

環境モデルにおける物体の見え方の知識

松原 仁* 伊庭 斉志** 井上 博允**

+ 電子技術総合研究所

++ 東京大学工学部

要約

視点の位置と視線の方向を定めた際に環境を構成する物体がそこからどのように見えるはずかという知識を物体の見え方の知識と名付ける。ここでは知能ロボットシステムCOSMOSにおける物体の見え方の知識の表現方法とそれを自動的に推論する過程について述べる。知能ロボットシステムを実現するためには視覚系と行動系に共通の環境モデルを構築することが必要であるという考え方にに基づき、COSMOSの環境モデルでは視覚系のための知識と行動系のための知識をFRLで統一的に表現している。アーチの例を詳しく示すことによって二つの系の知識の間の結び付きを具体的に説明する。

Representation of Visual Knowledge
in Environment Model

Hitoshi MATSUBARA* Hitoshi IBA** Hirochika INOUE**

+ Electrotechnical Laboratory

1-1-4 Umezono, Sakura-mura, Niihari-gun, Ibaraki, 305 Japan

++ Department of Mechanical Engineering, Faculty of Engineering,

University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo, 113 Japan

ABSTRACT

Visual knowledge is that of how objects in some task environment are seen from a given viewpoint. This paper describes the representation method and the automatic reasoning process of the visual knowledge in the environment model of the intelligent robot system COSMOS. For the purpose of construction of intelligent robot systems, we think it necessary to build the environment model which is shared by both vision system and action system. In our COSMOS, knowledge of both systems is uniformly represented in FRL. By showing an example of 'arch', we present interrelationship between their knowledge in detail.

1. はじめに

知能ロボットシステムを構築するには感覚・思考・行動系を有機的に統合することが必要であるが、環境モデルに関する従来の研究では各系のモデルが独立に扱われている場合が多い。感覚系として代表的な視覚系の研究では特にモデルの果たす役割が重視されるようになってきたが、他の系とは独立の専用モデルを用意する場合はほとんどである。しかしロボットに三つの系にまたがるような知的作業を行なわせるためには、三つの系に共通の環境モデルの存在が重要であると考えられる。たとえば、視覚系と行動系との間の関係を例にとることとする。行動系ではロボットが次にどのような行動を起こすかあらかじめわかっているのに、視覚系ではその情報に基づいてモデル駆動型の画像理解を試みることができる。逆に視覚系で得られる画像情報によって、その時点の環境の状態が行動系の予想したものとは異なることが判明しモデルの内容が訂正される可能性もある。このように二つの系の間では知的作業のために極めて重要な情報が伝達される。それぞれの系が専用モデルを用意して両者の間でいちいち形式を変換する方式では情報の伝達にとってボトルネックになる危険がある。共通の環境モデル中に同一の形式で知識を表現しておけばそのような問題は起きない。我々は以上のような考え方にに基づき、三つの系に共通の環境モデルの構築に取り組んでいる。視点の位置および視線の方向を定めた際に環境を構成する物体がそこからどのように見えるはずかという知識を物体の見え方の知識と呼ぶことにする。すなわち、環境モデル中の視覚系のための知識が物体の見え方の知識である。ここでは現在開発中の知能ロボットプロトタイプシステムCOSMOS^{1, 2)}の環境モデルにおける物体の見え方の知識の表現方法とそれを自動的に推論する過程について述べる。なお、ロボットとしてはハンドアイシステムを、また作業環境としては積み木の世界を考える。さらに視覚系が対象とする画像情報は線画のレベルを想定する。

2. COSMOSの環境モデルにおける知識表現

ロボットの内部で対象となっている物体や作業の情報を管理し、作業環境の変化をシミュレートしているのが環境モデルである。環境モデルは、モデル駆動型の画像理解、行動計画の自動生成、高水準のロボット言語などのロボットにとって高度な機能を実現するために鍵となる重要な部分である。基本的にCOSMOSの環境モデルは、FRL³⁾というLisp組み込みのフレーム記述言語で書かれている。知識の階層構造、性質の継承メカニズム、付加手続きの埋め込みなどのフレームの特徴がロボットの環境モデルの知識を表現するのに適していると考えているためである。また、物体間の相対関係はRAPT⁴⁾にならって物体の面と軸を単位にして表現している。COSMOSの環境モデルの行動系の部分の詳細についてはここでは省略するので他の文献^{1, 2, 5)}を参照されたい。

