

## 時系列ステレオ画像を用いた環境モデルの構築

津戸明広 谷内田正彦

大阪大学基礎工学部

視覚センサを備えたロボットが環境内を移動すると、様々な視点で大量の視覚情報が得られる。その視覚情報より、ロボットの移動経路周囲の環境モデルを構築する際に問題となるのは、複数の画像間での対応関係を求めることと、新たに得られた環境のデータを用いてモデルを逐次更新、追加していくことである。

本論文では、ステレオカメラを搭載した移動ロボットが環境モデルを構築、更新する一連のシステムについて提案する。このシステムは移動ロボットが獲得した時系列ステレオ画像に3眼視の原理を適用し、移動前後の画像間で線分単位の時空間的対応付けを行ない、環境のワイヤフレームによる3次元幾何モデルを構築する。また、ロボットの移動時にはモデルを用いた効率的な対応付けを行ない、モデルの更新を行なう。

### Reconstructing the Environmental Model by using Sequential Stereo Images

Akihiro Tsudo , Masahiko Yachida

Faculty of Engineering Science , Osaka university  
1-1 Machikaneyama , Toyonaka , Osaka 560

In this paper, we propose a method for a mobile robot to reconstruct the environmental model from a sequence of stereo images. The theory of line-based trinocular stereo is applied to establish spatio-temporal correspondences among the sequential stereo images which have been obtained by a mobile robot with stereo system. Then, wire-framed environmental model is constructed based on the triangulation principle. This 3-D model is used to guide the correspondence process of the next frames which are obtained by the robot's motion, so that efficient processing of motion images of frames is possible. The model is updated when new 3D information is obtained

## 1. はじめに

環境を判断し自律的に動き回る移動ロボット、特に視覚センサを備えた知能ロボットの研究が幅広く行なわれている<sup>(1)(2)</sup>。移動視においてはその移動に伴い様々な視点から大量の視覚情報が得られるが、それらの情報を総合的に解釈して環境を理解するためには異なった視点での画像間でどの部分とどの部分が対応しているのかを決定する、対応付けの問題が非常に重要となる。この問題を解決する有効な方法として3眼視の手法がある<sup>(3)(4)(5)</sup>。

そこで我々はステレオカメラを搭載したロボットが獲得した時系列ステレオ画像に対して3眼視の原理を適用し、線分単位の時空間の対応付けを行なった。線分を用いる利点としては方向、長さといった線分の持つ性質を対応付けに利用できること、エッジ点と比べて画像中の特徴量が非常に少ないため、幾何学的拘束が有効に利用できること等があげられる。

また、対応付けを行なった後、ワイヤフレームによる環境の3次元モデルを構築する。このとき3次元情報を得るためにはロボットの正確な位置を知る必要がある。ロボットの位置は内部センサである程度見積ることができるが、さらに視覚情報を用いて補正を行なう。ロボットが移動を繰り返す過程では、得られた環境モデルを用いて効率的に対応付けを行ないモデルを更新し、新しく発見された線分はモデルに加える。

本報告では、以上の手法について述べ、実験結果を示す。

## 2. 3眼視の移動視への応用

### 2.1. 3眼視の原理

3眼視における線分の対応付けの原理を図1に示す。3枚の画面を基準画面、対応画面、検証画面と呼ぶ。基準画面中の線分 $p_1q_1$ に関して対応をとる場合、まず線分の端点 $p_1, q_1$ のエピポーラライン $e_p, e_q$ を対応画面に引く。対応画面中で少なくともその1部が2本のエピポーラライン間に存在し、似通った属性を持つ線分を対応候補とする。その候補線分に対して互いに端点がエピポーラ拘束を満たすように両線分

のANDをとる。そうして求めた上下の端点の組み合わせ $[p_1, p_2], [q_1, q_2]$ から3角測量で3次元線分の端点 $P, Q$ を求める。その2点を検証画面へ投影した時に投影された2点間 $p_3, q_3$ に線分が存在しなければならない。また、線分を対応の単位とすることにより線分の持つ、長さや傾きといった情報を利用し、対応候補の探索量を減少させることができる。

