

マルチセンサ情報を動的に統合することによる 道路シーンの解釈とモデリング†

浅田 稔†† 白井 良明††

距離画像・明度画像などのセンサ情報を解釈し、それらを世界モデルに統合することは、知能移動ロボットの重要な課題のひとつである。多種のセンサ情報を用いたこれまでのシステムのほとんどは、センサレベルの処理に留まり、世界モデルの構築に至っていない。本論文は、複数のセンサ情報を利用した世界モデル構築システムの一手法を提案する。本システムは、自然環境を含む複雑な屋外シーンを解釈するために、ボトムアップ・トップダウン解析の組合せによるセンサ統合を実現する。このことにより、多種のセンサ情報やそれまでの解析結果から、その場の状況に応じて、信頼性の高い情報を動的に選び出し、解釈を効率的かつ柔軟に進めることができる。また、解釈と同時に、個々の物体の階層的な幾何モデリングを行い、相互に中間結果を参照する。このことにより、効率的にシーンの解釈を進められると共に、解釈に応じた個々の物体の幾何モデルが生成され、物体間の関係が明らかになる。本システムを道路シーンのミニチュアモデルに適用した結果を示す。

1. はじめに

知能移動ロボットの開発は、人工知能とロボティクスの分野の中心課題のひとつであり、広域的かつ精力的に行われている(文献1)に多くの論文が発表されている)。これらの研究では、実世界で移動ロボットが様々なタスクを遂行するための基礎技術の確立が目標になっており、視覚航行、障害物検出および回避などが行われてきた。これまでの研究では、実システムとして稼働させるために、環境を単純なものに限定しており、一般のシーンを対象としたときには、多くの困難な問題を抱えている^{2),3)}。特に未知の環境下では、ロボット自身がセンサ情報を利用して、タスク達成に必要な環境に対する地図(世界モデル)を作成する必要があり、屋外の自然環境を含む複雑なシーンでは、1種類のセンサだけでは、必要とする情報を得ることが困難である。この問題を解決する一つの方法は、複数のセンサ情報を利用することである。

これまでも多種のセンサ情報を用いたマルチセンサシステムが研究されてきた。そこでは、入力されたセンサ情報を統計的に処理することによって、より信頼性の高いセンサ情報を得ることを目標としていたり⁴⁾、各センサ情報を個別に処理し、その最終結果を統合する手法⁵⁾などが提案されてきた。Gotoら¹³⁾は、距

離画像とカラー画像を相補的に利用する手法として、カラー距離画像を提案し、階段の検出に応用した。

我々の研究目標は、移動ロボットが様々なタスクを遂行する上で必要な環境の地図(世界モデル)を構築することであり、本論文では、なるべく多くの情報を利用(マルチセンサ統合)して、信頼できるシーンの解釈を実現させると共に、それを反映した環境の幾何モデルを構築する手法を提案する。以下に、本システムの特徴を述べる。

自然環境を含んだ複雑なシーンの解釈には、ボトムアップ解析とトップダウン解析の動的な組合せによる柔軟な制御が従来、考えられてきた⁶⁾⁻⁸⁾。これらの手法の利点は、処理手順をあらかじめ固定しておくのではなく、その場の状況に従って動的に解析手順を制御できることである。そこで本システムでは、この手法をマルチセンサ統合に応用し、複雑な自然環境を含むシーンに対する高度で柔軟な解釈を可能にしている。例えば、道路領域を発見するには、明度情報と距離情報が利用できるが、距離情報には、道路面に投影された影によるエッジなどが含まれないため、道路の候補領域を抽出するには、明度情報よりも距離情報の方が有利である。いったん、道路の候補領域が発見されれば、道路と路肩等の道路の外側の領域を区別するには、明度情報が有効である。このように、多種のセンサ情報やそれまでの解釈結果から得られる拘束条件の中から、その場の状況に応じて、最も信頼性の高い情報を選び出し、確実に、効率のよい解釈を実現させている。

柔軟な制御を用いたこれまでのシーンの解釈で

† Road Scene Interpretation and Modeling Based on Dynamic Integration of Multi-Sensory Data by MINORU ASADA and YOSHIKI SHIRAI (Department of Mechanical Engineering for Computer-Controlled Machinery, Faculty of Engineering, Osaka University).

†† 大阪大学工学部電子制御機械工学科

は^{7),8)}、2次元画像上の領域のラベルづけが最終出力であった。それらの研究では、対象世界が3次元の屋外シーンにもかかわらず、知識としては、2次元の領域に関する記述を用いており、3次元的な幾何学的推論は困難であった。自然環境を含む複雑な屋外シーンで、移動ロボットがタスクを遂行するには、シーン中の個々の物体を同定し、それらの間の幾何学的関係を推論する必要がある。本システムでは、これらを実現させるために、マルチセンサ統合を用いたシーンの解釈と同時に、個々の物体の幾何モデリングを行い、相互に中間結果を参照させる。このことにより、効率的にシーンの解釈を進めると共に、解釈に応じた個々の物体の幾何モデルを生成し、物体間の関係を明らかにする。

本論文では、道路シーンを対象として、明度画像、距離画像のマルチセンサ統合によるシーンの解釈および環境のモデリングを示す。以下、システムの前提条件と概要、センサ情報の解析、知識表現と道路シーンの解釈、幾何モデリング、解釈結果を述べ、最後に、考察と今後の課題を示す。

2. システムの前提条件および概要

2.1 システムの前提条件

本システムでは、道路シーンを対象とするが、以下の事柄を前提条件としている。

(1) 道路は、ほぼ平坦な地面上に存在するとし、クロスカントリーなどの急激な地面の凹凸はないものとする。この条件により、シーンは、大きな水平面とその上にある物体によって構成されていると考えることができる。

(2) 道路上やその近傍にある物体としては、人工物として自動車、建物、道路標識、郵便ポスト、自然物として高木、かん木を対象とし、それらに関する一般的な知識を前もって与える。これら以外にも、人間、自転車など数多くのものが道路シーンでは考えられるが、ここでは、それらについては、解釈せず(知識として与えず)、未知物体として取り扱う。

(3) センサ情報として、ある時刻に得られたTVカメラからの白黒濃淡画像とレンジファインダから得られる距離画像を扱う。TVカメラとレンジファイ

ンダのそれぞれの幾何パラメータや相互の幾何学的位置関係も既知とする。

2.2 システムの概要

図1に、システムの概観を示す。最初にセンサ情報(明度画像および距離画像)がボトムアップ制御により解析される。距離画像は、レンジファインダの幾何パラメータ(開口角度や移動ロボットに対する位置、向き)を用いて、高さ地図に変換される。高さ地図は、高さ情報を利用して、移動可能領域と障害物領域に分割され、それぞれの領域の属性を表す領域属性リストが作成される。

