

## 7. 建設ロボットにおけるナビゲーション用ビジョンシステムの開発\*

呂 健\*\*, 池田博康\*\*\*, 安田克己\*\*\*\*

### 7. Development of Navigation Vision System for Construction Robots\*

by Jian LU\*\*, Hiroyasu IKEDA\*\*\* and Katsumi YASUDA\*\*\*

**Abstract:** In this paper, a navigation vision system is proposed for the industrial mobile robots that are mainly used in unfixed work place, or planned to travel along changeable routes. Instead of the traditional navigation method in which magnetic tags or lanes are constructed on the floor to indicate the route, a series of navigation-signs (briefed as signs in the following) are placed along the route. By change the order and the types of signs in the series, the route can be planned and adjusted simply and flexibly on the spot. The navigation vision system recognizes these signs, detects large non-sign objects as obstacles, and sends these results to the travel control system.

The vision system is a hybrid system, consisted of two functions, i.e., the object detection function and the sign recognition function. The former is used to determine roughly the 3D position and the size of large objects, including signs, within the travel space of the robot. The later is used to recognize the specific type or meaning of a sign. This paper gives the overview of the vision system, describes details about the implementation approaches related to the two functions of the vision system. By experiments and discussions, it was verified that the design specification can be achieved practically by these approaches.

The results and significance of this study are mainly as the follows:

- (1) In the research field of robot travel control, for the requirements of the changeable work fields such as the construction work, where task and travel route often change, the navigation method with flexible and easy route-setting is proposed, and the approach for implementing the system is shown.
- (2) In the research field of robot vision, it is shown that the recognition method for 2D image is applicable even to wide 3D space, in case that the focus objective is effectively controlled according to the 3D information from stereo vision.
- (3) As the application of disparity image and stereo vision, the methods for 3D measurement and object extraction are developed. Although the example object used in this paper is the navigation sign of 30x30cm, the method is applicable to other types of objects simply by learning their features.

**Keywords;** Navigation, Robot vision, Disparity image, Stereo vision, Object detection, Template matching, 3D measurement

---

\* 本研究の一部は、日本機械学会第10回交通・物流部門大会（平成13年12月6日）に発表済<sup>1)</sup>、及び33rd International Symposium on Robotics（October 7-11, 2002）に発表済<sup>2)</sup>である。

\*\* 境界領域人間科学安全研究グループ Interdisciplinary and Human Science Safety Research Group

\*\*\* 機械システム安全研究グループ Mechanical and System Safety Research Group

\*\*\*\* 神鋼電機（株） Shinko Electric Co., Ltd.

## 1. はじめに

労働災害発生比率の依然高い建設業において、安全性を確保しながら、高度な自動化機能を持つ自動機械の導入により、人間の関与を減らして、安全化を推進することは重要である。建設ロボットはこのような手段の典型であり、多くの研究開発が行われているが、実用に至ったものは少ない<sup>3)</sup>。建設現場はFA工場のような完全自動化は困難である。

建設ロボットと作業者が混在する環境で作業を強いられることになるため、本研究では、施工作业に対応する建設ロボットの安全制御システムの開発研究を行い、その一環として、ロボットの自律走行ナビゲーション用ビジョンシステムの検討・開発を行った。本システムは建設用ロボットの早期危険回避と機能的な自律誘導の実現を目的とするものである。

従来、自律走行機能を持つ産業用ロボットのナビゲーション手法としては、磁気タッグ式<sup>4) 5)</sup>とレーン追跡式<sup>6) -8)</sup>の2種類が採用されてきた。磁気タッグ式ナビゲーションでは、ロボットの磁気センサーが走行経路に沿うように床面に設けた磁気タッグを検知し、走行経路を判断する。一方、レーン追跡式ナビゲーションでは、ロボットの画像センサーが床面に貼った色テープにより指示された走行レーンを認識し、追跡する。いずれも、事前にある程度の工事が必要であり、また、作業とルートが予め固定されたFA工場では、経路の設置の不便はないが、作業とルートが変化することが多い建設現場への適用は困難である。

そこで、本研究では、容易にルート変更が可能なロボットビジョンによるナビゲーション方法を提案し、そのシステムの開発を試みた。

## 2. ロボットビジョンの概要

提案するナビゲーション方法では、ロボットの走行経路は床に置いた標識によって指示され、ビジョンシステムはその指示を認識し走行制御部へ対応する走行指令を伝達する。標識の種類は左折、右折、停止、荷卸しなど約十種類がある (Fig. 1)。必要に応じ標識種類の追加も可能である。また、上記何れの標識でもないと判断された検出物は、走行に対する障害物と判断され、回避や停止など措置をとるための位置や大きさ情報を走行制御部へ伝達する。Fig. 2に障害物検出対象の一例を示す。

本研究は走行速度が 1 m/sのロボットを対象として、次の性能仕様を目指して、ロボットビジョンシ

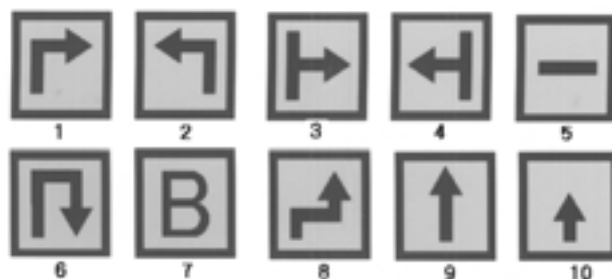


Fig. 1 10 types of navigation-signs.  
10種類のナビゲーション用標識



Fig. 2 An example scene in obstacle detection.  
障害物とする検出対象の一例

ステムの研究開発を行った。

- ・ センシング頻度: 5 Hz
- ・ 標識サイズ: 30 cm × 30 cm
- ・ 検出範囲: ロボット前方 1.6 m ~ 5 m
- ・ 検出視野: 前方 40 度

ロボットビジョンに関する研究やその応用は全く新しいことではないが、建設、機械など産業への応用は今まで主に 2 次元方式に限られている<sup>9) 10)</sup>。例えば、工場生産ラインにおける不良品の検出用センシング手段として良く知られている。これはセンシング距離が近く (空間が小さく)、または固定され、かつ対象物の形状と種類が 2 次元パターンで簡単に表現できるからである。しかし、本研究の開発対象とする建設ロボット用ナビゲーションの前記仕様では、センシング空間は相対的に広く、テンプレートマッチングなど伝統的な 2 次元画像処理・認識手法のみでは対応できないので、より高度化した 3 次元手法の適用が必要である。一方、3 次元手法を用いたロボットビジョンは研究段階に留まっており、産業へ普及応用に至らない主な理由は、2 次元の場合に比べると処理時間が遥かにかかるためである。

