

エネルギー最小化を用いた物体の不可視部分の予測

角 所 考[†] 淡 誠 一 郎[†]
 安 部 憲 広^{††} 北 橋 忠 宏[†]

本稿では、多面体を対象として、その単眼線画像から不可視部分を含む物体全体の3次元形状をボトムアップ的に復元するための方法について検討する。物体形状に関する情報を、面、辺、頂点の数およびそれらの間の接続関係などに関する位相情報と、個々の面、辺、頂点の定量的な方向や位置などに関する幾何情報とに分けて考えると、物体の不可視部分は可視部分とは異なり、位相情報が画像から得られないため、これを何らかの方法で決定する必要がある。そこでまず、この位相情報の決定のために有効であると考えられるいくつかの制約条件を導入し、それが利用可能なものに対象物体を限定することによって以降の処理を一種の組み合わせ最適化問題に帰着する。この解は複数となる場合があるため、次に、幾何情報に基づく優先順位を導入し、不可視部分形状を1つ決定することを考えるが、この段階で位相情報と幾何情報との間に相互依存性が生じる。一方われわれはすでに、可視部分の形状復元処理における仮説的な拘束条件とそれに基づく復元形状の間に生じる相互依存性の問題を解決するために、エネルギー最小化を導入し、両者を同時に決定する手法について提案している。そこで不可視部分の予測問題にこの手法を導入し、処理に関わる制約条件を表現したエネルギー関数を緩和的に最小化することにより、可視部分の形状復元、および不可視部分の位相情報と幾何情報の決定を同時に実行することを試みる。

Completion of a Polyhedral Object by Inferring Invisible Faces in the Formulation of Energy Minimization

KOH KAKUSHO,[†] SEIICHIRO DAN,[†] NORIHIRO ABE^{††} and TADAHIRO KITAHASHI[†]

The human vision system has the recognition ability of a completed three dimensional (3D) shape by filling up the invisible surfaces of an object inferring from its visible surfaces. In this paper, we propose a method of emulating this ability for polyhedral objects. The 3D shape of an object is determined by the two attributes: topological relation of faces, edges and vertices, and geometrical factors such as shapes, orientations and positions of surfaces. Shape recovery of visible surfaces can be viewed as determining the 1D geometrical factor, *depth*, consistent with an input image which provides the topological relation and the 2D geometrical factors. On the other hand, no input image is given in the problem of filling up invisible surfaces of an object because the surfaces do not produce their image in optical projection. Thus, we claim some constraints to determine the topological relation. The process of filling up invisible surfaces based on them results in a *chicken and egg problem*, that is the topological relation and the geometrical factors of invisible surfaces depend on each other. We apply an energy minimization scheme to solve the problem. This method determines both of the topological relation and the geometrical factors concurrently.

1. ま え が き

物体認識のための一連の処理において、2次元画像から物体の3次元形状を獲得する処理はその中核とな

る。これに関する従来の研究では、光学投影によって失われた奥行き情報を、与えられた画像をもとにどのように復元するかが焦点であり、物体の陰や裏側のような画像に現われない不可視部分は対象とならなかつた。しかし、投影過程において物体の約半分が視点から見えない部分となることから考えても、重要な形状特徴が不可視部分に存在し、可視部分だけからでは物体認識が困難となる場合も予測される。このような事態に対処するためには、物体の不可視部分の形状を何

[†] 大阪大学産業科学研究所
 The Institute of Scientific and Industrial Research, Osaka University

^{††} 九州工業大学情報工学部
 Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology

らかの方法で獲得する必要がある。

この問題に関する従来の代表的なアプローチには、可視部分の形状を既知の物体モデルと照合するというトップダウン処理によるものと、視点の移動によって新たに不可視部分の画像を獲得するという能動的視覚の利用によるものとの2つがある^{1),2)}。しかし、人間は、幾何学的知識を用いて可視部分の形状から不可視部分の形状をある程度予測できる場合がある。このような処理を実現することは、人間の有する形状理解能力の実現として興味深いばかりでなく、視点移動の際のプランニングや、必要な画像が環境的な制約から獲得できない場合への対処等にも利用できると思われる。

このような物体の不可視部分の予測に関する従来研究には2次元形状に関するものが多く^{3),4)}、3次元形状を扱ったものはまだあまり見られない。単眼画像からの3次元形状復元手法や距離データからの3次元記述生成手法に関する議論の中で、それらの手法の不可視部分への適用を試みた興味深い例もいくつか報告されているが^{5),6)}、隠れた頂点の存在を既知としていたり、可視部分からの外挿によって予測可能な部分のみを対象としている等の点で、不可視部分予測のための手法とは一線を画するものである。本稿では、単一の多面体人工物を対象として、その単眼線画像から不可視部分を含む物体全体の3次元形状を獲得するための方法について検討する。

2. 位相情報決定のための制約条件

本稿で扱う物体は多面体のみであるため、その形状を有限個の面、辺、頂点より成る面モデルによって表現する。これによって記述される情報は、面、辺、頂点の数およびそれら間の接続関係等の位相情報と、個々の面、辺、頂点の定量的な方向や位置等の幾何情報とに分類できる。通常の可視部分の形状復元は、入力画像に含まれる位相情報および2次元分の幾何情報に対して、残りの1次元分の幾何情報、すなわち奥行き情報を割り当てる処理であるといえる。これに対し、不可視部分の予測では、幾何情報の割り当ての対象となる位相情報自体が与えられないため、幾何情報の決定に先立ってこの位相情報を決定する必要がある。そこで、このような位相情報の決定のために有効であると考えられる制約条件について検討する。

2.1 多面体条件

図1において実線で示したような物体を人間が見た

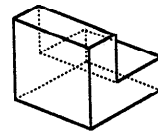


図1 不可視部分予測の例
Fig. 1 Probable invisible faces.

