

移動ロボットのためのステレオビジョンによる環境の幾何モデリング

杉本和英*、 高橋裕信*、 富田文明**

*三洋電機筑波研究所、 **電子技術総合研究所

未知の環境における移動ロボットの視覚としてステレオビジョンを用いる場合、異なる視点から得られるデータを照合し、ロボットの正確な移動量を算出し、そして部分的で不完全なデータを逐次統合して環境の面ベースの記述を生成することが問題となる。本論文では、ステレオの直接の出力として得られる三次元セグメントのワイヤフレームに対して矛盾なく面を内挿するアルゴリズムを適用し、環境の記述は、最も一般的な幾何モデルである境界表現 (b-rep) に基づいている。また、データが不完全で面の内挿ができない部分を局所的に解釈不可能な部分として残すことができ、その後の観測データを補間することによって解釈可能とすることができる。

Geometric Modeling of Environment by Stereo Vision for Mobile Robot

Kazuhide Sugimoto*, Hironobu Takahashi*, Fumiaki Tomita**

*Tsukuba Res. Cen. SANYO Ltd. **Electrotechnical Laboratory

*2-1 Koyadai, Tsukuba-shi, Ibaraki 305, Japan
**1-1-4 Umezono, Tsukuba-shi, Ibaraki 305, Japan

In this paper, we propose segment-based stereo vision on a mobile robot in unstructured environments. The problems here are how to match the partial stereo data which have been obtained from different views of the environment, how to estimate the motion of the robot from them, and how to generate the full surface-based descriptions of the environment by incrementally integrating incomplete stereo data. The descriptions of the environment are based on b-rep (boundary representation), the most common geometric model. Since the algorithm to interpolate surfaces to wire frames given by stereo vision has locality to detect which portion of data is incomplete, we can intensively make such incomplete portion complete by adding new data from another observation of the environment.

1. まえがき

環境を認識しながら自律的に移動するロボット (ALV, AGV) の実現は産業的に非常に期待されているがその課題も多く、特に視覚がネックとなっている。移動ロボットの視覚センサとしては、現在のところ、超音波やレーザによる能動センサがその安定性から優位を占めている [1]。将来は、ステレオ視が、人間と同じように、その簡便さと融通性から知能ロボットの共通的な視覚センサとなるであろう。しかし、ステレオ視を実用化するためには、従来の対応問題に加えて、測定精度、出力形式、面の内挿などステレオ固有の問題を解決する必要がある。

これまで、対応問題を解決するために多くのアルゴリズムが開発されてきた [2]。その多くは対応の単位として特徴点、特に、ゼロクロッシングを用いている。しかし、その実験結果として視差分布図に留まり、距離測定まで行なっているものは少ない。これは、視差を距離に変換すると誤差や雑音が増幅されるからであると思われる。そこで、最近では、線セグメントを対応の単位として用いる場合が多くなってきている [3, 4, 5]。これは、対応の容易さだけでなく、測定精度も優れているからである。

また、点単位のステレオの欠点として、その出力が単なる三次元の点の集合であることが挙げられる。すなわち、物体認識など高次の処理を行なうためには、線の抽出といった容易ではない前処理がさらに必要となる。この問題は、線単位のステレオを用いることで解決されている。線単位のステレオの出力は三次元のワイヤフレームであり、多くの後処理の入力としてそのまま利用することができる。

ステレオ視による距離測定精度は、原理的にはカメラの分解能とカメラ間距離によって定まるが、実際には、カメラパラメータの測定精度が問題となることが多い。カメラパラメータは既知のテストパターンを用いて事前に測定することができ、これをカメラキャリブレーションという。しかし、キャリブレーションには必ず誤差が存在するし、キャリブレーション後にいくつかのパラメータが変化してしまう。そこで、筆者らは、観測毎に未知の観測データからでも現在のカメラパラメータを自律的に計算することのできるセルフキャリブレーション法を開発した [6]。その結果、ステレオ視によっても飛躍的に高精度の距離データを得ることが可能になった。

ステレオ視の最大の欠点は、面上にシェイディングやテクスチャなどの特徴が存在しなければ、面の情報を直接には得られないことである。すなわち、たとえ移動ロボットがステレオ視により障害物のエッジを検出できても、エッジのどちら側に面があるのか、あるいはエッジのどちら側を通り抜けることができるのか決定できないことを示している。対象と背景が分離できる設定の場合に