平行に配置した2台のカメラでは、エピポーララインが水平になり、画像中で水平に近い線分に関しては対応候補が発見しにくいという問題があるが、対応画面を基準画面に対して垂直に近い位置に配置することにより解決することができる。従って、3眼視においては3台のカメラを直角3角形に配置し、線分の傾きにより対応画面を使い分ける方法が有効である。

そこで、移動視に3眼視の原理を適用する方法について次に述べる。

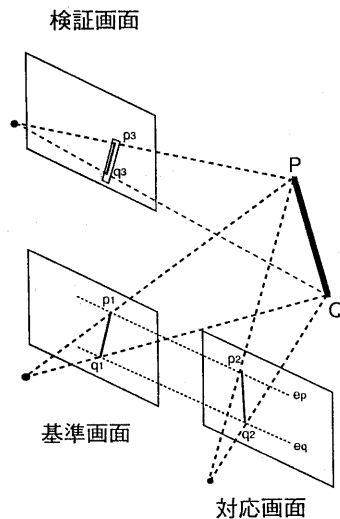


図1 3眼視における線分対応

### 2.2. 移動視への適用

2台のカメラを搭載した移動ロボットが移動前後で得た4枚の画像に3眼視の原理を適用する。この時のカメラの動きを図2に示す。カメラは光軸平行、ベースライン $d$ で、下向き角 $\alpha$ を持つ。カメラは下向き角を持っているので移動前後のカメラの位置関係は、3眼視における垂直方向の配置を実現している。よって、移動

前の2枚の画像と移動後の左右いずれか1枚の画像の計3枚の画像を用いることにより2組の3眼視を行なう。画像の組み合わせを図3に示す。それぞれの組み合わせの3眼視の結果より、各線分が正しく対応付けられている組を選択する。しかし、多眼視の場合各カメラの位置を正確に知る必要がある。移動視の場合ロボットに搭載した内部センサである程度の移動量や姿勢は見積ることができるが、移動前後のカメラ間には当然移動誤差が生じる。そこで視覚センサからの移動誤差の補正法を次に述べる。

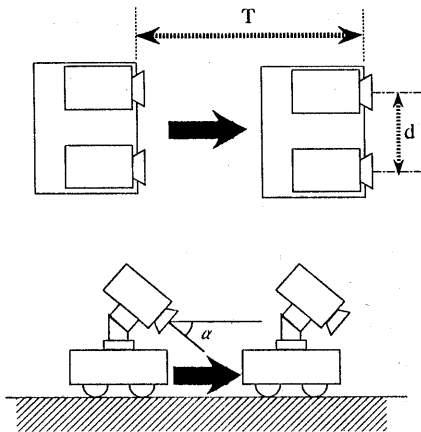


図2 カメラの動き

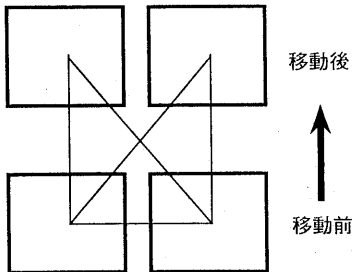


図3 3眼視の組み合わせ

### 2.3. 移動時における誤差の推定

移動視に3眼視を適用した場合、移動前の2台のカメラの位置、姿勢は前もってわかっており、その2台のカメラの視線の交点より、垂直線分の3次元位置を計算することができる。この値を基準値とし、移動の前後における画面間で対応がとれた垂直線分を利用し移動誤差を求める。

本研究では移動ロボットは平らな床面を移動するものと仮定する。運動から生じる移動誤差は平行移動が2成分、ロボットの垂直軸周りの回転移動が1成分の計3つであるとする。

平行移動誤差がなく、回転移動誤差のみが生じた場合を図4に示す。ここでP,Qはある3次元垂直線分の位置、Fはカメラの焦点、XOZは正しいカメラ座標系、X'O'Z'は誤差を含んだ座標系とする。Pのカメラ座標系に対する方位角 $\angle PFO$ と、誤差を含んだ座標系に対する方位角 $\angle PFO'$ との差 $\theta$ は、位置の異なる垂直線分Qの場合でも同じ値となる。しかし、平行移動誤差が生じた場合、これらの方位角の差は線分の位置によってばらつきが生じる。つまり理論的には、平行移動誤差がない場合は各垂直線分に対する方位角の差の分散値は0になり、平行移動誤差が大きくなれば分散値は大きくなる。そこで移動前後の画面間で対応がとれた各々の垂直線分に対してその方位角の差の分散値を計算すると、分散が最小になる位置が現在のカメラの位置である。また、そのときの方位角の差の平均値が回転移動誤差となる。本研究の仮定ではカメラは下向き角を持っており、この場合は投影された線分上の任意の点を用いて方位角を計算することができる。図5、6に線分上の任意の点を用いて方位角を計算する方法を示す。

