

移動ロボットの視覚に期待されるもの

油田 信一[†]

友納 正裕[†]

移動ロボットにとって、環境をセンシングし、それによって環境の様子と自分の状態を知ることは不可欠な機能である。とくに、画像を用いるロボットの視覚は、得られる情報量が大きいため、最も期待される環境センシング法である。本稿では、移動ロボットの視覚に求められる機能を述べ、従来の研究例、とくに、筆者らのグループで行ってきた方式を紹介する。

Functions expected to vision systems of mobile robots

Shin'ichi YUTA[†]

Masahiro TOMONO[†]

Sensing the environment is indispensable for a mobile robot to know the circumstances and its own states. Robot vision based on camera images is one of the most promising methods since it provides a large amount of information. This paper considers the functions expected to vision systems of mobile robots, and presents some of the approaches which have been developed by the authors.

1 はじめに

移動ロボットの特徴は、その行動空間が広いことである。据え置き型のマニピュレータの動作範囲がその本体の大きさにほぼ等しいの対して、移動ロボットの行動空間は自分自身の大きさの数倍から数千倍に達する。このように行動する空間が広いと、その行動空間についてあらかじめモデル化することが難しく、またモデル化したとしてもその精度が悪かったり環境が変化してしまったりする。このため、移動ロボットがある空間内で動く際に、環境を完全に既知なものとして扱うことはできず、何らかのセンサで環境の様子を知りつつ動くことが不可欠である。本稿では、移動ロボットが環境と自分自身の状態を知るための環境センシング方法のうち、とくに視覚について、それが期待される理由を述べ、これに求められる機能と、従来の研究、とくに、筆者らの研究室で行われてきた研究を紹介する。

2 移動ロボットの外界のセンシング

移動ロボットが時々刻々の環境の様子を知るためには、ロボットが外界を常時センシングすることが不可欠である。非接触にセンシングしようとする、電磁波または音響波を用いるしかない。一般に、波によ

て外界を計測・認識しようとする時、以下のことが重要となる。

- 検出を必要とする距離を波が伝播するか。
- 物体の検出に必要な分解能が得られるか。
- 波を情報として取り込むための変換器があるか。
- センシングがアクティブかパッシブか。

ここで、アクティブなセンシングとは、ロボットから電磁波や音響波を放射しその反射を検出して環境を知ることであり、パッシブなセンシングとは、自然に存在する電磁波や音響波を受けて環境を知ることである。

以下、地上(の空気中で)動く移動ロボットについて、上記の観点から、波による外界センシングの利用可能性を考える。

(1) 音響波によるセンシング

音響波では、人間の可聴周波数を超えた超音波が移動ロボットの外界センシングによく用いられる。超音波は、波長が短くなるほど空気中の伝播損失が大きくなり、100KHz(波長 3.4mm)では数メートル程度の伝播距離が限度である。

分解能については、超音波は電磁波に比べて波長が長く指向角が大きいため、方位分解能は良くない。たとえば、波長を 8.5mm(40KHz)とすると、センサデバイスの開口を 30mmとしても、指向角は 8.5 度程度の値となる。これが、超音波センシングの最大の弱点である。一方、音速は遅いので音波の伝播時間が精度

[†] 筑波大学 機能工学系 知能ロボット研究室
Intelligent Robot Laboratory, University of Tsukuba

よく計測できるため、距離分解能はよい。パルス波を発射して対象物体までの距離を計測する超音波パルスエコー方式では、1マイクロ秒の精度で時間を計れば0.17mmの分解能が得られる。

また、自然界には適切な超音波源はふつう存在しないので、アクティブなセンシング法が一般的である。

(2) 光および電磁波によるセンシング

光および電磁波の伝播距離は波長によるが、一般に超音波より長い。ただし、可視光は霧やホコリの中では散乱してしまう。波長の長いマイクロ波等は伝播損失がはるかに小さく、霧やホコリでも伝播距離は長い。

光は超音波に比べて波長がはるかに短く、高い空間分解能を実現できる。たとえば、1cmの開口のレンズで可視光の波長を600mmとすると、原理的には空気中の超音波センシングに比して 10^6 オダ優れた方位分解能が得られる。一方、電磁波の伝播速度(光速)はきわめて速く(0.3×10^6 km/秒)、対象物体までの伝播時間を用いた距離計測は容易ではない。だが、レーザのようなコヒーレントな光を用いる方法が開発され、測量用のレーザ距離センサ、あるいはそれをスキャンして円周方向の距離データを測るレーザレンジセンサは、原理的に光の伝播時間による計測であり、約1cmの距離測定精度が実現されている。

光および電磁波は、アクティブ/パッシブ両方のセンシングが可能である。すなわち、レーザ等を対象物体に照射してその反射波を検出することも、自然光のもとで撮影した画像をセンシングデータとして利用することも行われる。

(3) 画像によるセンシング

カメラを用いて光をセンシングすると、いわゆる画像が得られる。可視光を用いれば方向分解能がきわめて高くなり、超音波やマイクロ波に比べて圧倒的に大きな情報量が得られる。

また、基本的にパッシブであるから、太陽光や屋内照明などの自然の光源を利用できる。センサデバイスも、一般に普及しているカメラを利用できる。また、人間が見ているとおりの原データに基づくため、処理方法を直観的に理解しやすいという利点もある。

