

## 大局視と局所視の統合による複合センシングシステムの構築

奥村 仁、八木 康史、谷内田 正彦  
大阪大学 基礎工学部 システム工学科

人が周囲を眺める時、一様に詳細な観察を行うのではなく、環境全体を大まかに把握する（大局視と呼ぶ）とともに、興味ある部分を注視する（局所視と呼ぶ）、本研究では、移動ロボットの視覚として、このように環境全体を大まかに把握しながら、そこから興味ある物体を発見し、その詳細な3次元形状を認識できる視覚システムを開発する。

大局視では周囲360度を一度に観察できる全方位視覚センサを用い、そのセンサ情報（全方位画像）を処理して、周囲の環境が3次元的にどのような構成になっているのかを認識し、周囲環境の大局モデルを作成する。また、その中から注視すべき物体（作業対象、障害物等）を発見する。一方、局所視では、大局視で発見された興味ある物体を注視しながらいくつかの位置から観測し、それらの画像を総合的に解釈して注視物体の詳細な3次元形状を認識する。以上のように大局視と局所視を組合せることにより、環境理解の目的に応じた詳細さの環境モデルが能率良く構築できる。本報告では、大局視と局所視との機能を持つ複合視覚センサMISS(Multiple Image Sensing System)を提案し、両視覚機能の情報を相互利用することで3次元情報を獲得する方法について報告する。

Multiple Image Sensing System for Mobile Robot  
Hitoshi OKUMURA, Yasushi YAGI and Masahiko YACHIDA  
Department of Systems Engineering,  
Faculty of Engineering Science, Osaka University  
1-1 Machikaneyama-cho, Toyonaka, Osaka 560, JAPAN  
Tel. +81-6-844-1151 Ext. 4621, Fax. +81-6-854-7664  
E-mail: y-yagi@sys.es.osaka-u.ac.jp

In this paper, we propose a new multiple image sensing sensor (MISS), which combines with an omnidirectional image sensor COPIS (COncic Projection Image Sensor) and binocular vision, for navigating the robot and understanding interesting objects in an environment by integrating both sensory data. Since COPIS observes a 360 degree view around the robot, the robot can always estimate its own location and motion precisely. The location of unknown objects can also be estimated. COPIS can observe a global and precise information of features (vertical edges) in real-time, however, it is difficult to understand their details of shapes. On the other hand, The mobile robot has binocular vision and obtains a sequence of stereo images from the environment. We extend the principle of trinocular vision to establish correspondences between a sequence of binocular images. Although a view field of the binocular vision is limited by a visual angle of lens, the binocular vision is useful for understand spatial configuration of the environment. Therefore we integrate both merits and propose an efficient sensing system. The system has been evaluated on the prototype sensor in actual environment.

## 1. まえがき

近年、視覚付き移動ロボットに関する研究が広く行われている[1]-[4]。従来の視覚を用いた移動ロボットの研究では、一般的なカメラを採用したものが多く観測視野が前方環境のみに限られていた。しかし、一般環境では、側方、後方から接近する移動物体など、前方情報だけでは不十分な場合も起こり、ロボットが安全に走行するためには移動しながら全方位の情報を獲得することが望ましい。そこで、我々は移動ロボットの視覚誘導に対して円錐ミラーを用いた全方位視覚センサCOPIS(Conic Projection Image Sensor)を提案してきた[5][6]。視野領域はロボットを誘導するうえで最も重要な側方が中心であり、また周囲360度を一度に観測でき、実時間処理にも対応できる利点を持つ。また、全方位視覚センサCOPISにより撮像された画像面上では、環境内の垂直エッジが放射状に投影される。この特徴を利用して、ロボットの移動により撮像された画像から垂直エッジの抽出を行えば、物体の発見、物体の運動推定を容易に行うことができ、大局的な環境マップを作成できる。全方位視覚センサCOPISは視覚誘導に適したセンサであると言える。しかし、広い環境は限定された解像度(512×480)の画像に投影されるので、画像から物体の詳細な形を解析することは難しい。一方、目印や建造物、静止して障害物、移動物体などに対しては、従来のステレオ視で物体の構造認識ができる。しかし、環境中のすべての対象を解析するためには、視野がロボットの前方のみなので時間を要するという問題がある。移動ロボットにおいて必要とされる視覚機能は、(1)環境中でのロボットの自己位置同定や(2)未知物体の発見、及びその位置、動きの推定、さらに(3)興味物体に対する注視行動及びその物体の構造理解などさまざまである。すなわち、視覚機能によって、必要な画像の解像度や観測範囲、処理速度などは異なり、移動ロボットの視覚システムには複数の視覚機能の必要性が生じる。本研究では、性質の異なるセンサ系を統合一体化し、効率的に処理できる複合視覚センサシステムMISS(Multiple Image Sensing System)を提案する。

