  $B$  との照合を省略することができ，探索の高速化が図れる．照合を省略できる  $n$  の範囲は次の式で与えられる．

$$|M| \cdot S_{AM} + n < |N| \cdot \theta \quad n < |N| \cdot \theta - |M| \cdot S_{AM} \quad (3)$$

### 2.2.3 複数参照画像を考慮したアクティブ探索

学習段階では，物体の向きや照明の変化に対応するため，複数の向きや位置で物体を撮影し，探索段階においては，撮影した複数の画像を参照画像として高速に物体を探索する．

処理の流れを図 3 に示す．はじめに，それぞれの参照画像に対して各画素の色を符号化テーブルに基づき符号化し各符号に属する画素の出現頻度を計算することによってヒストグラム特徴 (reference histograms) を作成する．次に入力画像に対する符号化処理を行う．すなわち，参照画像に対する処理と同様に入力画像の各画素に対して符号化テーブルに基づく符号化を行う．この画像に対して，ヒストグラムが類似する局所領

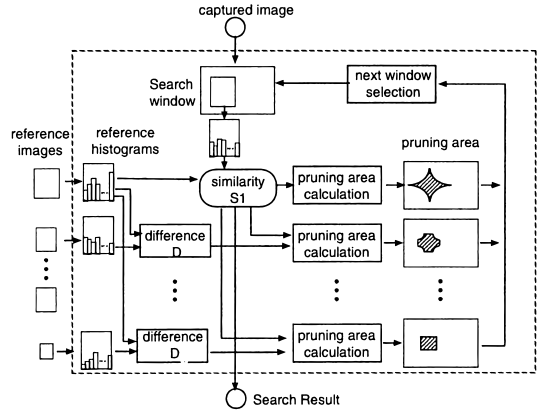


図 3 複数参照画像を用いた物体探索の処理の流れ  
Fig. 3 Object search for multiple reference images.

域を探索する原理は基本的にはアクティブ探索法と同じである．すなわち，ある局所領域に対して，符号を用いてヒストグラム (input histogram) を作成し，探索不要な領域 (pruning area) を計算しながら探索を行う．

更に，複数の参照画像を同時に効率良く探索するために，この図 3 の場合では，あらかじめ計算した参照画像間のヒストグラム特徴間の差 (difference) から，他の参照画像との類似値の上限値を求め，他の参照画像における照合をも省略する．これを並列探索と呼ぶ．これに加えて，多数のヒストグラムを併合したヒストグラムを用いて上限値を決定する手法も導入する．これを併合探索と呼ぶ．以下，これらの手法について説明する．

#### (a) 複数参照画像の並列探索

参照画像  $N$  に参照画像  $M$  と色の異なる画素が含まれる場合を扱えるよう式 (2), (3) を拡張する． $N$  の画素のうち， $M$  と異色な画素 ( $N$  の  $M$  に対する異色画素と呼ぶ) の数は  $|N|(1 - S_{MN})$  で表される (図 4)． $S_{AM}$  からは， $N$  の  $M$  に対する異色画素が  $A$  中に存在するかの情報は得られないため， $N$  の  $M$  に対する異色画素が  $A$  中に現れていると仮定して， $|N| \cdot S_{BN}$  の上限を推定し， $|N| \cdot \theta$  に達しない領域を求める必要がある．すなわち，以下の不等式が導かれる．

$$|N| \cdot S_{BN} < |M| \cdot S_{AM} + |N|(1 - S_{MN}) + n \\ n < |N| \cdot (\theta - 1 + S_{MN}) - |M| \cdot S_{AM} \quad (4)$$

#### (b) 複数参照画像の併合探索

探索する参照画像の間でヒストグラムの違い

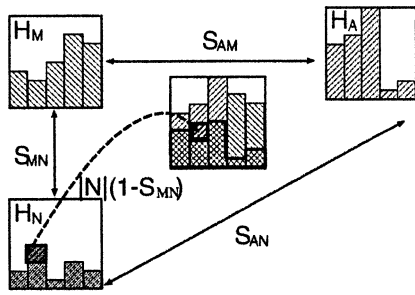


図 4 並列探索の仕組み  
Fig. 4 Relationship between two histograms.