このように画像は種々の点で優れた特性をもつが、問題点として次の2つがある。

- 得られる原データの情報は大きいですが、その中から必要な情報を抽出するのは容易ではない。
- 移動ロボットにとって、まず必要なことはロボットの周りの3次元的な空間構造を知ることであるが、画像情報はこれを直接与えるものではない。

この問題に対処して画像の優れた特性を生かすことが、移動ロボットのビジョンにも強く期待される。

3 移動ロボットのタスクと環境

移動ロボットに求められる動きは、目的の場所へ移動すること、あるいは、移動しながら何らかの作業を

遂行することである。この移動には、自律的な移動と遠隔操作による移動がある。いずれの場合も、その目的は、

- 地図情報とその地図上での目的地が与えられ、目的地まで走行する。
- 目的物を探しながら目的地まで移動する。
- 決められた領域をくまなく走行するなどして、その環境の情報を得る。

などである。ただし、遠隔操作による移動では、操作者が移動体の車輪や脚の動作を直接操作したり、比較的近い位置に目的地を指示してロボットがそこまで自律的に移動することを繰り返したりして、所定の目的を果たす。

一方、移動体が動く環境には、(1)世界がロボットが走行すべき「道」によって構成され、ロボットにはその道に沿って移動することが求められるもの、(2)部屋の中や広場あるいは荒野のような世界で、どこでも自由に移動できるし、そうすることが求められるもの、がある。後者については、さらに(2a)世界がロボットの移動できるほぼ平らな床面とそうでない(ロボットが入り込むことのできない)領域に分けられる屋内や街路のようなもの、(2b)屋外の荒地のように地面には凸凹があり、どこが通れるのか、あるいは通りやすいのかを考えなければならないもの、があるだろう。

さらに、いずれの環境についても、

- 正確な環境情報(地図)があらかじめ与えられている。
- 概略の環境情報は与えられているが、その精度が不足だったり、時々刻々に変化があったり他の通行人がいる。
- 環境情報が事前に与えられない。

などの場合がある。

移動ロボットの環境センシングに求められるものはこれらの各々の場合で異なっており、それぞれの目的に添って方式を検討する必要がある。だが、いずれの場合でも、その環境で目的を達成するためには、ロボットは最低限、常に、

- どこに向かって動くか。
- 行こうとする所が走行できるのか否か。

を判断し続ける必要がある。この判断は、あらかじめロボットが持つ環境情報と与えられた目的地、および、現在とそれまでに自らのセンサで得た環境情報によって行うこととなる。

4 移動ロボットの視覚に求められる機能

移動ロボットが遂行すべきタスクと働くべき環境を考えると、ロボットの視覚に求められる機能は次のように分類されよう。

- ロボットの自己位置推定のための視覚
ロボットが何らかの地図を有している時、地図上に記載されたランドマークを発見してその位置や方位を用いて自己の位置を推定する。
- ロボットが地図を生成するための視覚
ロボットが走行しながら、自分が環境から得た情報に基づいて、持っている地図の誤差を修正したり、新たに地図を生成する。
- ロボットが安全に走行するための視覚
ロボットが走行中、常時前方を監視し、衝突し得る障害物を検出する。また、脚ロボットが脚を地面におろす時、その位置が安全かどうかを判断する。このためには、事前にモデルを持つことは難しく、ラフでもよいが3次元の形状を検出することが不可欠である。
- ロボットが物に対して作業を行うための視覚
移動ロボットが対象物を発見・選別し、マニピュレータで何らかの操作をするためにその位置や姿勢を正確に検出する。移動ロボットに搭載するため、簡便なハードウェアと処理量で実現できることが望ましい。

以下の各章では、この各々についてやや詳しく述べ、従来研究、とくに筆者らの研究事例を紹介する。

5 自己位置推定

移動ロボットが目的地に正しく移動するには、ロボットが自分の位置姿勢を常に正しく把握する必要がある。本章では、自己位置推定におけるビジョンの役割と研究例を紹介する。以下では、ロボットは環境の地図をもち、平面上を移動するものとする。平面上を移動するロボットの位置姿勢は、位置と方位 (x, y, θ) の3次元で決まる。

5.1 移動ロボットの自己位置推定

移動ロボットでは、デッドレコニングにより自己位置推定を行う方法が一般的である。とくに、車輪型移動ロボットでは、オドメトリにより車輪の回転数を計測することでこれを実現する。だが、デッドレコニングでは誤差が累積するため、長距離を走行すると、ロボットの位置のずれが大きくなるという問題がある。このため、外界センサを用いて場所の目印であるランドマークを検出し、地図上の位置を参照することで自己位置を修正することが行われる。すなわち、ロボットがランドマークを観測して得るのはロボットから見たランドマークの相対位置姿勢であるが、地図からランドマークの絶対位置姿勢がわかれば、ロボットの絶対位置姿勢を計算することができる。

この枠組を実現する方法として、現在では確率的アプローチが主流となっている。そこでは、ロボットの位置姿勢を確率分布で表し、走行とともにその誤差分散が増大するフェーズと、ランドマークの観測データとそれまでのロボット位置姿勢の確率分布を融合して誤

差分散を更新するフェーズとから自己位置推定を実現する。これを Kalman Filter や Particle Filter を用いて、効率よく実現する方法が提案されている [4, 5, 38]。

