

移動ロボットの能動的視覚によるカメラ運動の制御と 物体／環境認識への応用

宮 脇 一 路[†] 李 仕 剛[†]
石 黒 浩[†] 辻 三 郎[†]

本論文では、移動ロボットによる能動的視覚の実現方法について述べる。従来の能動的視覚の研究では、カメラの正確な運動を前提にしていた。そのため、カメラの運動は正確に制御可能で動作範囲の限定されたマニピュレータにより実現されていた。この能動的視覚の適応範囲を拡大するには、移動ロボットの視覚として能動的視覚を用いることが考えられるが、それには幾つかの問題を克服しなければならない。まず、能動的視覚に必要なロボットの制御方法として、視覚フィードバックにより対象物体を中心とした円弧運動を行う方法を考案し、シミュレーションと移動ロボットによる実験を行った。そして、この制御方法を用いて、移動ロボットによる能動的視覚の可能性を示唆すべき二つの実験を行った。一つは注視による物体の3次元構造の獲得であり、内界センサの情報のみでは正確な制御が困難である移動ロボットにおいても、Aloimonosらによって明らかにされた能動的視覚の利点が見られ、実現できることを示した。もう一つは、移動ロボットの能動的視覚による新しい物体認識の手法を提案し、その基礎的な実験を行った。この方法は、計算コストが低く、安定した物体認識が行える。

Controlling Camera Motion by Active Vision of a Mobile Robot and Its Applications to Environment/Object Recognition

ICHIRO MIYAWAKI,[†] SINGANG LI,[†] HIROSHI ISHIGURO[†] and SABURO TSUJI[†]

Most active vision systems use a camera mounted on a manipulator, because the control of the camera motion is easy and accurate. A limited range of the camera motion by the fixed manipulator, however, often prevents the method from real applications. This paper explores a more general method that a camera on a mobile robot freely moves in the indoor environment and acquires the spatial information based on the active vision paradigm. The camera motion is controlled so as to fix its gaze upon a feature point and to keep the distance to the fixation point constant by visual feedback. That is, a robot with a camera moves along a circular path. By using this active-vision-based mobile robot, we propose and experiment a general method that a robot can acquire the spatial information based upon the active vision paradigm, and a simple and much less expensive active method for the object recognition.

1. はじめに

人間は、与えられたタスクに応じた注視点の選択と、その注視点への注視を繰り返すことで、実時間でロバストな視覚機能を実現している。近年、コンピュータビジョンの研究において、このような人間の視覚機能を基にした、能動的視覚 (active vision)^{1), 2), 11)} という概念が注目されている。能動的視覚の概念に基づく研究では、人間の網膜に似た CCD の開発に始

まり、注視点の選択 (attention), 注視の安定化 (gaze stabilization), 注視行動の実現方法 (gaze change), 目と手の協調動作, 注視行動を実現するハードウェアの開発に至るまで、さまざまな問題が議論されている¹²⁾。しかし、未だ未解決な問題が山積しており、工学的に明らかにされた能動的視覚の利点は数少ない。

これまでの、能動的視覚に関する最も重要な研究成果は、Aloimonos¹⁾ によって明らかにされた、「カメラを環境内のある一点を中心とする円弧軌道上を運動させながら、その点を注視 (その点の画像面上への投影が、常に画像中心に来るように、カメラの光軸を調整する) させることで、その点 (注視点と呼ぶ) を中

[†] 大阪大学基礎工学部システム工学科
Department of Systems Engineering, Faculty of
Engineering Science, Osaka University

心とする座標系で物体の構造記述が直接得られる」という事実である。彼らの研究成果について他の見方をすれば、「Shape from motion¹⁹⁾」における決定すべき未知のパラメータ（カメラの回転成分）を、カメラに正確な運動により既知のものとして扱えるようにした」ということになる。すなわち、彼の研究では、カメラの正確な運動が実現可能であることが前提条件なのである。

カメラの正確な運動を実現するには、正確に制御可能なマニピュレータなどの先端にカメラを取り付けなければならない。しかし、マニピュレータは正確な制御が容易である半面、動作範囲が比較的狭い領域に限定される。能動的視覚の概念をより広範囲の物体/環境認識の問題に適用するには、限られた範囲内で動作可能なマニピュレータを用いるだけでなく、より広範囲の環境で動作可能な移動ロボットによる能動的視覚の概念に基づく研究が期待される²⁾。

本論文では、移動ロボットによる注視制御の実現方法と、それをもとにした物体/環境認識の基礎的な手法について述べ、移動ロボットによる能動的視覚の実現方法を示す。

移動ロボットによりカメラ運動を実現する際の問題点は、移動ロボットが移動手段として用いる車輪が外界（床面）と接触しているため、スリップなどの外乱により、ロボットの内外界センサの情報のみでは、移動ロボットの運動を正確に制御できないことである。この問題を克服するには、視覚装置等の外界センサからの情報を、ロボットの運動制御にフィードバックする必要がある。外界センサの情報の利用方法は、さまざまなものが考えられる。例えば、そのような手法の一つとして、環境内の既知の位置に設置されたランドマークなどを頼りに、環境内におけるロボットの位置を検出し、ロボットの運動を正確に制御する方法がある^{4),5)}。このようなランドマークを用いる方法では、あらかじめ環境を計測し、ランドマークを設置または選択する必要がある。環境認識を目的とする能動的視覚にはふさわしくない方法である。環境内を自由に移動し探索するためには、ランドマークを用いない手法が望まれる。