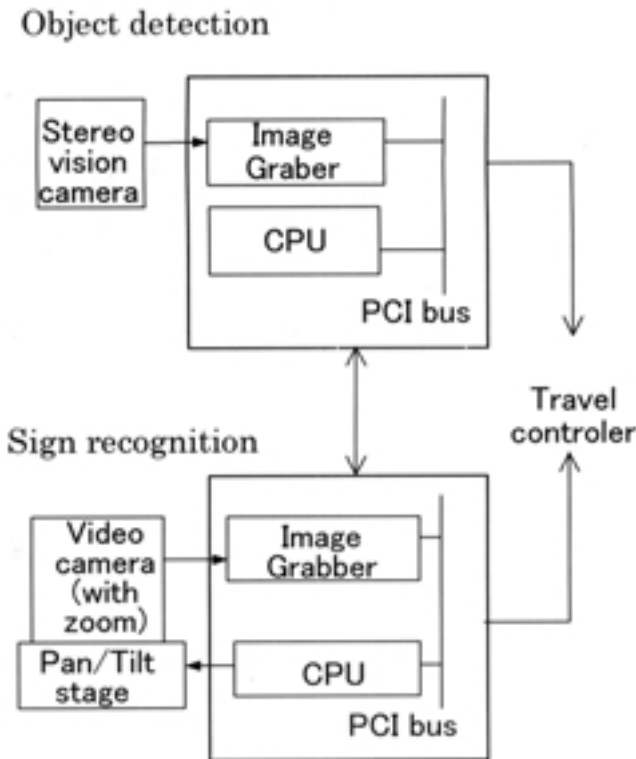


Fig. 3 Hardware structure of the vision system. ビジョンシステムのハードウェア構造

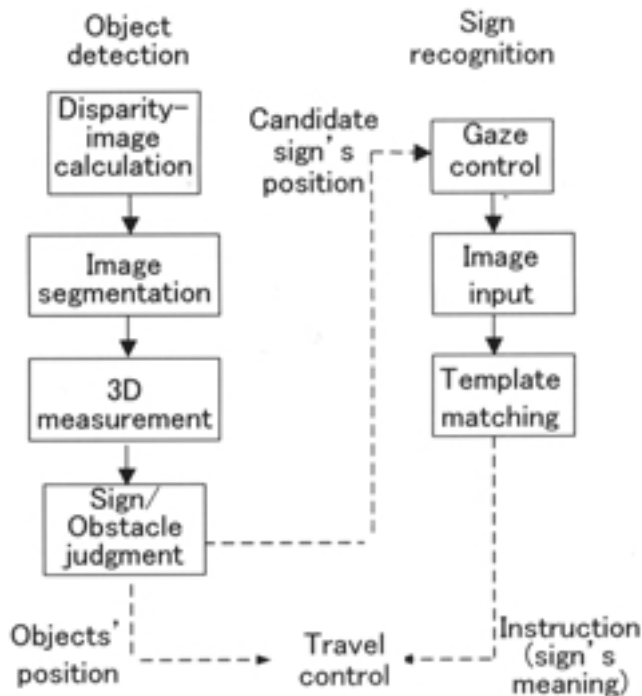


Fig. 4 Approach for the vision system. ビジョンシステムの実現方法

そこで、本研究での解決方法として、3次元手法の立体視を行う目標抽出部と伝統的な2次元処理を行う標識識別部の2部分の機能から構成するハイブリッドビジョンシステムを提案した (Fig. 3)。

目標抽出部は立体視によって、広範囲の走行前方に対して、対象とする物体の抽出及び粗い認識が高速に行える。標識認識部は目標抽出部からの情報に基づき、標識候補など重要目標に対し、カメラのズーム・パン・チルト機能により重要目標に対する注視及び詳細な認識が可能である。このような構造のハードウェアをベースにしたロボットビジョン処理の流れをFig. 4に示す。

### 3. 目標抽出部

Fig. 4の左側に示したように、目標抽出部処理は4つのステップからなり、視差画像の計算、画像の分割、抽出物体の3次元計測、及び標識判断という順番で行われる。以降、それぞれの機能について詳述する。

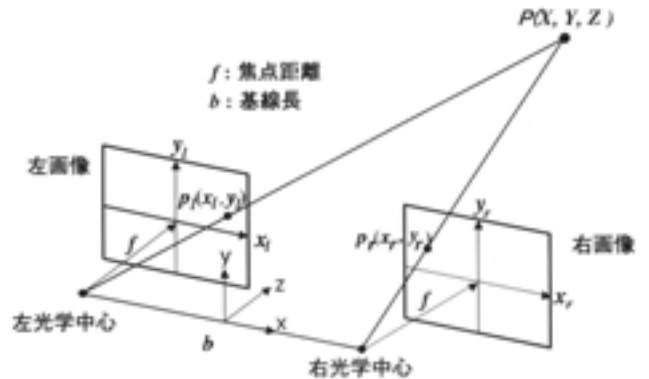


Fig. 5 Principle of stereo vision. 立体視原理

#### 3.1 視差画像の計算

立体視の基本原理は三角測量である<sup>4)</sup>。視差画像はこの基本原理に従い計算された視差値がピクセルとして構成される画像である。もっとも簡単な立体視の構成は、焦点距離が等しいカメラ2台が、光軸がお互いに平行で、かつ各々の画像面が同一平面上にあるように配置される場合である。Fig. 5に示すように、World座標系  $(x, y, z)$  と各々の画像における画像座標系を決めた時、空間中にある点  $P(X, Y, Z)$  が、各々の画像上  $(x_l, y_l)$ ,  $(x_r, y_r)$  に投影されたとすれば、以下の関係式が成立する。

$$X = \frac{b}{2d} (x_l + x_r) \tag{1}$$

$$Y = \frac{b (y_l + y_r)}{2d} \quad (2)$$

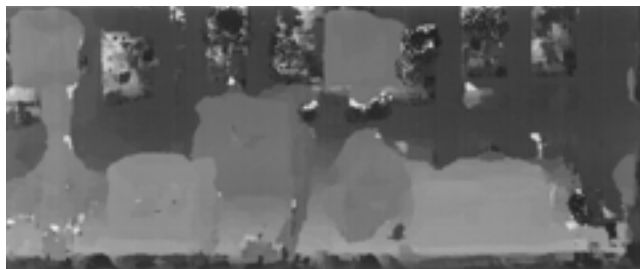
$$Z = \frac{bf}{d} \quad (3)$$

ただし、 $f$ は焦点距離 (focal length),  $b$ は基線長 (baseline),  $d$ は視差 (disparity) で、

$$d = x_l - x_r \quad (4)$$



(a) Reference Image  
参照画像



(b) Disparity Image  
視差画像

Fig. 6 Disparity image and reference image.  
視差画像とその参照画像

である。従って、 $f$ と $b$ が既知であれば、左右の画像上の投影座標から、元の三次元位置 ( $X, Y, Z$ ) が計算される。特に、奥行き $Z$ は、視差 $d$ のみから決定される。

Fig. 6 は視差画像を示す一例であり、同6 (b) は (a) のような場面に対し、式 (4) の計算から得られた視差画像である。

### 3.2 視差画像に基づく三次元計測

従来、運搬ロボット及び自動車用前方障害物検出のため、前述の立体視原理に基づく視差データの利用方法が研究されてきた<sup>5)</sup>。しかし、これらの用途は障害物の有無検出に限定され、位置計測を含まれていない。そのため、視差画像は幅と高さが各々数十ピクセル程度の低解像度なものでよく、ピクセル

ごとに三次元空間の点へ変換することで、直接障害物の判断が可能である。

本研究では、目標物の三次元空間位置と大きさも検出・計測の対象であるため、 $320 \times 240$ 以上のピクセル数を含む高解像度視差画像を用いている。しかし、このような高解像度画像の全ピクセルの三次元変換を求めることは処理時間がかかるため、ピクセルの分類 (即ち視差画像の分割) による効率的な手法が求められる。高解像度視差画像を対象とした画像の分割方法はいくつか提案されているが、それらはリモートセンシング等リアルタイム性を要求しない用途向けであるために、複雑な計算が必要となる<sup>6) 7)</sup>。そこで、ここでは、一般の二次元画像処理に用いられる領域拡張法<sup>8)</sup>と凸分割法<sup>9)</sup>の併用によって、簡易な視差画像の領域分割を実現する。分割された領域毎に3次元計測を行うこととした。