場合、さらに不可視部分の存在を感じるのは、与えられた部分だけではそれが閉じた立体とならず、点線で示したような補完的な部分を不可視部分として必要とするからであると考えられる。このように可視部分と不可視部分とを合わせたものが閉じた立体となるための制約条件を次のように定義する。

【制約条件1】(多面体条件)

すべての辺は2つの多角形面によって共有される。

これは位相幾何学における3次元多面体の定義に用いられている条件である。

さらに物体の各構成面が多角形面であるための条件を次のように定義する。

【制約条件2】(多角形条件)

- (1) 各面においてすべての頂点は2辺に共有される。
- (2) すべての辺は2頂点を端点に持つ。

なお本稿では、物体の構成面として穴のないものを考える。また、同じ2面が複数本の共有辺を持つような物体は考えない。このため、次の制約条件を利用する。

【制約条件3】(共有辺の一貫性)

任意の2面の共有辺数は、0または1である。

2.2 発見的な制約条件の利用

与えられた可視部分に追加することによって上述の制約条件1~3を満たし得る不可視部分は無数に存在する。したがって人間が一意的な不可視部分を予測する場合、何らかの発見的な制約条件を利用して可能な不可視部分をさらに制限していると考えられる。そこで、人間の予測するような不可視部分の導出に有効であると考えられる2つの発見的な制約条件の導入を考えると共に、これらの制約条件が利用可能なものに対象物体を限定して議論を進める。なお、これ以降の議論において、“面が可視である”とは、構成辺や構成頂点を除く少なくともその一部が可視である場合を、また“辺が可視である”とは、構成頂点を除く少なくともその一部が可視である場合を、それぞれ指すものとする。

第一に、すべての構成辺が不可視な面を多数含んだ不可視部分が一意に想起されることは経験的に稀である。これは、そのような面のとり得る幾何情報の自由度が高く、形状を絞りにくいため、これらを用いずにできるだけ単純な不可視部分を構成することが試みられるためであると考えられる。これに対応する制約条件として次のようなものを利用する。

〔制約条件 4〕(不可視面に対する制限)

すべての面は少なくとも一辺が可視である。

第二に、人間は可視部分が対象世界に多く見られる代表的な形状特徴を持つ場合に、それと同様の形状特徴を持つ不可視部分も最も可能性の高いものとして優先的に予測すると考えられる。本稿では、位相情報に関するこの形状特徴として、三面頂点を利用する。すなわち、実世界に存在する人工物体の頂点の多くは三面頂点であるため、可視部分を三面頂点のみから成るものに限定し、これに対する不可視部分として、三面頂点のみから成るものを要求する。

〔制約条件 5〕(三面頂点条件)

物体の頂点はすべて異なる 3 面およびそれらの 3 交線の交点である。

ただし、人間の予測する不可視部分形状にはこれらの制約条件が当てはまらないものも存在する。例えば正多角錐等は、4 面以上から成る頂点やすべての構成辺が不可視な面を含む不可視部分が予測される例の 1 つである。このような物体では部分形状間の対称性が明確であるため、形状特徴として制約条件 5 に代わってこの対称性が利用され、かつそれが制約条件 4 に優先しているものと考えられる。本稿では、このように利用すべき制約条件やそれらの間の優先順位が対象物体に応じて変化するような問題は取り扱わない。

2.3 制約条件に基づく不可視面の追加

上述の制約条件のうち、1, 4, 5 を利用すれば、次の 2 つの手続きによって、入力線画像の与える可視部分の位相情報から不可視部分を構成する面の数が決定できる(図 2)。ただし、自己遮蔽部分は T 形接合等に基づいてあらかじめ識別されているものとする。

(i) 面の生成

可視辺の中で制約条件 1 を満たさないもの(図 2 では $L_b \sim L_d$) に対して、新たに面(図 2 では F_c, F_a, F_a') を生成する。

(ii) 面の融合

制約条件 5 より、2 つの可視面に共有される頂点を介して隣接する 2 辺に対して生成された不可視面

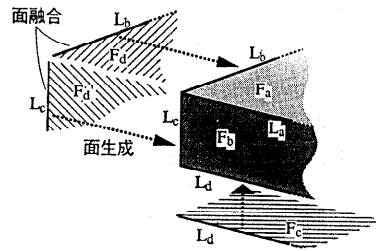


図 2 不可視面の生成と融合
Fig. 2 Addition and merger of invisible faces.

