

大局視と局所視の統合による移動ロボットのための 複合視覚センサ MISS の提案

奥村 仁[†] 八木 康 史[†] 谷内田 正彦[†]

人は道路などを歩く時、環境中のすべてのものに詳細な注意を払うのではなく、環境に対する自己位置推定や障害物の発見など歩行のために必要な情報を環境全体から大まかに把握する（大局視と呼ぶ）。一方、興味ある物体（作業対象、障害物、道しるべ、近づいてくる物体等）に対しては、人は興味部分を注視し、その部分についての詳細な情報を収集する（局所視と呼ぶ）。人と同様に自律移動ロボットにおいても、目的地までの安全な誘導および目的地での作業など作業内容により、必要とされる観測情報の性質は異なる。本報告では、ロボットにとって必要とされる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、試作したプロトタイプ構成について報告する。さらに注視対象の発見からその物体の3次元情報獲得までの一連の処理を両視覚機能の情報を統合することで行う一方法について報告する。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、得られる全方位画像とあらかじめ与えられた周囲環境のモデルとの照合を行い、注視すべき物体（作業対象、障害物等）を発見する。一方、局所視ではその物体を注視しながらいくつかの位置から観測し、注視物体の3次元形状を認識する。

Multiple Image Sensing System MISS for Mobile Robot

HITOSHI OKUMURA,[†] YASUSHI YAGI[†] and MASAHICO YACHIDA[†]

In this paper, we propose a new multiple visual sensing sensor (MISS), which combines with an omnidirectional image sensor COPIS (COnic 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. はじめに

人は道路などを歩く時、環境中のすべてのものに詳細な注意を払うのではなく、環境に対する自己位置推定や障害物の発見など歩行のために必要な情報を環境全体から大まかに把握する（大局視と呼ぶ）。一方、興味ある物体（作業対象、障害物、道しるべ、近づいてくる物体等）に対しては、人は興味部分を注視し、その部分についての詳細な情報を収集する（局所視と呼ぶ）。また視覚において、主に網膜の中心付近（視野6

度程度）での視力は高く、詳細な情報を獲得する上で適している。それに対し視覚の周辺部は、視力および色を検出する能力は低いとされるが、移動物体の知覚などの情報は、網膜の周辺領域においても低下しない。すなわち、人は目的に応じて大局視・局所視と言った異なる視覚機能をうまく使い分け、効率よく環境理解を行っている。

一般に昆虫などの生物においても、その生物が生存する環境に合わせ、捕食行動などに適したセンサとその情報処理系が備わっている。たとえばハエトリグモは、ほぼ周囲360度を観察できる側方の四つの目と注目物体をじっくり観測するための前方の二つの目から成る複数の目を持つ。そして、側方の目の視野に動物

[†] 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

体が飛び込んできて1箇所にとどまると、まず側方の目でその位置を求め、次にその方向に体の正面を向ける。こうして正面にきた注目物体を前方の目で拡大しそれが何であるかをじっくり観察する。その結果、それが餌、仲間、敵のいずれであるかを知り環境に応じた行動を選択する。すなわち大局視である側方の目は、注視対象の発見とその位置推定に用いられ、局所視である前方の目は、注視対象の認識に利用され、これら異なった視覚機能を持つことで環境に適応し生存している¹⁾。

ところで自律移動ロボットの作業内容を考えると、目的地までの安全な誘導および目的地での作業などがあり、作業内容により必要とされる視覚機能は異なる。誘導においては、部分的に詳細な情報は不要で、環境中でのロボットの位置同定や未知障害物の発見とその位置、動きの推定などの大局視が必要となる。それに対し、目的地での作業では、観測範囲は局所的であっても詳細な対象物体の情報を獲得する局所視が必要である。このように作業内容により、必要な画像の解像度や観測範囲・処理速度などが異なるため、移動ロボットの視覚系では、性質の異なる複数の視覚機能の必要性が生じる。したがって自律移動ロボットにおいても人、生物の場合と同様に作業目的に合わせた視覚系を用意し、それらの情報を統合した処理（センサフュージョン）が必要と言える。

複数センサの統合としては、従来よりセンサフュージョンの分野において数多く研究されてきた²⁾。マニピュレーションの目的に対しては、視覚に触覚、力覚などの異種センサからの情報を統合した研究が行われている。また視覚誘導の目的に対しては、視覚情報にレンジファインダから得られた距離情報を統合する研究が数多くある^{3),4)}。米国における NAVLAB の研究では、カラーカメラと ERIM レーザレンジファインダ⁵⁾を主センサとしてもち、道路の識別、障害物の検出を行っている。道路の識別は、数種類の道路領域や非道路領域を示す色に関する知識を用いて、ベイズの定理により視覚センサから得られた色情報から道路領域を検出し、検出された道路領域を Hough 変換することで、道路の向きと道路境界の推定を行っている^{6),7)}。さらに障害物の検出は、ERIM の距離センサを用い、距離画像から各画素での面の法線ベクトルを求め、この法線ベクトルから障害物の有無を検出している⁸⁾。すなわち、目的に応じたセンサを用意することでロボットの誘導を行っている。

これらの研究は、異なるセンサ情報の統合に関する研究であり、大局視、局所視といった異種視覚センサ

間の統合という考え方の研究は筆者の知る限りない。また移動ロボットに搭載するという条件のもとでは、可搬重量および処理コストに制限があり、いかにコンパクトに目的に応じた視覚機能を盛り込むかが重要な課題と言える。しかし一般に従来研究では、個別にセンサを用意していたため、システム全体が大きく（情報処理を含めて）なる傾向があった。

本研究では、人や生物などと同様に大局視と局所視と言う性質の異なるセンサ系を統合一体化し、ロボットにとって効率的に複数の作業目的に対応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、大局視として円錐ミラーを用いた全方位視覚センサ COPIS (Conic Projection Image Sensor)^{9),10)}の光学系、局所視としてステレオ視の光学系を利用している。

大局視として用いる COPIS の視野領域はロボットを誘導する上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 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、大局的な環境マップの作成や未知物体の発見、移動物体の運動推定などを容易に行うことができる。すなわち全方位視覚センサ COPIS は視覚誘導に適したセンサ（大局視）であると言える。しかし、COPIS では、周囲 360 度の環境を限定された解像度 (512×480) の画像に投影するので、入力画像上では光学的歪み（特に画像中心部では光学歪みが大きくなる）があり、詳細に物体形状を解析する目的には適していない。

一方、ステレオ視では両眼の共有視野領域が狭く、環境全体の情報を獲得しようとした場合、カメラを回転する必要がある。したがって、ステレオ視は大局的情報を実時間で獲得する問題には適していない。しかし、一般に画像歪みは少なく均一な解像度を持ち、注視対象の詳細な構造を獲得する目的に対しては適している。

そこで本 MISS では、全方位視覚センサ COPIS およびステレオ視といった性質の異なる視覚系を一つの画像面（カメラ）上で実現することで、実時間で大局視、局所視の情報を獲得できかつ両視覚系の持つ長所を共に活かせる複合視覚センサを構築している。以下、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) ならびに両視覚機能の情報を相互利用することで3次元情報を効率的に獲得する方法について報告する。