次に、マルチセンサ統合を動的に用いるシーンの解釈が、ボトムアップの解析結果に適用される。道路シーンに関する知識は、シーンに存在すると予想される物体と、それらの物体間の関係を表すフレーム構造を用いて表現され、それらに付随した手続きは、現在の地図の内容に従って動的に制御される。制御戦略の基本原理は、各段階で、最も信頼性の高い情報を利用して、以降の処理を効率的に進める拘束条件を伝播させることである。ここでは、領域属性リストを利用して、各段階で解釈に必要とされる物体の属性や拘束条件を発見し、それらの拘束を伝播させる。

個々の物体に関する階層的な幾何学的なモデリングは、シーンの解釈と同時に終わる。それぞれの物体

解釈結果を表すネットワーク

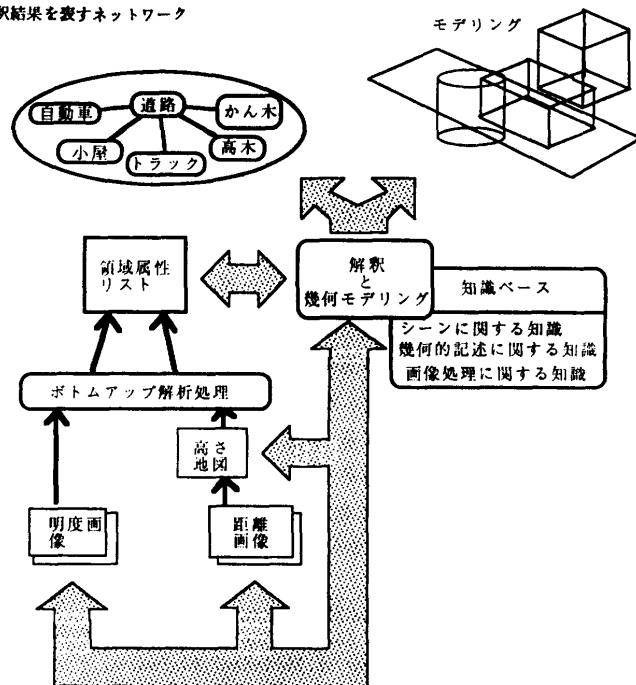


図1 システムの概要

Fig. 1 An overview of the system.

モデルは、属性リスト、他の物体モデルとの幾何学的関係、自身のモデルの幾何パラメータからなる。人工物に関しては、多面体モデルを、自然物に関しては一般円筒モデルを利用する。多面体モデルのパラメータは、解釈の段階に応じて高精度化され、具体化されたフレームが選ばれる（例えば、障害物→人工物→自動車→乗用車）。これに対し、自然物のモデリングは、本論文では、おおまかな位置と存在範囲を記述するにとどめる。これは、自然物に対するモデリング自身が非常に大きな問題であり、移動ロボットにとって自然物の詳細なモデリングが重要な意味を持つことが一般的にそんなに多くないと考えているからである。

このようにして、明度情報と距離情報を利用したシーンの解釈と幾何モデリングが進められる。

3. センサ情報の解析

3.1 センサ情報入力システム

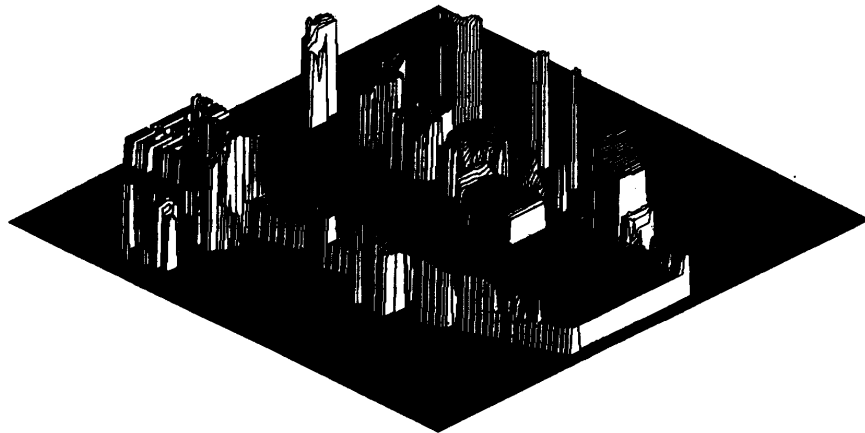
本論文では、実際の移動ロボットから屋外シーンを撮像したデータではなく、ロボットアームを利用し



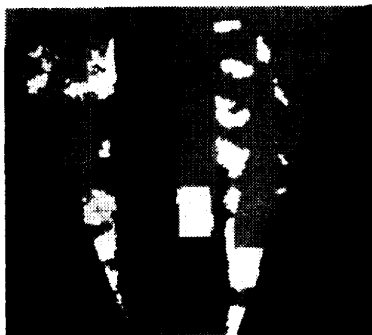
(a) 明度画像 (b) 距離画像
(a) Intensity image. (b) Range image.

図 2 センサ地図の例

Fig. 2 Examples of sensory data.



(a) 鳥瞰図
(a) Perspective view.



(b) 分割結果
(b) Segmented height map (top view).

図 3 高さ地図

Fig. 3 Height map.

た、物理的な移動ロボットシミュレータを用いて、明度画像、距離画像(図2参照)を得た。距離情報は、光切断法によるレンジファインダによって得ている¹⁰⁾。距離画像では、明るい点が観測者から近く、暗い点が遠い。黒い領域は、距離情報が得られなかった点で、隠ぺいや不十分な反射によって生じる。なお、シミュレータのスケール比は1:87である。

3.2 高さ地図の解析

3.1 節のシステムを利用して得られたビデオカメラからの明度画像とレンジファインダからの距離画像を、ここでは、センサ情報として扱う。センサ情報の解析に関しては、文献9)に詳細に述べているので、以下では、簡単に触れる。距離画像(図2(b))は、移動ロボットが存在すると想定される地面からの高さを表す高さ地図に変換される(図3参照)。高さ地図は、高さ情報を利用して、4種類の領域に分割される。すなわち、移動可能領域(図3(b)の暗い灰色)、障害物領域(同、白色)、未探索領域(同、黒色)、隠

ぺい領域(同、明るい灰色)である。移動可能領域は、想定される地面に近い点を集めて原始領域とし、その近傍で、傾斜や曲率がしきい値以下の点を併合することにより、得られる。高さ情報がある点で、移動可能領域以外が、障害物領域となる。

高さ地図の分割結果は、領域属性リストに変換される。領域属性リストには、ラベル(移動可能領域か、障害物領域)や幾何学的な属性(位置、高さ地図上に投影された面積、高さの平均、分散、および傾斜、曲率の平均、分散)などが、以降の解釈プロセスのために記述されている。図4に、高さ地図のボトムアップ解析により得られた領域属性リストの一部を示す。表の部分では、最大面積を持つ番号0の領域が移動可能領域を、その他が障害物領域を示す。ある領域は、複数の物体からなり、それらの境界部分で高い傾斜や曲率を持つので、それらの値を利用して、それぞれに細分割される。例えば、領域18は、小屋の屋根の部分、小屋の手前にある木箱と残りの部分にそれぞれ分割される。