そこで、本論文では、能動的視覚における基本的なカメラ運動である、注視制御に着目し、注視制御により移動ロボットの運動を制御する手法を提案する。本論文で述べる手法により、ロボットは、外界センサからのフィードバックにより、注視点を中心とした円弧

運動をすることができる。

次に、この移動ロボットによる注視制御を用いて、能動的視覚による物体/環境認識を実現する基礎的な手法について述べる。能動的視覚の利点は先に述べたように、注視制御により、直接物体中心の座標系で、物体の記述が得られることである。そこで、本論文では、カメラの運動制御が困難な場合、すなわち、移動ロボットによる注視制御においても、注視点を中心とした座標系で正確な物体の構造記述が得られることを示す。能動的視覚の手法¹⁾により、物体の構造記述を得るには、ロボットによりカメラの円弧運動を実現するだけでなく、カメラの回転角度を正確に知る必要がある。カメラの回転角度の測定には、屋内環境内に多数存在する、床面と平行な直線の画面上への投影を利用する。消失点を利用して、カメラの回転角度を測定する方法はすでに幾つか報告されているが⁸⁾⁻¹⁰⁾、床面と、それに平行な直線がなす消失点を利用した手法は、われわれ独自の手法である⁷⁾。

すなわち、画像情報を利用したカメラの回転角速度を測定する方法と、注視制御によるロボットの運動制御方法により、移動ロボットによる能動的視覚が実現できるのである。

注視制御を用いれば、物体中心の座標系で物体の正確な構造記述が得られる。しかし、物体/環境認識は、あらかじめコンピュータに記録されている3次元モデルとの対応のみにより実現されるものではない。注視制御により、3次元モデルを用いない手法も考えることができる。

物体の3次元モデルを用いた物体認識の一例として池内¹⁴⁾らの研究がある。池内らは、あらかじめコンピュータに記録されている物体の3次元モデルから、さまざまな視点からの見え方を生成し、生成された物体の見え方と、実際にカメラから得られる画像を対比することにより、物体認識を実現している。この手法の問題点は、物体の見え方を生成する際に高い計算コストがかかることである。

この問題点を解決する一つの方法として、本論文では、カメラが物体に注視しながら物体の回りを移動し、さまざまな角度から物体を観測する手法について述べる。この手法では、コンピュータに物体の特徴的な見え方を、2次元画像のモデルとして記録しておき、実際にロボットが移動して得られる画像データとモデルを比較しながら、物体認識を行う。

物体の2次元モデルと注視制御を用いた手法は、従

来の手法と比べ、計算コストを大幅に削減できるだけでなく、能動的視覚による物体認識の一例として意味のあるものである。本論文では、この手法に基づく基礎的な実験について述べる。

2. ロボットと環境における仮定

移動ロボットによる注視制御と後の物体/環境認識への応用において、ロボットと環境に関して、いくつかの仮説を設ける。

ロボット動作に関する仮定

- (1) ロボットは屋内環境の水平な床面上を移動する。このためロボットの運動は、床面上での二次元座標を表す (X, Y) と、方位を表す θ で表現される。
- (2) ロボットは微小時間移動し、停止した後画像を入力するといった動作を繰り返し、観測を行う（この仮定は、画像の処理速度がロボットの移動速度に比べかなり遅いために用いる）。
- (3) ロボットは二輪独立駆動方式の自律移動型ロボットを用いる。ロボットは環境内を自由に移動することができ、その運動は、車輪に取り付けられたエンコーダで測定することができる。しかし、車輪のスリップ等により、正確な情報は得られない。

カメラの動作に関する仮定

- (4) カメラはロボットに搭載されており、光軸はロボットの進行方向に対して直角に、床面と平行に設置されている。

環境に関する仮定

- (5) 環境内には多くの直線が存在する。例えば、壁と床・壁と天井の境界線や物体のエッジ等である。

3. 視覚フィードバックによるロボットの運動制御

本章では、注視制御中に得られる視覚情報を、ロボットの運動制御にフィードバックすることにより、対象物体を中心としたカメラの円弧運動を実現する方法について述べる。円弧運動の中心は対象物体上的一点（注視点）である。この点の画面中心からのずれを車輪の制御にフィードバックすることにより、円弧運動を実現する。本章では、ロボットの制御方法と、誤差・外乱を考慮したコンピュータによるシミュレーション、実際に移動ロボットを用いて行った実験結果

について述べる。

3.1 移動ロボットによる円弧運動の実現

本研究で用いる移動ロボットは、左右輪の独立した動きによって運動が制御される。

円弧運動を行うには、円弧運動において目標とする半径を与える必要がある（円弧の半径は正確に与える必要はないが、ロボットが環境内で、おおまかな経路計画を立てるとすると、おおよその半径を与えることは意味がある）。そこで、まず最初に、ロボットは微小距離直進し、車輪に取り付けられたエンコーダの情報をを用いて移動ステレオ視により注視点までの距離測定を行う（エンコーダの情報は、車輪のスリップ等の外乱に影響されるため、正確ではないが、円弧運動の目標半径の決定には十分である）。注視点までのおおまかな距離が得られると、その距離を半径とする円弧運動を行うための、左右輪の移動速度の理論値が決定される。半径を R 、車輪幅を B とすると、左右輪（右車輪を内輪とする）の移動速度の理論値 V_l 、 V_r は(1)式で表される（図1）。

$$V_l = \left(\frac{R+B}{R} \right) V, \quad V_r = V \quad (1)$$

B は既知であるため、 R を測定することにより V_l 、 V_r が決定できる。

しかし、実際にはロボットが左右輪の移動速 V_l 、 V_r 移動しても、理想軌道にそった正確な円弧運動を行うことはできない。実際の床の条件や物体までの距離の評価の誤差、車輪のスリップ等の外乱が発生するからである。そこで、理想軌道にそった正確な円弧運動を実現するために、視覚フィードバックを用いる。