COSMOSのこれまでの環境モデルはロボットの動作に伴う作業環境の変化のシミュレーションが主な用途であったので、知識の表現は2次元的なものにとどまっていた。視覚系では物体の見え方の知識を推論するのに3次元処理を行なう必要があるので、

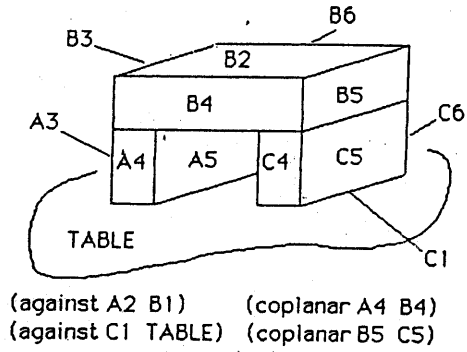


図1 アーチ

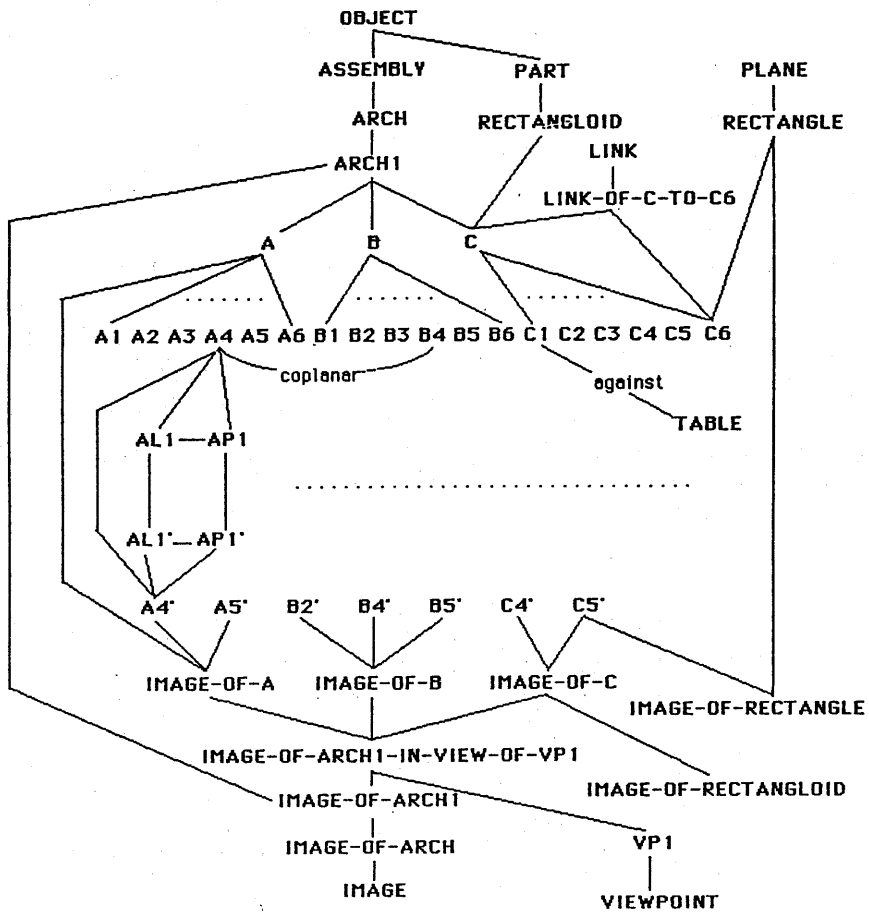


図2 アーチの知識の表現

ソリッドモデルの要素を取り入れた．そのデータ構造としては処理効率の面から境界表現 (boundary representation) ⁶⁾ を採用した．

アーチの例によって，COSMOSの環境モデルにおける知識表現の方法を示す．図1に示したテーブルの上にアーチが置いてある状態の知識（の一部）は，環境モデル内では図2のような階層構造で表現されている．図2は見やすさを考慮して実際のものよりかなり単純化してある．実際にはフレームの付加手続き（デモン）として動的に表現されている知識も，この図では仮に静的な知識として表現してある．また，フレーム間のポインターは一般に両方向とも存在しているが，この図では1本に代表させている．さらにポインターの名前も省略している．比較的単純な組み立て物であるアーチですら，環境モデル内ではこのように膨大な知識で構成されているのである．