本MISSでは、大局視と局所視の2つの光学系が一体化されている。大局視では周囲360度を一度に観察できる全方位視覚センサを用い、そのセンサ情報(全方位画像)を処理して、周囲の環境が3次元的にどのような構成になっているのかを認識し、周囲環境の大局モデルを作成する。また、そ

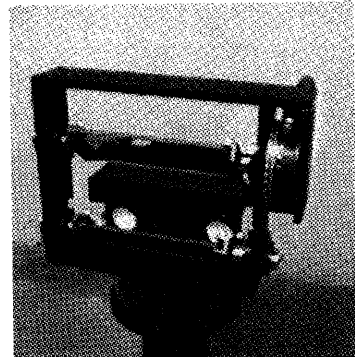
の中から注視すべき物体(作業対象、障害物、道しるべ、近づいて来る物体等)を発見する。一方、局所視では、大局視で発見された興味ある物体を注視しながらいくつかの位置から観測し、それらの画像を総合的に解釈して注視物体の詳細な3次元形状を認識する。以上のように大局視と局所視を組合せることにより、環境理解の目的に応じた詳細さの環境モデルが能率良く構築できる。

本報告では、大局視と局所視との機能を持つ複合化視覚センサMISS(Multiple Image Sensing System)を提案、両視覚機能の情報を相互利用することで3次元情報を獲得する方法について報告する。また、本システムでは双方のセンサ情報を融合し効率的にロボットの視覚誘導や環境中での興味物体の認識を行う手法も述べる。

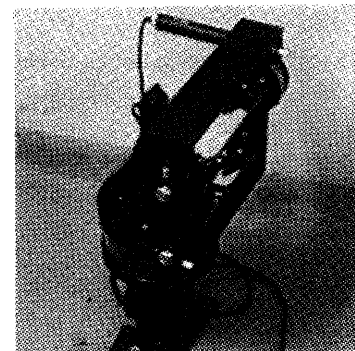
## 2. 複合視覚センサシステムMISS

図1(a),(b)に本報告で提案する複合視覚センサMISSのプロトタイプを示す。

図2に示すように、本システムは、大局視として全方位視覚(図2(a))、局所視としてステレオ視(図2(b)(c))を統合一体化し、両視覚を同一面に結像する。従って、実時間で大局視、局所視の情報を獲得できる。以下、各々のセンサについて説明する。



(a)



(b)