2. 複合視覚センサシステム MISS

図1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))を持ち、両視覚を統合一体化することで同一画像面に結像させている。実際には、画像面の周辺部を大局視、中央部分を局所視(ステレオ視)の視野領域として利用している。以下、各々のセンサについて説明する。

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)で示すようにカメラのレンズ中心を原点 O としカメラ光軸を 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。図3(b)は、上面図、図3(c)は、 Z 軸を通る鉛直面での断面図である。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。

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

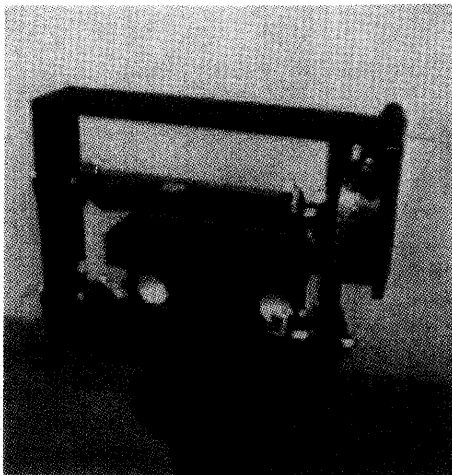


図1 複合視覚センサ MISS

Fig. 1 Prototype of multiple image sensing system MISS.

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。

一般に部屋や廊下などの人工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。COPISでは、垂直エッジが画像中心から放射状に延びた線分として現れる特徴を持つ(図4)。すなわち、垂直エッジの方位角 θ が、このエッジに対応する放射状線分の回転角を計算することで推定できる。このことより、COPISでは連続した画像中で放射状エッジを見つけることで容易に垂直エッジを発見、その方位を計測することができる。

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。方位の角度分解能は、画素単位で見た場合、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およそ 1 度である。特に、全方位画像の中心領域は方位の角度分解能が低くなるため、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)

前述のように局所視の視野領域を設けた場合、局所視の視野領域は制限される。しかし各画素の角度分解能は、通常のカメラとかわらない。また一般にレンズ系の光学歪みは、光軸中心から放射状に現れ、周辺部で最も大きく、画像中心で光学歪みが最も少なくなる。したがって、画像の中心領域は、ステレオ視として利用するのに最も適した領域と言える。

局所視はステレオ視部(図2(b))と光路変更部(図2(c))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(d)に示すように四つのミラーからなる。内側の二つのミラー面 M_{RI} , M_{LI} は垂直関係にあり、外側のミラー M_{RO} , M_{LO} と組み合わせて、円形に空けられた二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に示すように光路変更部の下で固定ミラーで上の可動ミラーへ向けられ、上のミラーが回転することで、ステレオ視の視線方向を上下に振ることができる。

2.3 複合視覚センサ MISS の特性

以上のように二つの光学系を一体化することで全方位画像と左右ステレオ画像は共に一つの画像面上に構築できる。図5(a)に示すような約 2.5×1.6 m の広さ

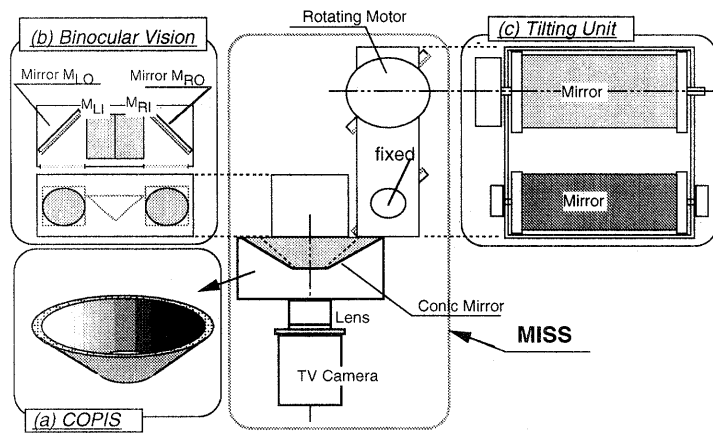


図 2 (a)~(c) 複合視覚センサシステム
Fig. 2 (a)~(c) Multiple image sensing system.

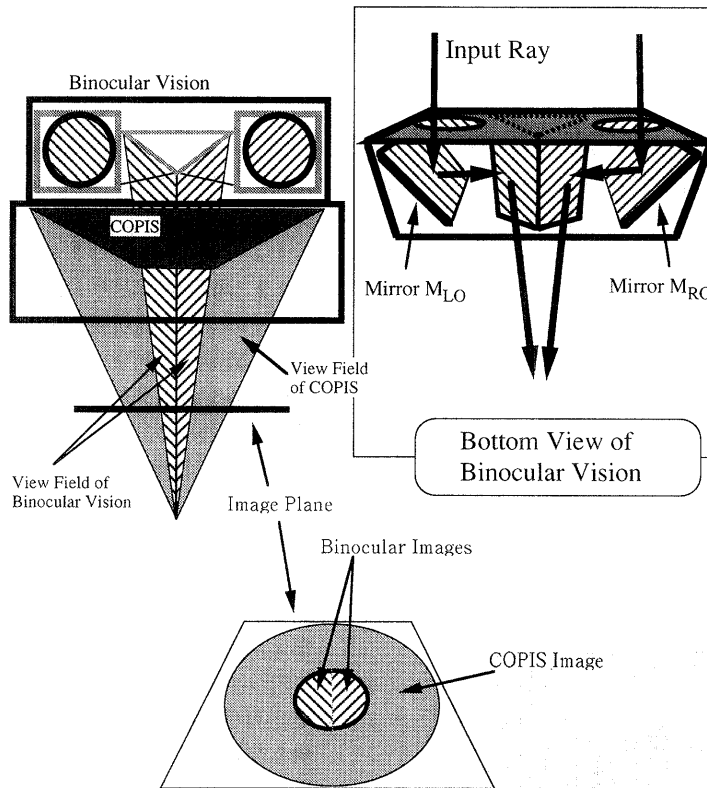
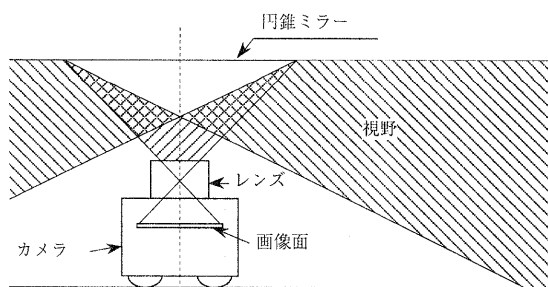


図 2 (d) ステレオ視部の構成
Fig. 2 (d) Binocular vision.