3. 物体の見え方の知識の表現

COSMOSの環境モデルにおける物体の見え方の知識は，Minskyがフレーム理論の説明に用いた例（図3参照）⁷⁾ とほぼ同じ形式で表現されている．図2の点線で囲んだ部分が図1のアーチの見え方に相当する．線画というレベルは，生の画像のレベルと人間が持っているはずの内部表現のレベルの中間に位置しているので，以後どちらの方向へ研究を進めていくにしても発展性の面から妥当な選択と考えている．視点（TVカメラの位置と姿勢に相当する）のフレームの下には複数の具体的な視点のフレームが存在する．一方，画像フレームの下には，具体的な画像の一つとして，ある視点から見たアーチの画像のフレームが存在する．視点のフレームとその視点から見たアーチの画像のフレームは，ポインターで結ばれている．また，そのアーチの画像のフレームを構成する面，辺，頂点のフレームと，アーチそのもののフレームを構成する面，辺，頂点のフレームとの間も，ポインターによって適切な対応が付けられている．なおアーチの見え方は，特殊な場合は考えないとしても，図1の場合の他に図4のような場合もあり得る（図1とは位相が異なる）．画像の位相に応じて物体の見え方の知識は別々に構築される．

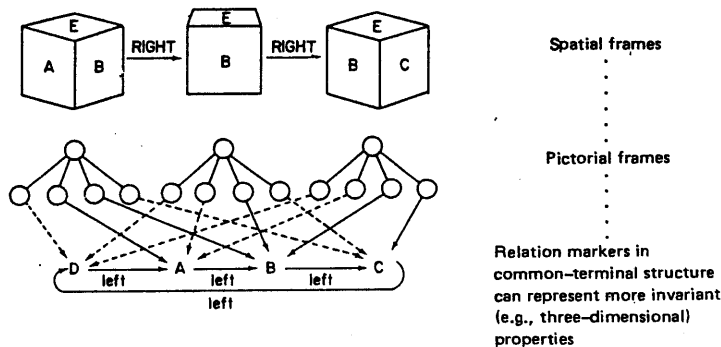


図3 フレーム理論⁷⁾

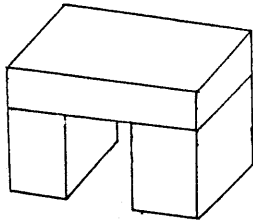


図4 アーチの別の見え方

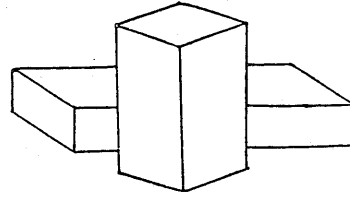


図5 左右に分かれて見える物体

4. 物体の見え方の知識の推論

環境モデルの中の物体の見え方の知識を、行動系のための知識、すなわち物体自体の知識から推論する過程について説明する。図2のアーチの画像の部分のフレームを作成して、なおかつ元のアーチ自体のフレームと適切な対応を付けることが推論に相当する。物体の見え方を推論するためには、基本的にはグラフィックスの手法を用いればよい。すなわち、各物体の位置と姿勢、視点の位置、視線の方向が既知であるならば、隠線消去を行なうことによってその視点から見た線画を抽出することができる。しかし、環境モデルにおける物体の見え方の知識という場合には、2次元画像のフレームと元の3次元のフレームとの対応がとれていなければならない。例えば、図5の線画において左右に分かれて見えている面が、実際には一つの面であることを認識していなくてはならない。この点でグラフィックスにおける隠線消去の手法の多くは適していない。なぜなら2次元の線画情報をディスプレイに出力するだけで、元の3次元情報との対応関係は保存していないからである。したがって、情報としては線画を対象にしているが、面の対応関係を保存するために、物体の見え方の推論には隠面消去の考え方を応用した手法を用いている。

