



技術解説

# 移動ロボットの視覚

谷内田 正彦\*

## Vision for Mobile Robot

**Key Words :** Mobile robot, Robot vision, Trinocular vision, Omnidirectional sensig

### 1. 知的移動ロボットには視覚が必要

従来のロボットはある一定の場所に据え付けられ、定型的な繰り返し作業をしていることが多かった。しかし、ロボットが動物のように移動できるとその行動範囲は一挙に広がり、応用分野も拡大する。また、研究面からも未開の分野で興味深いテーマが多い。このようなことから、近年、移動ロボットの研究が急速に盛んとなってきた。特に、米国では、国防総省先端研究計画局の自律走行車計画やNASAの火星探査ロボット計画などのプロジェクトを資金的な駆動力として、活発に研究が行われている。

一方、移動ロボットの実用化としては、搬送作業を自動化するための無人搬送車などがある。ロボットを誘導するため、床面に電磁誘導線や白線などを設置し、センサでこれらを検出して移動の制御を行っている。一定の経路上で定型作業を繰り返すには性能価格比の良い方法といえる。

しかし、オフィスや家庭、あるいは屋外など一般の人々が混在する環境で不定型な作業をするには、自分の周囲の環境を認識できる視覚を持った移動ロボットが必要となる。本稿では我々

の研究室で行っている移動ロボット用の視覚の研究の一端を述べる。

### 2. 三次元情報を獲得する

ロボット用の目としてはテレビカメラが良く用いられる。しかし、テレビカメラで得られるのは3次元世界を2次元の平面に投影した画像である。この2次元画像から元の3次元世界を認識するのは難しいので、通常はカメラを2台使い三角測量の原理を用いて3次元情報を獲得する。これを両眼ステレオ視と呼ぶ。話は少々ややこしくなるが、その原理を図1を用いて説明しよう。

#### 2.1 両眼視

図1で、 $F_1, F_2$ はそれぞれカメラ $C_1, C_2$ のレンズ中心(焦点に相当)である。これらのカメラを設定した時に、既知の位置に既知の物体を置き、それらの左右画面への映り方から、 $F_1, F_2$ の位置や $C_1, C_2$ の位置、向きを計測することができる。これらの値はカメラの設定を変えない限り一定である。

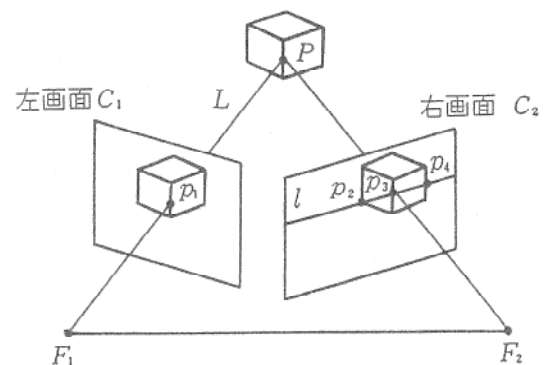


図1 両眼ステレオ視の原理



\*Masahiko YACHIDA  
1945年9月4日生  
昭和44年大阪大学工学部電気工学科卒業  
現在、大阪大学基礎工学部、システム工学科、教授、工学博士、ロボット工学、コンピュータビジョン、人工知能  
TEL 06-850-6360

今、空間中の点Pが左画面C<sub>1</sub>には点p<sub>1</sub>に映り、右画面C<sub>2</sub>には点p<sub>3</sub>に映っており、p<sub>1</sub>、p<sub>3</sub>の画面上での位置が画像処理によりわかったとする。そうすると、p<sub>1</sub>、p<sub>3</sub>の3次元位置が決まり、またF<sub>1</sub>、F<sub>2</sub>の3次元位置はわかっているので、点Pの3次元位置は三角測量の原理から二つの直線F<sub>1</sub>、p<sub>2</sub>とF<sub>2</sub>、p<sub>3</sub>の交点として求めることができる。