今、図5はカメラ座標系の側面図、図6は上面図である。図5より、3次元線分上の任意の点の投影点をP( $x_p, y_p$ )、カメラの焦点距離をf、カメラの下向き角を $\alpha$ とする。カメラ中心Oを世界座標系の水平面に投影した点を $D_0$ 、点Pを水平面に投影した後、直線 $\overline{FD_0}$ に投影した点を $D_1$ とすると、

$$\overline{FD_0} = f \cos(\alpha) \quad \dots (1)$$

$$\overline{FD_1} = \overline{FD_0} + y_p \sin(\alpha) \quad \dots (2)$$

となる。また、図6より求める方位角は

$$\theta = \tan^{-1} \frac{x_p}{\overline{FD_1}} \quad \dots (3)$$

となる。

### 3. 環境モデルの構築及び更新

#### 3.1. 初期モデルの構築

時系列ステレオ画像の最初の4枚のうち、前に述べた2組の画像の組み合わせを用いて3眼視による対応付けを行ない、初期モデルを構築する。この場合、環境に関する情報は何も無いわけであるから、可能性のある対応はすべて考慮し、検証する必要がある。基準画面中の垂直、水平線分の本数をそれぞれ  $n_v, n_h$ , 対応画面中の本数を  $m_v, m_h$  とすると、考えられる対応の組み合わせの数  $C_1$  は

$$C_1 = n_v m_v + n_h m_h \quad \dots (4)$$

となる。各組み合わせについて前に述べた3眼視の手法により検証及び3次元位置の計算をし、初期モデルの構築を行なう。

#### 3.2. モデルを用いた対応付けとモデルの更新

初期モデルを構築した後新たに移動を行ない、再び画面間で対応関係を求める。2回目以降の移動時には前回の移動までで得られた環境モデルがすでに存在している。そこで、そのモデルを用いて効率的に対応付けを行なう。

まず、垂直線分の対応を求める場合、基準画面、対応画面ともにカメラの位置は既知であるので、現在の環境モデルを投影することによりモデル中の線分のそれぞれの画面での位置を知ることができる。従って、画面間の対応関係は容易に求まる。

次に、モデルに存在しない線分について初期モデル構築時と同様に対応関係を求め、その後移動誤差の補正を行なったあと、垂直線分と同様に水平線分の対応を求める。

2回目以降  $n$  回目の移動時における対応の組み合わせの数  $C_n$  は、モデルを用いた場合容易に1対1の対応関係が求まるので、モデルの線分のうち画面中に存在する線分を  $n_m, m_m$  とし、それ以外の線分を  $n_0, m_0$  とすると

$$\begin{aligned} C_n &= n_{vm} + n_{hm} + n_{v0}m_{v0} + n_{h0}m_{h0} \\ &= C_1 - A \quad \dots (5) \end{aligned}$$

ただし、

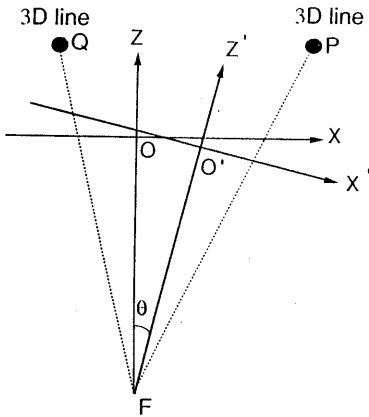


図4 回転移動誤差のみ生じたときのカメラ座標系(上面図)