5.2 自己位置推定のための視覚

自己位置推定において、ビジョンは主にランドマークの検出に用いられる¹。用いるビジョンの種類に応じて、それに適したランドマークを選択する必要がある。たとえば、画像エッジ、色、画像の主成分ベクトル、物体などがランドマークとして用いられる。ビジョンとランドマークの種類を選択する上で重要なのは、(1) どの程度の位置情報を得るか、(2) 地図情報とどう対応をつけるか、である。なお、ランドマークには、検出が容易のように人工的に作成・設置したものと、環境中に自然に存在するものがある。

ランドマークから得る位置情報は完全なものでなくともよい。1回のランドマーク観測からは (x, y, θ) の一部分しか得られなくても、複数のランドマークの観測結果を融合すれば、各次元で不確実性を減らすことができる。さらに、オドメトリの累積誤差は x, y, θ 間で相関をもつため、どれかの値を修正すれば、相関を利用して他の値も修正することができる。

このように、ランドマークから得られる位置姿勢が部分的でもよいので、原理的にはさまざまなものがランドマークとして利用可能である。たとえば、壁をランドマークとして、超音波センサやレーザレンジセンサを用いて壁までの距離データを測定することで自己位置を修正することも広く行われている。また、単眼ビジョンを用いて、ランドマークの方位情報だけを利用して、自己位置を修正することも行われる。だが、1回のランドマーク観測から得られる位置情報が豊富なほどよいのは言うまでもない。そこで、ランドマーク検出にかかるコストと、そこから得た位置情報による自己位置修正処理の精度や効率とのトレードオフにより、用いるセンサとランドマークを決めることになる。

観測データと地図情報との対応は、重要な問題である。自己位置を推定するには、いま観測したランドマークが地図上のどのランドマークであるかを同定しなくてはならない。このためには、一意に識別できるランドマークが望ましい。人工ランドマークでは比較的容易に一意性を実現できるが、作成・設置の手間がかかる。環境中に自然に存在するランドマークを用いる場合は、一意性はあまり保証できないが、環境に手を加えない点で望ましい。ランドマークを複雑な特徴から構成するほど、識別は容易になるが、検出処理も難しくなるという問題がある。

センサデータだけでは一意に識別できないランドマークは、その位置情報を識別の手がかりにすることが行われる。たとえば、画像エッジのような単純な特徴は、環境中に多数存在するため、それだけでは一意に識別できない。そのため、目標のエッジの環境内で

¹車輪の回転数でなくビジョンによりオドメトリを構成することも可能である。ビジョンを用いてもロボットの移動量を測定していることに変わりはないので、累積誤差の問題が生じる。

の位置を地図から取得し、ロボットの現在位置姿勢から画像中に見えるエッジ位置を予測することがよく行われる。ロボット位置姿勢の誤差が比較的小さく地図も正確であれば、この方法は有効であり、単純なセンサでも利用可能なことから、従来の移動ロボットで広く用いられて来た。

近年、自己位置推定は、単純なランドマークを用いるものから、複雑なランドマークを用いるものへと移っている。これは、CPUの処理能力の発達にも依存している。当然、用途に応じて使い分けるべきであるが、複雑なランドマークを用いると効率やロボスタ性の向上が期待できる。これが、他のセンサに比べ圧倒的な情報量をもつビジョンに期待されることと考える。

5.3 従来の研究例

ビジョンを用いた自己位置推定では、多くの方法が提案されている。屋内など人工環境では、垂直エッジが安定かつ容易に抽出できるため、これをランドマークとして利用したシステムが早くから開発されている[26, 35]。また、エッジの組合せによってドアや窓などの構造をもった物体をモデル化し、これをランドマークとして用いるシステムも多い[1, 2, 6, 8, 9, 11, 12, 31, 32, 33]。このような幾何学的モデルをランドマークとして用いる方法は、自己位置推定の精度が比較的良好という利点があるが、モデル作成に手間がかかり、ランドマーク検出処理も一般に複雑になるという問題がある。このため、幾何構造に関する複雑な処理を避けて、色やテクスチャ、画像の主成分分析など、画像の見え (appearance) のマッチングを用いる方法も多く提案されている[3, 7, 18, 24, 34]。

筆者らのグループでの初期の研究として、坪内・油田[33]の幾何モデルを用いたシステムがある。このシステムでは、屋内環境の3次元モデルを多角形の組合せによりあらかじめ構築しておき、ロボットは単眼カメラ画像から色情報を用いてドアなどの領域を抽出して、環境モデルとマッチングして自己位置の修正を行う。ただし、色情報だけでは正確なモデルマッチングは難しいという問題がある。

性質のよいランドマークならば、単純でも自己位置推定に利用できる。たとえば、Launay・大矢ら[14]は、単眼カメラを用いて天井の蛍光灯を認識し、ナビゲーションを行うシステムを開発した。ロボットは蛍光灯の位置と向きから自己位置を修正する。天井の蛍光灯は、オクルージョンで隠れる恐れが少ないうえ、発光していれば検出が容易なため、単眼ビジョンのランドマークとしてかなり適している。

また、部屋番号を一意的なすぐれたランドマークとして利用したシステムもいくつか開発されている[16, 25, 30]。友納・油田[30]は、部屋番号を認識してナビゲーションを行うロボットを開発した。ロボットは、廊下を走行中に部屋番号プレートを見つけ、その中に書かれた文字を抽出して部屋番号を認識する。部屋番号プレートは小さくて遠くからでは見つけにくいので、まずドアを色と垂直エッジを用いて認識し、次に、ドアの近くにある四角形を部屋番号プレートとし