しかし、以上の議論は、左画面の点p<sub>1</sub>に右画面の点p<sub>3</sub>が対応することがわかって初めてできる。対応点を求める一つの方法は明るさの相関を用いる方法である。すなわち、一方の画面で目印となるような特徴的な点を発見し、その点のまわりに「窓」と呼ぶ小領域(7×7点程度が多い)をとり、他の画面でその窓内の明るさとよく似た分布を持つ部分を相関により見つける。特徴点としては、角(カド)などを用いることもあるが、数多くの点の3次元情報を得たいという理由から、エッジ点(面と面との境界)がよく使われる。他方の画面で対応点を探るとき、全画面を探索したのでは処理時間もかかり、また、偽の対応づけを行う可能性が高くなるので、探索範囲を狭くすることが重要である。

両眼視の場合、対応点は必ず図1に示す視線像と呼ばれる直線l上にある。すなわち、空間中の点PはF<sub>1</sub>とp<sub>1</sub>を結ぶ直線L上にあるので、当然、Pの右画面C<sub>2</sub>への像p<sub>3</sub>は、直線Lの右画面への像l上にあるといえる。したがって、点p<sub>1</sub>の対応点を探索するには、画面C<sub>2</sub>の直線l上のみを探索すればよい。

探索範囲が限定されても、特徴点の局所的な性質(明るさの相関)のみで対応点を決定するのは容易ではない。たとえば、図1の左画面の点p<sub>1</sub>に対応する可能性のある点は、右画面でD<sub>2</sub>、D<sub>3</sub>、D<sub>4</sub>の3点がある。この例は簡単なので、局所的性質だけで対応づけができそうだが、対象が少し複雑になると、明るさの分布の似た所の数が増え、対応点の候補を絞り切れなくなる。このため、様々な方法が考案されてきたが、まだ決めてとなる方法はない。我々はカメラをもう1台使い、3眼にすることによりこの対応付け問題を解決した。その原理を図2を用いて説

明しよう。

## 2.2 三眼視

空間の点Pの画面1への像をp<sub>1</sub>とし、画面2への像をp<sub>2</sub>とする。p<sub>1</sub>に対応する点を画面2で求めようとする、直線F<sub>1</sub>、p<sub>1</sub>の画面2への像l<sub>2</sub>上で探せばよい。3点F<sub>1</sub>、F<sub>2</sub>、p<sub>1</sub>で構成する平面上に点の一つしかなければ対応はすぐ求まるが、図に示すように、その平面上に他の点Rがあると、その像r<sub>2</sub>も直線l<sub>2</sub>上に乗ってくる。このとき、点p<sub>1</sub>に対応する点がp<sub>2</sub>かr<sub>2</sub>かを判別することは難しい。実際の画像では、F<sub>1</sub>、F<sub>2</sub>、p<sub>1</sub>で構成される視線面上の点はすべて視線像l<sub>2</sub>上に乗ってくるのでなおさらである。これが両眼視における対応点決定の問題であった。

この問題はもう1台のカメラ3を導入することにより容易になる。図2でp<sub>3</sub>、r<sub>3</sub>は、点P、Rの画面3への像であり、当然p<sub>3</sub>は直線F<sub>1</sub>p<sub>1</sub>の像l<sub>3</sub>上にある。また、点p<sub>2</sub>の対応点を画面3で求めることを考えると、その対応点はF<sub>2</sub>p<sub>2</sub>の像m<sub>3</sub>上になくてならない。すなわち、もしp<sub>1</sub>とp<sub>2</sub>とが対応しているなら、直線F<sub>1</sub>p<sub>1</sub>とF<sub>2</sub>p<sub>2</sub>の画面3への像l<sub>3</sub>とm<sub>3</sub>の交点に、点Pの像がなくてはならない。図2の場合、p<sub>1</sub>とp<sub>2</sub>とが対応しているので、l<sub>3</sub>とm<sub>3</sub>の交点のところ