これら一連の処理を実現するためのステップは次の通りである。

- 1) 視差画像に対し、近傍の画素の間に、視差値の近いものを同一領域に属するような領域拡張処理を行う。これは、同一目標物に対応する画素を同一領域へ分類するための処理である。
- 2) 上述 1) で得られた領域に対し、凸領域分割を行う。距離の近い異なる目標物に対応する画素は 1) で同一領域に分類される可能性があるため、このステップは異なる目標物に対応する画素を異なる領域へ分類するための処理である。
- 3) 上述 2) で得られた面積が予め決められた閾値以上である各領域に対し、各々の二次元重心画素に対応する座標 ( $X, Y, Z$ ) を求め、これをその領域に対応する目標物の重心位置とす

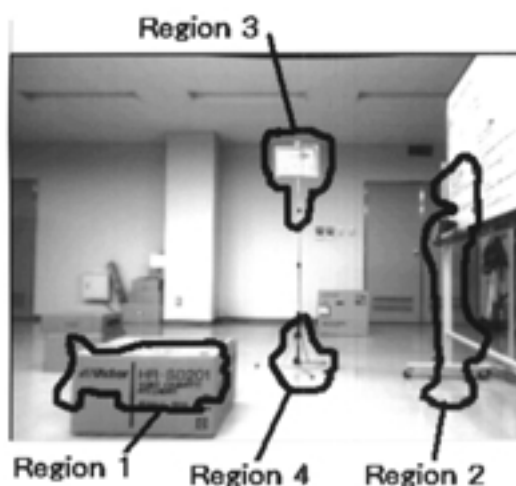


Fig. 7 Object/region extraction.  
領域分割結果の一例

Table 1 3D Measurement of Objects.  
目標物の3D計測値

| Region No. | Object Gravity Size [m]<br>(X, Y, Z) |       |      | Size [m] |       |
|------------|--------------------------------------|-------|------|----------|-------|
|            | X                                    | Y     | Z    | Xsize    | Ysize |
| 1          | -0.34                                | 0.24  | 0.96 | 0.38     | 0.27  |
| 2          | 0.58                                 | 0.12  | 1.67 | 0.29     | 0.95  |
| 3          | 0.0                                  | 0.42  | 1.92 | 0.25     | 0.30  |
| 4          | -0.01                                | -0.43 | 1.95 | 0.31     | 0.42  |

る。領域境界の全画素に対応する  $(X, Y)$  座標から、 $\min\{X\}$ ,  $\max\{X\}$ ,  $\min\{Y\}$ ,  $\max\{Y\}$  を求め、対応目標物の  $X, Y$  座標の最大値と最小値とする。さらに、領域に対応する目標物の幅、高さとして  $X\ size = \max\{X\} - \min\{X\}$ ,  $Y\ size = \max\{Y\} - \min\{Y\}$  を求める。

Fig. 7 には、上述の 1) と 2) のステップを行ってから得られた領域を示す。本来の処理によれば、数百の領域が得られたが、簡潔な表示とするため、ここでは面積が閾値 300 以上の領域のみを示した。実用上、近くにある、または大きな目標物を優先的に認識しなければならないため、大きな分割領域に対応するものとしている。Table 1 は Fig. 7 で抽出された領域に対して、上記 3) のステップを実行した 3D 計測結果である。

### 3.3 三次元計測値による標識・障害物検出

次に、Table 1 のように抽出された目標物ごとの 3 次元計測データに基づき、統計的パターン認識方法を用いて、予め学習で獲得した標識の大きさや設置高度など標識に関する標準特徴との距離を求め、標識検出を行う。

#### 3.3.1 標識の標準特徴

今回のビジョンシステムの設計では、標識の大きさは  $0.3\text{ m} \times 0.3\text{ m}$  であり、設置高度は立体視カメラの光学中心と同じ  $0.85\text{ m}$  である。Table 1 と同じ  $Y, X\ size, Y\ size$  の 3 つの特徴ベクターから構成する特徴空間を定義すると、標識としての特徴は次のように表すことができる。

$$Y = 0 \quad (5)$$

$$X\ size = 0.3 \quad (6)$$

$$Y\ size = 0.3 \quad (7)$$

さらに、標識判断を行う場合、次の評価関数を使う。

$$Wt = \text{sqr}(Y) + \text{sqr}(X\ size - 0.3) + \text{sqr}(Y\ size - 0.3) \quad (8)$$

ただし、 $\text{sqr}$  は平方値を求める関数である。

式 (8) から分かるように  $Wt$  は前述特徴空間における抽出物と標識の距離の平方値である。本来距離の計算として、式 (8) 右辺の平方根を使うべきものであるが、計算上の便利さを考慮し、平方根を求めずに、そのまま特徴空間の距離評価に使われる。

#### 3.3.2 標準特徴に対する校正とその学習

式 (8) において、 $Wt = 0$  の場合、対応する抽出物が標識であると判断できる。しかし、これは理想的な条件での特徴表示にすぎないため、実際の応用環境に応じ、3D 計測値  $X, Y, Z$  及び  $X\ size, Y\ size$  の誤差により、上記評価関数に使われる標識の標準特徴



Fig. 8 Experiment of measurement error on standard features.

#### 標準特徴の計測誤差に関する実験

値に対する校正が必要となってくる。

このため、検出対象空間の奥行き  $2\text{ m} \sim 4\text{ m}$  範囲のいくつかの代表的な位置においた標識に対する 3D 計測実験を行い、計測値と実際値の誤差を測った。Fig. 8 はこのような実験の一例 (標識位置  $(Z, X) = (3, 0)$ ) を示した。Fig. 9 にはこれら実験の結果を示し、同図において、横軸は実際のな標識設置位置  $(Z, X)$ 、縦軸は大きさ誤差 (Fig. 9 (a))、と位置誤差 (Fig. 9 (b)) をそれぞれ示した。同図に示しているように、奥行き  $Z = 2\text{ m} \sim 4\text{ m}$  の全実験の範囲において、 $X\ size, Y\ size$  の誤差の範囲は  $(0.14, 0.62)$ 、誤差の中心値は  $0.38$  である。標識の  $X\ size, Y\ size$  の標準値として、式 (6), (7) の  $0.3$  の代わりに  $0.3 + 0.38$  を使うことは誤差修正案の一つになる。しかし、実際には、ロボットが標識に向かって接近するので、 $4\text{ m} \sim 5\text{ m}$  の遠距離空間より、 $2\text{ m} \sim 3\text{ m}$  の近距離空間での検出をより重視すべきと考えられる。このため、 $Z$  値が  $2\text{ m} \sim 3\text{ m}$  の空間に対応する誤差範囲  $(0.14, 0.31)$  の中心値  $0.23$  を実際のサイ

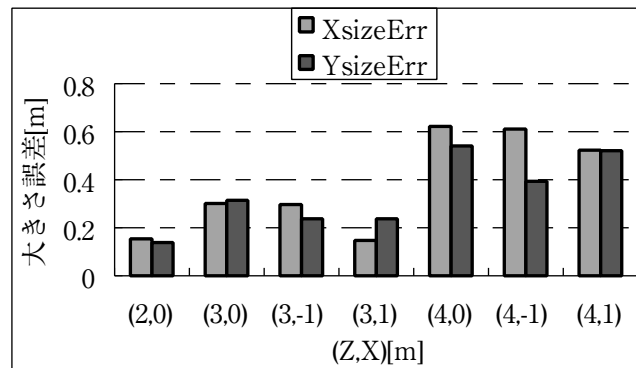


ズ 0.3 に加えて、標識の Xsize, Ysize の校正値とした。一方、位置座標値の Y においては、誤差があるものの、Xsize, Ysize と比べると小さく、その範囲も 0 を中心とするので、校正は不要と考える。そこで、今回開発したシステムでは、標識判断のために下記の評価関数を採用した。