#	area (pix.)	mean height	slope		curvature	
			mean	var.	mean	var.
0	11352	0.48	0.47	0.46	0.54	0.53
1	75	32.49	14.82	17.55	6.53	12.23
2	420	21.97	24.15	30.40	21.56	37.19
11	34	13.68	4.21	3.39	3.02	1.70
14	672	15.41	2.66	5.10	1.08	1.39
15	307	18.82	6.58	4.52	1.37	2.25
18	535	29.70	14.26	25.64	10.35	26.71
18-1	465	36.86	7.23	2.56	1.03	0.57
18-2	42	17.56	2.78	3.24	2.58	1.30
18-3	38	5.68	3.56	4.32	5.23	3.48
20	170	15.72	6.27	5.20	1.24	1.18

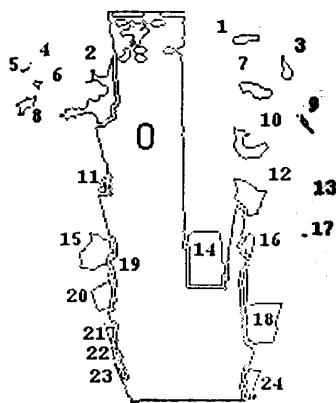


図4 領域属性リスト
Fig. 4 A region property list.

4. 知識表現と幾何モデリング

4.1 道路シーンに関する知識の表現

移動可能領域や障害物領域を道路、トラック、小屋、高木、かん木などのより具体的な物体として同定するためには、道路シーンに関する知識が必要である。ここでは、この知識を、シーンに存在すると思われる物体モデルのフレーム構造のネットワークとして表現している(図5にその一部を示す)。各フレームには、物体モデル自身や他の物体モデルとの関係に関する幾何的な拘束条件が記述されている。

図6に示すのは、道路、自動車フレームの例である。各スロットには、対応する属性とそれらをセンサ情報から求める手続きが含まれている。例えば、道路フレームでは、道路境界の持つ属性(平行で滑らかなカーブ)と共に、それを抽出するための、画像処理手

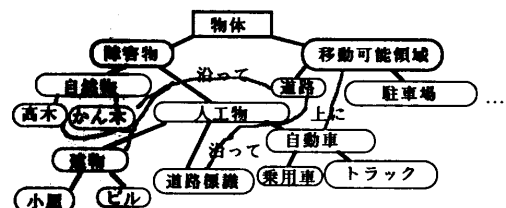


図5 道路シーンに現れる物体間の関係を表すネットワーク
Fig. 5 A network representing relations between objects expected in road scenes.

道路フレーム			
上位クラス	移動可能領域		
属性	一定幅 滑らかなカーブ		
サイズ	幅	最大幅	最小幅
	曲率	最大曲率	最小曲率
他の物体との関係			
	道路上	自動車	
	道路に沿って	かん木	高木 建物 自動車

自動車フレーム			
属性	移動可能		
上位クラス	人工物		
サブクラス	トラック		
構成物	エンジン室	キャビン	トランクまたは荷台
サイズ	幅	最大幅	
	長さ	最大	最小
	高さ	最高	最低
他の物体との関係			
	～の上で移動	移動可能領域	
	～の前で停止	停止サイン 障害物	

図 6 道路シーンに現れる物体に関する知識のフレーム表現
Fig. 6 Knowledge representation about objects in road scenes using frame representation.

続き、およびその評価ルーチンが含まれている。
評価ルーチンは、それぞれの処理内容に従って、三つのレベルの評価を下す。すなわち、ア) 信頼よく、対応する属性が検出できた場合、イ) 検出できなかったことが明確な場合、ウ) センサ情報の精度が低かったり、対応するセンサ情報が存在しなかったために明確な情報が得られない場合の3種類である。これらの評価方法は、対象とする属性やセンサ情報によってあらかじめ決定することが可能で、事前にシステムに与

えられている。
サイズスロットの値は、属性の範囲を示す値（例えば、最小幅、最大幅、最大曲率など）で表現されている。自動車フレームの構成部（エンジン室、運転室、荷室（トランクまたは荷台））は、それぞれの部分に対応するフレームへのポイントを持ち、それぞれが直方体モデルで近似されている。それらのサイズの違いにより、サブクラスの「乗用車」と「トラック」に分けられる。構成物のサイズとそれらの幾何的關係を表すために、幾何モデルを各フレームが持つ（4.2節、および図7参照）。

他の物体との関係（「近くに」、「沿って」、「の上に」など）は、それらの記述と共に、高さ地図上でそれらを確認するためのルーチンへのポイントを持つ。これらは、以下のように定義される。
「近くに」→高さ地図上で、二つの領域（高さは無視）の距離をそれらの境界点間距離の最短距離で定義し、それが、小さい方の面積（高さ地図上での投影面積）の平方根よりも小さいとき、それらが互いに「近くに」あるとする。

「沿って」→高さ地図上の道路領域外で、物体が道路領域の「近くに」存在するとき、その物体は、道路に「沿って」存在する。（道路領域と他の物体との関係を表現する語として定義）
「の上に」→ある領域の境界のほとんど（80%以上）が別の領域に囲まれ、その重心位置が、囲んでいる領域の重心位置よりも高いとき、その領域は、囲んでいる領域「の上に」存在すると定義される。

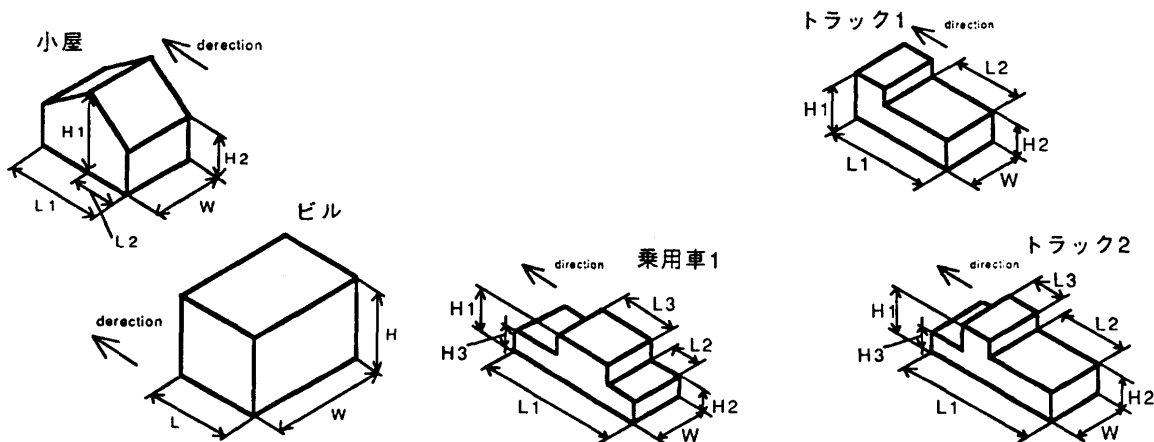


図 7 人工物に関する幾何モデル
Fig. 7 Geometric models for artificial objects.