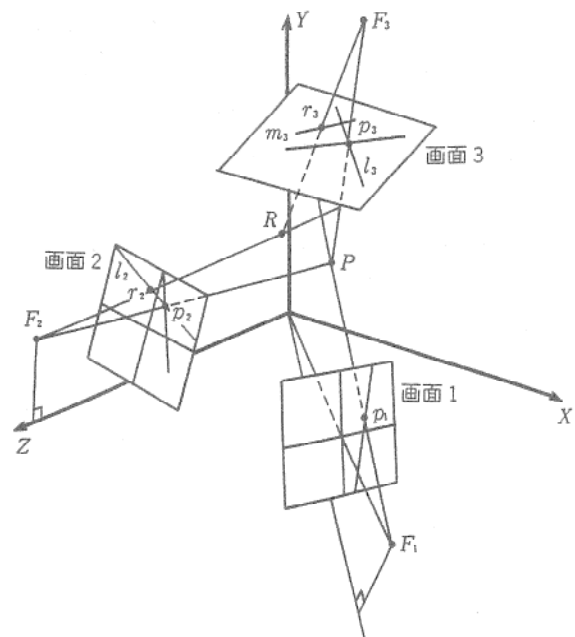


図2 三眼視の原理

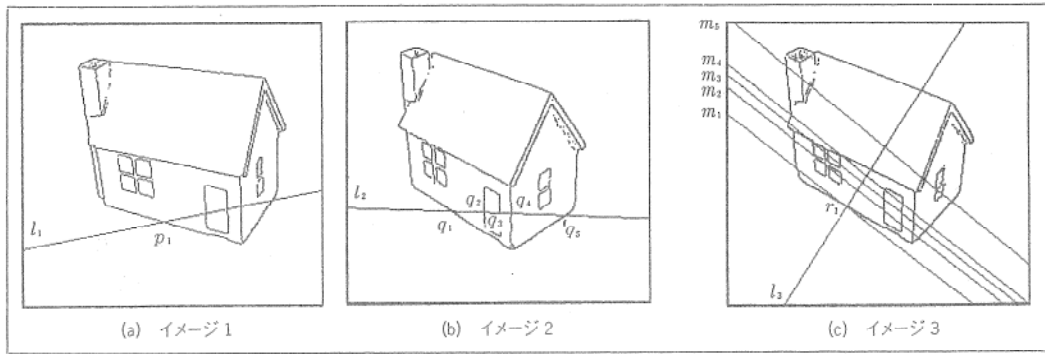


図3 3眼視による3次元情報の獲得

点Pの像 $p_3$ が存在し、このことから3点 $p_1, p_2, p_3$ が対応していることがわかり、点Pの3次元位置は3直線 $F_1p_1, F_2p_2, F_3p_3$ の交点として求めることができる。

実際にはもう少し、工夫しているが、入力画像を微分してエッジを検出し、画面1の点 $p_1$ に対し視線像を引いた結果を図3に示す。同図(c)から分かるように、視線像の交点上にエッジ点があるのはただ1点のみで、これからこの点が真の対応点であることが分かる。