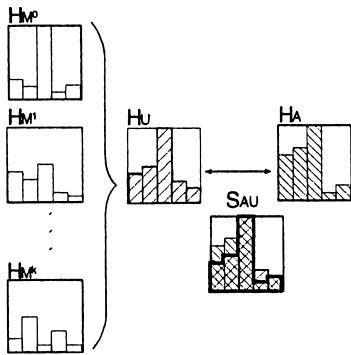


図 5 併合探索の仕組み  
Fig. 5 Relationships between multi-histograms.

が大きい場合には、式 (4) に含まれる加算部分  $|N|(1 - S_{MN}) \cdot |M|$  が大きくなり照合を省略する効果が少ない。そこで、併合探索では並列探索のように類似値にヒストグラムの違いを反映させるのではなくヒストグラム自身に反映させることによって不要な照合を省略させる (図 5)。参照画像のすべてのヒストグラムを  $H_{M^0}, H_{M^1}, \dots, H_{M^k}$  とする。このとき、併合ヒストグラム  $H_U$  を

$$H_{U_i} = \max(H_{M_i^0}, H_{M_i^1}, \dots, H_{M_i^k}) \quad (5)$$

と定義する。 $H_{U_i}$  は  $U$  の各符号の数である。 $H_{U_i}$  の各要素は、各参照画像の最大値から構成されるため、ある探索窓  $A$  と  $U$  との類似値は  $A$  と  $U$  に属する  $M^0, M^1, \dots, M^k$  の類似値よりも必ず大きい。したがって  $|U| \cdot S_{AU}$  は  $|M| \cdot S_{AM}^j (\forall j \in 1, \dots, k)$  の上限値となる。 $U$  を探索することによって  $M^0, M^1, \dots, M^k$  の照合を省略できる。ただし、類似値自体を求める必要がある場合には、 $U$  により検出した領域に関して、更に  $M^0, M^1, \dots, M^k$  を用いて探索する必要がある。

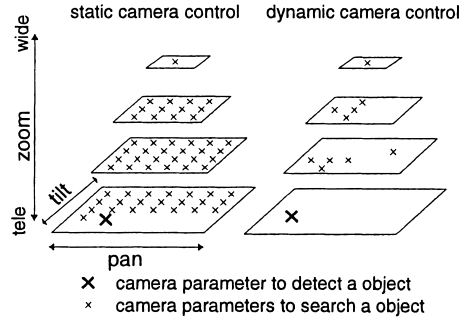


図 6 カメラ制御回数の削減の必要性  
Fig. 6 Necessity for camera control strategy.

### 2.3 高速探索法のカメラ制御への適用

本節では、2.2 で提案した高速探索の結果を利用して動的制御によってパン・チルト・ズームの制御回数を削減する手法を説明する。

ズームの総段数を  $d$ 、ズーム倍率を  $\delta$ 、各ズームの段階を  $i$  とすれば、パン・チルト・ズームで行って得られる総探索空間は、以下の式で表せる。

