

# 対象物把握のためのロボットビジョン技術

Robot Vision Technology for Target Recognition

西山 学

■ NISHIYAMA Manabu

オフィスや家庭など、人間が活動する複雑な環境でロボットが各種サービスを提供するためには、その対象物の相対的位置や姿勢を把握することがまず必要となる。画像を用いてこれを実現する場合には、対象物までの距離や相対姿勢の変化、隠れなど、対象物の状態によって生じる見え方の変化に対してロバストであることが求められる。反対に、対象物の見え方からその姿勢などを推定することも必要になる。

東芝は、このような要求に応えるロボットビジョン技術として、距離変動に対してロバストに平面ランドマークを検出するアルゴリズム、及び方向にかかわらず食器など円形状を持つ対象物を検出するアルゴリズムを開発した。

It is necessary for robots to be able to recognize the relative positions and postures of target objects, in order to offer various types of services in complex environments such as homes and offices where humans are acting. When visual information is used, it should be robust against changes in the appearance of objects. On the other hand, it is also necessary for robot vision to be able to estimate the posture of objects from changes in an object's appearance.

To meet these requirements, Toshiba has developed a robot vision algorithm for planar landmark detection that is robust against changes in distance, and for dish detection without dependence on the direction.

## 1 まえがき

ロボット技術には現在、工場など事前に整備が可能な環境だけでなく、施設や家庭など、より複雑な環境でも使用できるロバスト性が要求されるようになってきている。これらの複雑な環境は、ロボットのために整備されているという前提を設けられない。

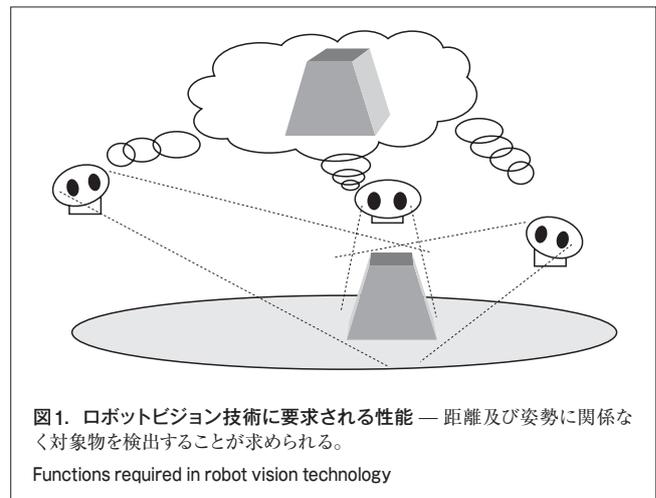
このような環境の中でロボットが物理的サービスを実現するには、周囲の状況把握が必要になる。例えば、ロボットが物をつかんでどこかに運ぶには、運搬対象物の位置や方向、ロボットの位置や方向などの情報を把握しておかなくてはならない。

人間が自分の周囲からの情報の多くを視覚から得ているのと同様、ロボットが周囲の状況を把握する場合も、視覚情報及びそれを処理するロボットビジョン技術が利用される例は数多い。

ここでは、東芝が開発した、視覚情報を用いて周囲の物体との位置関係や状態を把握する技術について述べる。

## 2 ロボットビジョン技術に要求されること

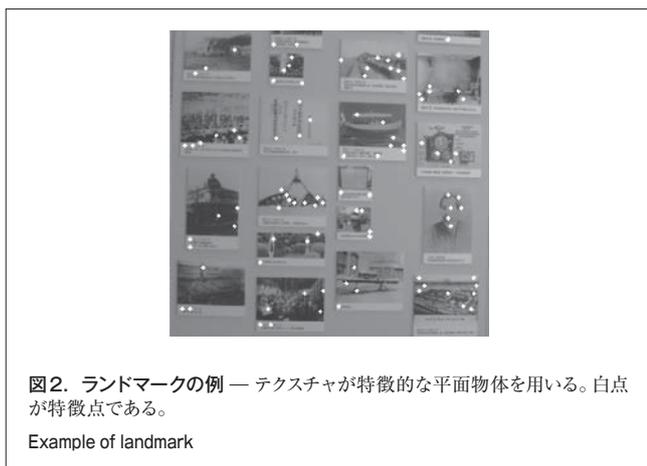
人間が活動する環境は、多くの場合ロボットのために整備されていない。つまり、作業しようとする対象物が必ずしも決まった位置や向きで置かれているとは限らず、照明環境も一定であるとは限らない。したがって、視覚情報を利用して対象物