(図中では F_a, F_a') を同一とみなす。

制約条件 4 より、さらに新たな面を追加することはできないため、この段階で物体全体の構成面数 n が決定される。次にこれらの面に対して制約条件 1, 2, 3, 5 を満たす位相情報を求める必要があるが、ここで制約条件 3, 5 が満たされるように面、辺、頂点を、

F_i : i 番目の面

$L_{ij}(=L_{ji})$: 2 面 F_i, F_j の交線

$P_{ijk}(=P_{ikj}=P_{jik}=P_{jki}=P_{kij}=P_{kji})$:

3 面 F_i, F_j, F_k の交点 (3 辺 L_{ij}, L_{jk}, L_{ki} の交点) ($i, j, k=1, \dots, n$)

のように表すと、この表現によって記述される物体は制約条件 1 および 2 (1) も満足するので、結局、これ以降の処理は、頂点集合

$$\{P_{ijk} | (i, j, k) \in \Pi\}$$

$$\Pi = \{(i, j, k) | i, j, k=1, \dots, n, i \neq j, i \neq k, j \neq k\}$$

の中から残る制約条件 2 (2) を満たすものの組を選択するという一種の組み合わせ最適化問題に帰着される。

3. 幾何情報の利用

2.3 節で示した手続きによって不可視部分の位相情報が一意に定まるならば、次に可視部分と同様の形状復元処理を用いて、可視部分と整合する幾何情報を決定すればよく、位相情報と幾何情報の決定は逐次的に実行できる。実際、線画解釈の研究において、不可視面数が与えられればその他の位相情報および不可視頂点のスクリーン座標が一意に決定できる例が示されている⁷⁾。しかし、可視部分の形状によっては可能な不可視部分の位相情報が一意とはならない場合もある。例えば図 3 (a) のような面 $F_1 \sim F_3$ によって構成される物体の可視部分に対しては、2.3 節で示した手続きによって 5 つの不可視面 $F_4 \sim F_8$ が追加され、これらの組み合わせから得られる不可視部分の可能な位相情報

報は図3(b)~(f)のように複数となる(この段階ではまだ物体の幾何情報は問題としてはいないが、位相情報のわかりやすい例示のために、図では代表的な幾何情報を割り当てたものを示してある)。この中には、例えば図3(d)、(f)等のように、もし可視部分の幾何情報として直角のみから構成されるようなものが割り当てられた場合、不可視部分には多面体としての幾何情報の割り当てが不可能となるものも含まれる。

このように不可視部分の位相情報が複数となる場合、それらをすべて求め、得られたそれぞれに対して個別に幾何情報の割り当ての可否の判定を行うと共に、可能な場合には、その幾何情報を決定し、複数の不可視部分を導出するという方法も考えられる。しかし本稿では、(1)2.3節で述べた組み合わせ最適化問題の解として得られる位相情報の数は、追加される不可視面数の増加と共に増えていくこと、(2)人間が予測する不可視部分には幾何情報に基づく優先順位があると考えられること、(3)不可視部分の予測結果を視点移動等に利用する場合、次の段階で実際に移動できる場所は1カ所であり、複数の形状を考慮するとしても、それらの間の競合解消のために何らかの優先順位が必要であること、の3点を考慮し、位相情報に対して割り当て可能な幾何情報に関する制約条件および評価基準を併せて導入、利用することにより、最も優先順位の高い不可視部分形状を1つ決定することを考える。

3.1 幾何情報の決定のための制約条件

はじめに、与えられた位相情報に対して割り当て可能な幾何情報を決定することを考える。この処理は通

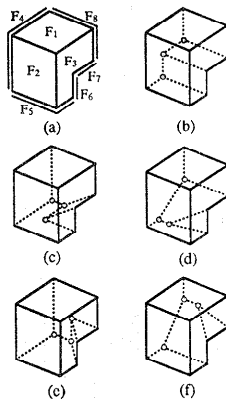


図3 不可視部分の位相情報が複数となる例

Fig. 3 An example of invisible faces with plural topologies.

常の可視部分の形状復元と同一であるため、可視部分と不可視部分を同列に扱う。

物体の幾何情報は、位相情報の定める物体の全構成頂点の3次元位置集合によって表現する。これによって表現される形状を多面体画像からの復元形状として妥当なものにするために、次の制約条件を用いる。

【制約条件6】(画像との整合性)

可視部分の復元形状のスクリーンへの再投影像は入力線画像と一致する。

【制約条件7】(平面性)

各面の構成頂点は同一平面上に位置する。

さらに、奥行き情報を復元するためには、対象物体の3次元形状に関する何らかの拘束条件が必要となる。対象物体に対する汎用性を考えれば、この拘束条件には、できるだけ多くの物体に見られる形状特徴を利用することが望ましいといえる。本稿では、人工物体の代表的な形状特徴であり、通常の可視部分の形状復元においても盛んに利用されている角の直角性を利用する^{8),9)}。ただし、物体のすべての角が直角であるとは限らないため、この拘束条件を、画像中の場所によってそれを適用すべきか否かが異なる一種の仮説として扱い、その適用内容は、復元形状に基づいて決定する¹⁰⁾。

【直角仮説】

“対象とする角は3次元では直角である。”

【制約条件8】(仮説の適用条件)

(1) 直角仮説が適用された角は直角となる。

(2) 直角とならない角には直角仮説は適用されない。

3.2 不可視部分の一意決定のための評価基準

次に不可視部分の一意決定のための評価基準について考える。前節において3次元形状復元のための拘束条件として直角性を導入した理由は、それが人工物の形状に多く見られる形状特徴であるためである。この前提より、直角仮説がより多く適用された復元形状ほど、対象世界により多く見られる形状であるということが出来る。そこで、可能な不可視部分形状に優先順位をつけるためにこのことを利用し、次のような評価基準を導入する。

【評価基準】(適用仮説数の多さ)