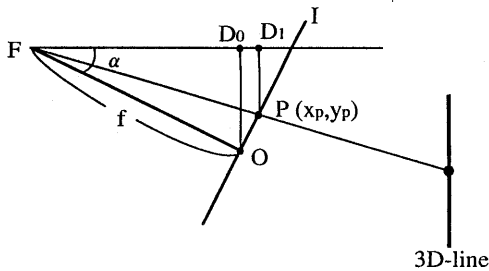


図5 垂直線分の投影(側面図)

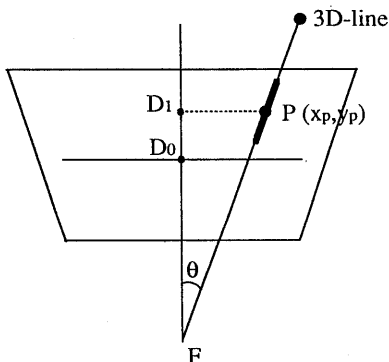


図6 垂直線分の投影(上面図)

$$A = n_{vm}(m_{v0}-1) + n_{hm}(m_{h0}-1) + (n_v m_{vm} + n_h m_{hm})$$

... (6)

となり、対応の可能性がある組み合わせを減少させることができる。

現在のモデル中に存在する線分のうち新たに対応づけられた線分は、3次元位置が得られた場合、より近い位置からの情報と考えられる。そこで、ロボットの移動により随時得られる3次元情報に重み付けを施し、環境モデルの更新を行なう。n回移動した時に得られたモデル中のある線分の位置を  $M_n$ 、このとき更新されてモデルに与えられる位置を  $P_n$  とする。このとき位置  $P_n$  は、新たに得られた位置に対する重みを  $a$ 、それまでに得られている位置の重みを  $b$  とすることにより次式で求める。

$$P_n = aM_n + bP_{n-1}$$

$$= a \sum_{i=2}^n b^{n-i} M_i + b^{n-1} M_1 \quad \dots (7)$$

ただし  $a+b=1.0$  で  $M_1$  は最初に得られた位置である。以上の処理をモデル中の各線分について繰り返す。また現在のモデル中に存在せず、新たに発見された線分に関してはモデルに加えていく。このようにして環境モデルの構築、更新を行なう。モデルの更新の様子を図7に示す。ここで、 $M$  は現在得られているモデル、 $N$  は新たに得られた線分とする。 $M$  と  $N$  の共通部分が更新すべき線分として、式(7)に従って更新され、 $N$  の残りの部分が新しく発見された線分であるので、モデルに加える。これを新しい画像が得られる毎に繰り返し、モデルを更新する。

#### 4. 実験結果

手先にCCDカメラを搭載したマニピュレータで移動ロボットの動きを実現し、処理はSUN SPARC-station330で行なった。得られる画像は256×240画素、各点8bitである。

2台のカメラ間は50mmで、ロボットが50mm移動したら新しく撮像する。なお、カメラの下向き角は30°とする。

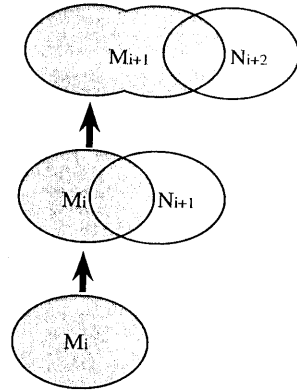


図7 モデルの更新

原画像を図8、抽出された線分を図9に示す。線分抽出のプロセスは以下の通りである。まず、原画像を微分し、2つのしきい値で2値化する。それぞれの2値化画像について微分値の極大点を残して細線化し、しきい値の大きいほうの画像でエッジ追跡し、とぎれたらしきい値の小さいほうの画像でさらに追跡する。得られた点列を最小二乗法で直線近似し、線分を得る。

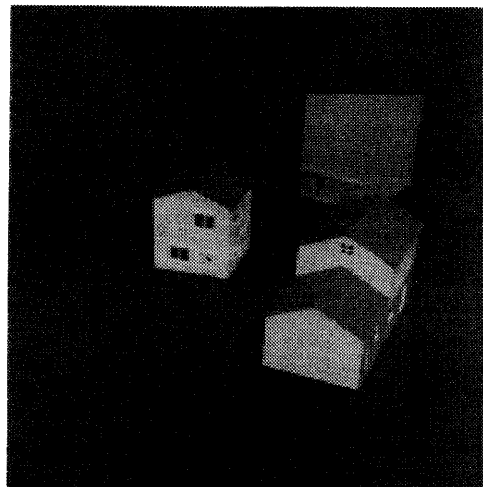


図8 原画像 (移動前左)