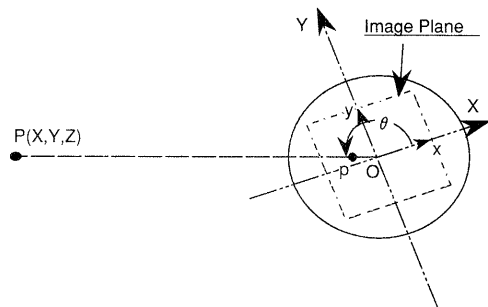
の室内環境における入力画像の一例を図 5 (b) に示す。

図 5 (a) では、センサは垂直エッジで囲まれた環境にあり、そして両眼視で環境内の物体を注視している。ステレオ視での注視物体は全方位画像中では右方面に映っている。全方位視の上下方向の視野範囲は、仰角がほぼ水平方向、俯角が約 13 度である。両眼視での

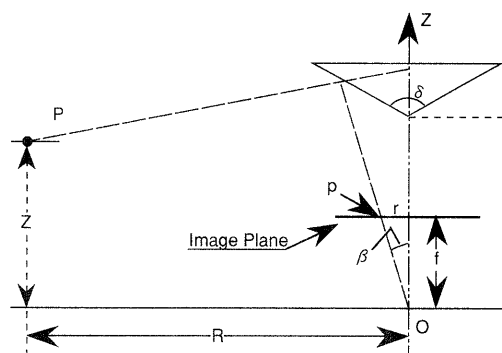
水平方向の視野角は、約 13 度、垂直方向は約 6 度である。本センサの局所視の視野は決して広くはない。しかし 1 章でも述べたように一般に人の中心視と周辺視の関係を考えても、詳細な情報の獲得は、主に網膜の中心付近で行われ、視野 6 度程度に制限されている。そして人間は眼球運動により注視方向を変えるこ



(a) Visual field of COPIS



(b) Invariability of azimuth angle in COPIS



(c) Linear relation of tilt angle in COPIS

図3 COPIS のシステム構成
Fig. 3 COPIS imaging system.

とで得られる局所的詳細情報間を周辺視がとりもつことで注視物体全体の情報を得ている。すなわち、人間のよう複数のタスクを実行するために中心視と周辺視は重要な機能となる。このような枠組みで考えた場合、大局視と局所視の両機能を持つ本センサでは、ロボット自身が回転することで、水平方向に視野を広げることができ、また垂直方向については光路変更部のミラーを走査することで広い視野角を得ることができる。そしてこれら局所視間の情報をつなぐ情報として全方位視覚の情報は利用できるもので、制限された視野でも、部分を詳細に見るという目的に対し適合してい

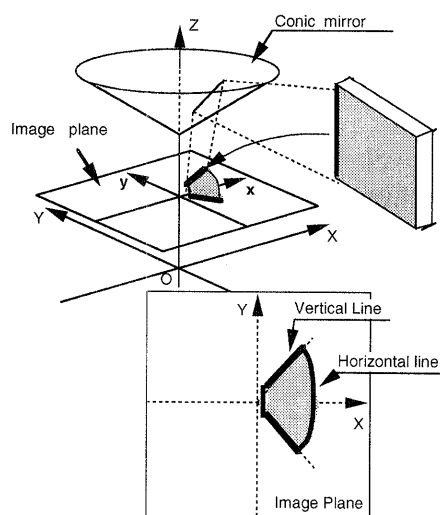


図4 円錐射影

Fig. 4 Conic projection of 3-D lines.

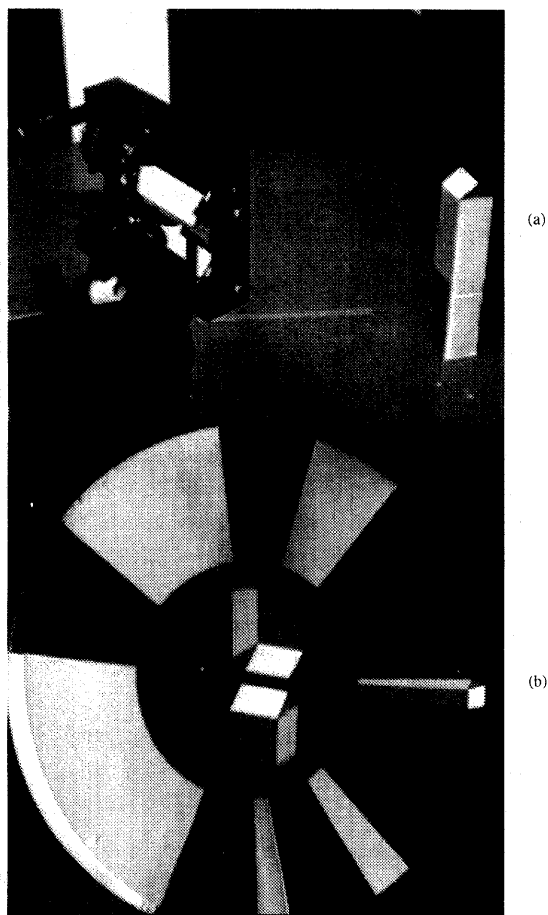


図5 入力画像と実験環境の一例

Fig. 5 An example of input image & experimental environment.

ると言える。なお局所視情報の大局視による統合に関しては、別報にて報告する予定である。

大局視（全方位視覚）と局所視（両眼視）をもしそれぞれ独立した視覚として構築した場合、入力カメラは両眼視用に2台、全方位視用に1台が必要で、さらに実時間で画像撮像を行うためには、それぞれのカメラの映像信号を入力（AD変換）するための画像入力装置も各々用意する必要がある。その結果、大きさや値段などシステムの全体コストが上がる。それに対し、本センサでは1台のカメラですべての画像を得ることができるため、処理系も含めてコンパクトになる利点がある。また本センサでは、一つの画像面上で入力から前処理の段階まで一括処理ができるため、処理の効率化が可能となる。例えば、次章で述べる複合センシングの一方法では、入力、平滑化、微分、2値化までの一連の前処理を同一画像面上で一括で処理している。

さらに複数のカメラを利用した場合、カメラキャリブレーションの問題が発生するが、本センサでは、カメラが1台のため三つ（全方位+両眼視）の光学系でカメラの焦点距離が同じまたカメラ中心はすべて同じなどの特徴があるため、これらの制約関係を考慮することで別々のカメラを3台用意する場合と比べ、キャリブレーション時のパラメータを減らすことができる。キャリブレーションの結果については、4.1節に示す。

3. 複合センシング方法

3.1 概要

本章では、ロボットが移動しながら、大局視により環境全体の中から未知物体を発見、その位置を推定し、注視しながら局所視により、この未知物体のより詳細な3次元情報を獲得するための大局視と局所視の利用方法について述べる。すなわち注視行動における発見、追跡、認識の一連の行動を大局視と局所視の観点から考える。

本章で示すセンシング方法では、以下に示す条件をおく。ロボットが移動する環境は既知とする（環境マップが既知）が、部分的に未知物体が存在する。ただしこの物体の位置は未知である。ロボットは環境内のある基準地点から出発し、与えた経路に沿って走行する。この時、ロボット自身の移動量は、エンコードからの情報により得られるが、実際にはロボットが移動する際の揺れ、車輪のすべりなどの原因によって、計測誤差が含まれるものとする。

複合センシングの概要を図6に示す。環境マップと大局視の全方位入力画像との照合によりロボットの

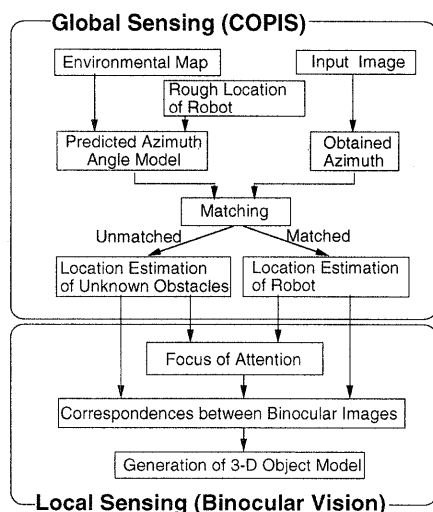


図6 複合視覚 MISS におけるセンシングアルゴリズム
Fig. 6 Outline of navigation algorithm by multiple sensing.