不可視部分は、直角仮説ができるだけ多く適用されるものを優先する。

4. エネルギー最小化

2章と3章で述べた制約条件と評価基準に基づく不可視部分の予測では位相情報の決定に幾何情報が必要となるが、逆に幾何情報の決定には位相情報が必要であり、この段階で両者の間に相互依存性が生じる。例えば、図3(a)に示した物体の可視部分に対応する不可視部分の例として、図4(a)~(c)のようなものが考えられ、3.2節で述べた評価基準に基づけば、これらは、(a)、(b)、(c)の順序で優先順位が高くなる。しかし、このうち(a)と(c)の位相情報は同じであり、図3(b)~(f)に示された位相情報の中から最適なものを選択するには、それらに割り当て可能な幾何情報を考慮する必要がある。さらに幾何情報は定量的な情報であり、ある位相情報に対して望ましい幾何情報の割り当てが可能か否かを判定する際には、これに必然的に含まれる誤差や雑音をある程度許容できる必要がある。

一方、われわれはすでに、3.1節で述べた仮説的な拘束条件に基づく形状復元処理において、仮説の適用内容とそれに基づく復元形状との間に生じる同様の問題を解決するためにエネルギー最小化を導入し、両者を緩和的かつ同時に決定する手法について提案している¹⁰⁾。エネルギー最小化では問題を関係指向的に記述するため、異種情報間の相互作用の実現が安易である。また誤差や雑音を含んだデータに関する制約充足処理の実現に有効であり¹¹⁾、多面体の幾何情報の決定にも利用されている¹²⁾。また、2章で位相情報の決定処理が組み合わせ最適化問題に帰着されており、これを何らかの方法によって解く必要があるが、エネルギー最小化はこのための手法としても利用できる。そこで、以下ではまず、位相的制約条件に基づく組み合わせ最適化問題を解くためのエネルギー関数を定義する。次にこれを以前より提案している可視部分の幾何情報決定のためのエネルギー関数と組み合わせ、その最小化によって、位相情報と幾何情報との同時決定を試みる。

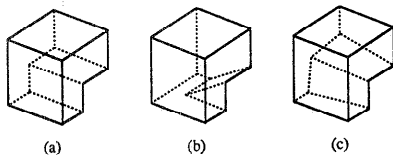


図4 位相情報が幾何情報に依存する例

Fig. 4 Dependence of the topology of an object on its geometrical shape.

4.1 位相的制約条件のエネルギー関数表現

頂点 P_{ijk} が物体を構成する頂点として選択される場合を1で、されない場合を0によって表現する変数 ε_{ijk} (これを位相変数と呼ぶ) を考え ($\varepsilon_{ijk} = \varepsilon_{ikj} = \varepsilon_{jik} = \varepsilon_{jki} = \varepsilon_{kij} = \varepsilon_{kji}$)、制約条件2(2)を次のエネルギー関数 E_{com} によって表現する。

$$E_{com} = \sum_{(i,j,k) \in \Pi} \varepsilon_{ijk} \left\{ \left(\sum_{\substack{s=1 \\ s \neq i,j,k}}^n \varepsilon_{ijs} - 1 \right)^2 + \left(\sum_{\substack{s=1 \\ s \neq i,j,k}}^n \varepsilon_{isk} - 1 \right)^2 + \left(\sum_{\substack{s=1 \\ s \neq i,j,k}}^n \varepsilon_{ijs} - 1 \right)^2 \right\} \quad (1)$$

このエネルギー関数は、ある頂点 P_{ijk} が存在する場合、そこから延びる3辺 L_{jk} , L_{ik} , L_{ij} に対してそれぞれもう一方の端点が決まっているときに最小値0をとる。

4.2 幾何学的制約条件のエネルギー関数表現

カメラモデルとして中心投影を考え、焦点距離 f 等のカメラパラメータを既知として、図5のようなカメラ中心座標系における頂点 P_{ijk} の位置ベクトルを、

$$\mathbf{p}_{ijk} = (p_{ijk}^x, p_{ijk}^y, p_{ijk}^z) \quad (2)$$

で表す ($\mathbf{p}_{ijk} = \mathbf{p}_{ikj} = \mathbf{p}_{jik} = \mathbf{p}_{jki} = \mathbf{p}_{kij} = \mathbf{p}_{kji}$)。

4.2.1 画像との整合性のエネルギー関数表現

物体形状が制約条件6を満たすためには、可視頂点 P_{ijk} 、それに対応する画像上の点 P_{ijk}' 、視点位置 O の3点は一直線上に並ばなければならない。そこで P_{ijk}' の位置ベクトルを \mathbf{p}_{ijk}' で表し、エネルギー関数 E_{img1} を次のように定義する。ただし、 \cdot はベクトルの内積を表す。

$$E_{img1} = \sum_{(i,j,k) \in V} \text{para}(\mathbf{p}_{ijk}, \mathbf{p}_{ijk}') \quad (3)$$

$$(V = \{(i, j, k) | P_{ijk} \text{ は可視頂点}\}) \quad (3)$$

$$\text{para}(\mathbf{v}_1, \mathbf{v}_2) = 1 - \text{orth}(\mathbf{v}_1, \mathbf{v}_2) \quad (4)$$

$$\text{orth}(\mathbf{v}_1, \mathbf{v}_2) = \frac{(\mathbf{v}_1 \cdot \mathbf{v}_2)^2}{|\mathbf{v}_1|^2 |\mathbf{v}_2|^2} \quad (5)$$

$\text{para}(\mathbf{v}_1, \mathbf{v}_2)$, $\text{orth}(\mathbf{v}_1, \mathbf{v}_2)$ は2つのベクトル \mathbf{v}_1 , \mathbf{v}_2 が平行となるとおよび直交するときそれぞれ最小値0をとる関数である。エネルギー関数 E_{img1} は、すべ

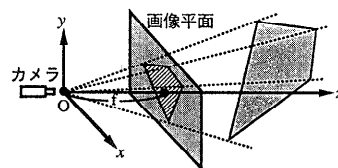


図5 カメラ中心座標系

Fig. 5 A camera-centered coordinate system.