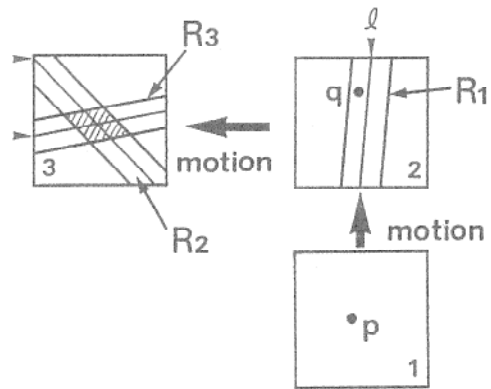
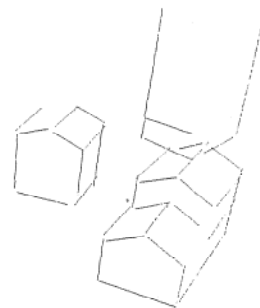
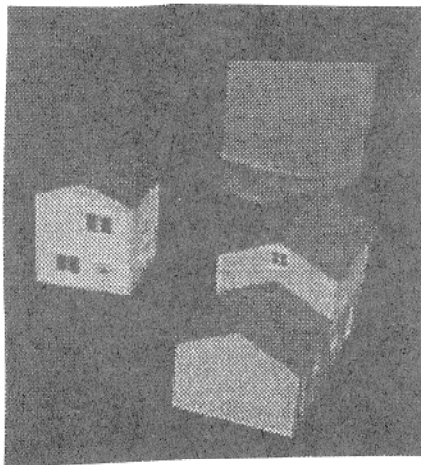


図4 3眼視の移動ロボットへの応用

### 3. 移動ロボットへの応用

この3眼視の方法を移動ロボットの視覚に応用した例について述べよう。もちろん、移動ロボット上にカメラを3台乗せてもよいが、それでは重くなる。移動ロボットに位置センサを付けカメラの位置は移動中既知であるとする、カメラ1台でも3眼視の方法を使える。すなわち、ロボットが図4のように3回移動し、カメ

ラの位置が分かっているならば、3眼視と同様の方法で、環境の3次元情報が得られる。しかし、ロボットが直線運動をすると入力画像も直線上に並び、理想的な3眼視の配置にならない。理想的な配置の画像を得るためには、ロボットは蛇行しながら進まないといけないので、不便である。このため、実際はロボットに2台のカメラを搭載し、直進運動をしても理想的な配置の



(a) (b)

図5 移動ロボットによる3次元情報の獲得

画像を得られるようにしている。

ロボットが移動するに伴い、ロボット前方の3次元情報が新たに得られ環境のモデルが更新される。図5(a)に入力画像の一部を、同図(b)に8枚の画像入力後に構築された環境モデルを示す。

以上で作成した環境のモデルでは、線分が3次元空間中に浮かんでいるだけのものにすぎない。これから次に、面を認識し、またそれらの間の空間的な接続関係を解釈することにより、環境の空間構成を認識する。この時、3次元世界において一般的に成立する幾何学的・物理的な知識をルールの形で知識ベースに入れておき、この知識と環境モデルを用いて推論することにより環境の空間構成を認識する。図6に面の認識結果を示す。後方の物体は十分な情報がなかったため誤認識されているが、他は正確に認識されている。

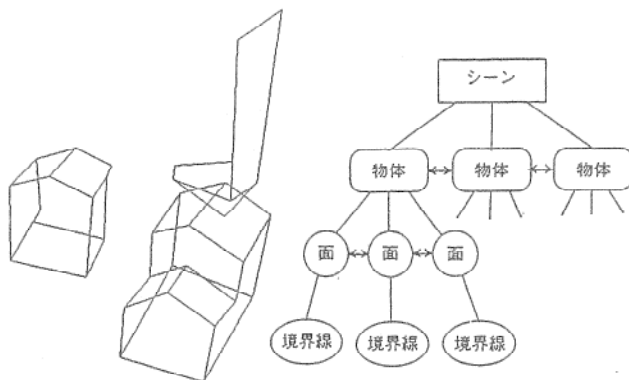


図6 環境認識の結果

#### 4. 大 局 視

以上は、通常のテレビカメラを入力としていたが、テレビカメラでは視野が限られている。このため、その視野内を詳細に解析するのに良い方法であるが(局所視と呼ぶ)、その視野内のことしかわからない。

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(ナビゲーション)と目的地での作業に分けることができるが、これらの作業内容により必要とされる視覚機能は異なる。ナビゲーションには目的地までの移動経路の計画や側方や後方からの移動体との衝突回避など