は、エッジから単純に面を内挿するアルゴリズムはあるが、一般的な場合に適用すると偽の面が生成されてしまう。これに対して、著者らは、線単位のステレオ視によって得られるワイヤフレームに対して物理的に矛盾のない面 (平面) だけを生成する方法を開発した [7]。

以上のように、ステレオ固有の問題をある程度解決することができたので、本論文では、ステレオ視により未知の環境内を自律的に移動するロボットを考える。ロボットの行なう作業のシナリオは以下の通りである。ロボットは、世界地図 (環境の幾何モデルであるが、はじめは不完全である) によって定義された自由空間をある程度移動した後、新たに環境を観測することによって、新たな視点からの環境の局所地図を作成する。そして、局所地図とこれまでの世界地図を照合し、局所地図を世界地図に統合することによって世界地図を更新する。以上の処理を与えられた目標を達成するまで繰り返すことによって、世界地図を次第に完全なものとしていく。

同様な研究として、Bras-Mehlmanら [8] は、異なる視点から得られた環境の複数のステレオデータからドロネー三角網を用いて環境の面表現を生成する方法を開発しているが、計算コストが非常に高く、また偽の面も生成される。

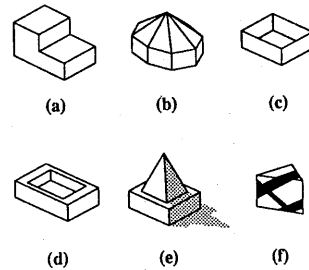


図1: シーン中に存在する物体

2. 世界と表現

ロボットの移動する環境に存在する物体は、平面でのみ構成されていることだけを前提とする。勿論、ロボットが事前に環境内に存在する物体を知ることにはない。むしろ、どのような物体がどこに存在するかを知ることが目的とする。図1に、環境内に存在してもよい物体の例を示す。物体は、(a)のようにエッジが凸または凹の多面体であってもよいし、(b)のように三面 (以下) 頂点に限定されることもなく、(c)のように紙で構成されていてもよい。また、(d)のように穴のある物体でもよいし、(f)のように物体の表面に影や模様があってもよい。勿論、(e)のように複数の物体が重なってもよい。

物体を記述するためのデータ構造としては境界表現 (b-rep) を採用した。これは、図2に示すように、物体の面 (F) とエッジ (E) の関係を表すグラフ構造となっている。各エッジは互いに反対の方向

を持つ二本のセグメントによって表され、互いにポイントを持ついわゆるウィングドエッジ構造となっている。各セグメントの方向はセグメントが構成する面を右に見るように設定されている。したがって、紙は表裏二つの面を持つことになり、各面は、図3に示すように、互いに反対方向を持った異なるセグメントのリストによって構成される。

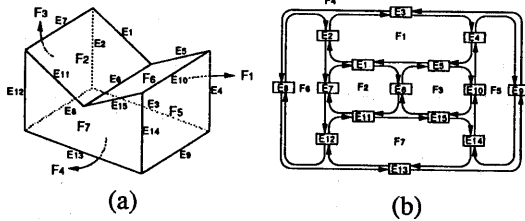


図2: 物体の境界表現: (a) 物体の構成要素; (b) データ構造

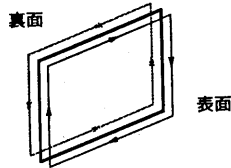


図3: 紙の裏と表の面を表す2つのループ

3. ステレオ解析

筆者らは、画像の中間表現として、立体の(三次元)境界表現の仕様にしがたう画像の(二次元)境界表現を開発しており、ステレオ画像の対応はこのデータ構造上で行なわれる[5]。画像の境界表現では、画像は一様な領域の集合で表される。各領域 R は、図4のように、一つの外周境界 B と穴が存在する場合にはその内周境界 H によって表される。各境界は、それを構成する境界線セグメント S のリストによって表される。各セグメントは、領域を右に見るような方向を持ち、また隣接する領域を構成する反対の方向を持つセグメントへのポイントを持ちウィングドエッジを構成している。対応の単位として境界線セグメントを用いる。対応探索の際には、エビポーラ拘束の他に、領域の明るさ、セグメントの方向、他のセグメントとの連結性を利用する。以下に、その概要を述べる。