Fig.8 Prototype of Multiple Image Sensing System MISS

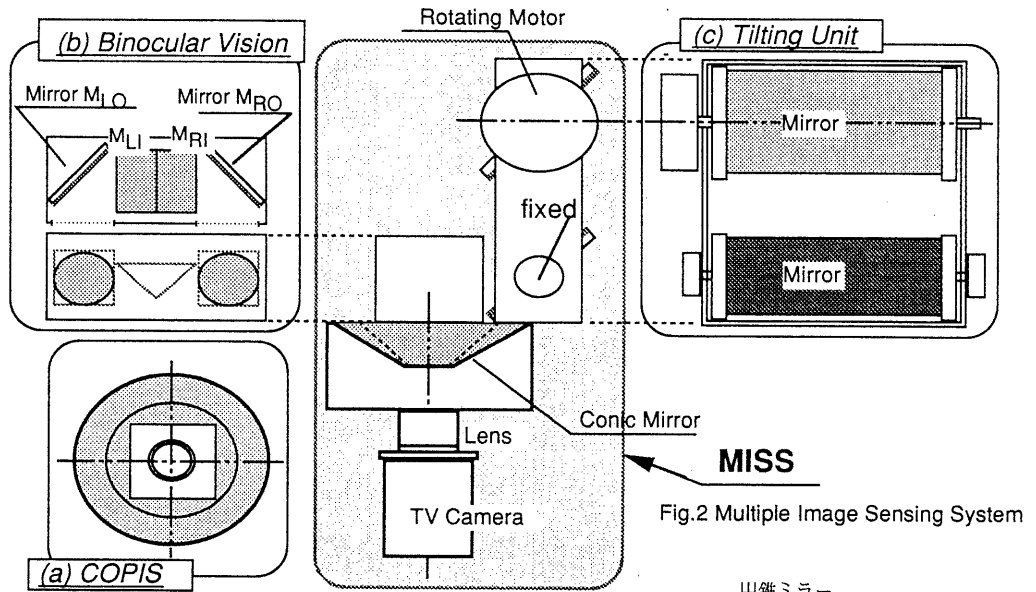


Fig.2 Multiple Image Sensing System

## 2.1 大局視（全方位視覚センサCOPIS）

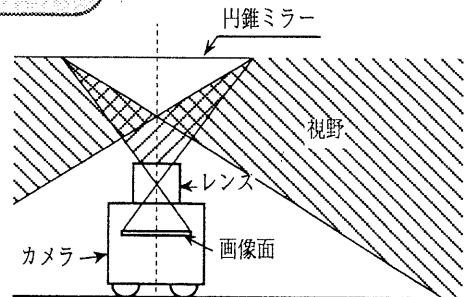
COPISは円錐ミラーとTVカメラを組み合わせた視覚センサで環境の全方位画像を実時間（TVカメラのframe時間）で撮像できる。図3(a)のように鉛直下向きに設置された円錐ミラーにより、水平面下方鉛直軸周り360度全方位が視野範囲となる。

### 2.1.1 円錐投影

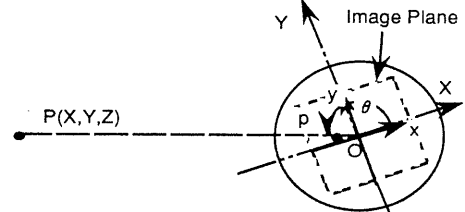
COPISでは、シーンは円錐ミラーとレンズを通して画像面に投影される。この投影を円錐投影と呼ぶ。今、図3(b),(c)で示めすようにカメラ光軸をZ軸（鉛直方向）とするカメラに固定したO-XYZ 3次元座標系を考える。画像面 $xy$ は、XY平面に平行でカメラ焦点距離 $f$ だけ離れたZ軸上の点を通る面とする。このとき、任意の点 $P(X,Y,Z)$ に対する画像面上の写像点を $p(x,y)$ とすると次の関係がある。

$$\tan \theta = \frac{Y}{X} = \frac{y}{x} \quad (1)$$

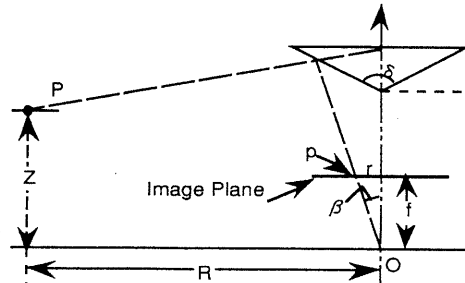
すなわち、 $Y/X$ で定まる点Pの方位角は、 $y/x$ で定まる画像面内の像点の回転角を算出することで得られる。部屋や廊下などの人工の環境では垂直エッジを持つ物体（例えば、ドア、机）が数多くあり垂直エッジは有用な特徴量となりうる。図4で示すように、一本の垂直エッジ上の点は高さに関わりなく同一の方向性を示し、垂直エッジは画像中心から放射状に延びた線分として現れる。すなわち、対象物体の方位角 $\theta$ がその物体の画像上の回転角として直接現われる。このことより、COPISでは連続した画像中での放射状のエッジを見つけることで容易に垂直エッジを発見できる。