て検出する。ドアという大きな物体を手がかりとすることで、正確な地図を用いなくても部屋番号の検出に成功した。図1に処理例を示す。

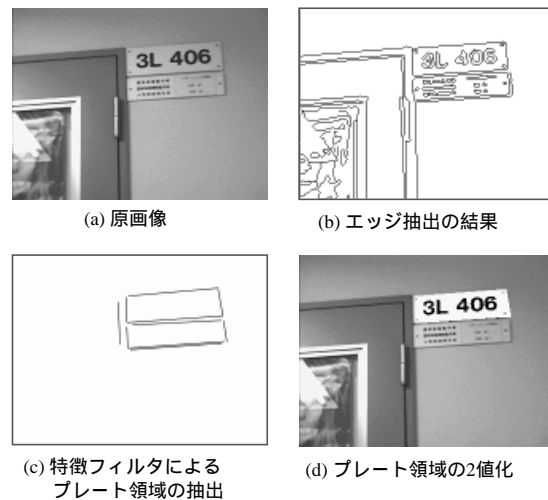
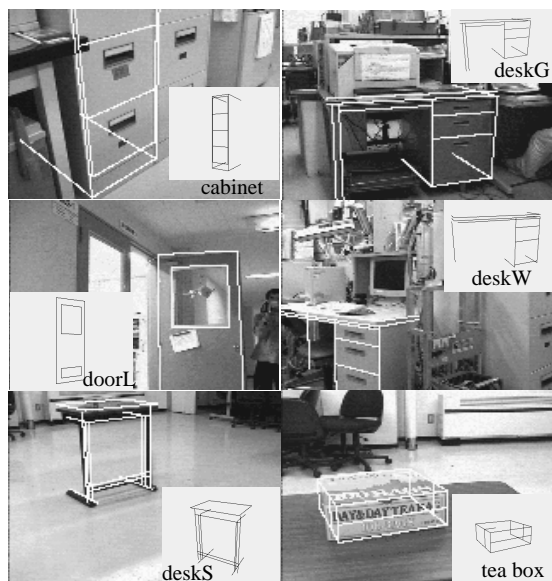


図 1: 部屋番号の検出処理

エッジの組合せによって環境中の物体をモデル化し、構造をもったランドマークとして利用すれば、観測データと地図情報との対応づけの一意性の問題は、単純なランドマークの場合よりもはるかに軽減される。また、物体を認識できるので、それに応じた動作も計画しやすくなる。このアプローチでは、ロボットの現在位置姿勢と地図情報から、画像中でランドマークがどのように見えるかを予測して効率と精度を高めることが多い[12, 2, 6, 8]。

友納・油田[31, 32]は、物体認識技術を用いて、画像中のランドマークの位置予測がなくてもロボスタに物体を認識して自己位置推定を行うシステムを開発した。この認識手法は、エッジの並進不変量や特徴グルーピングを利用した汎用性の高いものであるため、直線分のワイアーフレームでモデル化できるものならば、ドアや机、キャビネット、廊下の角など環境中の多くの物体に適用できる。図2に、物体認識の例とナビゲーションの様子を示す。

全方位ビジョンは、ロボットの全周囲が見えるため、ランドマークを撮影するためのカメラ操作が必要ないという利点がある。このため、ロボットのカメラ制御が簡単になるとともに、動作の反応も速くできる。また、視野が広く複数のランドマークを一度に取得できるため、多少のオクルージョンに強いという利点もある。全方位ビジョンのこれらの利点を生かすために、ランドマークとして、環境に広く分布する特徴(点、エッジ、色)のパターンを用いることが多い。たとえば、唐・油田[26]は、ロボット周囲の環境中に存在する垂直エッジのパターンをランドマークとして使い、エッジの照合をロボスタにするために、エッジの両側の色も利用した。この方法では、ロボットは明示的な地図を持たず、Teaching-and-playbackの枠組で、走行する。すなわち、まず、ユーザに教示された経路を走行しながら、各地点で環境の画像を撮影し、そこで



(a) 物体認識の例 (物体モデルを併記)



(b) 物体認識によるナビゲーションの様子

図 2: 物体認識によるナビゲーション

のオドメトリデータと垂直エッジパターンを保存する。次からは、その教示データをもとに、自己位置を推定しながら目的地まで走行する。

6 環境モデルの構築

環境モデルは、移動ロボットが活動する環境中の物体の形状や位置姿勢を示したものである。移動ロボットにおいては、地図と呼ぶことも多いため、本稿でもそのように呼ぶ。環境の地図を手で作成するには多大な工数がかかるため、ロボットが環境を探索してその地図を構築する研究は盛んになされている。また、環境は時々刻々変化するため、既知環境であっても、その変化を地図に反映する必要があり、そこでも同様の技術が必要になる。

6.1 移動ロボットによる地図の生成

移動ロボットのための地図に最低限必要なのは、走行可能領域とランドマークの登録である。地図の構築においては、ロボットは環境内を自律あるいは遠隔操縦で走行しながら、外界センサによって環境の計測データを収集する。そして、走行中オンライン、あるいは、走行後オフラインで、その計測データをつなぎ合わせて走行可能領域を求め、また、目だった特徴をランドマークとして登録する。