1) まず、領域間で対応の候補を求める。対応の必要条件是以下の通りである。

- 領域の明るさが類似している。
- 領域内にエビポーラ拘束を満足する対応部分が存在する。

2) 次に、対応する領域の境界線を構成するセグメント間で対応の候補を求める。対応の必要条件是以下の通りである。

- セグメントの方向が類似している。
 - セグメント上にエビポーラ拘束を満足する対応部分が存在する。
- 3) 一つのセグメントが同一のエビポーラ線上で複数のセグメントと多重対応する場合には、対応するセグメントの空間的な連続性を調べ、その結果得られるセグメントの系列(ワイヤフレーム)の中で、最も長くかつ形状が最も類似する系列を選択する。特に、平面を構成するアフィン変換可能なセグメントの系列を優先する。

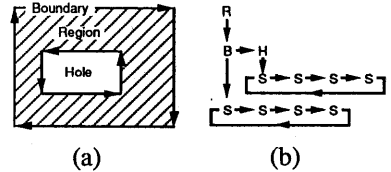


図4: 画像の境界表現: (a) 穴の空いた領域 (b) データ構造

以上のステレオの出力は三次元セグメントの三次元ループである。ループは右(もしくは左)画像中の領域の境界線を三次元空間中に投影することにより生成される。例えば、図5(a)のステレオ画像からは同図(b)に示すように7つのセグメントのループが得られる。各セグメントは実際にはエッジ上に乗っているが、各エッジ上には互いに反対の向きを持った二つのセグメントが存在することを示すために離して示してある。各セグメントは矢印で示してあるように面を右手に見るような方向を持っており、同一エッジ上に存在する反対向きのセグメントへのポイントを持つウィングドエッジを構成している。

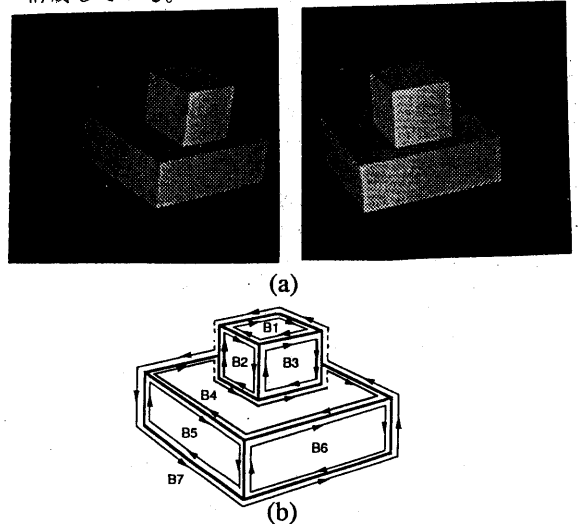


図5: ステレオの出力: (a) 原画像; (b) セグメントのループ

左回りのループ (例えば、穴を表すループ) が存在する場合、第6図に示すように二本のダミーセグメントを作ることによってその外側の右回りのループと統合する。

セグメントの位置の正確さは、視点からの距離と傾きによる。各セグメントには、この正確さの情報も持たせてある。原画像中で対応する二次元のセグメントを考える。ここで、左画像のセグメントの両端点を (x_s, y_s) 、 (x_e, y_e) 、左画像のセグメントの両端点を (x'_s, y'_s) 、 (x'_e, y'_e) とする。対応の結果得られる三次元セグメントの精度の指標を、次のように二つの対応するセグメントの間の面積によって定義する。

$$\frac{1}{2} \{ (x_s - x'_s) + (x_e - x'_e) \} \cdot |y_s - y_e|$$

したがって、この指標は視点により近くより垂直な三次元セグメントほど大きな値を持つ。

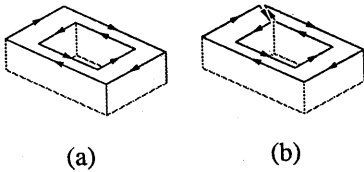


図6: (a)面の外周ループと穴のループ; (b)2つのループの統合

4. 面の内挿

図5(b)の各ループに対する面の内挿を考える。ループ上のセグメントの矢印は、その右側に面が存在することを示している。しかし、ループ B_4 の破線で表示されたセグメントと、ループ B_7 の全てのセグメントはこの視点からは見えていない面上のセグメントである。そこで、ループ上の各セグメントを、見えている面上のセグメント(実セグメント)もしくは見えていない面上のセグメント(虚セグメント)に分類する実在テストを開発した。実セグメントだけがループに対する面の内挿に貢献することになる。基本的な方法は参考文献 [7] を参照されたい。