(a) Visual field of COPIS



(b) Invariability of azimuth angle in COPIS



(c) Linear relation of tilt angle in COPIS

Fig.3 COPIS Imaging System

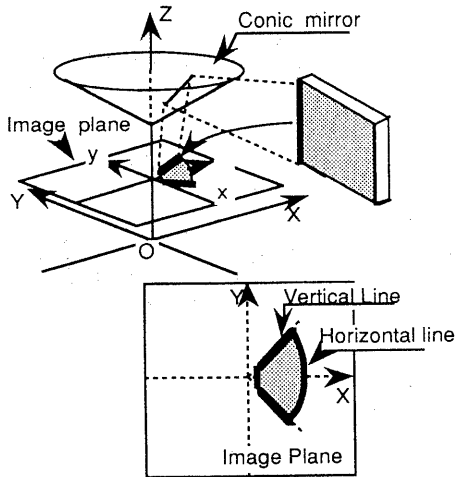


Fig.4 Conic Projection of 3-D Lines

全方位視覚センサCOPISは実時間で周囲360度の視野領域を撮像できる利点を持つが、円錐ミラーの幾何学的な性質から画像歪が発生する。また、角度分解能は画像の周辺部でおよそ0.25度で、画像中心から60ピクセル半径内の領域ではおよそ1度である。特に、全方位画像の中心領域は解像度が低く垂直エッジの方位角計測に適していない。そこで、本システムでは図2(a)に示すように、画像歪の大きかった円錐ミラーの中心域をくり抜きドーナツ型の円錐台とし、円錐台内部の穴を図2(b),(c)で示すように局所視の視野領域として有効に活用する。

## 2.2 局所視 (ステレオ視)

局所視はステレオ視部と光路変更部とからなり、同一の画像面上に一对のステレオ画像が得られる。ステレオ視部は図2(b)に示すように4つのミラーからなる。内側の2つのミラーMRI,MLIはそれぞれ垂直に取り付けてあり、外側のミラーMRO,MLOと組み合わせて、円形に空けられた2つの窓からの入射光を画像面上に取り込む仕組みになっている。また、ステレオ視の視線方向は図2(c)で示すように光路変更部の下の固定ミラーで上のミラーへ向けられ、上のミラーが回転することで、上下角を変えることができ、ステレオ視の視線方向を上下に振ることができる。

## 3. 複合センシングアルゴリズム

### 3.1 概要

ロボットは最初基準地点に置き、部屋やビルの廊下を与えたパスに沿って走行させる。ロボットは基準地点 (スタート地点) とロボット自身の移

動量を与えられるが、ロボットが移動する際の揺れによって、計測誤差が生じる。それゆえ、ロボットが移動する環境は既知とし、環境マップと入力画像の両方から得られる方位角を使ってロボットの位置と移動量を補正する。環境マップと観測物体の対応付け終了後、環境マップで見つからなかった垂直エッジは未知物体として認識し、その位置を算出する。未知物体の認識のためにステレオ視では未知の垂直エッジの近傍を注視する。ロボットが回転したときも全方位視覚センサCOPISにより全方位画像が得られるので未知の垂直エッジの位置は算出できる。求められた未知の垂直エッジの位置情報はステレオ視での注視動作の制御に利用する。時系列ステレオ画像間での対応付け問題に対しては全方位視覚センサCOPISから得られるセンサ情報を利用する。すなわち、全方位視により計測した垂直エッジの位置情報をステレオ画像面上に投影変換することで、垂直エッジに関する対応関係を容易に見つけることができる。また、ステレオ画像間のそれ以外の対応付けには、3台のカメラにより生じる幾何学的な拘束を利用した三眼視の原理を拡張して用いる。以下、両視覚系の処理について述べる。

### 3.2 大局視による環境モデルの構築

大局視では全方位視覚センサCOPISのセンサ情報 (全方位画像) を処理して、周囲の環境が3次元的にどのような構成になっているのかを認識し、周囲環境の大局モデルを作成する。また、その中から注視すべき物体を発見する。

#### 3.2.1 ロボットの自己位置同定

自己位置を求めるには、観測された垂直エッジをあらかじめ与えられている環境マップに対応付ける必要がある。対応付けは次のように行う。

ロボットの移動量を $(\Delta x, \Delta y)$ とし、前フレームで同定された位置を $(X_{robot}, Y_{robot})$ とすると (開始時はスタート位置とする)、環境マップに記されている垂直エッジ  $(X, Y)$  は次式の方法 $\theta$ に観測されるはずである。

$$\tan \theta = \frac{Y - (Y_{robot} + \Delta Y)}{X - (X_{robot} + \Delta X)} \quad (2)$$

この関係を利用し、環境マップ中の物体の方位角の予測モデルを作成する。観測された垂直エッジがあらかじめ与えられた環境マップに対応付いたなら図5のように対応する環境マップのエッジ位置と観測方向からロボットの自己位置が算出でき