ここで、GPS などの絶対位置姿勢を知るセンサを持たない場合、前述のようにロボットの自己位置推定は地図を参照して行うことになる。そのため、地図がない状態ではロボットの位置を正確に知ることはできない。一方、外界センサの計測データはロボットの位置を基準とするから、ロボットの位置がわからなければ正確な地図は作れない。このように、自己位置推定と地図作成は鶏と卵の関係にあり、両者を同時に推定する SLAM (Simultaneous Localization and Mapping) が盛んに研究されている [15, 27, 28, 37]。

走行可能領域を知るには、ロボットが走行しながら、環境中の物体との距離を即座に測定できることが望ましい。このため、レーザレンジセンサや超音波センサなどの距離センサがよく用いられる。とくに、レーザレンジセンサは分解能や精度に優れるため、現在、地図生成用センサの主流となっている。これらのセンサを用いる場合は、センサが備え付けられた高さの平面でスキャンした距離データを床面に投影して 2 次元地図を作成することが多い。

なお、レーザレンジセンサを 3 次元的にスキャンして環境の 3 次元モデルを作成し、その上にカメラで撮影した画像をマッピングしてリアルな地図を生成する研究もある [17, 28]。

6.2 地図生成のための視覚

走行可能領域の検出にビジョンを用いる場合は、距離を直接測定できるステレオビジョンが適しているが、高い精度を実現するのはやや難しい。また、通常のカメラでは、視野が限定されるため広い範囲を一度に見ることはできず、走行しながら周囲の距離データを効率よく収集するのは難しい。これらのことから、走行可能領域を求めるためには、ビジョンの利用は、まだ課題が多いと言える。

一方、ランドマーク登録のためには、その特徴の抽出と位置姿勢の推定ができればよく、自己位置推定でのランドマーク検出と同様の処理で実現できる。そのため、この処理はビジョンが一般に適している。ただし、特定のモデルに基づくランドマークを用いる場合は、ロボットが自律的に適切なランドマークを選択するのは困難であり、人間が何らかの支援をする必要がある。ランドマークとして単に画像の見えを用いる場合は、ロボットが定期的に観測した画像をそのまま用いることもできるため、ロボットが自律的にランドマークを登録することが可能である。

上述の地図は環境の正確な形状を示すことを目的としていたが、走行経路だけを示した地図 (トポロジカルマップ) も移動ロボットではよく使われる。この場合、地図は経路をグラフ化したものであり、その各ノード (場所を示す) においてロボットが撮影した画像を保存しておく。そして、次に走行するときは、各時点で撮った画像と保存しておいた画像とのマッチングにより自己位置推定を行う。画像データをそのまま用いるのは効率が悪いので、何らかの特徴抽出を行うことが多い。この地図は、人間が一度ロボットを経路に沿って走らせて「道」を教え、その後は自律的に走行させ

る Teaching-and-playback でよく用いられる。

6.3 従来の研究例

純粹にビジョンだけを用いて定量的なマップを生成した研究は、上述の理由により少ない。八木ら [36] は、全方位カメラで垂直エッジを抽出し、ロボットの走行に伴うエッジ方位角の変化から地図を生成するシステムを開発した。また、レーザレンジセンサを併用した研究として、三浦ら [19] は、全方位カメラを縦に2つ並べてステレオ視を実現し、これにレーザレンジセンサを組み合わせて、互いの弱点を補完した局所地図を生成するシステムを開発した。

経路地図であれば、上述の Teaching-and-playback が使えるので、ビジョンによる方式も多数提案されている [18, 23, 26]。筆者らのグループでも、宮崎・大矢ら [23] は、3眼ステレオを用いて環境中のエッジをランドマークとして抽出し、Teaching-and-playback により経路地図を覚えて、一度教えられた経路をナビゲーションするシステムを開発した。また、5.3節で述べた唐・油田 [26] のシステムも Teaching-and-playback で経路地図を生成する。

一方、あらかじめ人手で環境の形状情報を記述した地図をロボットに与え、ロボットがランドマークの観測結果に基づいて、地図を修正するアプローチもある。5.3節で述べた友納・油田 [31, 32] のシステムでは、人手でおおまかに作った地図をロボットが走行しながら修正する。地図は物体単位で構成され、ロボットは物体認識で推定した物体の位置姿勢に基づいて地図の誤差を修正する。

また、地図生成でなくても、屋外の道は不定型であるため、多くの場合、走行しながら走行可能領域を常に探す必要がある。大野・坪内ら [21] は、色分布を用いて画像から道路領域を抽出し、その中心線に沿って走行するシステムを開発した。

7 障害物検出

移動ロボットにおける重要な課題として、障害物の検出と回避がある。障害物とは、文字通り、ロボットの走行の障害となる物体であるが、基本的に地図からは予測しにくいという性質をもつ。ロボットが地図をもち自己位置推定をしながら走行する場合は、地図レベルで経路計画をしているので、地図が正確であれば、地図に載っている物体が障害物になることは原理的にはない。しかし、地図が不正確だったり、環境変化がある場合は、経路上に予期せぬ物体が存在することが起こる。移動物体も多くの場合、障害物となりうる。また、ロボットが地図をもたずにナビゲーションを行う場合は、あらゆる物が障害物になりうる。