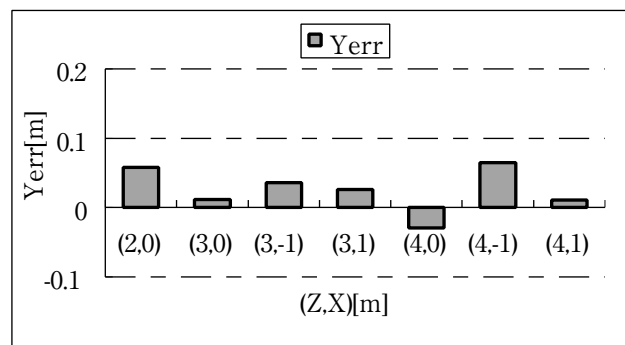
$$Wt = \text{sqr}(Y) + \text{sqr}(Xsize - 0.53) + \text{sqr}(Ysize - 0.53) \quad (9)$$

### 3.3.3 閾値による目標物の判断

式 (9) の評価関数を用いた判断では、直接  $Wt = 0$  を条件とすると、判断ミスの許容範囲が小さすぎるので、 $Wt = 0$  として判断できる閾値を使う。この閾値を得るため、前述と同じ実験において、式 (9) により  $Wt$  を計算した。結果を Fig. 10 に示す。 $Wt$  は高々 0.3 であるので、閾値を 0.3 に設定した。すなわち、 $Wt < 0.3$  の抽出物を標識とし、他の抽出物が障害物と判断する。



(a) Error on size  
大きさ誤差



(b) Error on Y  
位置の Y 座標誤差

Fig. 9 Measurement Error on standard Navigation-signs.  
標識の標準特徴に対する計測誤差

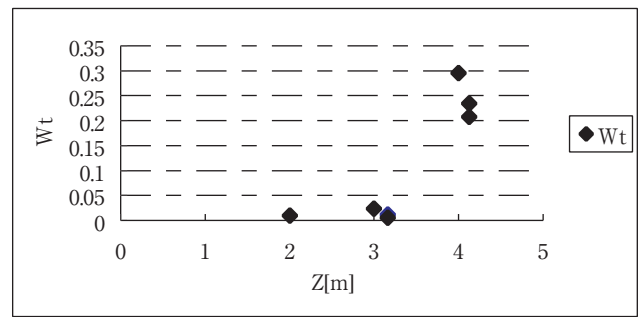


Fig. 10 Distribution of evaluation function  $Wt$  For standard signs.  
標準的な標識に対する評価関数  $Wt$  の分布

## 3.4 目標抽出の性能評価

### 3.4.1 実験結果

開発した目標抽出部の標識検出における性能評価実験を行った。Fig. 11 のような環境に対して行った実験の結果を Table 2 に示す。各行はそれぞれ異なる位置に設置された標識に対する計測値である。(Z, X) 列は標識の Z 座標と X 座標を表す。なお、すべて実験において、標識設置の高さは立体視カメラの光学中心と同じ 0.85 m なので、 $Y = 0$  となる。Table 2 に示したように、奥行き  $Z = 2 \text{ m} \sim 4 \text{ m}$  の範囲に対し、すべての実験に抽出された標識候補領域に対する評価関数  $Wt$  は 0.3 以下の値が得られ、漏れなく標識が検出できている。

### 3.4.2 検出漏れと誤検出について

奥行き座標  $Z$  が 5 m 以上においては、標識部分対応の視差画像が小さすぎるため、使われる面積閾値が大きめに設定された場合、視差画像から標識領域の抽出が正しくできず、標識の検出も不可能（いわゆる検出漏れ）である。逆に、使われる閾値が小さめに設定された場合、標識の正しい検出が可能になるが、非標識目標を標識として検出・判断する誤検出の恐れがある。しかし、これらいずれの場合にも、

Table 2 3D measurement of detected signs.  
検出された標識の 3D 計測値

| (Z, X) | $Wt$  | X     | Y     | Z     | Xsize | Ysize |
|--------|-------|-------|-------|-------|-------|-------|
| (2,0)  | 0.008 | 0.073 | -0.02 | 2.003 | 0.459 | 0.422 |
| (2,-1) | 0.019 | -0.91 | 0.009 | 1.997 | 0.37  | 0.455 |
| (3,0)  | 0.009 | 0.096 | 0.025 | 3.12  | 0.575 | 0.549 |
| (3,-1) | 0.059 | -0.95 | 0.009 | 3.172 | 0.743 | 0.488 |
| (3,1)  | 0.023 | 1.089 | 0.013 | 2.959 | 0.583 | 0.627 |
| (4,0)  | 0.137 | 0.152 | 0.047 | 4.308 | 0.821 | 0.678 |
| (4,-1) | 0.229 | -1    | 0.1   | 4.448 | 0.74  | 0.902 |
| (4,1)  | 0.204 | 1.213 | 0.064 | 4.33  | 0.831 | 0.801 |

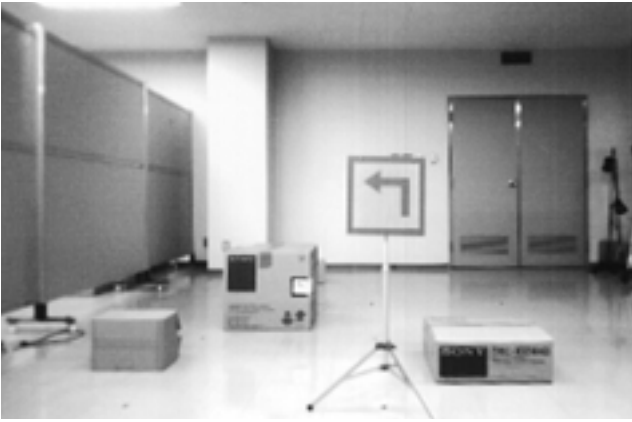


Fig. 11 Experiment of sing detection.  
標識検出実験の一例

実際のナビゲーションでは問題にならないと考える。なぜなら、検出漏れは 5 m 以上の遠距離の場合に発生するが、4 m 以内の範囲に入ると、漏れなく検出できるようになり、ロボットの走行制御へ伝達して走行制御が機能する時間的余裕が得られる。一方、抽出物が標識として誤検出する場合、次の段階の標識識別部によって、誤りを発見できる。