ロボットが正確な円弧運動を行うための条件としては、次の二つが上げられる。

- (1) 注視点に対してロボットの進行方向は常に直角である。
- (2) ロボットと注視点の距離は常に一定である。

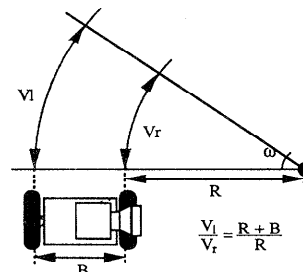


図1 左右の車輪の速度

Fig. 1 The angular velocities of two wheels.

この2つの条件を満たすようにロボットの制御を行う。

(1)を満たすためには、カメラの光軸はロボットの進行方向に対して直角に設置されているので、注視点が常に画面中心を通る垂直軸上に投影されるように、ロボットの制御を行えばよい。ここで、注視点は画面上の特徴的な一点でもよいし、画面上の領域が容易に切り出せる場合はその領域の重心であってもよい。本実験では後者を用いる。

(2)を満たすためには、注視する特徴(領域)の画像上での見かけの高さを常に一定に保てばよい。

3.2 制御方法

座標系は、カメラ中心の座標 $O-XY$ を用いる。その原点はカメラ中心におき、 X 軸はカメラの光軸と一致させる。ロボットが時間 t から $t+\Delta t$ の間に左右輪の速度 V_l, V_r で移動したとすると、その回転半径 r と回転角 θ は(2)式で示される。

$$r = \frac{V_l}{V_l - V_r} B, \theta = \frac{V_l - V_r}{B} \Delta t \quad (2)$$

ここで、 B はロボットの車輪幅であり、 Δt はロボットが移動する時間である。

時間 t と Δt で注視点の座標をそれぞれ $(X_i, Y_i), (X_{i+1}, Y_{i+1})$ とする。 (X_{i+1}, Y_{i+1}) は図2より(3)式で求められる。

$$\begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \end{bmatrix} + \begin{bmatrix} r(1 - \cos \theta) \\ r \sin \theta \end{bmatrix} \quad (3)$$

ロボットの運動制御は、式(1)より右車輪の速度 V_r を V に固定し、左車輪の速度 V_l を $(R+B)V/R$ の近傍で操作することによって行う。ここではロボットの左車輪への入力を u とし、 u を操作する。これより、 V_l, V_r は次の(4)式になる。

$$V_l = \left(\frac{R+B}{R} \right) V + u, \quad V_r = V \quad (4)$$

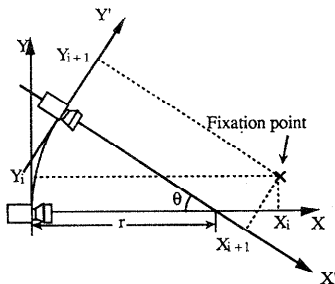


図2 座標変換
Fig. 2 Coordinate transformations.

(4)式を(2)式に代入して(5)式を得る。

$$r = \frac{V}{\frac{B}{R} + u} B, \quad \theta = \frac{\frac{B}{R} V + u}{B} \Delta t \quad (5)$$

式(5)を式(3)に代入し、平衡状態 $X=R, Y=0, u=0$ の近傍で微小変化分 $[\delta X \ \delta Y]^T$ について線形化を行うと、(6)式が得られる。

$$\begin{bmatrix} \delta X_{i+1} \\ \delta Y_{i+1} \end{bmatrix} = \begin{bmatrix} \cos \frac{V}{R} \Delta t & -\sin \frac{V}{R} \Delta t \\ \sin \frac{V}{R} \Delta t & \cos \frac{V}{R} \Delta t \end{bmatrix} \begin{bmatrix} \delta X_i \\ \delta Y_i \end{bmatrix} + \begin{bmatrix} -\frac{R^2}{BV} \left(1 - \cos \frac{V}{R} \Delta t \right) \\ \frac{R^2}{BV \sin \frac{V}{R} \Delta t} \end{bmatrix} u \quad (6)$$

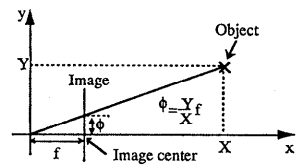
一方、ロボットの位置 (X, Y) における注視点の画面中心からのずれ ϕ と見かけの高さ l は、図3より(7)式で表される。

$$\phi = \frac{Y}{X} f, \quad l = \frac{fL}{X} \quad (7)$$

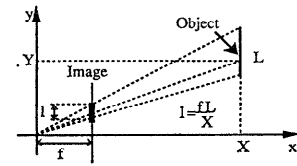
ここで、 f はカメラの焦点距離、 L は実際の領域の高さである。

同様に、(7)式を $X=R, Y=0$ の近傍で微小変化分 $[\delta \phi \ \delta l]^T$ について線形化を行うと、(8)式が得られる。

$$\begin{bmatrix} \delta \phi \\ \delta l \end{bmatrix} = \begin{bmatrix} 0 & \frac{f}{R} \\ -\frac{fR}{R^2} & 0 \end{bmatrix} \begin{bmatrix} \delta X \\ \delta Y \end{bmatrix} \quad (8)$$



(a) 画像中心からのずれ
(a) Error from image center



(b) 観測される物体の高さ
(b) Height of the observed object

図3 画像中心からのずれと物体の高さ

Fig. 3 Error from image center and height of the observed object.