$$\text{総探索空間} = S \sum_{i=0}^d \delta^{2i} \quad (6)$$

ここで、 $S$  はズーム倍率 1 で撮影した 1 枚の撮影画像を探索するための探索空間である。このとき倍率  $\delta$  を小さくすれば、段数  $d$  が非常に大きくなる。このため、 $\delta$  を物体検出には十分な精度が保証できる範囲で総探索空間を最小にする  $\delta$  を選ぶ。例えば実験で用いた物体の場合には  $\delta = 2 \sim 3$  となる。また、探索精度は参照画像が大きいほど良くなり、例えば 1 辺の画素数が 30 以上の物体の場合には十分な精度があることがわかっている。つまり、ここでは 1 辺の画素数が 30 以上の物体を探索し検出した場合には、探索を終了し、検出できなかった場合にはもう 1 段階ズームインし入力画像をより拡大する。更にパン・チルトを順に動かして視野内をもれなく探索している。これにより大きく写る近い物体から小さく写る遠い物体までを探索することができる (図 6 左)。しかしながら、ズームインを行うと遠い物体が大きく見える反面、一度に撮影できる視野が狭くなるため数多くのパン・チルト・ズーム回数を必要とし、ばく大な探索回数になるために単純な順序で探索する手法では計算時間が非常にかかる。そこであらかじめ物体がありそうな領域を優先してカメラを向ける。つまり、小さい参照画像を探索

した結果を用いて、探索する順序を物体がありそうな領域が最初になるように動的に切り換えることによって物体を検出するまでの探索回数を小さくする(図 6 右)。

以下詳細なアルゴリズムを述べる。なお、以下のアルゴリズム中、1 辺 30 画素以上の参照画像のサイズを「検出サイズ」と呼ぶ。一方、1 辺 10 ~ 30 画素の領域はズームしないと検出物体として確定できないためにこのサイズを「候補サイズ」と呼ぶことにする。  
[初期化]

(1) 最適な総探索空間を実現するパン・チルト・ズーム位置をズーム倍率の昇順に並べる。これをカメラ位置リストと呼ぶ。

(2) 手続き: 探索を実行する。

[手続き: 探索]

(1) カメラ位置リストの先頭の値を、パン・チルト・ズームのパラメータにセットする。カメラ位置リストの先頭の値を削除する。画像を取得する。複数参照画像を用いたアクティブ探索を実行する。

(2) 検出サイズで検出した場合、確かに物体を検出したことを意味するので、探索を終了する。

(3) 候補サイズで検出した場合、検出した領域を含むパン・チルト・ズームパラメータをカメラ位置リストの先頭に並べ換える。複数検出した場合は、候補サイズの降順に並べる(1)へ戻る。

(4) 物体を検出しない場合、カメラ位置リストが空ならば(5)へ。そうでなければ手続き: 探索を実行。

(5) 物体が存在しなかったとして、終了。

### 3. 実験

前章に述べた探索法を用いて、室内環境から小物体を探索する実験を行い、探索速度と探索精度を評価した。

#### 3.1 実験条件

実験に用いたワークステーション、カメラの仕様を表 1 に示す。探索する物体として、図 7 に示すような 3 種類の物体を選んだ。物体の大きさは、高さ約 5cm

表 1 実験に用いた計算機及びカメラの仕様  
Table 1 Specifications of the computer and camera used in the experiments.

計算機	SGI 社 O <sub>2</sub>
CPU	R10000(250 MHz)
カメラ	Sony 社 EVI-D30
解像度	320 × 240

から 10cm 程度である。

まず、学習段階において、対象とする物体を各 100 枚程度適当に撮影し、参照画像とした。具体的には、照明条件を変えるために図 8 の A ~ E の 5 箇所に配置し、物体の正面、左右の計 3 方向を、1.2 倍ずつカメラズームを変え撮影した。

探索段階では、物体をカメラから 5m 以内の近くから遠くまで、図 8 の a ~ e の 5 箇所に配置した。更に、物体の向きを左右に 2 種類変化させた。このとき物体の輪郭形状が異なる場合がある。この場合にもアクティブ探索法は適用可能であるが、より高速に探索を行うため、輪郭形状が相似な物体内部の方形を参照画像として利用した。また、ヒストグラム区画生成に使用した  $\sigma$  は、向きの変化により色の変動が大きいものに対しては、大きい  $\sigma$  を、色の変動が少ないものに対しては小さい  $\sigma$  を物体ごとに選択した ( $\sigma = 1.5 \sim 3$ )。  $\sigma$  は計算速度にはあまり影響を与えなかった。その理由は符号数の減少による計算速度の向上の効果と、背景類似値の上昇による枝刈り効率の低下の効果が相殺しているためと推察される。しきい値は、予備実験により物体ごとに設定した。



図 7 参照画像の例

Fig. 7 Examples of reference images.