る。すなわち、3つ以上のエッジが観測できれば、各エッジ位置から観測方位の方向に引いた線分の交点として求まる。

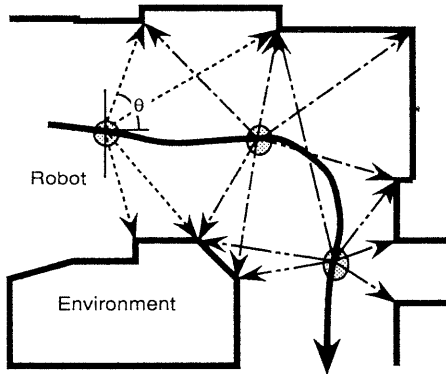


Fig.5 Estimation of Location of Robot

しかし、測定された方位角はロボットの揺れのために誤差を含んでいる。そこで、最小二乗法により、より正確な位置を算出している。なお、COPISは全方位画像を撮像するのでロボットが回転してもロボットの自己位置は同定できる。詳細に関しては参考文献[10]を参照されたい。なお、自己位置はステレオ移動視におけるカメラ位置更新に利用される。

### 3.2.2 未知物体位置の算出

観測された物体と環境マップの対応付けの際に、対応付かなかった物体は未知物体として新たに環境マップに登録される。未知物体の位置は、ロボット移動に伴う異なった観測点での物体方位角を計測することで、三角測量の原理から求められる。

未知物体の座標位置を $P_0(X_0, Y_0)$ とする。今、ロボットの位置が $(X(t), Y(t))$ で表されたとすると、時刻 $t$ における点 $P$ の相対位置は $(-X(t)+X_0, -Y(t)+Y_0)$ となる。従って、時刻 $t$ における未知物体の方位角 $\theta(t)$ は次式となる。

$$\tan \theta(t) = \frac{-Y(t) + Y_0}{-X(t) + X_0} \quad (3)$$

ここで、ある時刻 $t_1, t_2$ での方位角が観測できれば、両時刻での式(3)の関係から得られる連立方程式を解くことで未知物体の座標位置 $(X_0, Y_0)$ は計算できる。

なお、実際には、垂直エッジの方位角に計測誤差があるため、2点から計算するだけでは精度が悪い。そこで、計測精度を上げるため、図6に示すように、ロボットの移動とともに撮像した連続画像から、方位角の変化を観測し、最小二乗法により算出する。詳細に関しては参考文献[10]を参照さ

れたい。算出された垂直エッジ位置情報はステレオ移動視の対応付け問題に利用する。

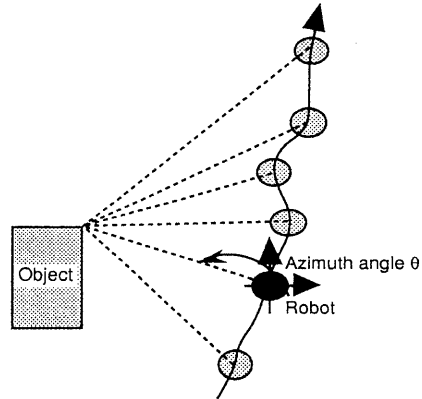


Fig.6 Estimation of Location of an Unknown Object

### 3.3 局所視における物体構造の獲得

本手法では、ステレオ移動視で獲得されたステレオ画像を3枚組み合わせることで、3つのカメラによる幾何学的な拘束を利用する三眼視の原理を拡張しステレオ画像間の対応付けを行う。

#### 3.3.1 三眼視の原理

三眼視における線分の対応付けの原理を図7に示す。画面1の線分 $p_1q_1$ の対応をとるために、まずエピポーラライン $ep, eq$ を引き、画面2における対応候補を見つける。次に、画面2でエピポーラライン $ep, eq$ の間に存在する、線分 $p_1q_1$ と同じような特徴を持つ線分を候補として選択する。それぞれの候補に対して、三角測量で算出した三次元位置を画面3に投影したとき、候補が正しいものならば画面3での投影された場所に線分が存在する。正しくなければ存在しない。こうして、候補が正しいか正しくないかは画面3で検証できる。