を認識し、周囲の状況を把握するには、これらの要因による見え方の変化に関係なく対象物を認識できることが必要である(図1)。また、反対に、見え方の変化から対象物の位置や姿勢、ほかの物体との位置関係を推定することも、信頼性の高い動作を実現するのに必要とされる。

## 3 決定木を用いた平面ランドマークの認識

ロボットが移動するときの自己位置の把握は、荷物の運搬や施設内の案内といった作業を確実に遂行するために必要と



される。これを視覚情報で実現する、つまり、目印となるランドマークを検出し、それとの相対位置を計測するには、次の2点が求められる。

- (1) 距離や向きが変化しても、検出可能であること
- (2) 見え方の変化から距離や向きを推定できること

従来は、四角の枠など定型の図形をランドマークに利用することが行われてきたが、オフィスや家庭内にあるカレンダーやポスターなど特徴的な図柄をそのままランドマークとして利用できれば、ロボットのための環境整備の必要がなくなる。このようなランドマークを用いて主に施設内の案内作業<sup>(1)</sup>を実現するために、ステレオ視と決定木を用いたランドマークの検出及びロボットの位置姿勢の把握技術を開発した。

### 3.1 ランドマークの条件

ランドマークには、図2のように、平面で模様(テクスチャ)が特徴的な図形を利用する。模様のない壁面などは特徴が少なく、利用できない。

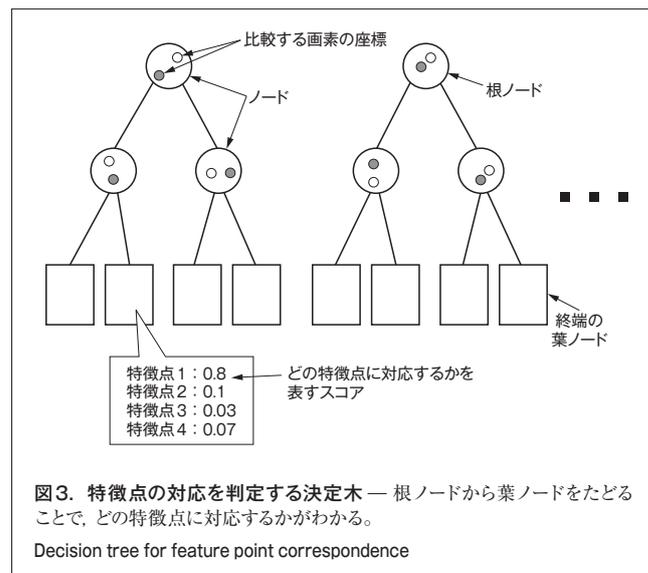
ランドマークは、この図中にプロットした点で表される特徴的な点<sup>(注1)</sup>(以下、特徴点と呼ぶ)を中心とする、小さな部分画像の集合として学習する。入力画像からのランドマークの検出は、画像からこの部分画像を見つけだした後、ランドマーク中の配置との適合をチェックすることで実現する。このときの部分画像の配置情報を使うことにより、ロボットからランドマークまでの距離と方向を把握する。

### 3.2 ランドマークの学習と識別

ランドマークから、事前に特徴点周辺の部分画像パターンを学習する。ランドマークを検出するときは、次の手順で行う。

- (1) 入力画像から特徴点を抽出
- (2) 特徴点周辺の部分画像を切り出して、ランドマーク中のどの特徴点に対応するかを判定
- (3) ランドマークと対応した部分画像の位置関係から、検出の成否の判定及び距離や姿勢の推定を実行