(6)式と(8)式より, ロボットの $u \rightarrow [\delta\phi \ \delta l]^T$ の伝達関数 G_p は, (9)式になる.

$$G_p = \frac{1}{BV \left\{ z^2 - 2 \cos \left(\frac{V}{R} \Delta t \right) z + 1 \right\}} \cdot \begin{bmatrix} Rf \left(\sin \frac{V}{R} \Delta t \right) (z-1) \\ Lf \left(1 - \cos \frac{V}{R} \Delta t \right) (z+1) \end{bmatrix} \quad (9)$$

制御則を(10)式とする.

$$u = -G_c(z)^T \begin{bmatrix} \delta\phi \\ \delta l \end{bmatrix}, \quad G_c(z) = \begin{bmatrix} K_1(z) \\ K_2(z) \end{bmatrix} \quad (10)$$

これは, 注視点の画面中心からのずれと領域の大きさの変化をフィードバックして, 入力 u を決定するものである.

この閉ループシステムのブロック図は図4で表される. これより, 外乱 $\omega \rightarrow [\delta\phi \ \delta l]^T$ の閉ループ伝達関数 G_o は, (11)式で表される.

$$G_o = [1 + G_p G_c]^{-1} G_p \quad (11)$$

G_o において, ステップ状の外乱 $\omega(z) = \frac{K_0}{1-z^{-1}}$ に対して, 最終値の定理より(12)式が成立する.

$$\begin{bmatrix} \delta\phi(\infty) \\ \delta l(\infty) \end{bmatrix} = \lim_{z \rightarrow 1} (1-z^{-1}) G_o \omega(z) = \lim_{z \rightarrow 1} G_o K_0 \quad (12)$$

これより, $\lim_{z \rightarrow 1} G_o = 0$ とすれば $\begin{bmatrix} \delta\phi(\infty) \\ \delta l(\infty) \end{bmatrix} = 0$ とできる. このことから, 制御則は(13)式のように決定する.

$$G_c(z) = \begin{bmatrix} K_1(z) \\ K_2(z) \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ 1-z^{-1} \end{bmatrix} \quad (13)$$

この制御則について次のことがいえる.

○画面中心からのずれ $\delta\phi$ については比例制御となる.

○大きさの変化 δl については積分制御となる.

次に, この G_c における安定性を検討する.

G_c を G_o に代入し, その分母多項式を $D(z)$ と置く. このシステムが安定であることは, $D(z)=0$ の根の絶対値がすべて1より小さいことと同値である. これより, (14)式の条件が導かれる.

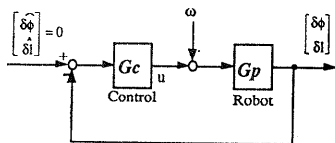


図4 フィードバックループ
Fig. 4 Feedback loop.

$$0 < k_1 < \frac{2BV \left(1 + \cos \frac{V}{R} \Delta t \right)}{Rf \left(1 + \sin \frac{V}{R} \Delta t \right)},$$

$$0 < k_2 < \frac{BV \cos \frac{V}{R} \Delta t}{Lf \left(1 + \cos \frac{V}{R} \Delta t \right)} \quad (14)$$

3.3 注視制御のための特徴追跡

視覚フィードバックによる注視制御において, 注視すべき特徴点を追跡する必要がある. 本実験では, 物体の四角形の面を注視対象とするため, ここでは, その4つの頂点を特徴点として追跡する.

注視点の追跡には SSDA (Sequential Similarity Detection Algorithm) を用い, coarse-to-fine の手法により高速化している(図5). これはまず, 撮像した 256×240 pixels の入力画像を 4×4 pixels の近傍で平均を取ることにによりそれぞれ 64×60 pixels の画像に縮小する. この縮小された画像において 5×5 pixels のウィンドウを設け, 5×3 pixels の探索範囲で SSDA により特徴点のおおまかな位置を決定する. 次に撮像した 256×240 pixels の画像に戻ってそのおおまかな位置をもとに 9×9 pixels のウィンドウを設け, 5×3 pixels の探索範囲で再び SSDA によって特徴点の正確な位置を検出する. この方法で特徴点を追跡することにより, 特徴点の移動量が大きくても追跡が正確になり, なおかつ大幅な高速化が実現できた. 処理時間は, 4つの頂点の追跡に対し, 画像の入力から特徴点の正確な位置の検出まで 0.385 sec であった.

3.4 実験結果

(1) シミュレーションによる実験結果

3.2節で決定した制御則に従って, 外乱を考慮したシミュレーションを行った. 実験で設定した環境を以下に示す.

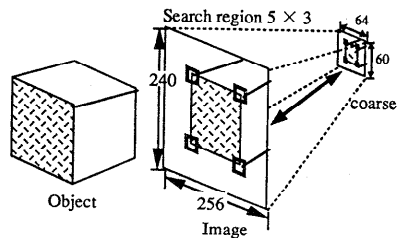


図5 Coarse-to-fine 法によるトラッキング
Fig. 5 Tracking.

- ロボットの移動速度 (V) 約 10 cm/s
- サンプリング周期 (Δt) 1 sec
- ロボットと対象物体の距離 (R) 約 200 cm
- 注視する領域の実際の高さ (L) 30 cm
- カメラの焦点距離 (f) 1.26 cm
- 車輪幅 (B) 33.6 cm
- 入力画像 512×480 pixels
- 外乱, 誤差 左右輪の移動速度に最大 0.3 cm/s の外乱を一様分布乱数にて発生. また 1/50 の確率で最大 3° の姿勢変化を発生.