での可視頂点 P_{ijk} ($(i, j, k) \in V$) に対して, その位置ベクトル \mathbf{p}_{ijk} および対応する画像中の点 P_{ijk}' の位置ベクトル \mathbf{p}_{ijk}' が互いに平行となるときに最小値 0 をとる (図 6 (a)).

さらに, 可視頂点 P_{ijk} を端点とする辺 L_{ij} の像 L_{ij}' が自己遮蔽によって画像中の位置 Q_{ij} で遮られ, 線分 $P_{ijk}'Q_{ij}$ として現われている場合, L_{ij} のもう一方の頂点の画像平面への再投影像は, この線分の延長線上になければならない. そこで Q_{ij} の位置ベクトルを \mathbf{q}_{ij} で表し, 次のようなエネルギー関数 E_{img2} を定義する. ただし, \times はベクトルの外積を表す.

$$E_{img2} = \sum_{((i,j),k) \in H} \sum_{s=1}^n \{ \varepsilon_{ijs} \text{orth}(\mathbf{p}_{ijs} - \mathbf{p}_{ijk}, \mathbf{p}_{ijk}' \times \mathbf{q}_{ij}) \}$$

$$H = \{((i,j),k) | L_{ij} \text{ は } P_{ijk} \text{ を可視頂点とする被遮蔽辺} \}$$

(6)

このエネルギー関数 E_{img2} は, 頂点 P_{ijk} から ε_{ijs} ($s=1, \dots, n, s \neq i, j, k$) の定める L_{ij} のもう一方の頂点に向かうベクトルが, 視点 O と線分 $P_{ijk}'Q_{ij}$ の作る平面の法線ベクトルに直交するとき最小値 0 をとる (図 6 (b)).

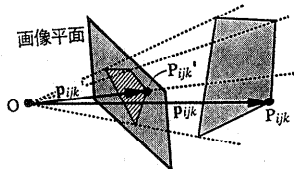
これらを用いて E_{img} を次式で定義する.

$$E_{img} = E_{img1} + E_{img2} \tag{7}$$

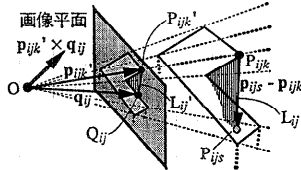
4.2.2 平面性のエネルギー関数表現

制約条件 7 は次のエネルギー関数 E_{flat} で表現する.

$$E_{flat} = \sum_{g=1}^n E_{flat}^g \tag{8}$$



(a) 可視頂点の場合
(a) For a visible vertex



(b) 被遮蔽辺の場合
(b) For the invisible vertex of an occluded edge

図 6 画像と復元形状との整合性
Fig. 6 Coincidence between an image and the projection of its recovered shape.

$$E_{flat}^g = \sum_{(h,i,j,k,l) \in N_g} \{ \varepsilon_{ghi} \varepsilon_{oij} \varepsilon_{ojk} \varepsilon_{okl} \text{orth}((\mathbf{p}_{ghi} - \mathbf{p}_{oij}) \times (\mathbf{p}_{ojk} - \mathbf{p}_{oij}), (\mathbf{p}_{okl} - \mathbf{p}_{ojk})) \}$$

$$(N_g = \{(h,i,j,k,l) | P_{ghi}, P_{oij}, P_{ojk}, P_{okl} \text{ は面 } F_g \text{ の連続 4 隣接頂点} \}) \tag{9}$$

E_{flat}^g は, ε_{oij} ($i, j=1, \dots, n, i, j \neq g, i \neq j$) の定める面 F_g の構成頂点のうち, 互いに連続して隣接する任意の 4 頂点 $P_{ghi}, P_{oij}, P_{ojk}, P_{okl}$ の定める 3 つのベクトル $\mathbf{p}_{ghi} - \mathbf{p}_{oij}, \mathbf{p}_{ojk} - \mathbf{p}_{oij}, \mathbf{p}_{okl} - \mathbf{p}_{ojk}$ が互いに共面であるとき (図 7) に最小値 0 をとり, E_{flat} はすべての面 F_g ($g=1, \dots, n$) に対して E_{flat}^g が 0 となるとき最小値 0 をとる.