4.1 オクルージョンの検出

まず奥行きが不連続となるオクルージョン点(丸で表示)を検出する。オクルージョン点では、明らかな実セグメント(太実線で表示)と明らかな虚セグメント(点線で表示)を区別することができる。

4.2 実在テスト 1

注目するループが「解釈可能」、すなわち唯一の面の内挿が可能であるためには以下の条件を満たさねばならない。

- 1) 可能性: 明らかな実セグメントを含むセグメントの組み合わせにより平面を定義することができ、他のセグメントがその平面の裏側に位置することにはならない。

- 2) 一意性: 上記の条件を満足する平面は唯一つである。

注目するループが上記の条件を満足する場合、定義した平面上に存在するセグメントが実セグメントであり、平面より手前側にある他のセグメントは虚セグメントであると判定される。その結果、図7(a)に示すように、物体の各エッジを、実および虚セグメントを共有するオクルージョンエッジ(→)と二本の実セグメントによって共有される稜エッジに分類可能である。さらに、稜エッジはその両側の面の法線のなす角度により、凸(+), 凹(-)、影や模様によってできる平坦(p)のエッジに分類することができる。

注目しているループに対して面を唯一決定できる場合、実セグメントには含まれている虚セグメントの代わりに、仮想セグメントを生成する。図7(b)に点線で示すように、それらは各虚セグメントに対する視線ベクトルを生成された面に投影することにより得られる。

オクルージョンや照明の具合により存在すべきエッジが検出できない場合には、注目するループは実在テストを満足せずに「解釈不可能」になる。つまり、そのループに対して唯一の面を内挿することができない。例えば、図8(a)に示すように、本のように凹に折った紙を考える。いま、同図(b)に示すように、凹のエッジが検出できなかったとする。その結果、内側の領域のループに対しては、(c)、(d)の陰影をつけた二つの面を生成することができるが、これは面の一意性の条件に反する。したがって、このループは解釈不可能となる。

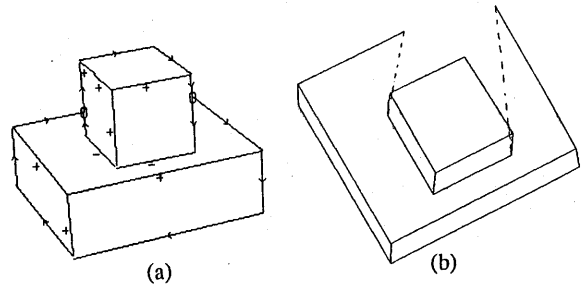


図7: (a)エッジのラベル付け; (b)仮想セグメント(点線)

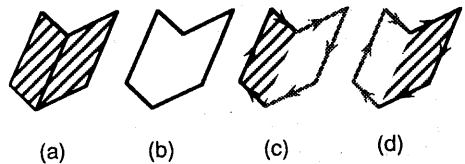


図8: 解釈不可能なループ: (a)凹状に折られた紙(本) (b)凹エッジの消失; (c)(d)内挿された面(斜線の部分)

4.3 実在テスト 2

注目しているループは、隣接するループの解釈の結果によっても、解釈不可能になる場合がある。例えば、図9(a)に示すように、屋根のように凸に折った紙を考える。いま、同図(b)のように、凸のエッジが検出できなかったとする。内側の領域のループに対して、同図(c)の陰影をつけた面が実在テスト1を満足する。したがって、その面より手前にあるセグメント(薄い線)は虚セグメントであると判定される。一方、同図(d)の外側の領域のループに対しては、ループ上の全てのセグメントは虚セグメントであると判定される。その結果、定義された面より手前のエッジ上には実セグメントは存在しない、つまり、エッジが存在しないことになる。これは、実際にエッジは検出されているので矛盾する。このような場合には、両方のループは解釈不可能となる。