ここで, k_1, k_2 は次のようになる.

$$0 < k_1 < 5.37, 0 < k_2 < 4.48$$

シミュレーションにおいて, 最適な応答を示した定

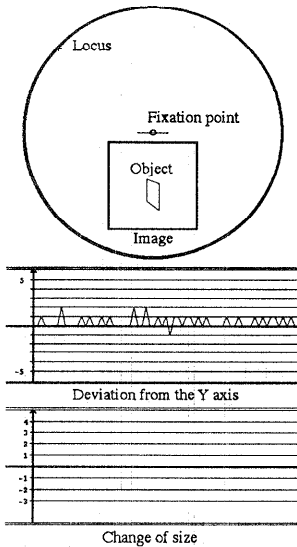


図 6 実験結果

Fig. 6 Experimental results.

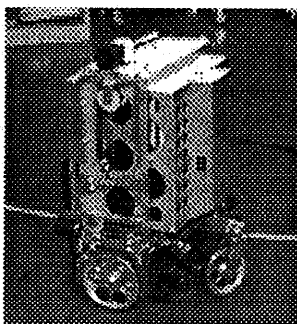


図 7 移動ロボット

Fig. 7 Mobile robot.

数は次のとおりである.

$$k_1=3.2, k_2=1.2$$

シミュレーションによる結果を図 6 に示す. グラフは, 物体の画面中心からのずれと, 大きさの変化を表している. 実験結果では, ずれは最大 3 pixel, 大きさの変化は 1 pixel 以内である.

(2) 移動ロボットを用いた実験結果

実験に用いた移動ロボットを図 7 に示す. また, システムのハードウェア構成を図 8 に示す. 実験環境は, シミュレーションとほぼ同じである. 対象物体は 20 cm×30 cm の黒い長方形の面を持ち, この面の重心を注視する. 実験結果を図 9 に示す. 図中の白線

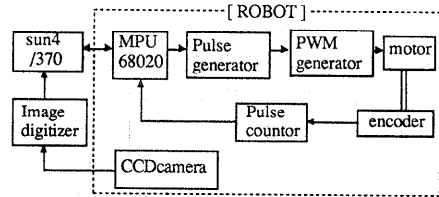
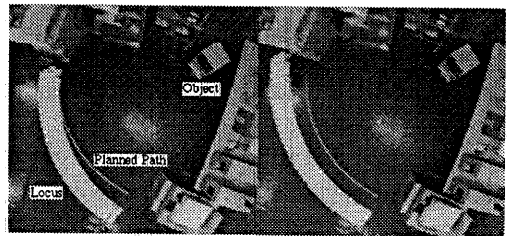
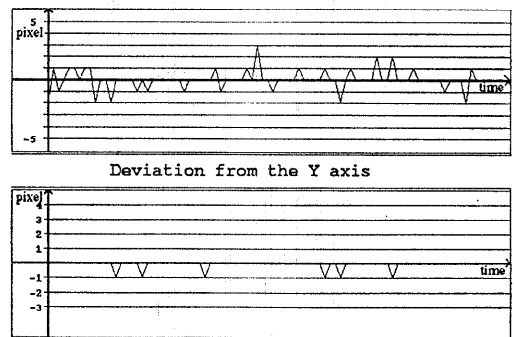


図 8 ハードウェア構成

Fig. 8 Hardware configuration.



- (a) 視覚フィードバックなし
- (b) 視覚フィードバックあり
- (a) Uncontrolled robot
- (b) Controlled robot



Change of size

- (c) 誤差
- (c) Error

図 9 実験結果

Fig. 9 Experimental results.

で書かれた円弧は理想軌道である。図9(a)は視覚フィードバックを行わなかった場合、図9(b)は、本制御則にしたがって視覚フィードバックを行った場合である。図(b)において、物体の画面中心からのずれと、大きさの変化を図9(c)に表す。長方形の重心の画面中心からのずれは最大3 pixel、大きさの変化は1 pixel 以内である。

4. 能動的視覚の物体/環境認識への応用

本章では、3章で述べた、視覚フィードバックによりカメラの円弧運動を実現する方法を用いた移動ロボットの注視制御による、物体の構造記述の獲得と、移動ロボット注視制御による新しい物体認識へのアプローチについて述べる。

4.1 能動的視覚による物体の3次元構造の復元

本手法の手順は次のようである。ロボットが、注視する物体を環境内から検出し、その物体上に注視点を定めた後、その注視点を中心とした円弧運動を行う。そして、カメラの回転角を求め、このカメラの回転角から、環境内で注視点付近に存在する垂直線の位置を、注視した物体を原点とする物体中心の座標系で記述する。本論文では、床面と平行に設置した2次元の直交座標系上で、垂直線の2次元位置を計測するとどめるが、本手法を物体の3次元の位置決め問題に拡張することは容易である。

環境内には多くの垂直線が存在するため、この垂直線を環境構造を記述する特徴とみなし、注視点を原点とする物体中心の座標系ではこの垂直線の位置を計測する⁶⁾。実際の環境内にはさまざまな外乱が存在するため、カメラは理想軌道からはずれることがある。しかし、画像上でその誤差を計測すれば、特徴点の3次元位置の誤差を補正することができる。

図10に示すように、注視点の画像中心からのずれは γ とし、カメラの回転角を θ とする。Sは環境内

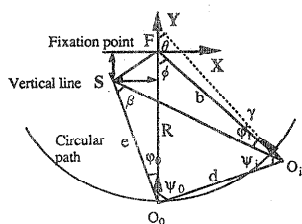


図10 物点の3次元位置の計算

Fig. 10 Geometry for computing the 3D position of scene points.