自己位置を推定しながら、与えた経路に沿って走行する。そして環境マップと対応しなかった垂直エッジは、未知物体の構成要素の候補とする。ロボットは移動しながら上記の処理を行い、連続して未知物体候補となった場合、それらのエッジを未知物体と同定し、位置の推定を行う。以上の処理は、大局視の情報により行う。

次に大局視で発見された未知物体を注視対象とし、局所視の視線を注視対象に向ける。さらにロボットは移動しながら、局所視で注視対象を見続け、その詳細な3次元情報を獲得する。この局所視による注視動作においては、大局視で推定した未知垂直エッジの位置情報を利用する。

局所視では、三角測量により大局視で観測されない垂直エッジ以外のエッジについても3次元情報を獲得する。一般に、ステレオ視により3次元情報を獲得するためには、左右の画像間での対応関係を得る必要がある。垂直エッジに関しては、大局視により位置情報が得られているので、この位置情報を利用することで左右ステレオ視の画面内での対応関係を得る。そして対応付いた垂直エッジを基準に他のエッジ間の対応関係を探索する。さらに3台のカメラにより生じる幾何学的拘束を利用した三眼視の原理を、移動視に拡張して用いることで、上記処理の結果に対する検証を行い、最終的に対応付いたエッジの3次元情報を獲得する。以下、両視覚系の処理について述べる。

3.2 大局視における処理

大局視では全方位視覚センサ COPIS のセンサ情報

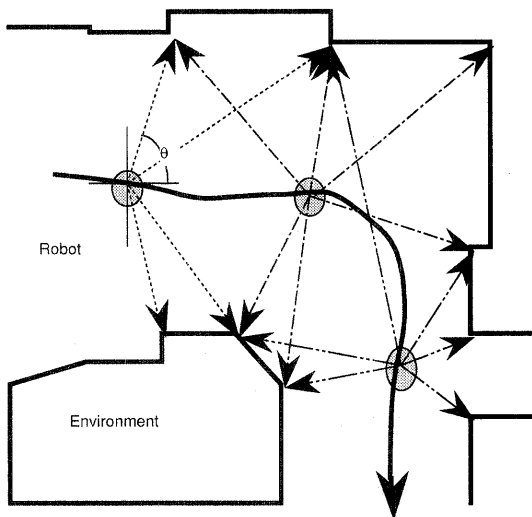


図7 環境マップを用いたロボットの自己位置同定
Fig. 7 Estimation of location of robot.

(全方位画像)を処理して、周囲環境に対する自己位置の推定を行い、また未知物体を発見し、その位置推定を行う。

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づける必要がある。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 $(X_{\text{robot}}, Y_{\text{robot}})$ 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 θ に観測されるはずである。

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

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7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。実際には、三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める。

しかし、実際の観測方位角には計測誤差が含まれている。そこで、本手法では、最小二乗法を用いることでより正確な位置を算出している。詳細に関しては文献11)を参照されたい。なお、自己位置は後述のステレオ移動視におけるカメラ位置更新に利用する。

3.2.2 未知物体位置の算出

観測物体と環境マップの対応付けの際に、対応付か

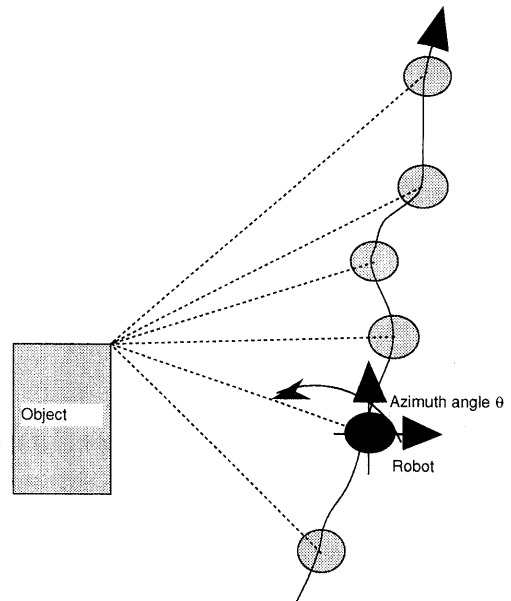


図8 未知物体の位置推定
Fig. 8 Estimation of location of an unknown object.

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る。

未知物体の座標位置を $P0(X0, Y0)$ 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 $(-X(t) + X0, -Y(t) + Y0)$ となる。したがって、時刻 t における未知物体の方位角 $\theta(t)$ は次式となる。

$$\tan \theta(t) = \frac{-Y(t) + Y0}{-X(t) + X0} \quad (3)$$

ここで、ある時刻 $t1, t2$ 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 $(X0, Y0)$ 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8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。詳細に関しては文献11)を参照されたい。算出された垂直エッジ位置情報は、注視制御ならびに後述の対応付け問題に利用する。注視制御では、大局視と局所視の関係があらかじめカメラキャリブレーションにより補正されているので、大局視に対する未知物体の位置情報から局所視と未知物体の関係が推定できる。よって局

所視の左右両眼の視野内に未知物体が観測されるようなロボットの姿勢（方位）を、推定することで注視行動を行う。なお、補正方法に関しては、4.1 節を参照されたい。

3.3 大局情報の利用による局所視での物体構造の獲得

3眼視の原理を適用した移動ステレオ視では、ロボットの移動量が求まっていることが前提となる。またステレオ視では、ステレオ画像間での対応付け問題が存在する。これらの問題に対し、大局視情報を利用する。以下、大局視情報の局所視での利用方法について述べる。

3.3.1 大局視情報のステレオ移動視での利用

移動ステレオ視での大局情報の利用に関する説明の前に、本手法で用いた移動ステレオ視の原理について説明する。

三眼視における線分の対応付けの原理を図 9 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_2q_2 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。詳細に関しては文献 12)~14) を参照されたい。

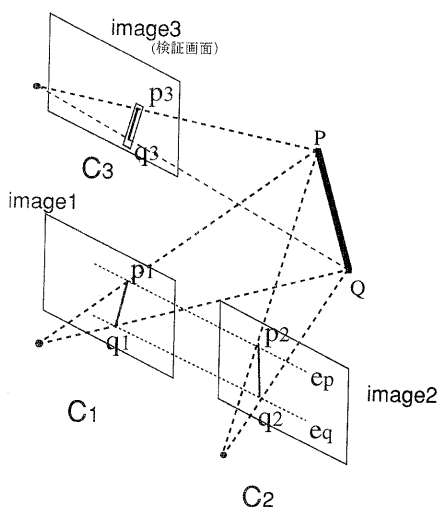


図 9 三眼視の原理

Fig. 9 Principle of segment-based trinocular stereo.