4.2 道路シーンの幾何モデリング

(1) 人工物のための多面体モデリング

人工物の多くは、平面で構成されている場合が多いので、ここでは人工物に対して、多面体モデリングを行う。物体の幾何モデリングの目的は、シーンの解釈を助ける幾何学的推論のために、物体の位置、大きさ、形状を与えることである。シーンの解釈の中間結果を反映させるために、システムは物体の解釈の具体化に従って現在、二つのレベルの多面体モデルを用意している。

最初のレベルでは、直方体モデルが物体の大きな位置と大きさを表すのに使われる。直方体モデルは、4枚の垂直面と1枚の水平面からなり、それらすべては物体のある部分と接している。換言すれば、このモデルは、高さ地図上で物体を囲む外接直方体になっている。直方体の向きは、高さ地図の座標軸に平行になるように取っている。

次のレベルは、すべての面が高さ地図上の物体のエッジや面に接している多面体モデルである。ここでは、人工物のモデルとして、自動車、建物があり、それぞれサブクラスとして「乗用車」、「トラック」、および「ビル」、「小屋」のモデルを用意している。図7にそれらを示す。前節で述べたように、それらは、サイズスロットに示されている範囲で、可変である。「乗用車」と「トラック」の区別は、それぞれの構成部分のサイズの範囲で識別可能である。また、「ビル」と「小屋」は、屋根の傾斜から区別される。これらのモデルは、想定される地面の上に存在すると仮定されるので、照合の際の位置と向きの自由度は、平面上の平行移動と回転の三つである。

(2) 一般円筒を用いた自然物に対する幾何モデリング

自然物に対する幾何モデリングの目的は、物体の大きな位置と大きさを表すことである。人工物と異なり、自然物に対して正確な幾何パラメータを求めることは困難であると同時に、移動ロボットにとって重要な意味を持つ場合は少ない。本システムでは、一般円筒モデルを自然物に対して適用する。一般円筒モデルは、自由度が多すぎるので、ここでは、自然物のラベルの時点では、直円筒を、高木、かん木のラベルが確定した時点では、回転体モデルを採用する。かん木や高木などの自然物は、通常幹に基づいて成長し、幹の方向は重力の影響により垂直である。そこで、垂直方向を直円筒や、回転体の軸の方向と一致させ、以下の

ように定める。

(a) 直円筒近似：自然物とラベル付けされた領域に対して、高さ地図上でその領域を囲む最小円を決定し、その領域の最高値を高さとする直円筒で近似する。

(b) 回転体近似：高木、かん木のラベルが決定した段階で、円筒の底面に平行にある間隔でスライスし、各スライス内でそれぞれ、もう一度、円近似し、回転体の各高さにおける半径を決定する。

上記の過程において、細長い領域は、近似円筒内もしくは回転体内で、多くの自由空間を含む可能性が高い。これは、障害物回避にとって、好ましいことではない。この領域の近傍を通過するとき、衝突の危険率が高ければ、その領域を数個に分離し、それぞれについて近似円筒を設けることにより、安全な経路を計画することも考えられる。

5. 道路シーンの解釈

5.1 制御戦略

解釈の制御戦略における基本原理は、各物体に対するラベルづけを、その時点で利用できる最も信頼性の高い情報を用いて具体化することである。情報の信頼性の評価基準は、以下の手続き(1)~(5)で示すように、

(a) 解釈過程の中で、より具体化された解釈ラベル(図5においてなるべく下位)を持つものからの情報が信頼性が高いとしている(定性的評価)。

(b) 同じレベルの場合には、観測者からの距離が小さいものからの情報の精度が高いとして優先させる(定量的評価)。

初期状態は、移動可能領域と障害物領域のラベルを持つ領域属性リストであり、目標は、このリストの各領域に対するラベルをより具体化することで、以下にその方法を述べる。ただし、対象となる領域は、小領域を取り除くために、高さ地図上である一定面積以上の領域であり、かつ、観測者から余り離れていない領域とする。

(1) 道路シーンの解釈においては、道路領域が解釈の核となるので、道路領域が最初の対象となり、その抽出(解釈)を試みる。

(2) 対象となる領域がなくなるまで(3)から(5)を繰り返す。

(3) 解釈が成功した場合、その領域の解釈から派生する拘束条件を伝播させ、それによって解釈可能と

期待できる対象領域を選択する。そのとき、解釈された領域に近い領域を選ぶ。候補が複数存在したとき、観測者に近いものを優先させる。

(4) 解釈が失敗すれば、その失敗から派生する拘束条件を伝播させ、それによって解釈可能と期待できる対象領域を選択する。有効な拘束条件が派生しないときには、任意の領域に対して、解釈を進める。そのときの選択基準として、観測者の近傍にある領域から、解釈を始める。

(5) 対象領域が、確定した時点で、その解釈の候補となる物体の属性が検出可能か、また、その物体を構成する部分が決定できるかを調べ、解釈可能かどうかを判定する。

(5-1) 抽出すべき属性もしくは、構成部分を決定するための情報が、3次元幾何情報である場合には、距離画像、または高さ地図に対して、特徴抽出ルーチンを適用し、評価する。

(5-2) 抽出すべき属性もしくは、構成部分を決定するための情報が、明度情報によって、抽出される特徴に対しては、明度画像に対して、特徴抽出ルーチンを適用し、評価する。

(5-3) 評価の結果、信頼できる属性もしくは、構成部分が、一つしかない場合には、それに対応する物体に解釈できたとする。複数存在し、それらが異なる物体に対応するとき、矛盾があるとして、解釈は、失敗したと見なす。信頼できる属性が、確認できなかったときにも、解釈は失敗したと見なす。

以下では、道路領域の抽出を出発点として解釈を進めた過程を示す。

5.2 道路領域の抽出

道路フレームの記述より、道路領域は、移動可能領域内に存在することは分かるが、移動可能領域フレームには道路境界に対する記述はない。道路境界に対する記述は、道路フレームに含まれており、道路境界抽出の画像処理ルーチンやその評価法も道路フレームに含まれている。

そこで、移動可能領域のインスタンスフレームの情報を利用するために、移動可能領域を明度画像上に射影し(図8(d)参照)、その領域内で道路境界の抽出を試みる。移動可能領域に対応する明度画像の部分にエッジ抽出オペレータを適用する。図8(a)と(b)は、Sobel オペレータによるエッジ点の大きさとその方向を表している。画像処理評価ルーチンは、明度画像内での移動可能領域でのエッジ密度と、エッジの方