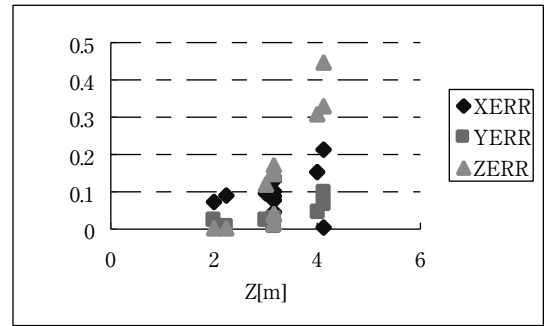
### 3.4.3 計測誤差

Fig. 12 は本研究で開発した目標抽出部で検出された標識の位置とサイズに対する計測値の誤差を示している。

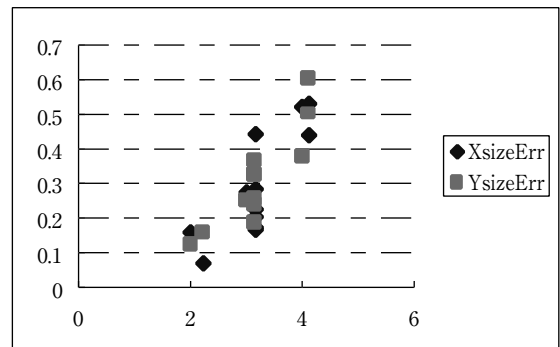
立体視カメラの仕様データによると、距離 2.0 m, 3.0 m, 4.0 m の分解能はそれぞれ約 0.15 m, 0.35 m, 0.55 m である<sup>15)</sup>。これに対し、本研究の実験結果は Fig. 12 に示すように、位置計測誤差 (XERR, YERR, ZERR を含む) はそれぞれ 0.1 m 以下, 0.2 m 以下, 0.45 m 以下であり、大きさ誤差 (Xsize, Ysize を含む) がそれぞれ 0.15 m 以下, 0.45 m 以下, 0.65 m 以下であった。カメラから得られた視差画像のピクセル毎の分解能と比較し、一連の処理ステップを行ってからの抽出物に対して、Fig. 12 に示したようなほぼ同じ計測精度を得られた。この結果から、提案した目標抽出方法の適切性が示されたと考えられる。

### 3.4.4 目標抽出の速度

実験には 700 MHz Pentium III の PC を使用し、4 Hz の目標抽出速度が得られた。開発環境に使われているソフトウェアのライセンスの関係で、今はこのタイプの PC での実験に留まっているが、実応用システムの実現のためには、より高速の PC を使うと、設計仕様の 5 Hz のセンシングレートは達成できると考えられる。



(a) Measurement error on position  
位置誤差



(b) Measurement error on size  
サイズ誤差

Fig. 12 Measurement error in sign detection.  
標識検出の計測誤差

## 4. 標識識別方法

本研究で、標識識別部の実現のための課題は、目標抽出部の出力情報に基づいて、テンプレートマッチング手法を標識識別に有効に適用するため、如何にズーム・パン・チルト等の注視制御を効率良く行うかということである。そこで、ここではテンプレートマッチングと注視制御についてそれぞれ説明する。

### 4.1 テンプレートマッチング

テンプレートマッチングは 2 次元画像処理に良く使われている方法である<sup>16)</sup>。その基本原理として、入力画像をあらかじめ登録されたパターン (テンプレート) とマッチングすることによって、入力画像のパターンを識別する。マッチングに対する評価は正規化相関値により行われる。入力画像があるパターンと一致する場合、相関値は最大の 100 % である

が、誤差を許容するため、実際には 100 %より小さい閾値を使う。

本研究に使われている 10 種類の標識画像の相互相関値（異なる種類の標識の類似性を表す）を **Table 3** に示す。表の見方は、例えば、表の 1 行目が **Fig. 1** に 1 番の標識（右折）をテンプレートとしたときに他の標識との相互相関値を表す。なお、相関値 10 %未満の場合は空欄にしている。ここでは、相互相関値は最大でも 42 %である。このため、42%より大きい値（例えば 50 %）を閾値として設定すれば、相関値が閾値より大きいである場合は、対応の標識とし、逆の場合は、対応の標識ではないとそれぞれ判断ができる。

Table 3 Correlation coefficients between signs.  
標識の相互相関値

|    | 1 | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9 | 10 |
|----|---|----|----|----|----|----|----|----|---|----|
| 1  |   | 14 | 15 |    |    | 15 |    |    |   |    |
| 2  |   |    |    |    |    |    |    |    |   |    |
| 3  |   |    |    | 12 | 37 | 18 |    | 15 |   |    |
| 4  |   |    |    |    | 42 |    |    |    |   |    |
| 5  |   |    |    |    |    |    |    |    |   |    |
| 6  |   |    |    |    |    |    | 27 | 18 |   |    |
| 7  |   |    |    |    |    |    |    |    |   |    |
| 8  |   |    |    |    |    |    |    |    |   |    |
| 9  |   |    |    |    |    |    |    |    |   | 33 |
| 10 |   |    |    |    |    |    |    |    |   |    |

#### 4.1.1 前処理

テンプレートマッチングを実行する前に、その準備段階の前処理が必要である。主な前処理として、カラー変換、二値化及び粒子切り出しの 3 つがある。これら前処理についての詳細は文献<sup>16)~18)</sup>にあるが、以下に簡単に説明する。

カラー画像の色表示方法として、三原色を用いる RGB 表示と色相 H、彩度 S、輝度 L を用いる HSL 表示がある。デジタル方式でカラー画像を処理するとき、CCD カメラからの入力画像は一般的に RGB 表示である。カラー変換は RGB 表示の入力画像を HSL 表示へ変換し、カラー特徴の抽出と色解析の簡単化を図る。カラー変換処理後の HSL 表示の標識画像において、バックグラウンドとする青色部分は色相角度が 90 ~ 270 範囲の値になり、矢印など赤色部分の色相角度はそれ以外の範囲の値になる。これで後の二値化処理が容易になる。

二値化処理は HSL 表示のカラー画像を 0 (黒) と

1 (白) の二種類の値で表す白黒画像へ変換する。二値化変換後、標識画像において、元々の青色のバックグラウンド部分は値が 1 のピクセル (白) になり、矢印等他の部分は値が 0 のピクセル (黒) になる。カメラからの入力画像はこれら前処理の後、二値画像としてテンプレートマッチングが行われる。

二値画像において、すべての画素はその値が 0 か 1 によって、前景と背景に分類される。前景は画像解析・認識の目標物など画像の重要な部分に対応し、背景は前景以外の部分に対応する。粒子切り出しは二値画像から前景を切り出す処理であり、切り出された部分は粒子画像とも呼ぶ。ここでの粒子切り出し処理は、標識部分を含む元々大きな入力画像から、標識の全体を含む最小領域を切り出す処理である。

#### 4.1.2 テンプレートマッチングの処理ステップ

前処理を行ってから、テンプレートマッチングの処理は次のステップからなる (**Fig. 13**)。

##### 1) 前準備

アプリケーション起動時に、登録してある標識画像データを読み込み、格子状に配列してひとつの画像にしてまとめておく (これを参照画像とする)。

##### 2) 識別処理

- a) 領域分割処理から渡された粒子画像を標識画像データと同じ大きさに拡大・縮小する (サイズ正規化)
- b) 正規化された粒子画像をテンプレートとして、参照画像をスキャンして最大相関値となる場所を検出する
- c) 検出された場所から、マッチした標識画像を特定する

ここに、相関最大値が設定した下限値を下回る場合は、該当する標識データなしとして、次の候補粒子画像でマッチングする (次がなければ標識無しと判定する)。

#### 4.2 注視制御の概要

**Fig. 14** に示したように、注視制御は標識識別のための第一ステップとして、パン・チルト角度の計算、ズームの計算、及びこれらの計算に基づいたパン・チルト、ズーム等動作の制御を行う。

##### 4.2.1 ビデオカメラ座標系への座標変換

目標抽出部から得られた抽出物に関する位置情報 (X, Y, Z) は、立体視カメラの光学中心が原点とする位置座標である。これを、標識識別用ビデオカメラの中心点が原点とする座標 (xc, yc, zc) への変換が必要であるため、以下の処理を行う。