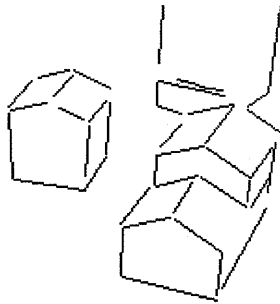


図9 線分抽出結果

こうして得られた最初の4枚の画像の線分を用いて対応付けを行ない、構築された初期モデルが図10である。

次に2回目、3回目の移動の後、更新されたモデルを図11、12に示す。カメラが物体に近づき、線分の位置精度が上がっていることがわかる。計測誤差は物体までの距離に対し約3%程度であった。

一般にステレオ法による距離計測の精度は、カメラ間のベースラインの長ささと物体までの距離、そしてカメラの分解能に依存する。今回の実験においてはカメラ間の距離が50mm、物体までの距離が約300mm～約550mmであった。この設定値を用いたのはカメラのレンズが焦点距離の長めのものを使っており、物体までの距離を固定し、平行な2台のカメラの共通の視野をある程度確保するには、ベースラインを短くする必要があったためである。しかしベースラインを長くし、レンズを広角系のものに変更した場合、カメラの分解能の問題やレンズの光学的なひずみの問題等が生じるため、これらの問題もあわせて検討していかねばならない。

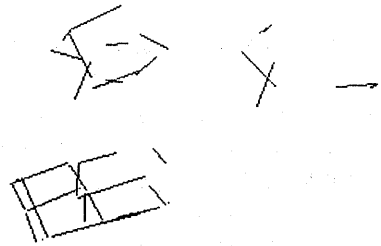


図10 構築された初期モデル(上面図)

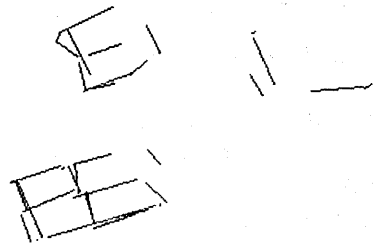


図11 2回目の移動後更新されたモデル(上面図)

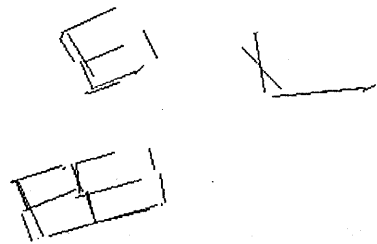


図12 3回目の移動後更新されたモデル(上面図)

## 5. まとめ

時系列ステレオ画像に3眼視の原理を利用し、線分間の対応付けを行ない、ロボットの移動に伴い環境モデルが更新される様子を示した。

本手法で得られる環境モデルはワイヤフレームモデルであり、環境の理解という点では不十分であるが、線分の性質を用いて面の構成の理解や、物体の認識へと発展できると考える。

なお、本研究の一部は文部省科学研究費重点領域（知的情報通信）63633009の援助を受けた。

## [参考文献]

- (1) A.M.Waxman, J.J.Lemoigne, and L.S. Davis  
"A Visual Navigation System for autonomous Land Vehicles" IEEE J. Robotics and Automation, vol.RA-3, No.2, pp.124-141 (1987)
- (2) 鄭、辻：“空間的な対応に基づいた移動ロボットの環境認識” 信学論D vol.J71 - D No.9 pp.1682 - 1692 (1988)
- (3) 来海、谷内田：“3眼視によるセグメント間の対応付けと3次元情報の獲得” 信学技報 PRU88 - 324 pp.13 - 20 (1988)
- (4) Y. Kitamura, M. Kimachi and M. Yachida  
"Trinocular Vision : New approach for correspondence problem" Proc. 8th. ICPR, pp1041-1044 (1986)
- (5) N.Ayache and F.Lustman " Trinocular stereo-vision : recent results" Proc.10th.IJCAI, pp 826-828 (1987)