向の分散に関する統計的解析から、道路境界のエッジを連結させていくことが困難であることを知る。すなわち、道路面上のテキスチャパターンにより、道路境界と同程度の強さのエッジが、あらゆる方向に向いており、道路境界の出発点が与えられたとしても、道路境界を正しく結び付けていくことは困難である。

そこで画像処理評価ルーチンは、滑らかで局所的に直線近似できると仮定される道路境界を発見するために、Hough 変換を選択する。Hough 変換は、カーブの場合には、Hough 空間上で連なったクラスタを形成すると考えられる。影やテキスチャの影響を受けるが、既に、候補が絞られているので、クラスタを識別する際には、あまり大きな影響は及ばないと仮定している。もし、影響を及ぼすようであれば、人間でも識別が困難だと考えられる。画面上の各エッジ点は、その位置と方向を持ち、Hough 空間で一点に射影されるので、計算時間は通常の Hough 変換に比べ短い。図8(c)と(d)に、Hough 空間と候補線をそれぞれ示す。Hough 空間内での二つのピークが道路境界に相当している。これらのガイドラインを利用して、道路境界上のエッジ点が結び付けられていく。連結方法は、求めたガイドラインの近傍で、ガイドラインとはほぼ同一方向の最も強いエッジを出発点とし、その両端から、エッジ強度と方向について許容範囲のものを連結していく。

連結された道路境界は、道路領域に対する平面近似を行うために、高さ地図上に投影される(図8(e)参照)。近似平面のパラメータと境界点の位置および道路幅、向きなどが、道路フレームのインスタンスフレームに記述される。

図9に、この時点での、環境のモデリングを示す。道路領域は、平板で表現されている。また、物体領域は、人工物、自然物の仮ラベルに従い、それぞれ、直方体、直円筒で近似されている。仮ラベルは、高さ地図上で各領域の曲率の分散に依存し、しきい値以上であれば、自然物、未満であれば、人工物としている。

5.3 道路領域周辺の物体の解釈

道路の抽出が成功すれば、移動可能領域のインスタンスフレームは、道路領域のインスタンスフレームと非道路領域のインスタンスフレームに分割かつ具体化される。次に、道路領域のインスタンスフレームに隣接している障害物領域のインスタンスフレームに対する解釈を道路からの拘束条件を利用し行う。

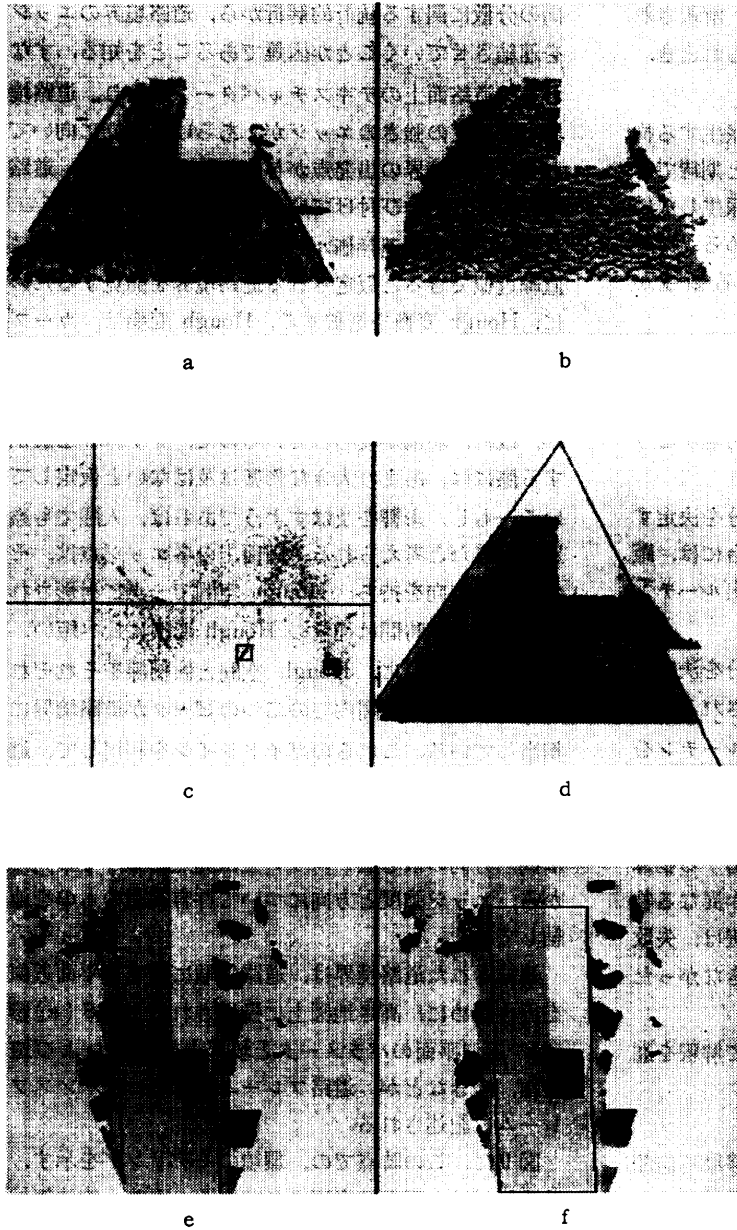


図 8 道路境界の抽出：(a) 道路候補領域のエッジ強度画像，(b) 道路候補領域のエッジ方向画像（明暗が方向の違いを示す），(c) Hough空間（図8(a, b)から射影），(d) 道路候補領域と検出された道路境界候補（実線），(e) 高さ地図上に射影された道路境界，(f) 道路幅一定の知識を利用して推論された道路領域

Fig. 8 Road boundary detection: (a) edge strength image of a candidate for road region, (b) edge direction image of a candidate for road region, (c) Hough space mapped from the edge points in (a) and (b), (d) candidate for road region and candidate lines for road boundaries obtained from Hough space, (e) detected road boundaries mapped on the height map, (f) inferred road region using the knowledge of constant width of the road.