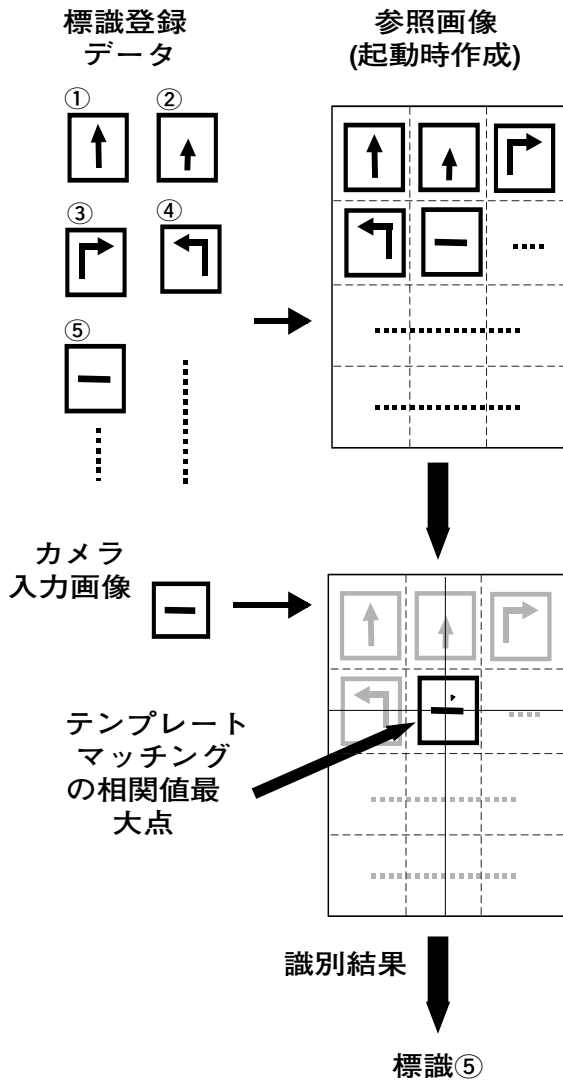


Fig. 13 Template matching. テンプレートマッチング

$$\begin{aligned} xc &= X + tx \\ yc &= Y + ty \\ zc &= Z + tz \end{aligned}$$

ここで、 $(tx, ty, tz)$  はオフセットである。

#### 4.2.2 パン・チルト角度の計算

前にも述べたように、標識識別用センサーとしてのビデオカメラはパン・チルト機能付きの雲台に設置される。対象物をこのカメラの視野中心に捕らえるために、目標物検出部により得られた標識の位置情報から、パン・チルト角度が計算され、パン・チルト動作が制御される。Fig. 15 に示したカメラ中心が座標原点の座標システムにおいて、標識位置は  $(xc, yc, zc)$  とすると、パン角度は

$$\theta = -\tan^{-1}(xc/zc) \tag{10}$$

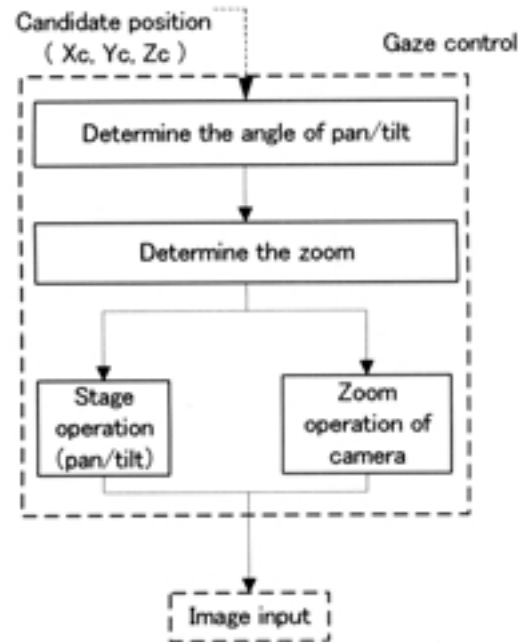


Fig. 14 Flowchart of gaze control. 注視制御の流れ

であり、チルト角度は

$$\phi = -\tan^{-1}(yc/\sqrt{xc^2 + zc^2}) \tag{11}$$

である。

#### 4.2.3 ズーム値の計算

雲台が標識画像がとれる適切な角度へ調整されたらならば、次のステップにおいては、カメラが所定の大きさの標識画像を入力するため、カメラのズームを適切に設定しなければならない。

Fig. 16 に示したように、標識のサイズに対応する画角  $B$  に加えて、マージン角  $Ma$  をとる必要がある。このため、目標値としての画角  $A$  は

$$A = B + 2Ma \tag{12}$$

である。なお、画角  $B$  は

$$B = 2 \tan^{-1}(S/2L) \tag{13}$$

である。

#### 4.3 注視制御における問題点

本節は、標識識別部において、仕様達成を目指すために解決すべき問題点を検討し、実験を通して検証する。

標識のサイズは  $30 \text{ cm} \times 30 \text{ cm}$  に固定され、距離が遠い場合 (最大  $5 \text{ m}$ ) にも、カメラのズーム機能によって、標識画像をテンプレートマッチン

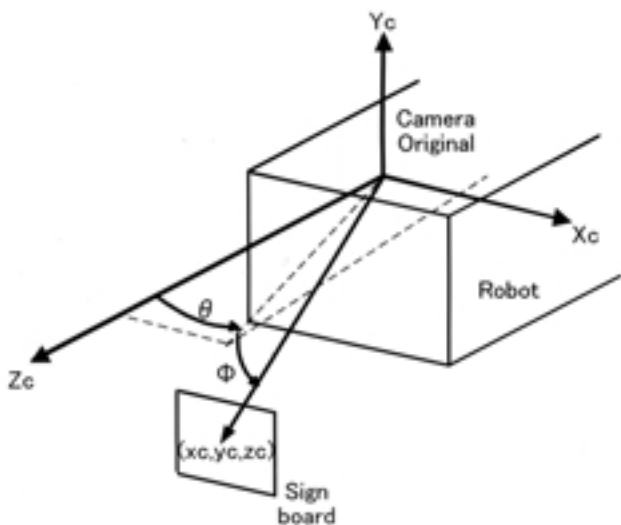


Fig. 15 Camera coordinate system.  
カメラの座標系

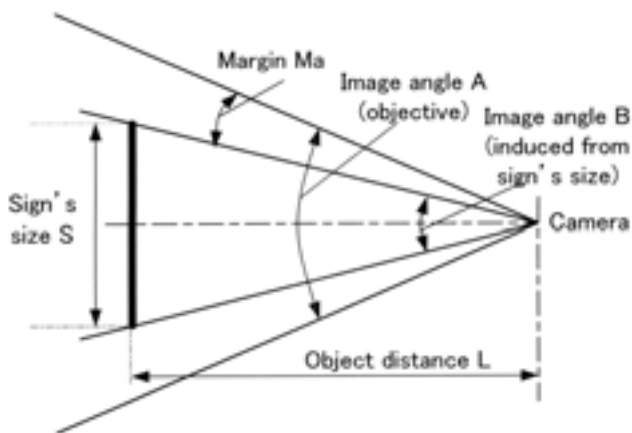


Fig. 16 Objective zoom value (image angle).  
目標ズーム値 (画角)

グ処理に適す大きさで獲得できるため、目標抽出部のような精度や分解能問題はない。その一方、カメラのズームや雲台のパン・チルト等の操作は機械動作であるため、これらの動作時間は仕様達成のボトルネックになる。ここではこれら操作時間の問題について考察する。