推論の過程は二つの段階から成る。第一の段階では物体の各面について可視か不可視かを面の間の拘束条件などから推測できるものについて判定する。第二段階では、第一段階で未決定として残った面について実際に刈り込みの手法を適用して可視・不可視を判定する。各々の過程で面が可視と判定された時点で元の面のフレームとの対応付けを行なう。以下で図1, 2のアーチを例にしてこの二つの段階の処理について説明する。視点の位置、視線の方向、及びアーチの各物体の位置と姿勢は既知と仮定する。また、直交射影を仮定しているが、透視射影にも拡張可能である。

推論の過程は実時間で処理することは想定していない。視覚系のために前もって物体の見え方の知識を生成しておくことが推論の目的である。

4.1 第一段階

(1) 後面消去(backface elimination)⁸⁾

各面の外向き法線ベクトルを計算して、それと視点から物体に向かうベクトルが同方向(差が90度以内)であれば、その面は不可視として除外する。アーチの例で

は、A 1, A 3, A 6, B 1, B 3, B 6, C 1, C 3, C 6の面がこの段階で不可視として除外される。他の面は未決定として残る。

(2) 面の間拘束条件による推論

RAPT⁴⁾流の関係記述を面の間拘束条件としてとらえ、それに基づいて可視か不可視かを判定する。記述子は、against が二つの面が同一平面上にあって反対向きに接している意味で、coplanarが二つの面が同一平面上にあって同じ向きに接している意味である。

中心となる原則は次の二つである。

P 1 :

二つの面がagainst の関係にあって一方が(1)の結果などから不可視であることが判明していれば、他方は完全可視面(面上の全ての点が可視となる面)とはなりえない。

P 2 :

二つの面がcoplanarの関係にあって一方が可視ならば、他方は前にさえぎる物体がない限り可視である。

以上の原則から次のような推論規則(全部で十数個)を作成した。

R 1 :

上に他の物体がのっていない物体をTOP とするとき、TOP の数がひとつであればTOP の未決定面は可視である。

R 2 :

互いにagainst な面をF 1, F 2とするとき、F 1が不可視面であり、かつF 2がF 1に含まれる(F 2内の任意の線分はF 1内にある)ならF 2は不可視面である。

これらの規則をアーチの例に適用すると、R 1からB 2, B 4, B 5が可視面となり、R 2からA 2とC 2が不可視面と判定される。さらにP 2からA 4, C 4, C 5が可視面と判定される。これによって未決定な面のみとなる。この段階において可視と判定された面は物体の見え方の知識として環境モデルに登録し同時に元の3次元情報の面のフレームと対応をつける。

4. 2 第二段階

第一段階で未決定として残った各面に対して以下の手続きを適用する。

(1) 重なりテスト(overlap test)⁸⁾

重なりテストで、問題の面を隠している可能性のある面を取り出す。これらは不可視面からは選ばない。アーチの例ではB 4, C 4が選ばれる。

(2) 刈り込み

(1)で選んだ面(前方面)と問題の面(後方面)に対して刈り込みのアルゴリズムを適用する。これは前方面と後方面の頂点の情報から、前方面に隠されていない部分の面の頂点の情報を返す。結果として複数の面の情報が得られることもあり(前方面が後方面を分割する場合)、何も得られないこともある(前方面が後方面を覆い隠す場合)。

アルゴリズムの基本的考え方は、前方面をF1、後方面をF2とすると以下のようになる。

まず二つのリンクlink1, link2を作成する。

link1はF2の辺上を巡回しながら、F1の外側もしくは辺上にある頂点、またはF1との交点を順に連結したリンクである。link2はF1の辺上を巡回しながら、F1の頂点、またはlink1に現われる頂点を順に連結したリンクである。