のある垂直線、 O_0 と O_1 はカメラの位置を示す。これより、垂直線の位置は(15)式のように計算される⁶⁾。

$$x = e \sin \varphi_0 - K_d \sin \varphi_0 \frac{\sin \Psi_i}{\sin \beta} R$$

$$y = e \cos \varphi_0 - R = \left(K_d \cos \varphi_0 \frac{\sin \Psi_i}{\sin \beta} - 1 \right) R$$

$$K_d = \sqrt{1 + \left(\frac{h_{r0}}{h_{r1}} \right)^2 - 2 \frac{h_{r0}}{h_{r1}} \cos \phi} \quad (15)$$

h_{r0} と h_{r1} は、それぞれカメラの位置 O_0 、 O_1 における注視物体の画像上の高さである。 $\phi = \theta - \gamma$ であるから、 x 、 y を求めるには、カメラの回転角を評価する必要がある。これは、環境内に存在する水平線を用いて評価する⁷⁾。

(1) カメラ回転角の評価

カメラの回転角の評価では、環境内の水平線が、ロボットが運動する床面と無限遠点で交わる点(消失点)を利用する。カメラ回転角の評価のための消失点の検出とその利用については多くの研究^{9)~10)}が報告されているが、床面と、それに平行な直線が無限遠点で交わることによって生じる消失点を利用した研究はこれまでにない。一般に屋内環境には多数の床面と平行な直線が存在し、それらを利用して、カメラの回転角が評価できることは、移動ロボットにとって大きな利点である。

この消失点(床面と床面に平行な直線が無限遠点で交わる点)を利用するには、環境内に存在する多数の直線の中から、床面に平行なもののみを抽出しなければならない。

2章で述べたロボットと環境に関する仮定により、カメラの回転軸は床面に対し、垂直であるため、カメラの回転に伴う消失点の画面上での軌跡は直線となる(これを消失線と呼ぶ)。消失線はカメラの回転に対して不変であり、真の消失点はカメラの回転運動に伴ってこの消失線上を移動するが、カメラの直進運動に対しては不動である。この性質を用いて、画像上で水平線とそうでない線を、その消失点の動きから検出する。すなわち、画面上のすべての直線と、消失線との交点を消失点と仮定し、その消失点の動きから、偽の消失点と真の消失点を区別する(偽の消失点、すなわち、床面と平行でない直線と消失線の交点は、カメラの直進運動によって移動する)。その後、検出された真の消失点(床面に平行な直線と消失線の交点)の動きからカメラの回転角を求める。

図11に示すように、時刻 t_1 から t_2 の間カメラ

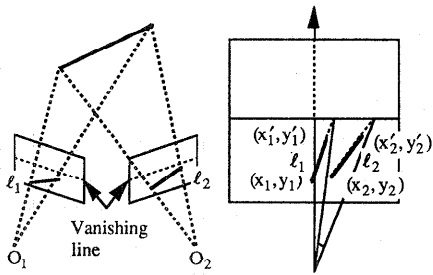


図 11 消失点の運動をもとにしたカメラの回転運動の評価

Fig. 11 Estimation of the camera rotation from the motion of the vanishing point.

の移動に対して、環境における平行線 L の画像上の直線は、 l_1 と l_2 になる。 α を L とカメラの光軸のなす角とすると、カメラの回転角 θ は、 l_1 と l_2 における α の違いによって決定される。結果として、カメラの回転角 θ は(16)式で表される。

$$\theta = \tan^{-1} \frac{x_1 y_1' - x_1' y_1}{(y_1' - y_1) f} - \tan^{-1} \frac{x_2 y_2' - x_2' y_2}{(y_2' - y_2) f} \quad (16)$$

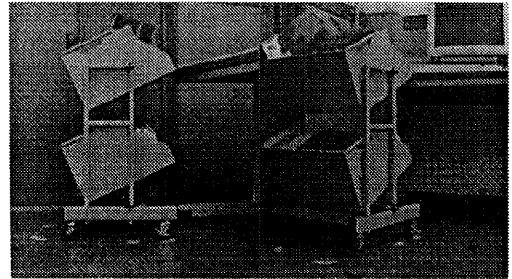
カメラの回転角を評価した後、各観測における垂直線の位置を決定する。垂直線の位置はカメラの注視制御中に連続的に得られるすべての画像について計算され、得られた複数の評価を最小2乗法によって統合し、より信頼性の高い垂直線の位置評価を得る。

(2) 実験結果

このアルゴリズムを、実際に移動ロボットを用いて検証した。実験は、研究室内の本棚を対象物体として行った(図 12 (a))。ロボットは、図中の黒の十字で示される垂直線を注視する。撮像された画像列は 113 フレームである。図 12 (b) に最後の画像を示す。この撮像された画像からそれぞれエッジ画像を得た後、Hough 変換を適用して直線を検出し、その直線の中から水平線のみを抽出する。そして、この水平線を用いてカメラの回転角を評価し、垂直線の 3 次元位置を評価する。ここで、短い直線ではカメラの回転角の評価の信頼性が低いため、長さ 30 pixels 以上のエッジのみを用いた。図 13 (a) に、図 12 (a) から得られたエッジ画像を示し、3 次元位置評価を行う垂直線を表す。実験の結果、得られた 3 次元位置を図 13 (b) に示す。黒の点は実際の垂直線の位置を示し、黒の十字は得られた結果を示す。注視点に近い所では誤差は小さいが、遠いところでは誤差は大きくなる。