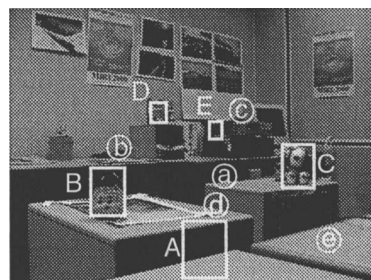


図 8 学習位置と探索位置

Fig. 8 Positions of learned objects and searched objects.

表 2 探索精度

Table 2 Accuracy of multi-template search.

	1 辺 30 画素以上	1 辺 10 ~ 30 画素
従来法	100%	57%
提案法	100%	89%

### 3.2 画像探索の精度評価

2.2.1 に述べた特徴による探索精度を評価するため、等分割のヒストグラム区画を用いた従来法との探索精度の違いを評価した。入力画像としては広角段階で撮影した画像 15 枚を用いた。

精度は適合率 (precision rate) と再現率 (recall rate) の平均値で評価した。ここで適合率とは、探索結果として出力されたもののうち正しいものの割合であり、再現率とは、探索されるべきもののうち探索結果として出力されたものの割合である。適合率と再現率がともに 100% であれば、検出もれや余分な検出がなかったことを意味する。

実験結果を表 2 に示す。この実験では、探索窓の大きさをパラメータとして精度を評価した。

従来法では、参照画像を多数撮影したにもかかわらず 1 辺の長さが 10 ~ 30 画素の範囲では、60% 程度の精度しか達成できなかった。一方、VQ の手法を用いた提案法は、1 辺 10 ~ 30 画素の物体に対しても、90% 近い精度を達成した。

### 3.3 画像探索の速度評価

2.2 に述べた探索法の高速化の効果を評価するため、3.2 と同じ条件で探索速度を測定した (表 3)。

比較した手法は、従来法、VQ による符号化を用いた VQ 法、VQ 法に加え並列探索を行った並列探索法、同様に併合探索を行った併合探索法の計 4 種類である。

まず、VQ により約 2 倍程度、探索時間が向上していることがわかる。並列探索法を用いることによって精度を保証しつつ速度は更に 10% 程度改善している。一方併合探索法では、速度は更に 80% 程度向上する。併合探索により速度が向上している理由は、5 ~ 10 枚の複数参照画像の照合をまとめて一つで探索しているためである。併合探索では、物体検出の取りこぼしはないが別の領域を誤検出する危険はある。しかし、今回の場合、複数の参照画像に対するヒストグラムは類似しているために、完全には精度を保証しないもののカメラの制御に利用するためには十分な精度が得られている。

表 3 入力画像 1 枚当りの探索時間

Table 3 Search time for an image.

	1 辺 30 画素以上	1 辺 10 ~ 30 画素
従来法	2.7s	9.1s
VQ 法	1.7s	3.9s
並列探索法	1.5s	3.5s
併合探索法	0.33s	0.35s

表 4 カメラ制御を含んだ探索時間

Table 4 Search time in room environment.

	時間
従来法	44.6s
VQ 予測法	22.6s
並列探索予測法	20.5s
併合探索予測法	9.6s

### 3.4 カメラ制御への適用

3.2, 3.3 において、実環境中の物体を精度良く、高速に探索できることを示した。本節では、これらの手法を 2.3 に適用した手法について評価する。今回の実験では、検出するために少なくとも 1 度のズームが必要な a ~ c にある物体を探索対象とした。

比較する手法は、予測なしの手法 (従来法)、従来のアクティブ探索に VQ を導入して物体の検出とカメラの動的制御を行う VQ 予測法、並列探索で物体の検出とカメラの動的制御を行う並列探索予測法、並列探索で物体の検出を行いカメラの動的制御を併合探索で行う併合探索予測法の 4 通りである (表 4)。