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 枚ずつ計 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 10 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。詳細に関しては文献 15), 16) を参照されたい。

三眼視の原理を移動ステレオ視に用いる場合、ロボットの移動前後におけるカメラ位置、姿勢を知る必要がある。そこで本システムでは、大局視により得られる環境に対する自己位置情報を移動ステレオ視における移動前後の位置関係の情報として利用する。

3.3.2 大局情報の利用による対応付けの効率化

MISS は、1 台のカメラ（撮像素子）からなるため、レンズ中心ならびにレンズの光軸などは本来一つしかない。しかし、二つの視覚系（全方位視、両眼視）をミラーの組み合わせにより実現しているため、仮想的に三つの座標系（全方位視、両眼視の左右）がある。ただし、局所視は両眼視のため二つの座標系があるが、大局視との関係は両者とも同様な表現系をとる。よって、以下、説明では、大局視ならびに局所視の座標系を $O\text{-}XYZ$, $o\text{-}xyz$ とする。局所視座標系の原点は、視覚系のレンズ中心とし、レンズの光軸と z 軸を一致させる。画像面は z 軸上で原点から焦点距離 f 離れた位置に垂直にあるとする。

大局、局所両座標系間の関係は、回転成分 R と並進成分 T で次式のように表せる。

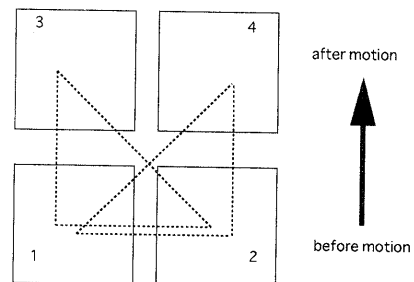


図 10 移動ステレオ視

Fig. 10 Two sets of three images.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + T \quad (4)$$

ここで、カメラ座標系の単位ベクトルを大局視座標系で表し Ex , Ey , Ez とすると、

$$Ex = \begin{bmatrix} Exx \\ Exy \\ Exz \end{bmatrix}, \quad Ey = \begin{bmatrix} Eyx \\ Eyy \\ Eyz \end{bmatrix}, \quad (5)$$

$$Ez = \begin{bmatrix} Ezx \\ Ezy \\ Ezz \end{bmatrix} \quad (5)$$

と表せる。このとき回転成分 R は

$$R = \begin{bmatrix} Exx & Eyx & Ezx \\ Exy & Eyy & Ezy \\ Exz & Eyz & Ezz \end{bmatrix} \quad (6)$$

で表される。一方、並進成分 T は、

$$T = \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix} \quad (7)$$

と表せる。

ここで大局視座標の点 $P(X, Y, Z)$ が局所視座標に変換された点を $p(x, y, z)$ とすると、カメラの画像面に投影される点 $p'(x', y')$ は、画像中心を (cx, cy) , アスペクト比を a とすると、

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{f}{z} \begin{bmatrix} \frac{x}{a} \\ y \end{bmatrix} + \begin{bmatrix} cx \\ cy \end{bmatrix} \quad (8)$$

と表せる。式 (4), (8) より、大局視座標上での点の局所視画像面上への投影点を決定できる。

したがって、大局視で観測された未知垂直エッジ位置が正確に求まっているとすると、この垂直エッジを含む直線を考え、式 (4) の関係より、局所視座標系でのこの直線方程式を求める。次に式 (8) より、局所視画像面上での投影直線を推定する。図 11 (a) のように実際の投影線分 (図中太線) は、この投影直線上 (図中点線) に現れる。よって、未知垂直エッジに関しては位置情報をステレオ画像の両画面に投影したとき、投影直線上に存在する線分が両画面で唯一存在するか否かにより、対応付けることができる。しかし、実際には大局視で観測された未知垂直エッジ位置には観測誤差が含まれるため、画像面への投影直線と各ステレオ

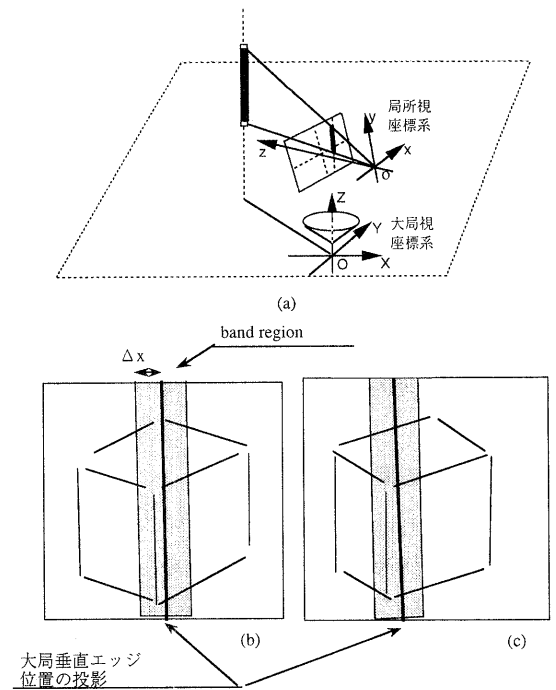


図 11 垂直エッジの対応付け

Fig. 11 Correspondence of vertical edges.

オ画像面上での対応線分は一致しない。そこで、ステレオ両画像面各々に投影される投影直線に、一定のバンド幅 (Δx) を持たせた領域を考え、図 11 (b), (c) に示すように垂直エッジの対応付けは、ステレオ画像の左右画面にバンド領域を投影し、その領域内における垂直エッジの有無を検証することで行う。現在バンド幅 (Δx) は実験的に決めている。

次に、垂直エッジの対応関係をもとにして帰納的に対応付けを行う方法について述べる。前記のように大局視情報を用いて局所視の両画像間で、垂直エッジの対応付けを行った後、一方のステレオ画像内で対応付いた垂直エッジに接続する未対応線分を探す。垂直エッジの端点近傍領域内にエッジ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も対応する垂直エッジに接続する線分を探し、すべての接続線分を対応付けの候補とする。そしてステレオ視におけるエッジポラ拘束を用い、対応候補となった線分候補から正しい対応点の探索を行う。このように対応する垂直エッジを手掛かりに探索範囲を限定することで対応付けを行う。それでも、複数候補が存在する場合などは、前節で述べた 3 眼視の原理を用い対応関係を検証する。検証画面でも検証されるべき線分は接続関係により限定