詳細に関しては参考文献[7],[8],[9]を参照されたい

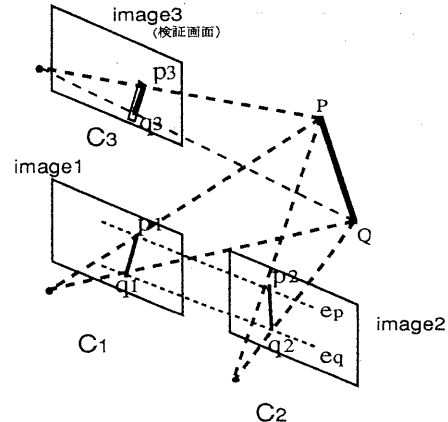


Fig.7 Principle of Segment-based Trinocular Stereo

### 3.3.2 ステレオ移動視

三眼視をステレオ移動視へ応用し、画像間での対応付けを行う。ステレオ視を搭載する移動ロボットが移動を行うと移動前後で各2枚づつ計4枚の画像が得られる。その4枚の画像に三眼視の原理を適用、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組の三眼視を行う。画像の組み合わせを図8に示す。それぞれの組み合わせの3眼視の結果より、画面1中の線分に関する対応付けが唯一に決定した3眼の組み合わせを各線分ごとに決定する。対応付けの単位として本手法では線分を用いているが、線分は画像中のエッジ点列から抽出されるのでエッジ点を対応付けの単位とするよりも、明らかに対応候補を減少させることができる。

詳細に関しては参考文献[11],[12]を参照されたい。

なお、三眼視の原理を移動視に用いる場合、ロボットの移動前後におけるカメラ位置を正確に知る必要がある。そこで本システムでは全方位視覚センサCOPISから大局的な位置情報を得ることができるため、この情報を利用することでカメラ位置が得られる。これら得られた注視物体の3次元線分情報から物体の面-稜線-頂点といった階層構造を獲得する[13]。

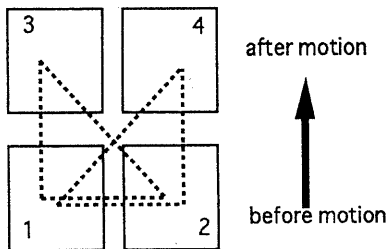


Fig.8 Two Sets of Three images

### 3.3.3 大局情報の利用による対応付けの効率化

COPIS情報を利用することで特定のエッジに関しては効率的に信頼よく正しい候補を見つけることができる。COPISでは未知物体の垂直エッジの正確な位置を算出できる。従って、位置情報をステレオ画像に投影することによりステレオ画像上での垂直エッジ位置が算出できる。

実際には、垂直エッジ位置には観測誤差が含まれるので、一定のバンド幅( $\Delta x$ )を持たせた領域を考える。図9(a),(b)にその領域を投影した様子を示す。ステレオ画像の左右画面に対して投影されたバンド領域内に垂直エッジが存在すれば垂直エッジの対応付けが行える。

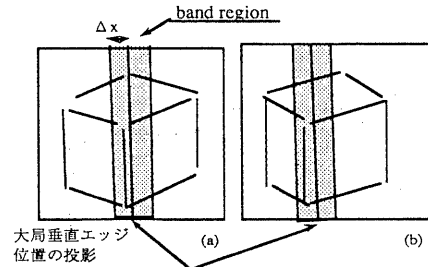


Fig.9 Correspondence of Vertical Edges

垂直エッジの対応付けの後、一方のステレオ画像内で垂直エッジ近傍領域内に他のエッジの端点が存在するなら、そのエッジは垂直エッジに接続しているものとする。そして、もう一方のステレオ画像では、対応する垂直エッジを手掛かりに探索範囲を限定し、エピソード拘束を用いることで対応点を決定する。各垂直エッジに対し同様の処理を行い対応関係を得る。さらに、垂直エッジに接続していないエッジに対しては、そのエッジをはさむ連続した垂直エッジを抽出し、この間を対応エッジ探索領域として限定する。

## 6.2 実験結果と検討

MISSを使って、室内環境でいくつかの実験を行った。画像は512×480画素(画素レベルは256)に量子化される。図10(a)は図10(b)で示す環境に対する入力画像の例である。画像中の物体は家のような箱である。それぞれの画像はロボットが2.5cm移動するごとに撮像した。図11(a),(b),(c)はそれぞれ1,7,29番目のフレームの全方位画像である。ロボットは70cm直進する。垂直エッジの方位角の軌跡を観測しながら、図12に示すように、COPISは自己位置を同定し、未知の垂直エッジ位置を算出する。図中、環境マップ上の既知垂直エッジは□印、算出したロボットの移動軌跡は細い黒線で示す。