4.2.3 直角仮説の適用のためのエネルギー関数

制約条件 8 は次のようなエネルギー関数 E_{hyp} として表現する. ただし, 頂点 P_{ijk} の周りの 3 つの角のうち, それぞれ面 F_i, F_j, F_k に含まれるものを $A_{ijk}^i, A_{ijk}^j, A_{ijk}^k$ で表し, 角 A_{ijk}^i に直角仮説が適用される場合を 1 で, されない場合を 0 によって表現する変数 (これを適用変数と呼ぶ) を r_{ijk}^i で表す.

$$E_{hyp} = \sum_{(i,j,k) \in \Pi} \varepsilon_{ijk} \{ R_{ijk}^i r_{ijk}^i + \text{th}_{\text{rect}}(1 - r_{ijk}^i) + R_{ijk}^j r_{ijk}^j + \text{th}_{\text{rect}}(1 - r_{ijk}^j) + R_{ijk}^k r_{ijk}^k + \text{th}_{\text{rect}}(1 - r_{ijk}^k) \} \tag{10}$$

$$R_{ijk}^i = \sum_{g=1}^n \sum_{h=1}^n \varepsilon_{ijg} \varepsilon_{ihk} \text{orth}((\mathbf{p}_{ijg} - \mathbf{p}_{ijk}), (\mathbf{p}_{ijk} - \mathbf{p}_{ijk})) \tag{11}$$

$$R_{ijk}^j = \sum_{g=1}^n \sum_{h=1}^n \varepsilon_{ojg} \varepsilon_{ojh} \text{orth}((\mathbf{p}_{ojg} - \mathbf{p}_{ijk}), (\mathbf{p}_{ijk} - \mathbf{p}_{ijk})) \tag{12}$$

$$R_{ijk}^k = \sum_{g=1}^n \sum_{h=1}^n \varepsilon_{okg} \varepsilon_{ohk} \text{orth}((\mathbf{p}_{okg} - \mathbf{p}_{ijk}), (\mathbf{p}_{ijk} - \mathbf{p}_{ijk})) \tag{13}$$

R_{ijk}^i は, $\varepsilon_{ijg}, \varepsilon_{ijg}, \varepsilon_{ihk}$ ($g, h=1, \dots, n, g, h \neq i, j, k$) の定める角 A_{ijk}^i の 3 つの構成頂点 $P_{ijg}, P_{ijk}, P_{ihk}$ の作る 2 つのベクトル $\mathbf{p}_{ijg} - \mathbf{p}_{ijk}, \mathbf{p}_{ihk} - \mathbf{p}_{ijk}$ (図 8)

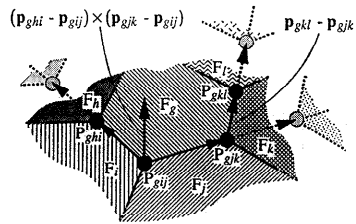


図 7 平面性
Fig. 7 Flattness of a surface.

が直交するとき最小値 0 をとる。E_{hyp} は、ε_{ijk} ((i, j, k) ∈ Π) の定める物体の構成頂点に対し、Rⁱ_{ijk}, R^j_{ijk}, R^k_{ijk} によって定義される角 Aⁱ_{ijk}, A^j_{ijk}, A^k_{ijk} の非直角性が、あらかじめ与えられたあるしきい値 th_{rect} より小さいものだけに直角仮説が適用され、かつ直角仮説が適用されているすべての角に対する Rⁱ_{ijk}, R^j_{ijk}, R^k_{ijk} の値が 0 となるとときに最小となる。

4.2.4 適用仮説数の多さに関するエネルギー関数

不可視部分の一意決定のための評価基準は次のエネルギー関数 E_{num} で表現する。

$$E_{num} = \sum_{((i, j, k), l) \in I} \{\varepsilon_{ijk}(1 - r_{ijk}^l)\}^2$$

$$I = \{((i, j, k), l) | l = i, j, k, \text{ 角 } A_{ijk}^l \text{ を構成する } 3 \text{ 頂点のうち少なくとも } 1 \text{ つが不可視}\} \quad (14)$$

4.3 最急降下法によるエネルギー最小化

4.1, 4.2 節で定義した 5 つのエネルギーを加えたものを最小化すべき全体のエネルギー関数として定義する。

$$E_{total} = C_{com}E_{com} + C_{img}E_{img} + C_{flat}E_{flat} + C_{hyp}E_{hyp} + C_{num}E_{num} \quad (15)$$

ただし、C_{com}, C_{img}, C_{flat}, C_{hyp}, C_{num} はそれぞれのエネルギーの重み付けを表す定数である。

このエネルギー関数を最小化するために、各変数の値を次のように最急降下方向に更新する (α, β, γ: 正定数)。ただし、離散時刻 t における ε_{ijk} 等の値を ε_{ijk}(t) 等によって表現する。なお、ε_{ijk}, rⁱ_{ijk} については 0, 1 の 2 値をとるようにするため、媒介変数 η_{ijk}, ρⁱ_{ijk} および式 (23) で定義される S 字型関数を利用している。ただしこの際、可視部分に関しては、その形状復元において位相情報はすでに決定しているため、これに対応する位相変数は 1 に固定して更新せず、また、その復元形状は不可視部分形状の影響を受けないことが当然であるため、式 (16)~(18) においては不可視頂点からの寄与分は無視するものとする。

$$p_{ijk}^x(t+1) = p_{ijk}^x(t) - \alpha \frac{\partial E_{total}}{\partial p_{ijk}^x(t)} \quad (16)$$

$$p_{ijk}^y(t+1) = p_{ijk}^y(t) - \alpha \frac{\partial E_{total}}{\partial p_{ijk}^y(t)} \quad (17)$$

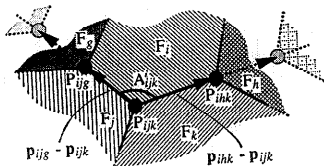


図 8 直角性
Fig. 8 Rectangularity of an angle.