探索精度はいずれも 100% であった。動的制御により物体検出までの時間約 50% 削減できたことがわかる。更に並列探索で 10%、併合探索で 50% の探索時間を削減できた。しかしながら併合探索で探索窓のサイズが 1 辺 10 画素以下の小さな領域の場合には一部の物体に対して探索候補の誤検出の増加のために、無駄なカメラ制御が多くなり探索時間が若干増加する場合もあった。

## 4. む す び

本論文では、パン・チルト・ズームカメラを用いて実環境からの物体探索を可能とする動的アクティブ探索法を提案した。本手法は、照明条件の変動、物体の向き、カメラパラメータの違いにより物体の見え方が大きく変わる場合においても高速に物体を検出する手法である。本手法はヒストグラムを用いた画像探索法であるアクティブ探索法を発展させたもので、あらかじめ収集した多数の参照画像に対して色分布を VQ を

用いて学習することによって、その精度・速度を改善する。更に、多数の参照画像でも高速に探索することが可能な並列探索法、併合探索法を提案した。これらをカメラの動的制御に適用しカメラの制御回数を削減した。約 5m の範囲内の最小 5cm 程度の大きさの物体をカメラを制御して探索した場合、平均約 10 秒で検出可能であった。これは従来のアクティブ探索法と単純なカメラ制御を行った場合に比較し 4 倍以上の速度向上に相当する。

本手法は照明変動に対処するために学習サンプルを増やす手法であるが、この学習サンプルをいかに効率的に取得するかについては今後検討する必要がある。また、幾何学的構造が異なるが類似した色ヒストグラムをもっている物体や、ある程度大きな隠ぺいなどがある場合にも精度が低下することがある。これらの問題に対しては参照画像を複数領域に分割してサブヒストグラムを構成し、サブヒストグラムごとで照合する手法などが考えられる。今後、本手法を単体のカメラによる探索から複数のカメラを用いた探索に拡張する。複数のカメラがとらえる物体の様々な見え方に対してアクティブ探索法の考えを導入し、複数のカメラが協調して高速に物体を探索する手法について検討を進める予定である。

謝辞 日ごろ御指導頂く NTT コミュニケーション科学基礎研究所、石井健一郎所長、管村昇部長、及び萩田紀博部長に感謝する。また日ごろ、御協力を頂く同研究所メディア認識研究グループの諸氏に感謝する。