ロボットの平均位置計測誤差はおおよそ10mmで未知の垂直エッジの平均位置計測誤差はおおよそ15mmであった。図13(a),(b),(c),(d)はそれぞれ7,8,9,10番目のフレームである。図14(a),(b)は7,8番目の直線抽出結果である。図15に示すように、ステレオ視で垂直エッジの探索領域を設定した後、図16で示すように垂直エッジの対応を見つける。黒の太線が対応付けられた垂直エッジである。対応付けは接続関係から帰納的に行われる。例えば、候補の両端点が接続されているものは図17のように対応付けされる。白丸は対応する頂点である。図18では、エッジの片側が垂直エッジに接続しているときの

対応付けである。黒丸は接続してない頂点である。図19は最後の検証を示す。図18で候補が垂直エッジに接続していない場合でも真の対応が求まる。ステレオ視でのエッジ位置の計測誤差はおよそ13mmであった。しかし、獲得されたロボット位置と未知物体の正確さは、ロボットの誘導と物体認識には十分だと考える。

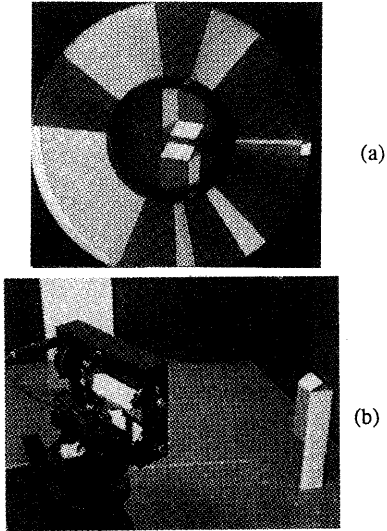


Fig.10 An Example of Input Image & Experimental Environment

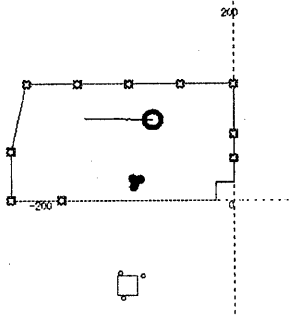


Fig.12 Results of Location Estimation

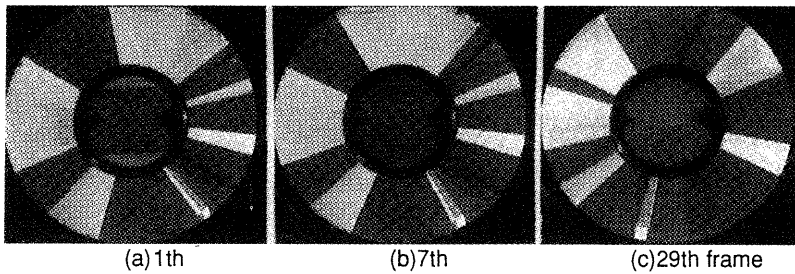


Fig.11 Example of Omnidirectional Images

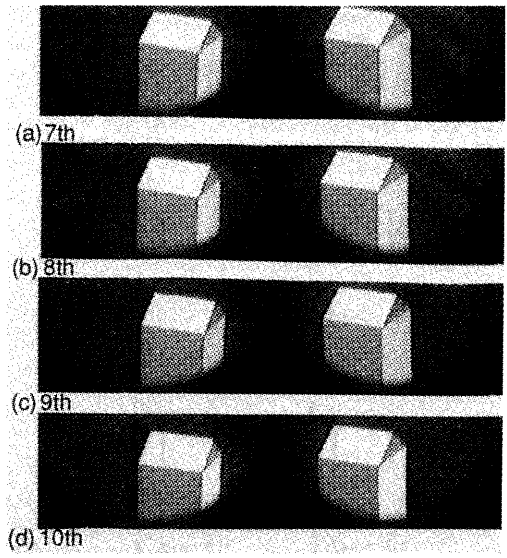


Fig.13 Examples of Binocular Motion Images

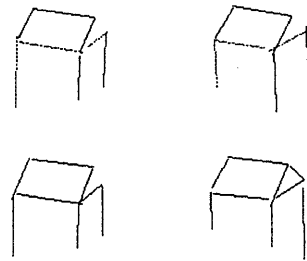


Fig.14 Edge Picture at the 7th & 8th frame

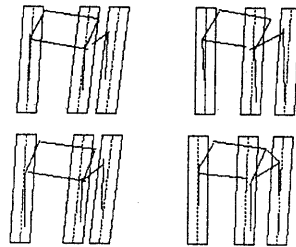


Fig.15 Setting Search Region at the 7th & 8th frame

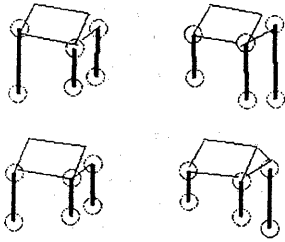


Fig.16 Correspondence of Vertical Edges

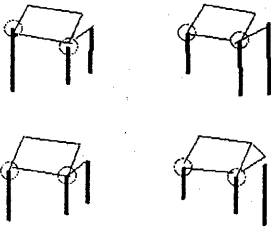


Fig.17 Verification of Candidates  
(Both vertices are connected)

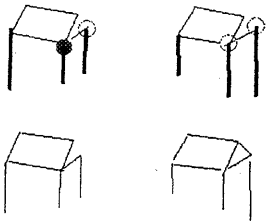


Fig.18 Verification of Candidates  
(One side of an edge is connected)

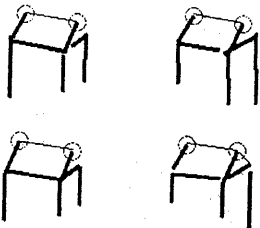


Fig.19 Results after Final Verification

## 7. 結論

本報告では、全方位視覚センサCOPISとステレオ視を組み合わせた複合視覚センサMISS(Multiple Image Sensing System)を提案し、両センサ情報を統合する手法について述べた。並びに今回は、大局視情報の局所視への利用について報告した。現在、局所視情報の大局視への利用について検討している。また、複合視覚センサMISSについてはレーザーレンジファインダとの統合を行う予定である。

## References

- [1] A.M.Waxman, J.J.LeMoigne and B.Scinvasan, A visual navigation system for autonomous land vehicles, *IEEE J.Robotics & Automation*, RA-3, 2, pp.124-141 (1987)