7.1 障害物検出のための視覚

障害物検出で必要とされる情報は、障害物の位置と大まかな形状・サイズである。移動障害物の場合は、移動方向と速度もあると望ましい。また、障害物との

衝突はかなりの危険を伴う可能性があるため、検出の確実性と迅速さが要求される。一方、障害物を回避することが目的であれば、それが何であるかは知る必要がない²。このため、ランドマークと違って地図情報との対応づけは基本的には必要がない。これらのことから、ランドマーク検出とは性質の異なる検出方法が要求される。

ビジョンセンサで障害物検出に最も有効なのは、ステレオビジョンである。ステレオでは距離情報が直接得られるため、障害物の位置と大まかな形状を即座に知ることができる。一方、地図からの予測がしにくいという性質から、単眼ビジョンを用いるのは難しい。前述のように、単眼ビジョンでは距離情報が得られないため、モデル照合でカバーする必要がある。しかし、予測ができなければ、照合用のモデルを特定できないため、検出は非常に難しくなる。ただし、現地点にありうる障害物のクラスがほぼ特定でき、たとえば、床面上に置かれているなどの仮定ができれば単眼ビジョンも適用可能ではある。

7.2 従来の研究例

ステレオビジョンは物体の位置情報を低コストで得ることができるため、移動ロボットの障害物検出では広く用いられている。とくに、平行ステレオは、原理も簡単なので、数多く用いられている。ここでは、筆者らのグループの研究例を簡単に紹介する。

熊野・大矢ら [13] は、ステレオビジョンと床面拘束を利用して障害物を検出するシステムを開発した。ここでは、平行ステレオの2枚の画像間で、床面にあれば同じ点となるはずの画素値を比較して、大きく食い違う場合、その点に障害物があると見なす。

また、河合・大矢ら [22] は、パンチルト機能のある平行ステレオカメラを用いて、ランドマーク検出と障害物検出を同時に行うシステムを構築した。ロボットは、廊下を走行中にステレオカメラで前方を監視して、障害物があれば回避する。障害物検出には、前述の床面拘束法を用いている。それと並行して、地図を参照しながら、ドアに近づくとその方向にカメラを向けて画像を撮影し、ドアの垂直エッジをランドマークとして自己位置を修正する。

8 物体操作

マニピュレータで物体を操作するロボットは、据え置き型では多くの例があるが、移動ロボットでの実現例は少ない。だが、ロボットが様々な有用な作業を行うためには、物体操作機能はきわめて重要である。

8.1 物体操作のための視覚

物体操作では、マニピュレータで把持するために、物体の位置姿勢と形状を正確に知ることが最低限必要

²もちろん、何であるかを知ったうえで、適切な回避動作をとるといって高度な障害物回避も考えられる

である。また、把持の際に、物体の硬さや安定性などの性質も知ることが望ましいため、物体の認識も重要な要件となりうる。この認識と位置姿勢の推定が、ランドマークや障害物と比べて一回り厳しい要件となる。

さらに、人間と同程度の大きさの移動ロボットに搭載するマニピュレータの把持対象物は、サイズが小さく、しかも、可搬であることが多い。可搬の場合は、地図に正確な位置を載せることは難しく、ランドマークに比べると地図からの予測はしにくい。ロボットは対象物があるはずの場所を地図で調べた後、その現場に行って対象物を探すという動作を行う必要がある。また、対象物が可搬でなくても、ロボットの自己位置の誤差があるため、地図から得た対象物の位置姿勢は十分な精度が期待できない。この点が、対象物の置き場所がほぼ決まっている据え置き型ロボットとの大きな違いである。

物体の位置姿勢と形状を知るには、ステレオビジョンが最も利用しやすい。しかし、対象物が決まっている場合は、モデル照合が可能なので、単眼ビジョンによる物体認識も利用できる。また、ステレオを用いる場合も、位置姿勢と形状の精度が要求されるため、エッジやテクスチャなどの高い精度が期待できる特徴を用いる必要がある。どのセンサを用いるにしても、物体操作では、高い精度を実現する処理の方に重点が置かれる。

8.2 従来の研究例

ステレオビジョンは物体操作を行うロボットに多く利用されており、瓶を掴むロボットや機の運搬を手伝うロボットも開発されている [10, 29]。一方、単眼ビジョンは、3次元情報を直接得られないため、物体を見つけやすいように環境を限定したり、地図からの予測を用いるものが多い。ここでは、筆者らの研究室で行った単眼ビジョンを用いた例を紹介する。

永谷・油田 [20] は、ドアを開けて部屋と廊下を走行する移動マニピュレータを開発した。ドアを開けるためにはロボットがノブを操作する必要があり、移動のためだけでも物体操作が必要な例になっている。ロボットはノブをカメラで認識し、マニピュレータで掴んで回してドアを開ける。ドアやノブの形状や位置は地図に書かれているが、ロボットの位置姿勢に誤差があるため、ビジョンを用いてノブの位置を確認する必要がある。このため、画像の垂直および水平方向の微分画素のヒストグラムの重心を求めてノブの中心とみなし、マニピュレータの目標位置を決める。図3に、ドア開け動作とノブの検出処理例を示す。