されている。同様に、残りの線分の対応付けでも、対応付いた線分（垂直エッジ、接続関係により対応付けられたものなど）の両画面での対応関係とそれらの線分に接続する未対応のエッジとの接続関係を調べた結果で対応候補を限定した対応付けが帰納的に行える。もちろん、この接続関係を用いた対応付けでは、接続関係のない線分などすべての線分が対応付けられるわけではない。しかし、接続関係で対応付いた線分は、その後の対応付けでは対応候補から除くことができ、また、すべての対応関係は未対応のエッジの対応付けの際の探索範囲の限定に利用できる。

以上、対応関係が得られた線分に関し、3次元情報を推定する。

4. 実験結果と検討

4.1 カメラキャリブレーション

大局視と局所視間の位置姿勢関係をあらかじめキャリブレーションする必要がある。図12に示すようにキャリブレーション環境は、四つの垂直な面で囲まれ、各面には等間隔の垂直状の模様がある。各面の長さ1mで垂直線の間隔は10cmである。また環境内の一つの角に立方体のキャリブレーション物体を配置する。この物体には、1cm等間隔の格子パターンがある。局所視（ステレオ視部）のキャリブレーションは、キャリブレーション物体との位置姿勢関係を求めることで行う。キャリブレーションは我々の研究室で以前より用いてきた方法を一部改良し用いた。文献17)の方法と基本原理は同じである。詳細は参考文献に譲るが、この方法ではカメラを1台ずつキャリブレーションする方法であるため、焦点距離が共通といった条件を考慮した処理がなされていない。そこで実際には、左右各々のキャリブレーションを行った際に求めた焦点距離の平均を共通の焦点距離とし、他のパラメータと

もに初期値として、各パラメータの値を変化させ、誤差最小化となるパラメータを探索している。個別に求めた左右の焦点距離の値は、655.157 (pixel) ならびに 655.115 (pixel) であり、これらの平均値を焦点距離としている。誤差は、キャリブレーション物体上の点 (x, y, z) の入力画像中での位置とその点を求めたパラメータを用い逆透視変換した時の画像中での位置との位置ずれにより定義した。表1はこの手法により求めた左右のカメラパラメータである。上記誤差に関しては、最終的に平均 0.21 (pixel) 最大 0.37 (pixel) であった。キャリブレーション物体とセンサ間の距離は 60 cm 程度である。

さらにキャリブレーション環境とキャリブレーション物体の関係はあらかじめ計測できるため、3.2.1項で述べた自己位置推定方法を利用することで、大局視とキャリブレーション環境との位置関係を求めることで、大局視と局所視のキャリブレーションが可能となる。大局視とキャリブレーション環境との相対位置推定結果は、(-531 mm, 579 mm) でこの時の最小二乗における偏差は 0.608 mm であった。

4.2 複合センシング方法

図5(a)に示した約 2.5×1.6 m の広さの室内環境において実験を行った。未知物体は、底面 8×8 cm の立方体形状である。画像サイズは、 512×480 画素（画素レベルは 256）である。図のように実験環境には、未知物体として家のような箱があり、回りの環境（他の垂直エッジ位置）は、あらかじめマップ情報として与えた。ロボットは 3.0 cm 移動するごとに画像を撮像し、141 cm 直進した。図13(a), (b), (c), (d) はそれぞれ 1, 13, 14, 39 フレームの入力画像である。一定以上対応付いた未知垂直エッジに対しては、次フレーム以降局所視により注視行動を行っている。本実験では、13 フレーム以上とした。図14は、垂直エッジの方位角を観測しながら、ロボットの自己位置を同定し、さらに未知の垂直エッジ位置を算出した結果である。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また未知物体の垂直エッジの位置を●印で示した。図中、未知物体上の垂直エッジが近接して存在するため位置関係が見づらい。よって、全体図の下に未知物体のみ引き出し結果を示した。ロボットの平均位置計測誤差はおよそ 7 mm で未知の垂直エッジの平均位置計測誤差は、最短観測距離が 67 cm に対しおよそ 7 mm であった。実測値は、巻き尺により行っているため、実測時の誤差が 1, 2 mm 考えられる。この点も考慮すると、a, b, c の3点に関しては、ほぼ正確に位置推定できた

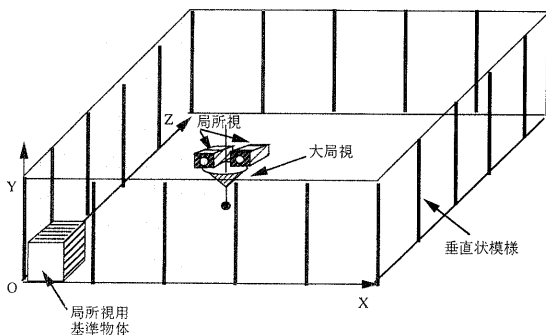
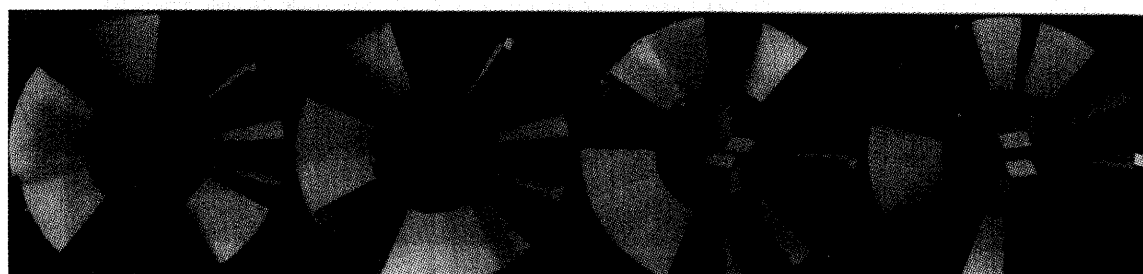


図12 カメラキャリブレーション
Fig. 12 Camera calibration.

表 1 大局視-局所視間のカメラパラメータ推定結果
Table 1 Results of estimated camera parameter.

	Rotational Matrix R			Translation T
Left	0.7580047	0.203287	-0.619709	532.449715
Camera	0.041213	-0.963214	-0.265556	481.826384
Coordinates	-0.650896	0.175764	-0.738540	663.020759
Right	0.713003	0.201832	-0.671483	619.686273
Camera	0.019331	-0.962969	-0.268920	482.827665
Coordinates	-0.700894	0.178761	-0.690501	583.891197



(a) 1th

(b) 13th

(c) 14th

(d) 39th

図 13 時系列全方位画像の一例

Fig. 13 Example of omnidirectional images.