(1) 道路上の物体の解釈

最初の候補は、道路領域と接している物体（図4の領域14）である。道路は幅一定という属性を持っているので、この物体は道路に接しているのではなく、道路上にあると仮定される。すなわち、高さ地図上では、左の道路境界に対する右の境界の一部は、この物体によって隠されて観測されていないが、右の境界が存在する部分に、他のものが存在することが明らかでない限り、この部分に境界が存在すると仮定される（図8(f)参照）。

道路フレームの「他の物体との関係」スロットより、道路上の物体は、現システムでは、自動車のみが解釈候補となり、自動車フレームと照合される。自動車フレームでは、サイズスロットの許容範囲から自動車であると確認する。自動車のサブクラスとしては、「乗用車」と「トラック」があり、それらの区別は、構成部分のサイズによって可能である。そこで、この領域に対して平面近似ルーチンを起動して、その大きさと面の方程式を求め、「トラック」または「乗用車」の仮説をたて、それに対応するフレームを呼び出して検証する。物体の全高は高さ地図から決定できるが、全長もしくは全幅は、観測できない部分があるので決定できない。この物体の平面領域は、面積と高さから、トラックの荷台と仮定され、トラックの仮説が立てられた。

向きに関しては、全幅の許容範囲と道路の進行方向から、縦方向（前向きもしくは後向き）であると判断される。それぞれについて、検証すると、前向きの場合、運転室の部分（荷台より高い）が高さ地図上で発見されるべきであるが、確認できない。後向きの場合、完全には、運転室が確認されないが、矛盾は生じない。よって、後向きと判断される。

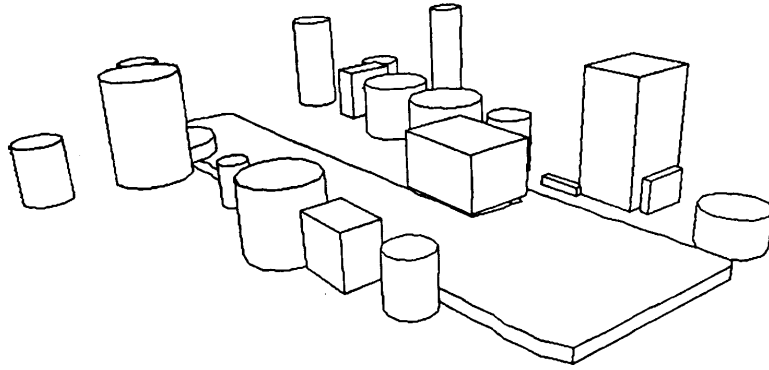


図 9 幾何モデリング：直方体と直円筒による第1次近似
 Fig. 9 Geometric modeling of the environment with solids and right cylinders.

この過程の幾何モデリングを図 10 に示す。人工物の仮ラベルの時点では、直方体であるが、平面近似から、トラックの荷台部分に相当する部分が検出され、その方向についても限定されるので、図に示すような、モデルとの照合が可能である。システムでは、トラックモデルとして、エンジンルームが独立してあるタイプと、縮退して運転室と融合しているタイプを持っており（図 7 参照）、この場合、運転室やエンジンルームは確認できないので、いずれとも照合可能であるが、図では、便宜的に後者の場合を示している。ここで、注意しなければならないのは、トラックの荷台の部分に関しては、高さ地図から、確認できたが、エンジンルームや、運転室に関しては、確認されておらず、それらの存在と矛盾するデータが存在しないために、この時点では、トラックと判断されていることである。

(2) 道路の傍にある物体の解釈

次の候補は、道路の右側にある物体（図 4 で領域番号 24 と 18-1）と左側にある物体（同 20）である。道路の傍にある物体の候補としては、道路フレームのスロットより、自然物としてはかん木、高木、人工物としては駐車している車、建物が考えられる。左側にある物体は、実際はかん木であるにもかかわらず、曲率の分散が小さい（図 4 参照）ので、最初、人工物と仮定された。しかしながら、この物体に対応する明度画像上での境界線の形状は多くの曲線部を含み（図 12(a) 参照）、また平面近似したときの誤差が大きく、複数の平面にも分割できず、自動車フレーム、建物フレームとの照合が失敗する。そこで自然物フレームと照合され、傾斜の分散が大きいくことなどから、自

然物と仮定が修正される。さらに、高さの最大値からかん木と解釈される。この修正により、影響を受ける物体は、この場合、存在しないが、修正以前のラベルからの拘束条件を利用して、既に解釈がなされたものが存在する場合には、それらの解釈をやり直す。

右側の手前の物体は、実際は乗用車であるが、観測されている部分が非常に小さく、また誤差により、高さの分散が小さくなり、自然物とラベル付けされており、高さからかん木と判断された。

小屋の屋根の一部は、曲率の分散が小さく（図 4 参照）、人工物と仮定されている。道路の近くにある人工物としては、建物や駐車している車などが考えられる。高さから建物と仮定される。建物フレームでは、そのサブクラスとして、「ビル」と「小屋」が存在し、それらの区別は、屋根の傾斜によって行っている。すなわち、水平な屋根を持つものが「ビル」であり、傾斜した屋根を持つものが「小屋」である。そこで、この物体に対する平面近似による傾斜の値から、小屋と仮定され小屋フレームと照合される。照合は成功し、その平面方程式による高さ、位置、傾き、境界点などのパラメータが、小屋のインスタンスフレームの屋根の部分に記述される。また、これらの値を利用して、

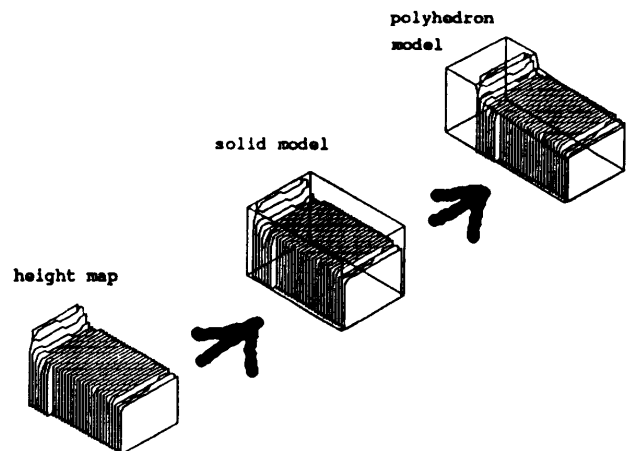


図 10 領域 14 に対する解釈過程に応じた物体のモデリング：トラックモデルとの照合

Fig. 10 Geometric modeling for region 14 according to the interpretation process : matching with a truck model.