虚セグメントと解釈不可能なループ上のセグメントには「未定」の印が付けられ、その接続関係とともに世界地図に保存される。

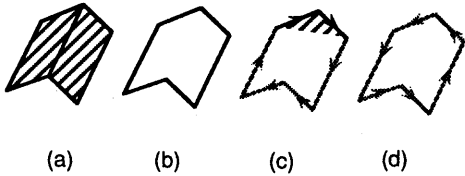


図9： 解釈不可能なループ： (a) 凸状に折られた紙(本) (b) 凸エッジの消失； (c) 内側の領域のループ； (d) 外側の領域のループ

5. データの統合

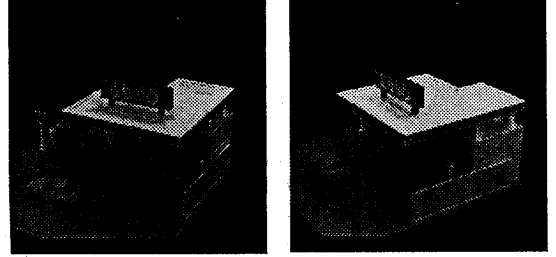
いま、移動ロボットは、初期位置での環境の観測から図10に示すような初期の世界地図を持つものと仮定する(定義された面に対しては陰影が施されている)。そして、ロボットが移動することによって、図11に示すように、新たな位置から観測した環境の局所地図を作成したとする。

5.1 対応探索

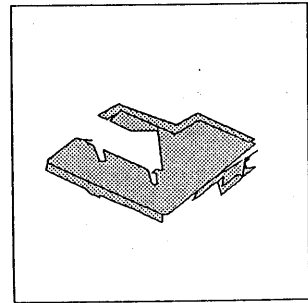
まず初めに移動前後の両地図間での対応を求める。移動前の座標系と移動後の座標系の移動量を回転行列 R 、平行移動ベクトル t とし、 x_i と x'_i をそれぞれ移動前後のカメラで観測された同一点の座標とすると、次の関係が成り立つ。

$$x_i = R^{-1}x'_i - t$$

移動パラメータはロボットの駆動系から大まかに求めることができる。これを移動パラメータの初期値として用いる。正確なパラメータは、世界地図と局所地図との対応を求めることにより、以下のように計算により求めることができる。

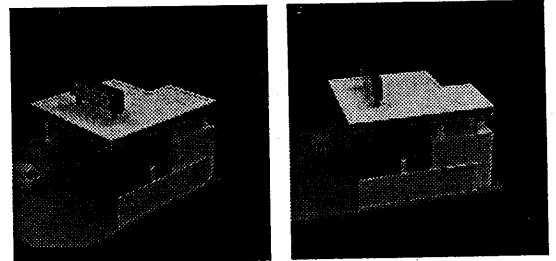


(a)

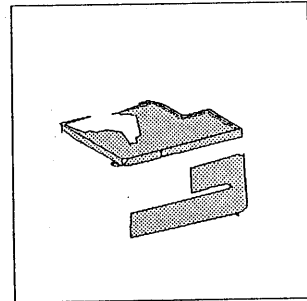


(b)

図10： 最初の視点から見たシーン： (a) ステレオ原画像； (b) 初期世界地図



(a)



(b)

図11： 2番目の視点から見たシーン： (a) ステレオ原画像； (b) 局所地図

- 1) 移動パラメータを用いて局所地図中のセグメントを座標変換して、世界地図と局所地図を重ね合わせる。移動パラメータには誤差が含まれるため、対応するセグメントも正確には重ならない。
- 2) 図12に示すように、局所地図中のセグメントの両端点(A', B')から世界地図中のセグメントを含む直線上の点に対して垂線を下ろすことにより、局所地図中の各セグメントに対して同じ方向を持つセグメントで最も近いものを選び出す。すなわち、世界地図中のセグメントのCDの部分に対して、局所地図中のセグメントのC'D'の部分が対応すると仮定した場合に、(CC'+DD')が最小となるセグメントを選ぶ。
- 3) しかし、CC'かDD'がしきい値 τ より大きい場合には、両セグメントは対応するとはみなさない。しきい値以下の場合には、両セグメントは対応するとみなされ、対応の誤差を(CC'+DD'×CD)によって定義する。そして、局所地図中の全てのセグメントの誤差の和をSとし、最小自乗法によりこの値Sを最小にするように移動パラメータを計算し直す。