$$p_{ijk}^z(t+1) = p_{ijk}^z(t) - \alpha \frac{\partial E_{total}}{\partial p_{ijk}^z(t)} \quad (18)$$

$$\rho_{ijk}^i(t+1) = \rho_{ijk}^i(t) - \beta \frac{\partial E_{total}}{\partial \rho_{ijk}^i(t)} \quad (19)$$

$$\eta_{ijk}(t+1) = \eta_{ijk}(t) - \gamma \frac{\partial E_{total}}{\partial \varepsilon_{ijk}(t)} \quad (20)$$

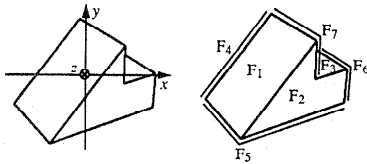
$$r_{ijk}^i(t) = \text{sigmoid}(\rho_{ijk}^i(t)) \quad (21)$$

$$\varepsilon_{ijk}(t) = \text{sigmoid}(\eta_{ijk}(t)) \quad (22)$$

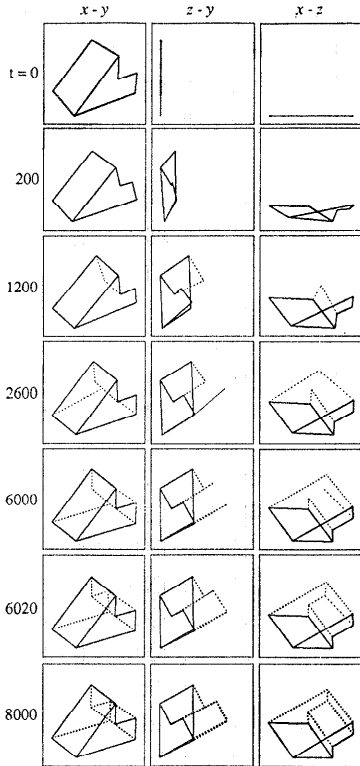
$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (23)$$

4.4 実験結果

図 9 (a) は矢印形の物体の人工的な中心投影線画像である。この線画像の持つ位相情報として領域、辺、頂点の関係が、幾何情報として各頂点のスクリーン座標が与えられるものとする。これをもとに、2.3 節で示した手続きによって不可視面 F₄~F₇ を生成し (図 9 (b)), 4.1~4.3 節で述べたエネルギー最小化を実行した。初期状態において、位相変数は可視頂点に対応するもの以外はすべて 0 とし、頂点の 3 次元位置は可視頂点を入力線画像の対応する頂点位置、それ以外は乱数によって定められた線画像の輪郭線内の中央付近のランダムな位置とした。さらに直角仮説の適用変数はすべての角に対して 0.5 とした。図 9 (c) は、エネルギー最小化の過程において次第に変化していく物体形状のうち代表的なものとして、t=0, 200, 1,200, 2,600, 6,000, 6,020, 8,000 のときのものを図 5 のカメラ中心座標系の x-y 平面、z-y 平面および x-z 平面にそれぞれ平行投影して表示したものである (図の各列において右方向をそれぞれ x 軸, z 軸, x 軸, 上方向をそれぞれ y 軸, y 軸, z 軸 としている)。実線は可視部分の復元形状を、点線は不可視部分の予測形状をそれぞれ表示しており、線の太さは、その両端の頂点の位相変数値の積に対応している。初期状態における 3 次元形状は入力線画像そのものであるため、その z-y 平面および x-z 平面への投影像は直線となっている。また、被遮蔽辺はもう一方の頂点が決まっていなかったため、表示されていない。各変数値の更新と共に、可視部分では奥行き情報が復元され、不可視部分では位相情報が決定されると同時に幾何情報も決定されていく (入力線画像および x-y 平面への投影図はそれぞれ中心投影、平行投影に基づくものであり、同一ではないことに注意されたい)。なお、仮説の適用変数の最終的な値は、rⁱ₂₅, rⁱ₂₇, rⁱ₄₅, rⁱ₄₇ が 0、それ以外は (ただし、位相変数の値が 1 である頂点の周りの角のみ) 1 となった。



(a) 入力画像 (b) 不可視面の追加結果
(a) An input image (b) Addition of faces



(c) エネルギー最小化の過程
(c) Tentative shapes in the energy minimization

図 9 実験結果

Fig. 9 An experimental result.

この実験において、エネルギー関数の各項の重み付け定数等は経験に基づいて、次のように設定している。

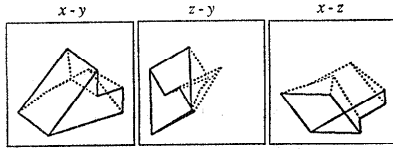
$$\begin{aligned} th_{rect} &= 0.005, c_{com} = 3.0, c_{img} = 8.0, c_{flat} = 0.3, \\ c_{hyp} &= 3.0, c_{num} = 0.1, \alpha = 100.0, \beta = 0.002, \\ \gamma &= 0.001 \end{aligned} \quad (24)$$

これらの値のうち、 α, β, γ は、3次元形状、仮説の適用変数、および位相変数が収束するまでの時間を左右し、これらを大きくするほど速やかに結果が得られる一方、あまり大きな値にすると結果が収束しなくなる（この実験ではさらに、幾何情報が位相情報の変