(注1) 輝度の変化が大きなコーナ点を画像から抽出するHarrisのコーナ点抽出法で抽出。



この中の(2)の判定を行うため、部分画像を入力として、対応するランドマーク中の特徴点を出力する識別器を作成する。

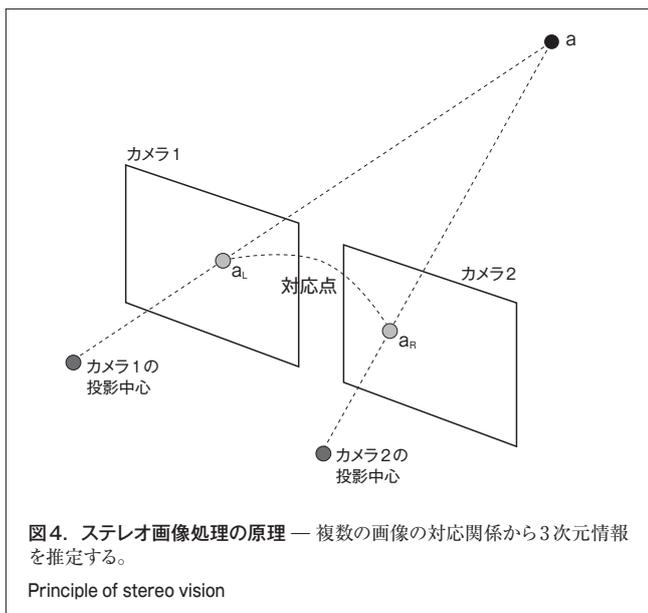
識別器には図3に示すような決定木による手法を用いている。決定木中のそれぞれのノード(節点)には、二つの点の座標値が記録されている。部分画像中のこの二つの画素の輝度の大小関係に応じて、下の二つのノードのどちらに遷移するのかが決定する。終端の葉ノードには、どの特徴点に対応するかを表すスコアが記録されている。複数の木を用意しておき、それぞれを合計したスコアで特徴点の対応を判定している。

決定木の学習は、ランドマークの特徴点周辺の部分画像をうまく切り分けられるノードを決めていくことに相当する。このとき、部分画像に対して回転、拡大又は縮小などの変換を施したものを加えて学習し、位置や向き、照明環境の変化に耐えられるようにする。学習にはノードの数(=木の高さに対する指数関数)に比例する時間を要するが、一方で、識別には根ノードから葉ノードをたどるため、木の高さに比例する時間で済む。識別時は高速に処理できるのがこの技術の特長で、ロボットが移動中のリアルタイム処理に適している。

### 3.3 距離変動に伴う見え方の変化への対処

以上のような決定木を利用して、距離や方向、明るさの変動に対してロボストにランドマークを検出できるようになった。しかし、対処できる変動にも限度があり、見え方の変化が大きくなると検出に失敗するようになる。特に、距離に反比例して見かけの大きさが変化するため、距離変動にあまり強くないということがわかってきた。

そこで、特徴点を抽出した後、ステレオ画像処理を行って特徴点までの距離を計測するようにした。ステレオ画像処理とは、図4のように、二つのカメラ間で同一箇所を撮影した対応点 $a_L$ ,  $a_R$ を求め、三角測量の原理で元の $a$ の3次元座標を推定することである。これを用いれば、画像中の各特徴点の3次



元座標を推定できる。一方、決定木はランドマークまでの距離を変化させた複数の画像を準備し、個別に決定木を作成した。そして、ステレオ画像処理によって得られる距離から、どの決定木を用いるか選択するように改良した。

ステレオ画像処理は、対応点を求めることができれば確実性が高い。一方、決定木を使った処理は、距離の変動が小さければ信頼性は高い。二つの処理を組み合わせることで、距離の大幅な変動があっても検出できるようになった。

### 3.4 処理結果の例

ランドマーク検出の処理結果例を図5に示す。画像の上半分がランドマーク、下半分が入力画像を表す。二つを結ぶ線分は特徴点の対応関係を表しており、入力画像中の枠はランドマークの検出結果を表している。このときに用いた決定木