Fig. 4 の右側に示したように、標識識別は注視制御、画像入力、画像処理（テンプレートマッチング）の3つのステップからなり、所用時間は3つのステップの合計時間である。これについて、次の実験データがある。

- 1) 画像入力にかかる時間は 66 ms である。
- 2) 画像処理にかかる時間は約 30 ms である (Pentium III 700 MHz での測定値)。

よって、Fig. 17 に示すように、標識識別にかかる時間は約 96 ms に注視制御時間を加えたものとな

る。注視制御用画角（ズーム値）やパン・チルト角度の計算にかかる時間は機械的なズーム、パン・チルト動作時間と比べると、無視できる。また、ズーム動作は雲台動作（パン・チルト）と並列実行できるため、注視制御時間はズーム動作とパン・チルト動作のなかの何れか時間の長いものと見なすことができる。

Table 4 は実験によるズーム時間測定値である。Table 4 (a) と 4 (b) の両方ともズーム目標距離を 5000 mm から 1600 mm へ設定し直す動作の測定値であるが、4 (a) は 500 mm の間隔のズーム動作に対し、4 (b) の方は 1000 mm の間隔のズーム動作に対する測定値である。4 (c) は一回のズーム操作で目標距離を 5000 mm から 1600 mm へ設定した結果である。Table 5 に示しているのはパン・チルトの操作速度に対する測定値である。

Table 4 と 5 から、次のことが注目される。

- 1) ズーム動作は相対的に時間がかかる。全てのズーム動作は一回に 200 ms 以上の時間がかかった。
- 2) パン・チルト動作の速度は 2.93 ~ 147.14 度/秒であり、平均値は約 72 度/秒である。通常のパン・チルト角は 10 度以下であることを考えると、 $10/72 = 138$  なので、パン・チルトにかかる時間は平均 138 ms 以下と見なせる。

仕様上の 5 Hz というセンシング頻度は、一連の動作が遅くとも 200 ms 以内に完成する必要がある。よって、ズーム動作時間は標識識別の全動作時間のボトルネックであるため、ズームの設定は一定の距離範囲に対して固定させ、再設定動作はできる限り少なくすることとする。

#### 4.4 ズーム動作の減少方法

頻繁なズーム動作を避けるために、ズーム設定は一定の目標距離範囲に対して固定することが望まし

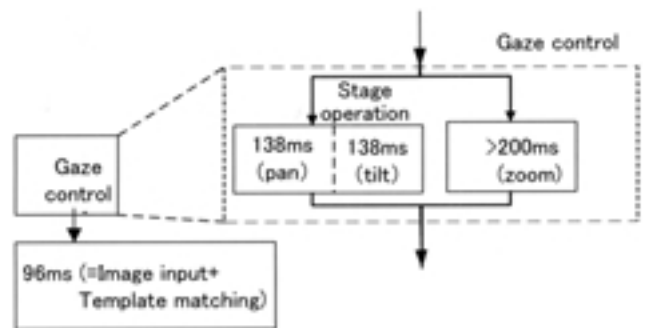


Fig. 17 Time for Sign-recognition.  
標識識別時間

Table 4 Measured time for zoom operation.  
ズーム動作の時間

| Step    | Objective Distance (mm) | Zoom value (%) | Time (ms) |
|---------|-------------------------|----------------|-----------|
| 0       | 5000                    | 73.3           |           |
| 1       | 4500                    | 72.3           | 260       |
| 2       | 4000                    | 71             | 261       |
| 3       | 3500                    | 67.5           | 310       |
| 4       | 3000                    | 62.3           | 351       |
| 5       | 2500                    | 55.7           | 361       |
| 6       | 2000                    | 46.6           | 380       |
| 7       | 1500                    | 33.4           | 470       |
| Average |                         | 342            |           |

(a) 500mm steps

| Step    | Objective Distance (mm) | Zoom value (%) | Time (ms) |
|---------|-------------------------|----------------|-----------|
| 0       | 5000                    | 73.3           |           |
| 1       | 4000                    | 71             | 281       |
| 2       | 3000                    | 62.3           | 381       |
| 3       | 2000                    | 46.6           | 521       |
| Average |                         | 394            |           |

(b) 1000mm steps

| Distance (mm) | time(ms) |
|---------------|----------|
| 1600 → 5000   | 852      |

(c) one step

い。しかし、一方で、カメラ入力 of 標識画像のサイズが適切でなくなる恐れがある。また、距離の変化に伴い、焦点ぼけ (defocus) になることもあり得る。このため、次は固定ズーム及び焦点ぼけの状況に対し、テンプレートマッチングによる標識識別の可能性を検証する必要がある。

#### 4.4.1 カメラの大画角における標識識別

カメラの大画角は低倍率のズームに対応する。この場合、カメラ入力画像の中に標識画像部分の割合は小さい。異なる標識画像の面積割合におけるテンプレートマッチング結果を比べるため、距離が 5000 mm の標識に対し、異なるズーム倍率で実験を行った。結果は Table 6 に示すように、次の点が注目される。第一行は最大標識サイズ (高さ = 100 ピクセル) の場合であり、最後の行は最小標識サイズ

(26 ピクセル) の場合である。標識サイズの高さが 100 ピクセルから 26 ピクセルへ (約 4:1) 変わると、テンプレートマッチングの結果の相関値は 87% から 66% まで減少したが、66% の相関値は Table 3 の結果より判断すると、まだ標識識別の能力程度と考えられる (Fig. 18)。一方、ロボットビジョンの検出範囲 1600 mm ~ 5000 mm での距離変化は約 1:3 であるので、ズーム値を (例えば 20% に) 固定した場合でも、使用は可能と考えられる。

#### 4.4.2 焦点ぼけにおける標識識別

焦点ぼけにおけるテンプレートマッチングの結果を調べるため、5000 mm 距離の標識に対し、焦点距

Table 5 Measured speed of pan/tilt operations.  
パン・チルト操作の速度測定値

| Operation | Parameter   | Maximum (deg./sec) | Minimum (deg./sec) |
|-----------|-------------|--------------------|--------------------|
| Pan       | Speed       | 147.14             | 2.93               |
|           | Start speed | 51.38              |                    |
| Tilt      | Speed       | 145.70             | 2.98               |
|           | Start speed | 57.14              |                    |

離を 5000 mm, 2000 mm, 及び 1000 mm にそれぞれ設定して実験を行った。結果は Table 7 に示すように、焦点距離がたとえ 1000 mm に設定されても、相関値が 66% を達成できる (Fig. 19)。これは前項と同様に標識識別が可能な値と見なせる。

## 5. まとめ

建設ロボットにおけるナビゲーション用として、2次元画像処理と3次元画像処理を併用するハイブリッド式ビジョンシステムを提案し、開発を試みた。本文では、このハイブリッド式ビジョンシステムの概要、3次元画像処理を用いた目標抽出部と2次元画像処理を用いた標識認識部の実現手法を説明し、設計仕様が達成できることを検証した。本研究の成果及び意義は次の通りである。

- 1) ロボット走行制御において、建設業など作業や経路などの変化が多い現場ニーズに対し、経路設定が柔軟・容易に行うことが可能なナビゲーションを提案し、そのためのシステムの実現方法を示した。
- 2) ロボットビジョンにおいて、立体視ベースの