以上の処理は、誤対応を取り除くために τ を小さくしながら、誤差が十分小さくなるまで繰り返される。図13は世界地図(第一の局所地図)中のセグメント(実線)に対して、第二の局所地図中のセグメント(破線)の重ね合わせをした結果である。

5.2 再解釈

図14(a)の世界地図および同図(b)の局所地図を考える。陰影をつけた部分は面が内挿されていることを表す。一方、陰影のついていないループは解釈不可能で、面が内挿されるのを待っている部分である。局所地図中の解釈不可能なループを、欠けているセグメントを世界地図から見つけることによって、解釈可能とすることができる。

- 1) 局所地図中の解釈不可能なループL'上のセグメントに対応するセグメントが世界地図中の解釈可能なループL上に存在する場合には、ループL上のすべてのセグメントと、ウィングドエッジを構成している反対方向のセグメントとをループL'に統合する。その結果、ループL'中にサブループが生成される。この時、少なくともループLに対応するサブループは、同図(c)のループAのように、解釈可能となる。
- 2) 残りの各サブループに対して実在テスト1を適用し、可能な場合には、同図(d)のループBのように、面を内挿する。
- 3) 同図(d)のループCのように、依然解釈不可能なループが残っている場合には、同図(e)のように、世界地図中の未定のセグメントの中から、ループの異なる二点を結ぶようなセグメントの系列をすべて求める。

- 4) 新たに生成された各サブループに対して実在テスト1を適用し、可能な場合には、同図(f)のループC₁とC₂のように、面を内挿する。
- 5) そして、すべての解釈可能なループに対して実在テスト2を適用する。

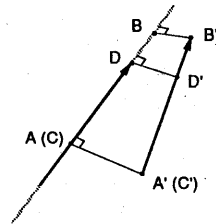


図12: セグメントの対応

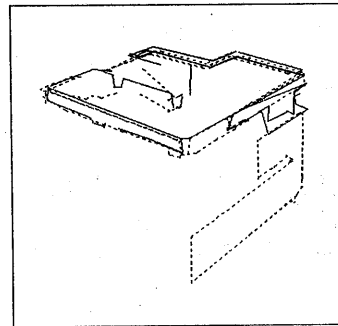


図13: 世界地図中のセグメント(実線)と第二の局所地図中のセグメント(破線)の対応の結果

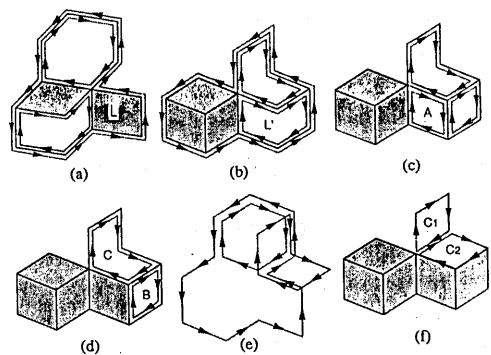


図14: 世界地図を用いた局所地図の再解釈

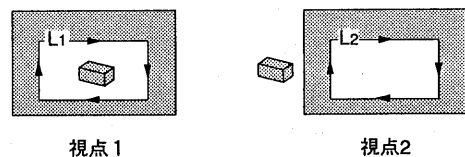


図15: 異なる視点から見た同一のループ

5.2 検証

図15(a)、(b)にそれぞれ示すように、異なる視点から同じ面を見た場合について考える。いま、それぞれ独立に解釈すると以下ようになる。ループ L_1 の内側後方には別の物体が見えるので、このループは穴であると判断され、面は内挿されない。一方、ループ L_2 の内側には何も見えていないので、このループには面を内挿する。したがって、視点1が先で視点2が後ならば、先に穴の情報が得られるので、この内側のループに対して面を内挿することは避けられる。しかし、視点2が先で視点1が後の場合には、後の視点で初めて穴であったことが分かるので、前の視点で内挿されていた面は誤りであり取り除かねばならない。