- [2] M.Turk, K.D.Morgenthaler, K.D.Gremban and M.Marra, VITS-A vision system for autonomous landvehicle navigation, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.PAMI-10, no.3 pp.342-360 (1988)
- [3] C. Thorpe, M.H. Hebert, T. Kanade and S.A. Shafer, Vision and navigation for the Carnegie-Mellon Navilab, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.PAMI-10, No.3, pp.362-373 (1988)
- [4] D. E. Dichmanns and V. Greafe, Dynamic monocular machine vision, *Machine Vision & Application*, 1, pp.223-240 (1988)
- [5] Y. Yagi and S. Kawato, Panorama scene analysis with conic projection, *Proc. IEEE Int. Workshop Intelligent Robots & Systems*, pp.181-187 (1990)
- [6] Y. Yagi, S. Kawato and S. Tsuji, Collision avoidance using omnidirectional image sensor (COPIS), *Proc. IEEE Int. Conf. Robotics & Automation*, pp.910-915 (1991)
- [7] M. Yachida, "3-D data acquisition by multiple views", 3rd ISRR, 1985.
- [8] N. Ayache and F. Lustman, "Fast and reliable passive trinocular stereo vision", *Proc. ICCV*, pp.422-427, 1987.
- [9] N. Ayache and F. Lustman "Trinocular stereovision : recent results" *Proc.10th.IJCAI*, pp. 826-828, 1987.
- [10] Y. Yagi, Y. Nishizawa and M. Yachida, Map Based Navigation of the Mobile Robot Using Omnidirectional Image Sensor COPIS, *Proc. IEEE Int. Conf. Robotics and Automation*, Vol.1, pp.47-52, Nice, (May 1992).
- [11] M. Kimachi and M. Yachida, "Getting 3-D information by edge-based trinocular stereo" (in Japanese), *信学会, PRU研究会*, 88, 324, pp.13-20, 1988.
- [12] M. Yachida and A. Tsudo, Understanding Spatial Configuration of Robot's Environment from Stereo Motion Images, *Proc.IEEE IROS'91*, pp.882-887, 1991
- [13] A. Utsumi, Y. Yagi and M. Yachida, Estimating Surface and Spatial Structure from Wire-frame Model Using Geometrical & Heuristical Relation, *Proc. IAPR Machine Vision and Appl.* pp.25-28 (1992)