が必要となるが、このためには、周囲環境の詳細な解析は必要ではなく、大まかで良いからロボットの周囲の環境全体を高速に把握することが必要となる(大局視と呼ぶ)。一方、目的地での作業では、観測範囲は局所的であっても、対象物体の詳細な情報を獲得できる必要がある。前に述べた3眼立体視は局所視には最適であるが、ナビゲーション用の大局視には向いていない。

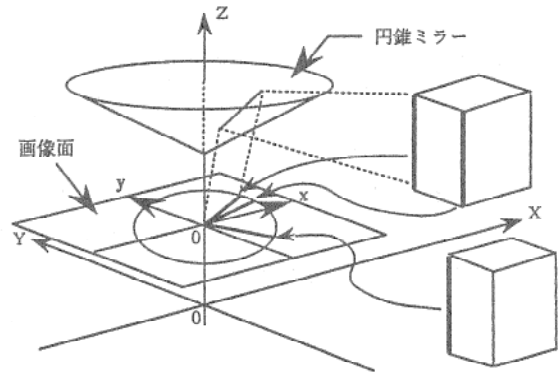
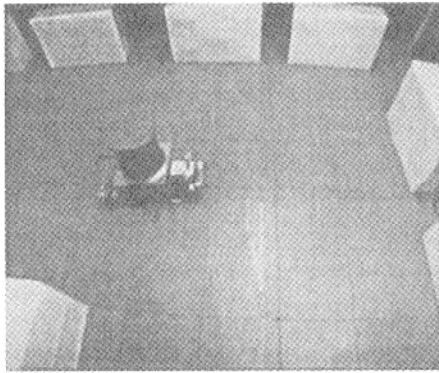


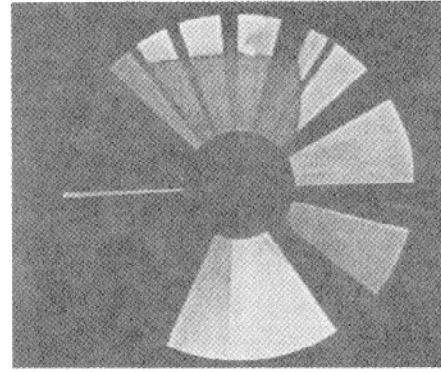
図7 全方位視覚の原理図

このため環境全体を一度に観測処理できる大局視用の視覚として、テレビカメラの前に円錐鏡を設置した全方位視覚センサを考案した。その原理を図7に示すが、円錐鏡を用いる方法はナビゲーションに必要な側方、下方を中心とした360度全周囲の情報を一度に入力できる。円錐鏡を用いる方法の興味ある点は、対象物体のロボットからの方位角が画像上での方位角として直接計測できることである。また、建物内で多く存在する垂直な線が、画像上で画像中心から放射状に伸びた線分として現われることである。従って、画像から放射状に伸びた直線を検出すれば、現在の環境で、どの方角に垂直線があるのかがわかる。一回の撮像では垂直線の方角はわかってもその絶対位置はわからないが、移動後にもう一度撮像して同様の処理を行えば、三角測量の原理からそれらの絶対位置もわかる。

図8(a)に示す環境での入力画像の例を同図(b)に示す。全方位の環境を一つの画像面に投影するため、方位角情報の分解能はあまり高くないが、全方位の画像を一度に入力できるため大局的に環境の構成を知るのに良い。実際の環



(a)



(b)