4.2 能動的視覚を用いた物体認識

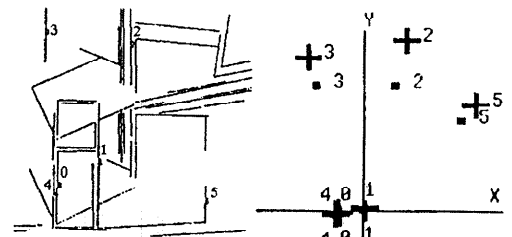
次に、能動的視覚の物体認識への応用について述べ



(a) 最初の画像 (b) 最後の画像
(a) First image (b) Last image

図 12 実験における 2 枚の画像例

Fig. 12 The first and last image of the image sequences.



(a) 抽出された直線 (b) 計算された位置
(a) Extracted lines (b) Computed position

図 13 実験結果

Fig. 13 Experimental results.

る。

すべての方向から物体を観察すると、その物体の完全な 3 次元形状情報が得られる。しかし、認識という目的においては、完全な 3 次元形状情報は必要なく、数少ない 2 次元の画像でも十分な情報を与える。そこで、われわれはロボットがこのような十分な情報を与える画像、いわゆる物体の特徴的な見え方をモデルとして記憶し、ロボットが能動的に移動して未知の物体の特徴的な見え方を探索することにより、物体認識を行う方法を提案する。

物体の特徴的な見え方やその見方はさまざま考えることができるが、ここでは、“物体の各面が円や長方形等の基本的な図形で構成される場合、その基本的な図形が画像上に現れる見え方”という物体の見え方の定義を考える。その見方として、“物体のある面の面積が最大になる視点”を選択する。この見方の定義に従い、ロボットは次のように物体を認識する。まず始めに、ロボットは物体の各面を 2 次元画像としてあらかじめ記憶しておく。次に、環境内から未知の物体の一つを選択し、その物体の一つの面の重心を中心とする

円弧運動を行いながら、その面の面積が最大になる視点に移動する。最後に、その地点で得られた画像と、あらかじめ記憶されている2次元画像との対応を求めることで、物体を認識する。

この方法の利点は、蓄えておくモデルは物体の2次元画像であるため、モデルの生成コストが非常に低く、処理も簡単になることである。また、物体の見かけの形状により認識を誤ることが少なくなる。例えば、ある視点から見た物体の面が楕円であったとする。この視点からだけでは、この面は楕円なのか、円を斜めから観測したために楕円に見えるのかの判断は困難である。しかし、本手法では、ロボットが実際に視点を変えて物体を観測するため、物体の形状を正しく認識して対応付けを行うことができる。

実験結果

実験は、2段の木箱の上段に5つの異なった箱を組み込んで異なった5つの物体とし、その箱の面を木箱

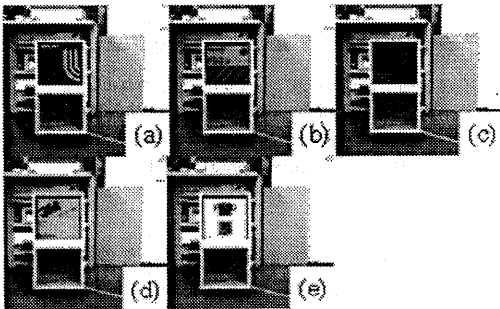
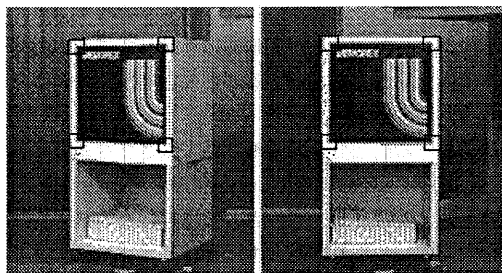


図 14 あらかじめ記録されているモデル
Fig. 14 Stored models.



(a) 円弧運動中に得られた画像
(a) An image taken while the circular motion
(b) 物体の大きさが最大になる視点における画像
(b) Image taken at the view point with maximum target size

図 15 カメラから得られた画像
Fig. 15 Images taken by the camera.

の特徴的な面として行った。モデルとして、この面を正面から撮像した5つの2次元画像を用いた。与えたモデルを図14に示す。実験の対象物体には、図14(a)を用いた。ロボットはこの箱の面の重心を中心とした円弧運動を行い、この面の面積が最大になる位置でモデルとの照合を行った。このときの、物体からロボットまでの距離(回転半径)は約2mである。

円弧運動を行っているときの入力画像を図15(a)に示し、照合を行っている画像を図15(b)に示す。本実験では、図15(b)に示すように、対象物体の4つの頂点を追跡し、注視制御を実現している。モデルと入力画像との照合は、この4つの頂点をもとに対象物を長方形として切り出して得られた長方形の画像と、モデル画像(モデル画像もあらかじめ物体を長方形で切り出してある)を比較することで行う。まず、得られた長方形の入力画像を拡大縮小することで、モデル画像との大きさを合わせ、その後、モデルと入力画像における各画素の輝度の差の和を計算することにより、モデルと入力画像が一致するかどうかの判断を行う。

実験の結果、ロボットは、対象物体は図14(a)に示す物体であると、正しく認識を行った。この実験では、画像内の物体を長方形で切り出すという方法を用いたが、より正確なセグメンテーションを行うことにより、より一般的な認識が可能になる。