このように、新たな局所データを用いて世界地図を更新する際、その視点で見えている世界地図中の面に対して、局所データ中のセグメントを隠すような面があるかどうか常に調べなければならない。そして、そのような面が存在すれば世界地図から取り除かなければならないし、その面を構成していたセグメントは現在の視点からは見えない面上のセグメントでなければならない。

最初の視点1では何らかの都合により解釈不可能なループが、次の視点2で解釈可能となる場合には、問題はより難しくなる。このループが穴であるという情報はどこにもないので、面が内挿されてしまう。したがって、以前の視点1に戻り、この面が他の如何なるセグメントをも隠さないことを調べる必要がある。

このように、移動ロボットは、それまでの世界地図中に存在しなかった新たな面を生成する際には常に、今までのすべての視点位置において、その面が何も隠さないことを調べなければならない。このことは、すべての視点情報とすべてのセグメントがどの視点から見えたかという情報を持っていないければならないことを意味する。しかしながら、実際の移動ロボットへの応用を考えた場合、すべてのデータを持つ必要はないと考える。なぜなら、ロボットが連続的に環境を観測しているのであればこのようなことが起こるのは稀であるし、最新の数ヶ所での視点でのデータを持っていれば実用上は十分である。

世界地図は新たに得られた局所地図を統合することにより更新される。局所地図中のセグメントで世界地図中に対応するセグメントが存在しない場合には、このセグメントは世界地図中に追加される。世界地図中に対応するセグメントが存在する場合には精度の良いものを世界地図中に統合する。

図16は、第二の観測後に更新された世界地図を示している。図17(a)は、第三の視点から観測した環境であり、同図(b)は、世界地図中のセグメント(実線)と第三の視点で得られた局所地図中のセグメント(破線)との対応を示している。そして、図18は、第三の観測後に更新された世界地図を示したものである。

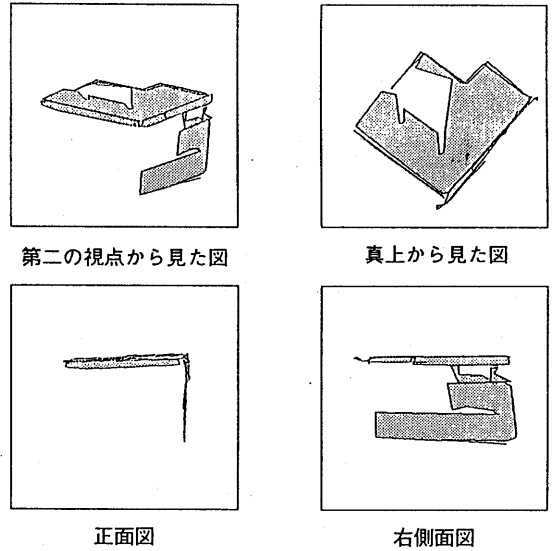
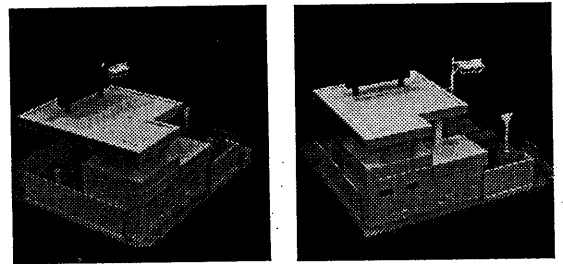
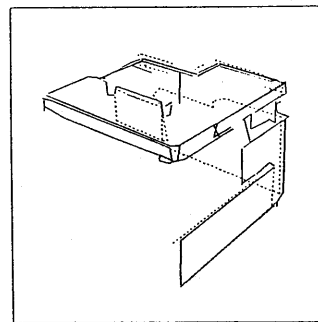


図16: 第二の観測後に更新された世界地図



(a)



(b)