は4段階で、各段階が20本、高さ15の木で構成されている。距離が1.4～4.0 mの範囲でのランドマークの検出と、距離及び姿勢の計測が可能であることを確認した。

## 4 だ円当てはめを用いた食器検出処理

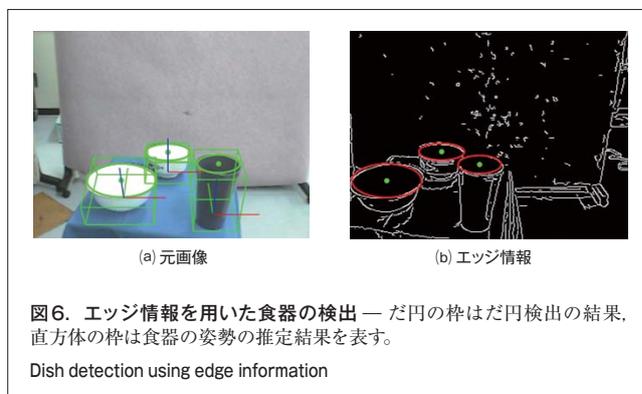
ロボットでは、自己位置の把握と並んで、マニピュレーション（操作）作業時の把持物体の位置と姿勢の把握にも視覚情報が多く用いられる。当社のロボット技術は、人間に身近な食器などの物体を作業対象としており、人間の近くで活動するロボットの実現を目指している。そのためには、食器類の位置と姿勢を把握して、ロボットアーム及びハンドを正しく制御することが必要となる。

やはりここでも、対象物の位置や姿勢、照明条件の変化などにロバストに対応する必要がある。更に、食器類を対象とすることから、テクスチャの少ない物体の検出を実現しなくてはならない。そこで、エッジ情報を使ってだ円抽出とステレオ処理を行うことで食器の3次元位置を検出し、そのだ円の形状から食器の姿勢を推定する技術を開発した（図6）。

### 4.1 画像中のエッジを用いただ円物体検出

だ円の方程式は二つの変数の2次式で表され、パラメータ空間は5次元となる。多数の点からある形の方程式のパラメータを求める場合、Hough変換<sup>(注2)</sup>のような多数決に基づく手法が有効である。しかし、求めるパラメータが多いため、ロボット作業にとって現実的な時間で処理するのは難しい。そこで、以下のような手順で処理している。

- (1) ステレオ画像を撮影し、エッジ画像を算出（エッジ抽出は一般的なCannyオペレータ<sup>(注3)</sup>を使用）
- (2) エッジ画像中のエッジ集合を連結成分ごとに分割
- (3) エッジ部分集合に対し、だ円の方程式を当てはめ、誤差の小さくなるパラメータを算出
- (4) 求められただ円の方程式の誤差を評価



(注2) デジタル画像処理で特定の図形を検出する手法の一つ。

(注3) 隣接する画素の輝度差に基づきエッジ（輪郭線）部分を抽出する手法。

- (5) 評価値の高いだ円に対してはステレオ視で食器の3次元位置を求め、だ円形状からその姿勢を算出
- (6) (3)~(5)の繰返し

この手順(4)において、算出されただ円パラメータ候補の“良さ”を定量的に評価する必要がある。だ円検出の場合、パラメータの良さとは、画像中のエッジ部分とどれだけ一致するかということである。そこで、求めただ円に対し、その軌跡上の複数の点から最近傍エッジまでの距離を測定し、この距離の総計が小さいほど適合度は高いと評価する。この評価値に基づいてだ円パラメータを評価し、あらかじめ決めたしきい値よりも小さければ、だ円が検出されたと判断する。

#### 4.2 だ円からの位置姿勢推定

以上により、画像中のだ円形の物体を検出できる。円形の縁は姿勢に関係なく画像上でだ円形として投影されるため、円形の縁を持った食器であれば姿勢や距離に関係なく検出できる。こうして検出されただ円から、更に3次元の位置や姿勢の情報を推定することで、ロボット作業に必要な情報を得られるようにする。