次にlink2の情報に基づき、link1をたどることでF1に隠されていない部分のF2の頂点列の情報を得る。これをアーチの例に適用したのが図6である。初めにC4（前方面）とA5（後方面）について刈り込みを行なう。すなわち、link1=1-2-3-4-5, link2=a-b-c-3-d-4 というリンクを得る。ここでlink2から3->d->4 というパスが見つかる。従ってlink1をたどりながら3まで来ると、d->4へと向かうことで1-2-3-d-4-5 という新しい図形（A5'）の頂点列の情報を得る。次に前方面B4と後方面A5'について再び刈り込みを行ない、その結果A5''の頂点の情報を得る。重なりテストで見つかった面はC4とB4だけなので、結局A5''が求める面である。

4.3 拘束条件としての相対関係

上記の推論は理論的には第二段階のみで十分である。しかし第二段階の処理は計算量が多いのでなるべく処理する面の数を減らしておくことが望ましい。第一段階の（1）の後面消去の処理は一般に行なわれていることであるが、（2）の面の間の拘束条件を用いた処理は新たに考案した手法である。物体の相対関係を表現するためのRAPT流の記述方法を、処理する面の数を減らす拘束条件として活用している。アーチのように拘束条件の多い組み立て物については、この処理が非常に有効であることを確認した。物体の数が増加すれば一般に拘束の数もそれに伴って増加するので、この処理は複雑な組み立て物に対しても有効に機能する。

5. おわりに

COSMOSの環境モデルにおける物体の見え方の知識の表現方法とそれを自動的に推論する過程について述べた。知能ロボットシステムを実現するには三つの系に共通の環

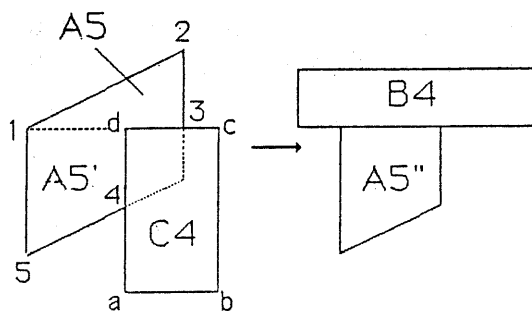


図6 刈り込み

境モデルを構築することが必要であるという考え方にに基づき、COSMOSの環境モデルでは視覚系のための知識と行動系のための知識をFRLを用いて統一的に表現している。

ここで述べた物体の見え方の知識の推論は、行動系のための知識から視覚系のための知識への方向の情報の伝達に相当している。逆方向の情報の伝達に相当する、画像情報から環境の状態を推論し必要に応じてモデルの内容を訂正する試みも行なっている⁹⁾。複雑な組み立て物について環境モデルにあらかじめ完全な情報を取めておくことは困難なので、不完全なモデルを感覚系からの情報で補うような方式をこれからは目指すべきだと思われる。

謝辞

日頃御討論いただく東京大学工学部機械工学科井上・稲葉研究室の皆様感謝します。

参考文献

- 1) 小笠原 司, 井上 博允: " 知能ロボット・プログラミングシステムCOSMOS" , 日本ロボット学会誌, vol.2,no.6, pp.507-525,1984
- 2) 松原 仁, 岡野 彰, 井上 博允: " 作業目標レベルのロボット言語の設計と試作" 日本ロボット学会誌, vol.3,no.3, pp.220-228,1985
- 3) R.Roberts and I.Goldstein:"The FRL manual",MIT AI Memo.409,1977
- 4) R.Popplestone:"RAPT:a language for describing assemblies",The Industrial Robot,pp.131-137,1978
- 5) 小笠原 司: " ロボットのソフトウェアのためのデータの表現と利用" , 日本ロボット学会誌, vol.2,no.2,pp.97-103,1984
- 6) A.G.Requicha: "Representation for rigid solids: theory, method and systems",ACM Computing Surveys, vol.12,no.4,pp.437-464,1980
- 7) M.Minsky: "A framework for representing knowledge", in P.Winston (ed.) "The psychology of computer vision", McGraw-Hill,1975
- 8) W.Newman and R.Sproull: "Principles of computer graphics", McGraw-Hill, 1979
- 9) 伊庭 斉志, 井上 博允: " 環境モデルの自動生成と非単調論理に基づくモデルの管理" , 日本ロボット学会第4回学術講演会予稿集, 1986 (発表予定)