5.3節で述べた友納・油田 [32] のシステムは、機の上に乗った紅茶箱の認識も実現している。このシステムでは、ロボットは人手で大まかに作られた地図を与えられるため、機の位置も、機の上の紅茶箱の位置も正確ではない。そのため、ロボットはまず機を認識して、その近くに寄った後、机上の画像を撮影して、そこから紅茶箱を見つける処理を行う。このシステムでは、地図はカメラを向ける方向を決めるためのおおまかな予測に使うだけで、机や紅茶箱の位置はすべて物

体認識により推定する。なお、マニピュレータは搭載していないので紅茶箱を運ぶ動作は実現していない。

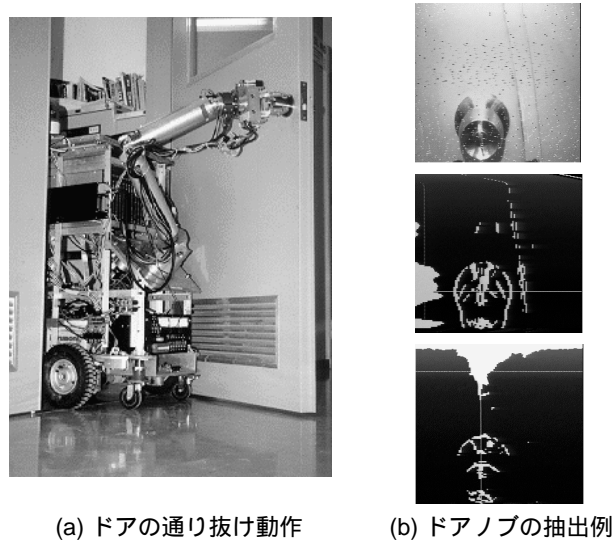


図 3: ドアの通り抜け

9 まとめ

本稿では、移動ロボットのための環境センシングにおいて、視覚がとくに期待される理由と、そこに求められる機能について述べ、これまでの研究例を紹介した。可視光による画像センシングは、他のセンシング法に比べ圧倒的な情報量を市販の簡単なデバイスで得ることができる利点がある。反面、その中から有効な情報を取り出す処理が難しいという問題がある。

人間や動物が移動する際には環境情報の大部分は視覚から得ており、それと同等の処理が実現できれば移動ロボットの行動能力は著しく向上するであろう。移動の際に必要なデータは自分自身や対象物あるいは障害物の位置姿勢であるが、人間はこれを高度な認識能力も用いて行っている。今後も CPU の処理能力は向上すると予想されるので、モデルや知識を総動員した高度な処理によって、移動ロボットのためのさらに正確な環境センシングが実現されてゆくと期待している。

参考文献

- [1] R. Basri and E. Rivlin: "Homing using Combinations of Model Views," *Proc. of IJCAI'93*, pp. 1586-1591, 1993.
- [2] H. I. Christensen, N. O. Kirkeby, S. Kristensen, L. Knudsen, and E. Granum: "Model-driven vision for in-door navigation," *Robotics and Autonomous Systems* 12 (1994) 199-207.
- [3] G. Dudek and D. Jucsur: "Robust Place Recognition using Local Appearance based Methods," *Proc. of ICRA2000*, pp. 1030-1035, 2000.
- [4] H. F. Durrant-Whyte and J. J. Leonard: "Navigation by Correlating Geometric Sensor Data," *Proc. of IROS'89*, pp. 440-447, 1989.