小屋の幾何モデルが決定される。

この過程の幾何モデリングを図 11 に示す。モデルの記述では、屋根の大きさや傾斜については、不等式で表現されているだけであり、特定のサイズの小屋を持っているわけではない。よって、センサから得られた情報を利用して、それらの値を決定している。

5.4 解釈結果

図 12 に解釈結果を示す。(a)は解釈結果のラベル付けを、(b)は(a)に対応する環境のモデリングを、(c)はシーンの解釈結果を物体間の関係を表すネットワークとして表現している。太線で囲まれたノードは、具体化が最後まで進んだラベルを示す。細線で囲まれたノードは、遠方のため対象領域から除外されている領域や、信頼できる情報がないために、解釈が未定の領域を示す。それぞれのノードには、それらに対応するインスタンスフレームへのポイントを持つ。インスタンスフレームには、高さ地図や明度画像、距離画像の解析結果から得られている属性(大きさや、明るさ)が各スロットに記述されている。太線のアークはノード間の関係を示し、ラベルの確定に利用されたことを示す。

なお、プログラムは、Sun 4 上の C 言語と EUSLISP¹⁴⁾ を用いた。

6. 考察および今後の課題

本論文では、移動ロボットのための環境モデル構築を目的とした明度画像と距離画像のマルチセンサ統合

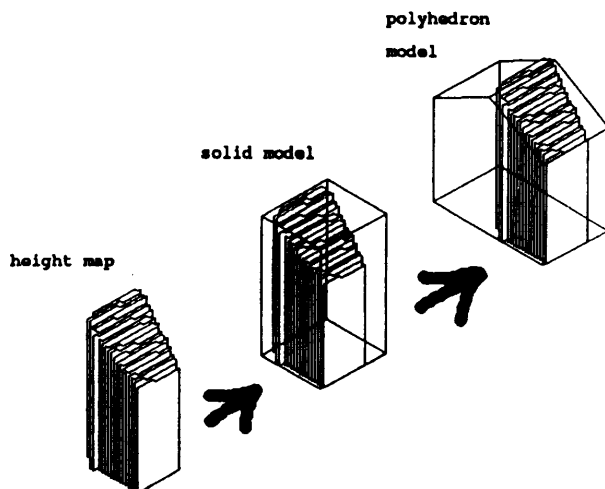


図 11 領域 18 に対する解釈過程に応じた物体のモデリング: 小屋モデルとの照合

Fig. 11 Geometric modeling for region 18 according to the interpretation process: matching with a cabin model.

によるシーンの解釈および幾何モデリングについて述べた。多種センサ情報の統合利用により、道路境界の抽出にみられたように、探索領域が限定され、効率的かつ正しくシーンの解釈が行えた。またトラックの同定プロセスで示したように、解釈の途中結果が幾何モデリングに反映され、高精度化された幾何パラメータがシーンの解釈を効率的かつ信頼性の高いものにしていく。

解釈の制御戦略では、道路領域を抽出することが、道路シーンを解釈する上で、重要と考え、最初の候補領域としたが、以下のように一般化することも考えられる。

(1) 図 5, 6 に示す道路シーンに関する知識のフレーム表現より、道路領域が他の物体(自動車、かん木、建物など)への拘束条件を伝える元の領域になっている(「他の物体との関係」スロットの物体の数が一番多い)。また既に得られている領域(移動可能領域)から、その物体のラベルまでの経路(ネットワークのアークの数)が最も短い。

(2) 領域属性リストを利用して各領域を調べると、一つの移動可能領域が最大面積を占め、かつ、傾斜、曲率の分散が共に、最も小さく、最も安定した平面領域として得られていることが分かる。

これらのことから、このシーンの解釈を信頼性良くかつ効率的に進めるには、道路領域を最初に抽出することであると決定できる。もし、他の物体、例えば建物などがセンサ情報の中で大きく捉えられ、最も安定した特徴として得られれば、その物体から解釈を進めることもできる。すなわち、最初にどの領域を出発点とするかは、シーンに関する知識と、実際のセンサ情報に依存していると考えられる。

今後の課題としては、以下のものが挙げられる。

(1) シーンの解釈に対する信頼度の表現は、多種多視点の観測データを統合する上で重要な問題である。物理的な次元が同じ場合、確率を用いて、表現、更新している例がある¹²⁾、物体のより高度な解釈に関してその信頼度を明確に定義することは難しい。本システムでも、各段階で最も信頼性のある特徴を選択するが、その方法は、特徴を求めたセンサ情報の性質(明度情報もしくは距離情報)と、その特徴の属性から事前に定義している3段階の評価を利用し、判断している。ま

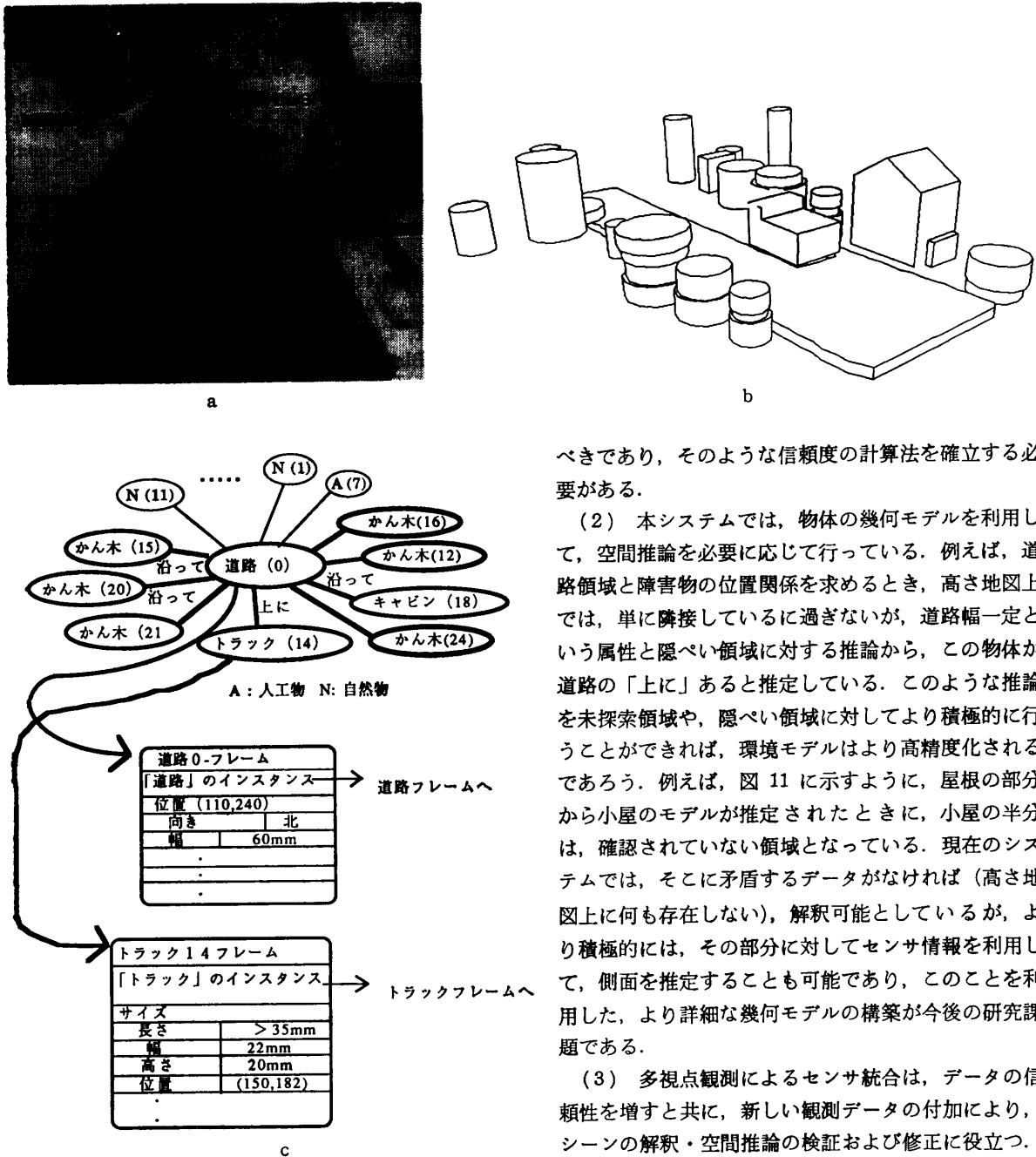


図 12 解釈結果: (a) ラベル付け, (b) (a)に対応する環境の幾何モデリング, (c) 解釈結果を表すネットワーク
 Fig. 12 Result of interpretation: (a) labeling, (b) geometric model representing the interpretation, (c) a network representing the interpretation results.