まず、食器の3次元位置はステレオ画像処理を利用して算出する。ただし、3章の技術を含めた一般的なステレオ画像処理では、左右画像の画素値の差を用いて対応点を探索するのにに対し、テクスチャの少ない食器の場合、画素値を使った対応点の探索は成功率が低い。そのため、一方で抽出されただ円をもう一方のエッジ画像上で動かしながら、適合度が最もなる位置を求めることで対応点の探索を実現している。

一方、姿勢の推定は図7のように、円形の物体が画像に投影されるプロセスを考え、だ円形のへん平率と軸の傾きから姿勢のパラメータを逆算することにより求めることができる。

#### 4.3 処理結果の例

以上で述べた処理結果の例を図8に示す。画像中の直方体の枠は推定された食器の姿勢を表している。ステレオ画像

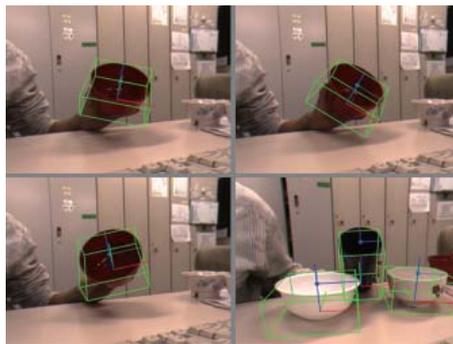
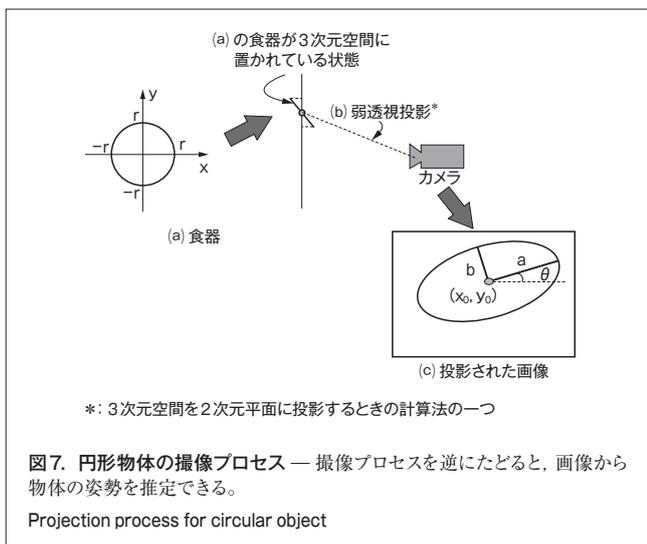


図8. 食器検出結果の例 — 姿勢の変化や複数の物体も検出できる。  
Example of dish detection

処理により食器の実サイズを取得できるため、大きさに応じてロボットハンドでの把持方法を変えることが可能になっている。

## 5 あとがき

ロボットビジョンでは、位置と姿勢などの変動に伴う対象物の見え方の変化に対応できることが求められている。そのような要求に応えるロボットビジョン技術として、移動やマニピュレーションにおける技術を説明した。現状は平面や円形など対象物の形状に制約を設けているため、今後は制約のない一般的な物体に適用できるように技術を発展させていく。

なお、この研究の一部は、独立行政法人 新エネルギー・産業技術総合開発機構 (NEDO) の次世代ロボット共通基盤開発プロジェクト「運動制御用デバイス及びモジュールの開発」の支援により行われた。

また、別の一部は、内閣府総合科学技術会議の推進する科学技術連携施策群 次世代ロボット連携群における補完的課題として、文部科学省の科学技術振興調整費により「科学技術連携施策群の効果的・効率的な推進」の一環として実施された。

## 文献

- (1) 十倉征司, ほか. “画像処理RTCと移動制御RTCを用いた案内ロボット”. ロボティクス・メカトロニクス講演会2008予稿集. 長野, 2008-06, 日本機械学会. 2008, 1P1-E17. (DVD-ROM).



西山 学 NISHIYAMA Manabu

研究開発センター マルチメディアラボラトリー。  
画像認識, ロボットビジョンの研究・開発に従事。日本ロボット学会会員。  
Multimedia Lab.