図17: (a) 第三の視点から得られた環境のステレオ原画像; (b) 世界地図中のセグメント(実線)と第三の局所地図中のセグメント(破線)の対応の結果

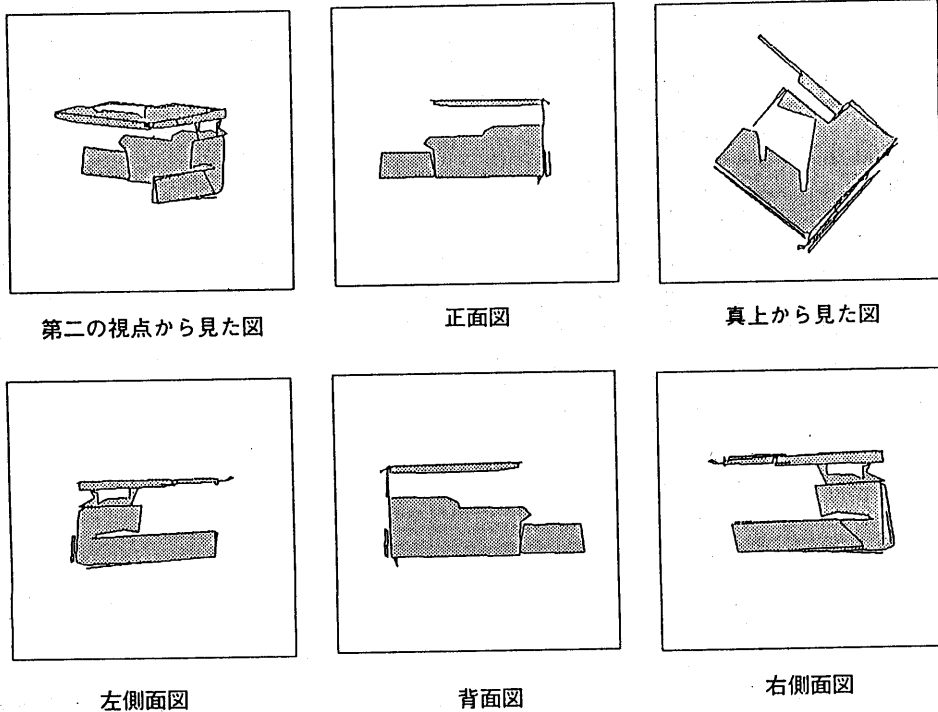


図18: 第三の視点からの局所地図を統合して更新された世界地図

6. まとめ

未知の環境における移動ロボットのためのステレオビジョンについて述べた。不完全なステレオデータを逐次統合しながら、環境の境界表現一面ベースの記述を生成した。各視点において、セグメントの欠けがなくすべてでのループの解釈が可能となる完全なステレオデータを得ることは難しいが、検出できないセグメントをできるだけ検出することにより解釈不可能なループを減らす努力をしていかなばならない。本論文で述べたアルゴリズムによれば、面を内挿する際にデータのどの部分が不完全であるかを知ることができるので、各観測において注視の手法を用いることができ、さらに次の観測位置の計画をすることもできる。

参考文献

- [1] Hebert, M. and Kanade, T., Outdoor scene analysis using range data, Proc. IEEE Int. Conf. Robotics and Automation, 1426-1532, 1986.
- [2] Barnard, S.T. and Fischler, M.A., Computational stereo, Computing Surveys, 14, 553-572, 1982.
- [3] Medioni, G.G. and Nevatia, R., Segment-based stereo matching, Proc. Image Understanding Workshop, 128-136, 1983.
- [4] Ayache, N. and Lustman, F., Trinocular stereovision: recent results, Proc. 10th Int. Joint Conf. Artificial Intelligence, 826-828, 1987.
- [5] Takahashi, H. and Tomita, F., Planarity constraint in stereo matching, Proc. 9th Int. Conf. Pattern Recognition, 446-449, 1988.
- [6] Takahashi, H. and Tomita, F., Self-calibration of stereo cameras, Proc. 2nd Int. Conf. Computer Vision, 123-128, 1988.
- [7] 杉本、高橋、富田: “ステレオによるシーンの解釈”, 情報学会 CV 研資 55-4
- [8] Le Bras-Mehlman, E.L., Schmitt, M., Faugeras, O.D., and Boissonnat, J.D. How the Delaunay triangulation can be used for representing stereo Data, Proc. 2nd Int. Conf. Computer Vision, 54-63, 1988.