5. 結 論

本論文では、環境内を自由に移動するロボットにカメラを搭載することにより、能動的視覚の適応範囲を拡大することを試みた。

まず、能動的視覚に必要なロボットの制御方法として、視覚フィードバックにより対象物体を中心とした円弧運動を行う方法を考案し、シミュレーションと移動ロボットによる実験を行った。

そして、この制御方法を用いて、移動ロボットによる能動的視覚の可能性を示唆すべき二つの実験を行った。一つは注視による物体の3次元構造の獲得であり、内界センサの情報のみでは正確な制御が困難である移動ロボットにおいても、Aloimonosらによって明らかにされた能動的視覚の利点¹⁾が、実現できることを示した。もう一つは、移動ロボットの能動的視覚による新しい物体認識の一手法を提案し、その基礎的な実験を行った。この方法は、計算コストが低く、安定した認識が行えることが実験によって示された。

本論文では、移動ロボットによる能動的視覚の実現

方法について述べたが、今後、能動的視覚による移動ロボットの環境/物体認識において、さらなる研究が必要とされることはいうまでもない。

参 考 文 献

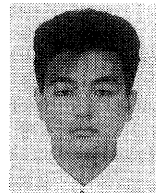
- 1) Aloimonos, J., Bandopdhay, A. and Weiss, I.: Active Vision, *Proc. 1st Int. Conf. Computer Vision*, pp. 35-54 (1987).
- 2) Ballard, D. H.: Reference Frames for Animate Vision, *Proc. 11th Int. Joint Conf. Artificial Intelligence*, pp. 1635-1641 (1989).
- 3) Stelmazyk, P., Ishiguro, H. and Tsuji, S.: Mobile Robot Navigation by an Active Control of the Vision System, *Proc. 12th Int. Joint Conf. on Artificial Intelligence*, pp. 1241-1246 (1991).
- 4) Yuta, S. and Kanayama, Y.: An Implementation of MITCHI—A Locomotion Command System for Intelligent Mobile Robots, *Proc. of 2nd Int. Conf. on Advanced Robotics*, pp. 127-134 (1985).
- 5) Kanayama, Y. and Yuta, S.: Vehicle Path Specification by a Sequence of Straight Lines, *IEEE Trans. on Robotics and Automation*, Vol. 4, No. 3, pp. 265-276 (1988).
- 6) Li, S., Miyawaki, I., Ishiguro, H. and Tsuji, S.: Finding 3D Structure by an Active-vision-based Mobile Robot, *Proc. Int. Conf. on Robotics and Automation*, Vol. 2, pp. 1812-1817 (1992).
- 7) Li, S., Tsuji, S. and Imai, M.: Determining of Camera Rotation from Vanishing Points of Lines on Horizontal Planes, *Proc. 3rd Int. Conf. Computer Vision*, pp. 499-502 (1990).
- 8) Barnard, S. T.: Methods for Interpreting Perspective Images, *Proc. Image Understanding Workshop*, pp. 193-203 (1982).
- 9) Fishler, M. A. et al.: Modeling and Using Physical Constraints in Scene Analysis, *Proc. AAAI-82*, pp. 30-35 (1982).
- 10) Magee, M. J. and Aggarwal, J. K.: Determining Vanishing Points from Perspective Images, *CVGIP*, Vol. 26, pp. 256-267 (1984).
- 11) 石黒 浩: 移動ロボットの視覚システム, 日本ロボット学会学会誌, Vol. 10, No. 2, pp. 159-164 (1992).
- 12) Swain, M. J. and Stricker, M. (eds.): *Promising Directions in Active Vision*, University of Chicago Technical Report CS91-27 (Proc. NFS Active Vision Workshop).

13) Horn, B. K. P.: *Robot Vision*, The MIT Press (1987).

14) Ikeuchi, K.: Generating an Interpretation Tree from a CAD Model for 3-D Object Recognition, *Int. J. Computer Vision*, Vol. 1, No. 2, pp. 145-165 (1987).

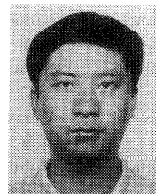
(平成4年12月24日受付)

(平成5年7月8日採録)



宮脇 一路

1991年大阪大学基礎工学部制御工学科卒業。1993年同大学大学院修士課程修了。同年三菱重工業株式会社入社。現在、同社高砂研究所勤務。ロボットビジョンの研究に従事。



李 仕剛

1985年中国清华大学電機工程卒業。1987年来日、1993年大阪大学大学院博士課程修了。工学博士。同年大阪大学基礎工学部助手、現在に至る。ロボットビジョンの研究に従事。



石黒 浩 (正会員)

1986年山梨大学工学部卒業。1988年同大学院修士課程修了。1991年大阪大学大学院博士課程修了。工学博士。同年山梨大学工学部助手、1992年より大阪大学基礎工学部助手、現在に至る。ロボットビジョンの研究に従事。電子情報通信学会、日本ロボット学会、人工知能学会、IEEE、AAAI 各会員。



辻 三郎 (正会員)

1955年大阪大学大学院修士課程修了。工学博士。電子技術総合研究所を経て、1970年より大阪大学基礎工学部教授。現在、同学部長。1990年より1992年まで人工知能学会会長。IJCAI-79, 2nd ICCV の Co-Program Chairman。人工知能、ロボティクス、コンピュータビジョンの研究に従事。計測自動制御学会、人工知能学会、IEEE、AAAI 各会員。