た、観測者からの距離の小さい方が精度が高いとして、遠方の特徴より近方の特徴を選択している。より一般的には、各特徴とその評価ルーチンに従って信頼度を計算し、その値に従って選択すべき特徴を決定す

べきであり、そのような信頼度の計算法を確立する必要がある。

(2) 本システムでは、物体の幾何モデルを利用して、空間推論を必要に応じて行っている。例えば、道路領域と障害物の位置関係を求めるとき、高さ地図上では、単に隣接しているに過ぎないが、道路幅一定という属性と隠ぺい領域に対する推論から、この物体が道路の「上に」と推定している。このような推論を未探索領域や、隠ぺい領域に対してより積極的に行うことができれば、環境モデルはより高精度化されるであろう。例えば、図 11 に示すように、屋根の部分から小屋のモデルが推定されたときに、小屋の半分は、確認されていない領域となっている。現在のシステムでは、そこに矛盾するデータがなければ（高さ地図上に何も存在しない）、解釈可能としているが、より積極的には、その部分に対してセンサ情報を利用して、側面を推定することも可能であり、このことを利用した、より詳細な幾何モデルの構築が今後の研究課題である。

(3) 多視点観測によるセンサ統合は、データの信頼性を増すと共に、新しい観測データの付加により、シーンの解釈・空間推論の検証および修正に役立つ。

特に、自動車などの移動物体は、画面間での対応づけにより、自動車であることの確認と、その運動パラメータが求められる。著者らは、領域属性リストを用いて、効率的に2枚の高さ地図間の物体の対応づけを行い、動物体を追跡する手法について既に検討した¹¹⁾。ここでは、「高さ」を物理的に共通な次元として統合しており、本論文で提案しているような高度なシーンの解釈結果は利用していない。多視点での観測データも含めて、本システムを拡張し、より高度で柔

軟なダイナミックシーンの解釈を行うことが考えられる。

謝辞 レンジファインダを作成し、本研究で使用しているデータの取得を可能にしてくれた米国メリーランド大学のダニエル・ディメンソン氏に、また本プログラムの一部作成に協力してくれた大阪大学大学院工学研究科修士課程在学中の門野啓一郎君、谷口恭弘君にならびに有益なコメントを頂いた査読者に感謝する(なお、第一著者の次男、浅田龍は、本研究で使った道路シーンの多くのパーツを塗装し組み立ててくれた)。

参 考 文 献

- 1) *Proceedings of the DARPA Image Understanding Workshop*, Vol. I and II (1987).
- 2) 辻: 知識処理とロボティクス, 人工知能学会誌, Vol. 1, No. 1, pp. 30-37 (1986).
- 3) Hermon, S. Y.: A Report on the NATO Workshop on Mobile Robot Implementation, *Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 604-610 (1988).
- 4) Ayache, N. and Faugeras, O.: Building, Registering, and Fusing Noisy Visual Maps, *Proc. of 1st Int. Conf. on Computer Vision*, pp. 73-82 (1987).
- 5) Thorpe, C., Hebert, M. H., Kanade, T. and Shafer, S. A.: Vision and Navigation for the Carnegie-Mellon Navlab, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 10, pp. 362-373 (1988).
- 6) 白井: シーンに関する知識を用いて明るさの配列を解析する, 白井, 杉原(共訳), コンピュータビジョンの心理, pp. 101-124, 産業図書 (1979).
- 7) Ohta, Y.: Knowledge-Based Interpretation of Outdoor Natural Color Scenes, *Research Note in Artificial Intelligence 4*, Pitman Advanced Publishing Program (1985).
- 8) 松山, 長尾: 航空写真の構造解析, 情報処理, Vol. 21, No. 5, pp. 468-480 (1980).
- 9) 浅田: センサ情報の統合と理解による移動ロボットのための世界モデルの構築, 日本ロボット学会誌, Vol. 8, pp. 160-170 (1990).
- 10) DeMenthon, D., Siddalingaiah, T. and Davis, L. S.: Production of Dense Range Images with the CVL Light-Stripe Range Scanner, Center for Automation Research Technical Report, CAR-TR-337, University of Maryland (1987).
- 11) Asada, M., Ikeda, E. and Shirai, Y.: Interpretation and Integration of Height Maps from a Range Image Sequence, *Proc. of IEEE Workshop on Intelligent Robot and Systems* pp. 54-61 (1989).
- 12) Matthis, L. and Elfes, A.: Integration of Sonar and Stereo Range Data, *Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 727-732 (1988).
- 13) Goto, Y., Matsuzaki, K., Kwen, I. and Obatake, T.: CMU Sidewalk Navigation System: A Blackboard-Based Outdoor Navigation System Using Sensor Fusion with Colored-Range Images, *Proc. Fall Joint Computer Conference*, pp. 105-113 (1986).
- 14) Matsui, T.: An Object Based Robot Programming System EUSLISP, Electrotechnical Laboratory Memo, ETL-RM-87-06 E (1987).

(平成元年8月31日受付)
(平成2年9月11日採録)



浅田 稔 (正会員)

昭和28年生。昭和52年大阪大学基礎工学部制御工学科卒業。昭和57年同大学院基礎工学研究科後期課程修了。同年同大学基礎工学部助手。昭和63年同大学工学部講師。平成元年度同大学助教授となり現在に至る。この間昭和61年より1年間米国メリーランド大学自動化研究センターにて客員研究員。コンピュータビジョンの研究に従事。人工知能、ロボティクスに興味を持つ。本会平成元年度研究賞。電子情報通信学会、日本ロボット学会、IEEE Computer Society 各会員。



白井 良明 (正会員)

昭和39年名古屋大学工学部機械工学科卒業。昭和44年東京大学大学院工学系博士課程修了。工学博士。同年電気試験所入所。以来コンピュータビジョンを研究。昭和63年大阪大学工学部電子制御機械工学科教授。著書「Three Dimensional Computer Vision」(Springer)など。