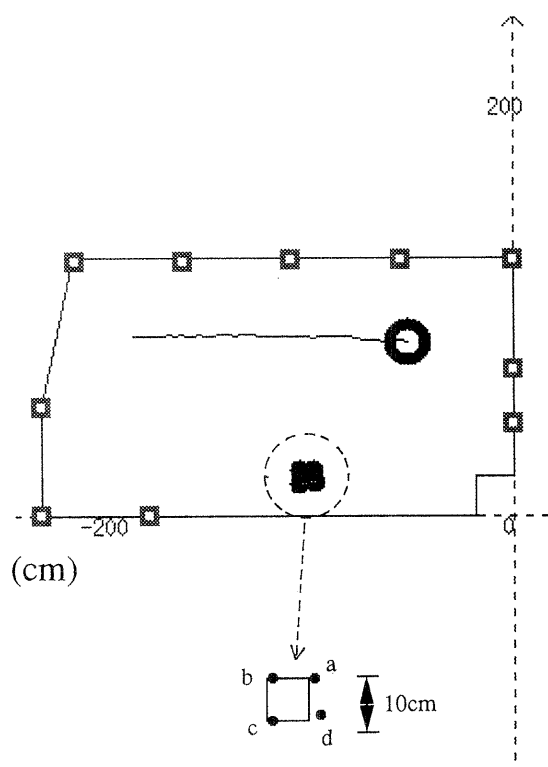


図 14 自己位置および未知垂直エッジ位置の推定結果

Fig. 14 Results of location estimation.

と言える。点 d に関しては、他の 3 点と比べ観測フレーム数が少なく、また観測方向がかたよっているため、誤差が大きかったと考えられる。この誤差は、観測を継続して行ったり、観測誤差を考慮したロボットの行動計画を行うことで改善できるものとする。図 16 は 16, 17 フレームでの局所視で観測した物体画像 (図 15 (a), (b)) に対する直線抽出結果である。図 17 に示すように COPIS で推定された垂直エッジ位置をステレオ視内に逆透視変換し、垂直エッジの探索領域 (バンド領域) を設定した。図 18 中、黒の太線が、対応付いた垂直エッジで、○印がエッジの端点位置である。この場合、両眼視の視野の関係から下方の端点に関しては両眼で対応位置が存在しない。よって両視野で共通する端点を基準に、他のエッジについての対応付けを行う。図 19 は、水平エッジの両端点が垂直エッジと接続されている場合の対応結果である。白丸は対応付いた頂点である。図 20 は、水平エッジの片側が垂直エッジに接続していないときの対応付け結果である。黒丸は接続していない頂点である。図 21 では、すでに対応が決定された線分 (図中太線) との接続関係で真の対応が求まった結果である。上記対応探索においては、移動ステレオ視により対応関係の検証も行ったが、シーンが単純であったため、移動ステレオ視を有効に利用した結果は示せなかった。しかし、発見、追跡、認識の一連行動を大局視と視野領域の限られた局所視の組み合わせで実現できることは示され

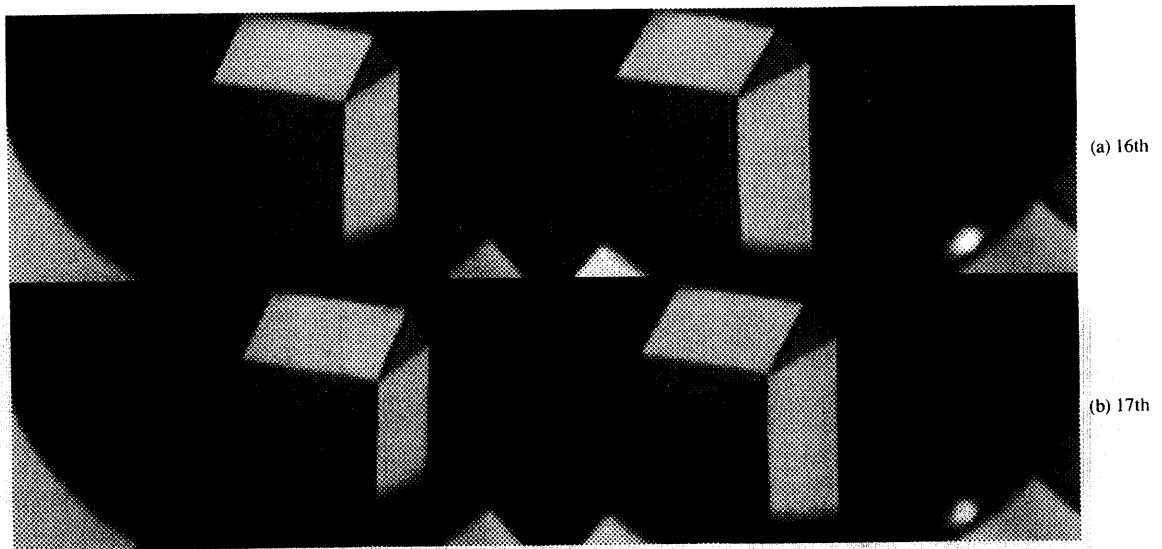


図 15 時系列ステレオ画像の一例

Fig. 15 Examples of binocular motion images.

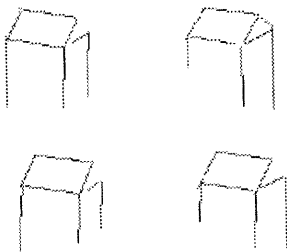


図 16 エッジ画像

Fig. 16 Edge picture at the 16th & 17th frame.

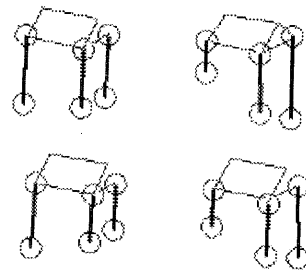


図 18 垂直エッジの対応関係

Fig. 18 Correspondence of vertical edges.

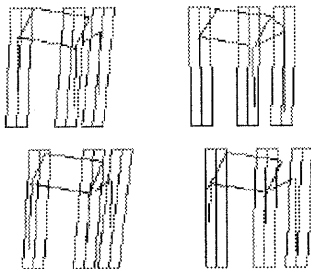


図 17 探索領域の設定

Fig. 17 Setting search region at the 16th & 17th frame.

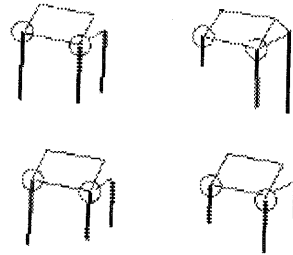


図 19 候補の検証

Fig. 19 Verification of candidates (Both vertices are connected).

たと思う。なお、ステレオ視でのエッジ位置の計測誤差はおおよそ7mmであった。計測誤差は大局視による垂直エッジの場合とほぼ同程度である。これは、大局視では時系列データから最小自乗法により位置推定を行ったことで方位角分解能が1度でも精度良く求まったのに対し、局所視ではステレオ視のベースラインの間隔が狭かったために大局視と同程度の精度となった

と考えられる。局所視においても時系列データを用いればよりよい精度で計測できるものとする。

また、今回の実験では、大局視での位置計測精度が比較的よかったため、大局視で得られた情報が局所視で有効に利用できた。しかし屋外環境などの実環境では、大局視での位置誤差が大きくなるものと考えられ、大局視から局所視へ投影変換した場合、必ずしもバン



図 20 候補の検証

Fig. 20 Verification of candidates (One side of an edge is connected).