Table 6 Correlation coefficients for different image magnifications.  
異なるズーム値に対する相関値

| Zoom value (%) | Height of sign image (pixels) | Coefficiency (%) |
|----------------|-------------------------------|------------------|
| 73             | 100                           | 87               |
| 60             | 66                            | 84               |
| 50             | 50                            | 87               |
| 40             | 39                            | 77               |
| 30             | 32                            | 66               |
| 20             | 26                            | 66               |

※distance 5000mm  
※zomm value=73% is used as the reference condition

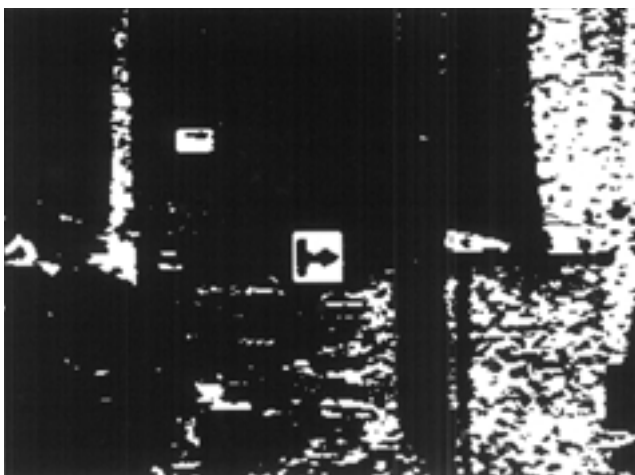


Fig. 18 Example for small sign image (zoom: 30%, distance: 5000mm).  
小さい標識に対する実験

Table 7 Correlation coefficients under defocusing condition.  
焦点ぼけの場合の相関値

| Focus value (%) | Focused distance (mm) | Correlation coefficient (%) |
|-----------------|-----------------------|-----------------------------|
| 12.7            | 5000                  | 84                          |
| 32              | 2000                  | 85                          |
| 48              | 1000                  | 66                          |

※distance 5000mm

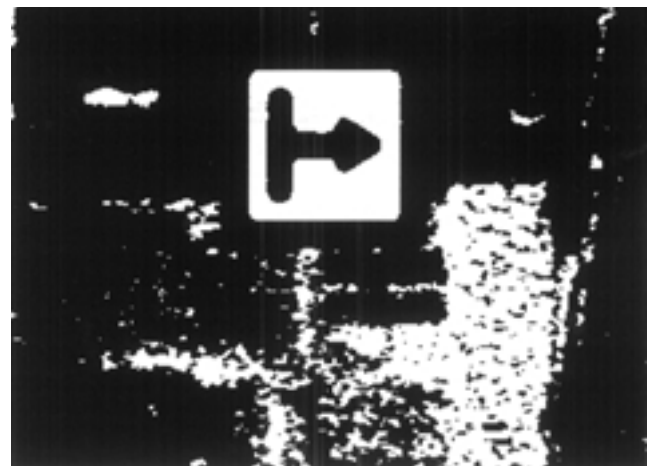


Fig. 19 Example of defocusing condition (focus value: 48%).  
焦点ぼけにおける実験

概略的 3次元情報に基づいて、カメラの注視制御により、広大空間においても、2次元画像処理方法が適用できる。

3) 立体視視差画像の応用方法として、領域分割に基づく 3D計測及び目標抽出方法を開発した。

今回は 30 cm x 30 cm の標識を検出対象としたが、他の対象目標に対しても、特徴抽出・学習を通して、簡単に適用できる。

今回は実際のロボット走行制御に適用して、走行機能が検証されるまでには至らなかったが、作業者と混在する環境下で、安全かつ機能的な誘導を実現可能であることが明らかになった。従来の赤外線センサーと超音波センサーによる人体（障害物）検知のみでは機能的な誘導のための情報が足りないので、人間と共存する自動機械の実用化が困難である。ビジョンシステムでは、このような環境下で、早期に危険状態を把握しつつ、停止に至らず、効率的な移動を維持できる有効な手段となり得る。

今後、カメラやPCハードウェアの改良によって、より実用的な技術が成熟することを期待したい。

### 参考文献

- 1) 呂健・池田博康・梅崎重夫・濱島京子・清水尚憲: 立体視レンジファインダーを用いたロボット障害物検出機能の実現, 日本機械学会第10回交通・物流部門大会講演論文集, p.405-406 (2001).
- 2) J. Lu, K. Yasuda, S. Nakagawa & H. Ikeda: A Navigation vision system for industrial mobile robots, Proceedings of the 33rd International Symposium on Robotics, October 7-11, 2002, Stockholm, Sweden.
- 3) 建設業労働惨害防止協会: 最近の建設機械の安全装置をめぐる現状と課題, 建設機械の安全性の確保に関する研修会 (1999).
- 4) 橋本秀一・日比均・植山剛・肥後徳仁・西山強志: 移動型デンスーロボット, 日本ロボット学会学術講演会予稿集, Vol.18, No.2, p953-954, 2000
- 5) K.M. Moulton, A. Cornell & E. Petriu: A Fuzzy Error Correction Control System, Proceedings of IEEE International Conference on Instrument Measurement Technology, Vol.16, No.2, p733-738, 1999.
- 6) O.A. Montasser & M. Abdellatif: Development of a computer vision measurement system to control a lane following mobile robot, Proceedings of IASTED International Conference on Model Identification Control, Vol.2001, No.2, p947-952, 2001.
- 7) X. An, W. Chang & X. Chen: Multi-layer Template Correlation Neural Network for Recognition of Lane Mark Based on Pipelined Image Processing Structure, Proceedings of IEEE international Conference on Robot Automation, Vol.1999, No.3, p2410-2415, 1999.
- 8) 竹内一雅・太田順・池田和生・新井民夫: 移動ロボットのナビゲーションのための人工ランドマーク配置設計, 日本ロボット学会学術講演会予稿集, Vol.17, No.1, p.247-24, 1999.
- 9) 中沢秀樹: 半導体市場から実装技術までの最新動向-高度画像センシングの有効性, エレクトロニクス実装技術, Vol15, No.12, p.64-68, 1999.
- 10) 荻野洋平・田中孝之・山藤和男: 新画像認識技術のコンセプトとシミュレーション, 日本機械学会ロボティクス・メカトロニクス講演会講演論文集, Vol.1997, No.B, p.1069-1070, 1997.
- 11) 奥富正敏: ステレオ視コンピュータービジョン, 新技術コミュニケーションズ, pp.123-137 (1998).
- 12) T. Williamson and C. Thorpe: A Trinocular Stereo System for Highway Obstacle Detection, Proceedings of IEEE International Conference on Robot & Automation; Vol.3, pp.2267- 2273, 1999.
- 13) K. Koster and M. Spann: MIR: An Approach to Robust Clustering- Application to Range Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22 (5) : p430-444, 2000.
- 14) O.R.P. Bellon, A.I. Direne and L. Silva: Edge Detection to Guide Range Image Segmentation by Clustering Techniques, Proceedings of IEEE International Conference on Image Processing, 1999, Vol.2, p725-729.
- 15) (株) ビュープラス: Triclopsの性能について, [www.viewplus.co.jp/products/triclops/perf\\_main.html](http://www.viewplus.co.jp/products/triclops/perf_main.html)
- 16) 高木幹雄・下田陽久: 画像解析ハンドブック, 東京大学出版会 (1991).
- 17) C. H. Chen, L. F. Pau & P. S. P. Wang: Handbook of Pattern Recognition & Computer Vision-2nd Edition, World Scientific Press, p932-933, 1999.
- 18) D.H. Ballard etc.: Computer Vision, Prentice-Hall Inc. (1982).

(平成 14 年 8 月 22 日 受理)