図 8 全方位画像

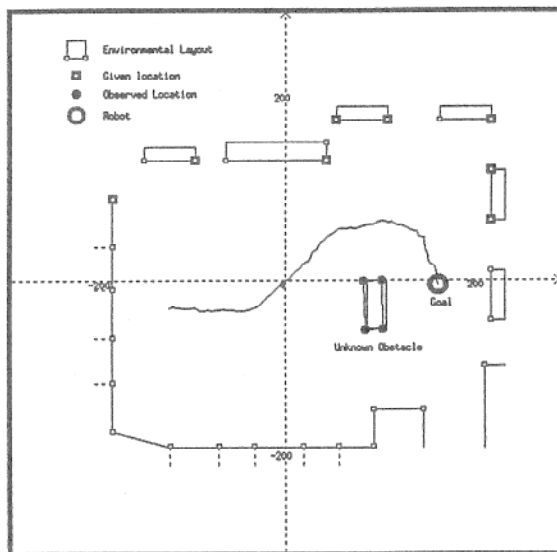


図 9 全方位画像から作成された環境の地図

境でロボットの移動量を既知として垂直線の位置を計算した結果をレイアウト図と重ねて表示したものを図9に示す。

この円錐鏡を用いる方法の欠点は、垂直線は容易に見つかるが、その他の線は円錐鏡による変換を受けて、検出しにくいことである。このため、双曲面鏡をテレビカメラの前に設置した新たな全方位視覚センサの開発した。双曲面鏡を用いれば、垂直線を検出しやすいという円錐鏡の利点を残しつつ、それ以外の線も簡単な逆変換をほどこすことにより容易に検出できる。

その原理を図10に示すが、双曲面の特徴は焦点を持つことである。空間中の点 $P(X, Y, Z)$ は双曲面ミラーにより反射されて画像上の点 $p(x, y)$ に写像される。レンズ中心 $O_c$ (既知)と点 $p(x, y)$ を結ぶ直線と双曲面ミラーとの交点を $P_1$ とすると、ミラーの焦点 $O_M$ と $P_1$ を

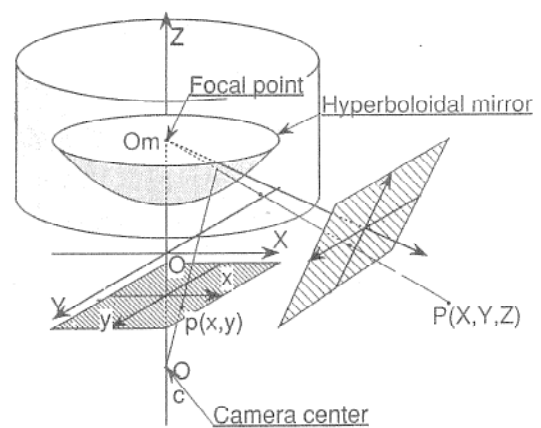
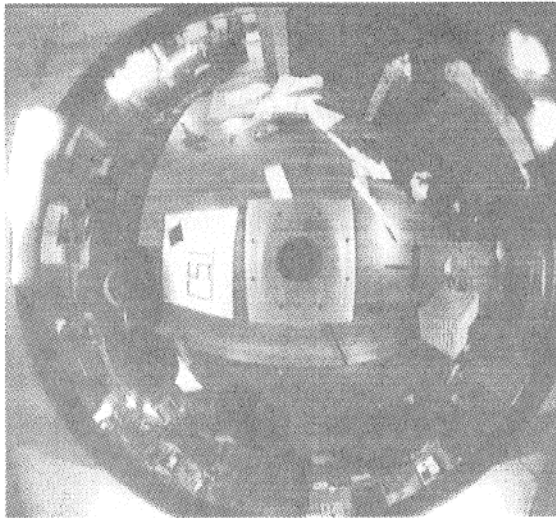