## 文 献

- [1] H. Ishiguro, "Distributed vision system: A perceptual information infrastructure for robot navigation," IJCAI-97, pp.36-41, 1997.
- [2] T. Wada and T. Matsuyama, "Appearance sphere: Background model for pan-tilt-zoom camera," Proc. ICPR'96, pp.718-722, 1996.
- [3] K. Yamazawa, Y. Yagi, and M. Yachida, "Obstacle detection with omnidirectional image sensor hyper-omnivision," Proc. ICPR'95, pp.1062-1067, 1995.
- [4] Y. Onoe, K. Yamazawa, N. Yokoya, and H. Take-mura, "Visual surveillance and monitoring system using an omnidirectional video camera," Proc. ICPR'98, pp.588-592, 1998.
- [5] 谷内清剛, 和田俊和, 松山隆司, "視点固定型パン・チルト・ズームカメラを用いた適応的見え方モデルに基づく人物頭部の検出・追跡," Proc. MIRU2000, vol.I, pp.9-14, 2000.
- [6] 伊藤 渡, 上田博唯, "広域監視用親カメラと追尾監視用子カメラによる協調監視システム," Proc. MIRU2000, vol.II, pp.235-240, 2000.
- [7] Y. Abe, M. Shikano, T. Fukuda, F. Arai, and Y. Tanaka, "Vision based navigation system by variable template matching for autonomous mobile robot," Proc. ICRA'98, pp.952-957, 1998.
- [8] G.-Y. Jiang and T.Y. Choi, "Robust detection of landmarks in color image based on fuzzy set theory," Proc. ICSP'98, pp.968-971, 1998.
- [9] 村瀬 洋, S.K. Nayar, "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," 信学論 (D-II), vol.J77-D-II, no.11, pp.2179-2187, Nov. 1994.
- [10] 田中弘一, 佐野睦夫, 大原秀一, 奥平雅士, "パラメトリックテンプレート法とその高精度ロバスタッチングへの応用," 信学論 (D-II), vol.J83-D-II, no.4, pp.1119-1130, April 2000.
- [11] 村松彰二, 大塚裕史, 小林芳樹, 清水英志, "テンプレートマッチングの高速化戦略と ga による最適化," 信学論 (D-II), vol.J83-D-II, no.6, pp.1487-1497, June 2000.
- [12] 村瀬 洋, V.V. Vinod, "ヒストグラム特徴を用いた高速物体探索法—アクティブ探索法," 信学論 (D-II), vol.J81-D-II, no.9, pp.2035-2042, Sept. 1998.
- [13] 柏野邦夫, ガビン スミス, 村瀬 洋, "ヒストグラム特徴を用いた高速探索法—時系列アクティブ探索法," 信学論 (D-II), vol.J82-D-II, no.9, pp.1365-1374, Sept. 1999.
- [14] 柏野邦夫, 黒住隆行, 村瀬 洋, "ヒストグラム特徴を用いた音や映像の高速 and/or 探索法," 信学論 (D-II), vol.J83-D-II, no.12, pp.2735-2744, Dec. 2000.
- [15] 黒住隆行, 柏野邦夫, 村瀬 洋, "時系列アクティブ探索法のための特徴変動に頑健な確率ディザボータリング," Proc. MIRU2000, vol.I, pp.313-318, 2000.
- [16] B. Schiele and J.L. Crowley, "Probabilistic object recognition using multidimensional receptive field histograms," Proc. ICPR'96, pp.50-54, 1996.
- [17] 川西隆仁, 村瀬 洋, "アクティブ探索法のパン・チルト・ズームカメラを用いた物体探索への適用," Proc. MIRU2000, vol.II, pp.31-36, 2000.
- [18] N.M. Nasrabadi and R.A. King, "Image coding using vector quantization: A review," IEEE Trans. Commun., vol.36, no.8, pp.957-971, 1988.
- [19] C.M. Huang and R.W. Harris, "A comparison of several vector quantization codebook generation approaches," IEEE Trans. Image Processing, vol.2, no.1, pp.108-112, 1993.
- [20] G.F. McLean, "Vector quantization for texture classification," IEEE Trans. Syst., Man & Cybern., vol.23, no.3, pp.637-649, 1993.

(平成 12 年 10 月 2 日受付, 13 年 1 月 30 日再受付)





川西 隆仁 (正員)

平 8 京大・工・情報卒．平 10 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，日本電信電話(株)入社．物体探索，アクティブビジョンの研究に従事．



村瀬 洋 (正員)

昭 53 名大・工・電子卒．昭 55 同大学院修士課程了．同年日本電信電話公社(現 NTT)入社．以来，文字・図形認識，コンピュータビジョン，マルチメディア認識の研究に従事．平 4 から 1 年間米国コロンビア大客員研究員．現在，日本電信電話(株)NTT コミュニケーション科学基礎研究所メディア認識研究グループリーダー．工博．昭 60 本会学術奨励賞，平 4 電気通信普及財団テレコムシステム技術賞，平 6 IEEE-CVPR 国際会議最優秀論文賞，平 7 情報処理学会山下記念研究賞，平 8 IEEE-ICRA 国際会議最優秀ビデオ賞，平 13 高柳記念電子科学技術振興財団高柳記念奨励賞各受賞．情報処理学会，IEEE 各会員．