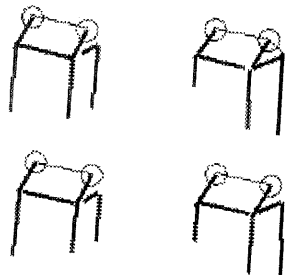


図 21 最終結果

Fig. 21 Results after final verification.

ド域内にエッジが入るとは限らなくなるものと考えられる。したがって、大局視からの情報によりある程度限定された範囲内でロボットが自ら左右に回転することで、注視対象を発見する必要があると考える。このような能動的観測および注視動作のプランニングに関しては今後の課題とする。

5. 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、注視行動における発見、追跡、認識の一連の行動を大局視と局所視の統合によって実現する方法について述べた。

本センサは、大局視と局所視という性質の異なる視覚機能の一つの撮像素子上に結像させた複合視覚センサである。一体化により局所視の視野領域は、通常のカメラと比べ、水平約 6 度と狭くなったが、人の中心視の視野領域も 6 度程度と考えると、実用上問題ないのではと考える。なぜなら、局所視間の情報をつなぐ情報として全方位視覚の情報は利用できるもので、制限された視野でも、部分を詳細に見るという目的に対し適合していると言える。なお局所視情報の大局視による統合に関しては、別報にて報告する予定である。

大局視-局所視での注視では、大局視により、未知対象の発見、追跡、局所視にて認識(対応付け、3 次元情報の獲得)を行い、さらに大局視で得られた情報を活用することができた。したがって、大局視と局所視という複合視覚は、注視行動の一連処理に有用であ

ると思われる。

試作したシステムでは、絞り、フォーカス、ズーム機能などをリモートコントロールする機能を有していない。したがって、フォーカスに関しては、現システムのカメラの焦点深度が深いため、ある程度の奥行き変化に対して適用できるが、絞りに関しては、ステレオ視部に改良を施す必要がある。またステレオ視には、ズーム機能がないため、注視対象をズームアップするなど能動的行動を行うためには、ロボットが近づくなどのロボット自身の行動と合わせて考える必要がある。

今後の課題としては、試作システムの改良、局所視情報の大局視への利用、注視動作のプランニング、局所視で計測誤差を小さくするための時系列ステレオ画像の有効利用、より複雑な実環境内での有効性評価などを考えている。

参 考 文 献

- 1) 清水 編: 生物の目とセンサ, 情報調査会 (1985).
- 2) 特集: センサフュージョン, 日本ロボット学会誌, Vol.12, No.5, pp.645-728 (1994).
- 3) Asada, M. and Shirai, Y.: Building a World Model for a Mobile Robot Using Dynamic Semantic Constraints, *Proc. of Int. Joint Conf. Artificial Intelligence*, pp.1629-1634 (1989).
- 4) Grandjean, P. and Vincent, A.R.: 3-D Modeling of Indoor Scenes by Fusion of Noisy Range and Stereo Data, *Proc. IEEE Int. Conf. Robotics & Automation*, pp.681-687 (1989).
- 5) Zuk, D. and Dell'eva, M.: 3D Sensor for Autonomous Land Vehicle, Technical Report, Environmental Research Institute of Michigan (1985).
- 6) Thorpe, C., Hebert, M., Kanade, T. and Shafer, S.: Vision and Navigation for the Carnegie Mellon NAVLAB, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.10, No.3, pp.362-373 (1988).
- 7) Waxman, A.M., LeMoigne, J.J. and Seinsivasan, B.A.: A Visual Navigation System for Autonomous Land Vehicles, *IEEE Journal of Robotics and Automation*, Vol.RA-3, No.2, pp.124-141 (1987).
- 8) Hebert, M.: Building and Navigating Maps of Road Scenes Using An Active Sensor, *Proc. of Int. Conf. Robotics and Automation*, pp.1136-1142 (1989).
- 9) Yagi, Y. and Kawato, S.: Panorama Scene Analysis with Conic Projection, *Proc. IEEE Int. Workshop Intelligent Robots & Systems*, pp.181-187 (1990).
- 10) Yagi, Y., Kawato, S. and Tsuji, S.: Collision

Avoidance Using Omnidirectional Image Sensor (COPIS), *Proc. IEEE Int. Conf. Robotics & Automation*, pp.910-915 (1991).

- 11) Yagi, Y., Nishizawa, Y. and Yachida, M.: Map Based Navigation of the Mobile Robot Using Omnidirectional Image Sensor COPIS, *Proc. IEEE Int. Conf. Robotics and Automation*, Nice, Vol.1, pp.47-52 (May 1992).
- 12) Yachida, M.: 3-D Data Acquisition by Multiple Views, *3rd ISRR* (1985).
- 13) Ayache, N. and Lustman, F.: Fast and Reliable Passive Trinocular Stereo Vision, *Proc. ICCV*, pp.422-427 (1987).
- 14) Ayache, N. and Lustman, F.: Trinocular Stereovision: Recent Results, *Proc. 10th IJCAI*, pp.826-828 (1987).
- 15) Kimachi, M. and Yachida, M.: Getting 3-D Information by Edge-Based Trinocular Stereo, 信学会, PRU研究会, Vol.88, No.324, pp.13-20 (1988).
- 16) Yachida, M. and Tsudo, A.: Understanding Spatial Configuration of Robot's Environment from Stereo Motion Images, *Proc. IEEE IROS'91*, pp.882-887 (1991).
- 17) 越後富夫: 三組の平行線群を用いたカメラパラメータの較正, 情報処理学会研究会報告, Vol.88, No.52, pp.1-8 (1988).

(平成6年10月24日受付)

(平成7年5月12日採録)



奥村 仁

1970年生。1992年大阪大学基礎工学部情報工学科中退。1994年同大学院制御工学分野修士課程修了。同年富士通(株)入社。ヒューマンインタフェース関連の開発に従事。



八木 康史(正会員)

1959年生。1983年大阪大学基礎工学部制御工学科卒業。1985年同大学院修士課程修了。同年三菱電機(株)入社。同社産業システム研究所にてロボットビジョンの研究に従事。1990年大阪大学基礎工学部情報工学科助手。1993年同学部システム工学科講師。IEEE, 電子情報通信学会, 日本ロボット学会, システム制御情報学会等各会員。工学博士。



谷内田正彦(正会員)

昭和46年大阪大学大学院工学研究科修士課程修了。同年同大基礎工学部制御工学科助手。同助教授を経て同学部情報工学科教授。平成6年同学部システム工学科教授。昭和42~43年デンマーク原子力研究所留学。昭和47~48年米イリノイ大学にてResearch Associate。昭和55~56年西独ハンブルグ大学Research Fellow。昭和57年米ミネソタ大学CDC Professor。ロボット学会, 人工知能学会等会員。著書: ロボットビジョン(昭見堂), コンピュータビジョン(丸善, 編著)等。コンピュータ・ビジョン, 画像処理, 人工知能, 移動ロボットなどの研究を行っている。工学博士。