図 10 双曲面ミラーを用いた全方位視覚センサ

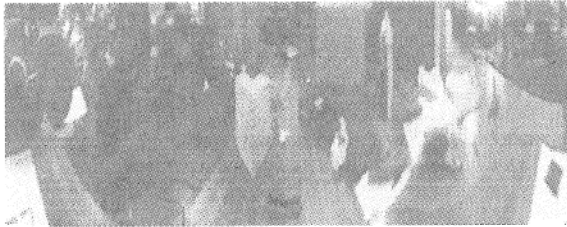
結ぶ直線は空間の点 $P(X, Y, Z)$ を通る。従って、写像点 $p(x, y)$ の位置から空間の点 $P(X, Y, Z)$ の方位角が一意に求められる。この性質から、任意の場所に任意の形状をした画像面を設定したときの画像を生成することができる。例えば、図10では双曲面ミラーの周りに一定の距離でリング状に画像面を設定した場合(パノラマ画像と呼ぶ)と一般のカメラと同様に平面的な画像面を設定した場合(透視変換画像と呼ぶ)を示している。図11にこの双曲面ミラーを使って得られた入力画像(a)、パノラマ画像(b)、透視変換画像(c)を示す。それぞれの画像は異なる利点を持っており、これらの画像の処理結果を総合することにより、ロボットの周囲の環境認識が容易となる。

### 5. 複 合 視

以上、移動ロボット用の大局視と局所視について述べてきたが、これらを機能的に組み合わせることにより、ロボットの周囲の環境を必要



(a)



(b)



(c)

図11 双曲面ミラーを使って得られた入力画像(a), パノラマ画像(b), 透視変換画像(c)

かつ十分な精度で能率良く認識できる。すなわち、目的地までの移動中は、ロボットの周囲360°を大局視で常に監視し、移動経路の計画とその更新、障害物や移動体との衝突回避を行なう。また同時に、道しるべ、作業対象、不審物、狭い通路などの発見を行ない、それらの候補が見つければ、局所視を起動する。大局視では、分解能が悪いためこれらの候補らしいというこ

とが分かる程度なので、局所視で詳細に解析し検証する。

これら2つの視覚を別々にロボットに搭載してもよいが、これらを一体化し軽量化を図ったほうが移動ロボットの視覚としては望ましい。このため、大局視と局所視のための光学系を合わせ持つ複合視覚センサシステムを考案した。この複合視覚センサでは、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。分かりにくいので、その構成図を図12に示す。円錐ミラーの中央をくり抜き、ここを通過して2つのミラーからの光が画像面中央に投影される。画像面周辺には円錐ミラーからの全方位画像が投影される。なお、ステレオ視用のミラーはモータで回転できるので、注視する場所を高速に変えることができる。

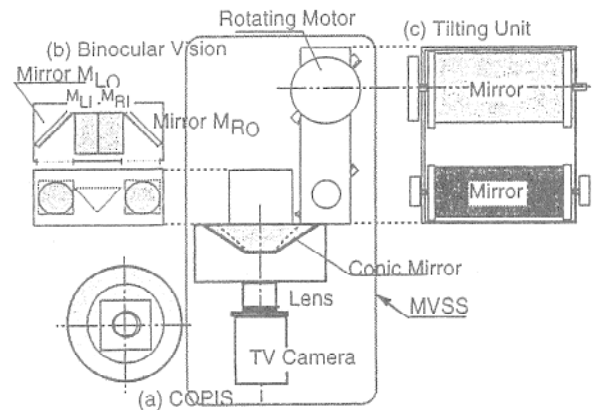


図12 複合視覚センサの構成図

## 6. おわりに

我々の研究室で行っている移動ロボットの視覚の研究の一端として、局所視、大局視、そして複合視について述べた。複合視を考えると、人間の目は中央付近の分解能が高く、周辺は広いところを見ているが分解能は低いということもヒントにした。しかし、最近、生物の本を調べていると、ハエトリグモは、ほぼ周囲360度を観察できる側方の4つの目と注目物体をじっくり観測するための前方の2つの目から成る複数の目を持っていることが分かった。我々の開発した複合視と考え方がよく似ており、やはり実世界で何千年を経過して進化してきた生物に見習う点が多いなと実感した次第です。