- [5] D. Fox, W. Burgard, F. Dellaert, and S. Thrun: "Monte Carlo Localization: Efficient Position Estimation for Mobile Robots," *Proc. of AAAI'99*, 1999.
- [6] A. Georgiev and P. K. Allen: "Vision for Mobile Robot Localization in Urban Environments," *Proc. of IROS2002*, pp. 472-477, 2002.
- [7] J. Gonzalez-Barbosa and S. Lacroix: "Rover localization in natural environments by indexing panoramic images," *Proc. of ICRA2002*, pp. 1365-1370, 2002.
- [8] A. Hauck and N. O. Stöfler: "A Hierarchic World Model Supporting Video-Based Localization, Exploration and Object Identification," *Proc. of ACCV'95*, pp. III-176-III-180, Dec. 1995.
- [9] J. B. Hayet, F. Lerasle, and M. Devy: "A Visual Landmark Framework for Indoor Mobile Robot Navigation," *Proc. of ICRA2002*, pp. 3942-3947, 2002.
- [10] S. Kagami, K. Nishiwaki, J. J. Kuffner, Y. Kuniyoshi, M. Inaba, and H. Inoue: "Online 3D Vision, Motion Planning and Bipedal Locomotion Control Coupling System of Humanoid Robot: H7," *Proc. of IROS2002*, pp. 2557-2562, 2002.
- [11] T. Kanbara, J. Miura, and Y. Shirai: "Selection of Efficient Landmarks for an Autonomous Vehicle," *Proc. of IROS'93*, pp. 1332-1338, 1993.
- [12] A. Kosaka and J. Pan: "Purdue Experiments in Model-Based Vision for Hallway Navigation," *Proc. of Workshop on Vision for Robots in IROS'95*, pp. 87-96, (1995).
- [13] M. Kumano, A. Ohya, S. Yuta: "Obstacle Avoidance of Autonomous Mobile Robot using Stereo Vision Sensor," *Proc. of the Second International Symposium on Robotics and Automation (ISRA'2000)*, pp. 497-502, 2000.
- [14] F. Launay, A. Ohya, and S. Yuta: "Autonomous Indoor Mobile Robot Navigation by Detecting Fluorescent Tubes," *Proc. of ICAR2001*, 2001.
- [15] J. J. Leonard and H. F. Durrant-Whyte: "Simultaneous Map Building and Localization for an Autonomous Mobile Robot," *Proc. of IROS'91*, pp. 1442-1447, 1991.
- [16] Y. Liu, T. Yamamura, T. Tanaka, and N. Ohnishi, Character-Based Mobile Robot Navigation, *Proc. of IROS'99*, pp.610-616, 1999.
- [17] C. Martin and S. Thrun: "Real-Time Acquisition of Compact Volumetric 3D Maps with Mobile Robots," *Proc. of ICRA2002*, pp. 311-316, 2002.
- [18] Y. Matsumoto, K. Ikeda, M. Inaba, and H. Inoue: "Exploration and Navigation in Corridor Environment Based on Omni-View Sequence," *Proc. of IROS2000*, pp. 1505-1510, 2000.
- [19] J. Miura, Y. Negishi, and Y. Shirai: "Mobile Robot Map Generation by Integrating Ominidirectional Stereo and Laser Range Finder," *Proc. of IROS2002*, pp. 250-255, 2002.
- [20] 永谷圭二, 油田信一: "タスクオリエンティッドアプローチによる自律移動マニピュレータの研究-ドアの通り抜けを含む屋内の自律走行の実現-", *日本ロボット学会誌*, Vol.17, No.6, pp.111-121, 1999.
- [21] K. Ohno, T. Tsubouchi, S. Maeyama, S. Yuta: "A Mobile Robot Campus Walkway Following with Daylight-Change-Proof Walkway Color Image Segmentation," *Proc. of IROS2001*, pp.77-83, 2001.
- [22] A. Ohya, D. Kawai, S. Yuta: "Autonomous Mobile Robot Navigation using Active Stereo Vision," *Proc. of IAS-6*, pp.1041-1048, 2000.
- [23] A. Ohya, Y. Miyazaki, and S. Yuta: "Autonomous Navigation of Mobile Robot based on Teaching and Playback Using Trinocular Vision," *Proc. of IECON'01*, pp.398-403, 2002.
- [24] S. Se, D. Lowe, and J. Little: "Vision-based Mobile Robot Localization And Mapping using Scale-Invariant Features," *Proc. of ICRA2001*, pp. 2051-2058, 2001.
- [25] T. Tanaka, Y. Yamazaki, H. Watanabe, T. Katae, and K. Yamafuji, Self-Position Detection System Using a Visual-Sensor for Mobile Robots, *Journal of Robotics and Mechatronics*, Vol. 11, No. 1, pp. 17-24, 1999.
- [26] L. Tang and S. Yuta: "Vision Based Navigation for Mobile Robots in Indoor Environment by Teaching and Playing-Back Scheme," *Proc. of ICRA2001*, 2001.
- [27] S. Thrun, W. Burgard, and D. Fox: "A Probabilistic Approach to Concurrent Mapping and Localization for Mobile Robots," *Machine Learning and Autonomous Robots*, 31/5, 1-25 (1998).
- [28] S. Thrun, W. Burgard, and D. Fox: "A Real-Time Algorithm for Mobile Robot Mapping with Applications to Multi-Robot and 3D Mapping," *Proc. of ICRA2000*, 2000.
- [29] 富田文明: "高機能3次元視覚システムVVV"、情報処理、42巻4号、pp. 370-375、2001.
- [30] M. Tomono and S. Yuta: "Mobile Robot Navigation in Indoor Environments using Object and Character Recognition," *Proc. of ICRA2000*, pp. 313-320, 2000.
- [31] 友納正裕, 油田信一: "不正確さを許すマップと物体認識による屋内ナビゲーションシステム"、第7回ロボティクスシンポジウム、pp. 1-8、2002.
- [32] M. Tomono and S. Yuta: "Indoor Navigation based on an Inaccurate Map using Object Recognition," *Proc. of IROS2002*, 2002.
- [33] T. Tsubouchi and S. Yuta: "Map Assisted Vision System of Mobile Robots for Reckoning in a Building Environment," *Proc. of ICRA'87*, pp. 1978-1984, 1987.
- [34] I. Ulrich and I. Nourbakhsh: "Appearance-Based Place Recognition for Topological Localization," *Proc. of ICRA2000*, pp. 1023-1029, 2000.
- [35] M. Yachida, T. Ichinose, and S. Tsuji: "Model-guided Monitoring of a Building Environment by a Mobile Robot," *Proc. of IJCAI'83*, pp. 1125-1127, 1983.
- [36] Y. Yagi, K. Shyouya, and M. Yachida: "Environmental Map Generation and Egomotion Estimation in a Dynamic Environment for an Ominidirectional Image Sensor," *Proc. of ICRA2000*, pp. 3493-3498, 2000.
- [37] B. Yamauchi, A. Schultz, and W. Adams: "Mobile Robot Exploration and Map-Building with Continuous Localization," *Proc. of ICRA'98*, pp. 3715-3720, 1998.
- [38] Y. Watanabe and S. Yuta: "Position Estimation of Mobile Robots with Internal and External Sensors Using Uncertainty Evolution Technique," *Proc. of ICRA'90*, pp. 2011-2016, 1990.