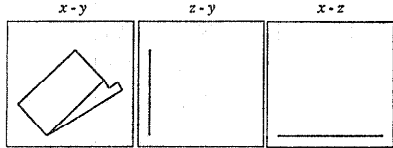
化に、また3次元形状が仮説の適用内容の変化にそれぞれ追従できるだけの時間的な余裕を与えるために、 β, γ を α に対して小さくしている）。 th_{rect} は直角仮説の適否の基準であり、これを大きくとれば任意の角に対して仮説が適用されるようになる。なお、実際にこの値を大きくとった場合の幾何情報の決定処理に関しては、すでに可視部分の形状復元問題を対象として同様の実験を行い、適用仮説に基づく定性的な3次元形状を復元結果として得ている¹³⁾。 $c_{com}, c_{img}, c_{flat}, c_{hyp}, c_{num}$ は、処理に導入された個々の制約条件をどの程度重視するかを決定する値であり、これらの比率は、式(15)で定義したエネルギー関数の解空間における超曲面の山や谷の大きさの比率を左右する。エネルギー関数は入力画像に対して定義されるため、この超曲面の本来の形状は入力画像によって決まる。したがって上記の定数のうちのどれをどの程度変化させた場合、得られた結果がどの程度の影響を受けるかは、厳密にはエネルギー関数の超曲面の本来の形状、すなわち入力画像によって異なる。しかしその一般的な傾向として、次のようなものが見られた。

(1) 不可視部分における位相情報の決定処理および幾何情報の決定処理にそれぞれ主として関与する定数 c_{com} および定数群 $c_{flat}, c_{hyp}, c_{num}$ との間の比率は2つの処理間の相互作用における両者の力関係を左右するため、得られる不可視部分形状は最も敏感にその変化の影響を受けた。具体的には、上記の値に対して $\pm 10\%$ 未満程度の比率の変化に対しては得られる結果は安定していたが、 c_{com} を $c_{flat}, c_{num}, c_{hyp}$ に対してそれ以上大きくするに従って、得られる不可視部分の位相情報が不可視頂点の初期位置のみに依存するようになり、得られた結果が平面性や適用仮説数等の幾何学的制約条件を満足しない結果が次第に生じやすくなった。この場合の例として、図9と同じ画像に対して式(24)で示した定数値のうち c_{com} だけを5.0に変え、5,000回の更新を行った結果を図10(a)に示す。また逆に c_{com} を $c_{flat}, c_{hyp}, c_{num}$ に対して小さくした場合には、何らかの位相情報が生成されはじめた段階で、その位相情報に含まれる不可視頂点の初期位置に対する幾何学的制約条件のエネルギーが、さらに位相情報の生成を促すための位相的制約条件のエネルギーを上回るようになるため、位相情報の生成が抑えられ、不可視部分自体が生成されなくなった。

(2) 主に幾何情報の決定に関与する定数 $c_{img}, c_{flat}, c_{hyp}, c_{num}$ 同士の間の比率の変化に対しては、



(a) 幾何学的制約条件が満足されなかった結果
 (a) A result violating the geometrical constraints



(b) 画像との整合性が無視された結果
 (b) A result violating the coincidence between an image and the projection of its recovered shape

図 10 重み付け定数の値を変えた場合の結果
 Fig. 10 Results with different values of the constants.

得られる結果は(1)に比べかなり鈍感であり、±30%程度の比率の変化に対して安定していた。しかしそれ以上の変化に対しては、まず、 c_{flat} を他に対して大きくとり過ぎると、初期状態である入力画像から3次元形状が立ち上がりにくくなり、小さく過ぎると、平面性条件を満たさない形状が復元されることが多くなった。また、 c_{img} を小さくすると、画像の整合性が満たされにくくなり、この値を0としたような極端な場合には、平面のままいくつかの角が直角になった形状が得られた。この場合の例として、図9と同じ画像に対して式(24)で示した定数値のうち c_{img} だけを0に変え、5,000回の更新を行った結果を図10(b)に示す。なおこのような場合には、物体を構成するすべての面が同じ向きであり、不可視部分を構成するいずれの3面に対してもその交点が定まらないため、不可視部分形状は決定できない。

以上のような現象は対象物体が複雑になるにつれて頻繁に見られるようになった。

5. むすびと考察

本稿では、多面体物体の不可視部分の予測問題を取り上げ、位相情報と幾何情報に関するいくつかの制約条件を利用した一意な不可視部分を決定する方法について提案した。また、この際に生じる位相情報と幾何情報の相互依存性を解決するため、問題をエネルギー最小化として定式化し、位相情報と幾何情報の同時決定を実現した。本手法において、可視部分の形状復元

処理と不可視部分の幾何情報の決定処理に利用される制約条件は一部を除いて同一であり、2つの処理にほぼ同じエネルギー関数が利用できたため、本稿で示した実験においては、これら2つの処理は同時に進行している。しかし、4.3節で述べたように可視頂点に関しては不可視頂点からの寄与分を無視しており、さらに実験結果からもわかるように、不可視部分の決定は可視部分の復元に比べて非常に遅いため、実行される処理は、実質的には、可視部分を復元した後に不可視部分を復元する処理とほぼ等価である。

以下では、本研究に関する今後の課題について考察する。

5.1 形状推論および視点移動への利用

われわれが日常、物体の3次元形状を他人に伝達しようとする場合、その線画を描いて示すことが多く、この際、単純な形状の物体についてはその可視部分のみを描いて済ませる場合も少なくない。線画の持つこのような物体形状の表現手段としての役割に注目すれば、人間が線画像からその3次元形状を推論する能力を実現することは、マン・マシン・インタフェースの向上のために重要な課題であるといえる¹⁴⁾。また、不可視部分の情報獲得のために視点を移動する場合、移動に伴う時間的、空間的な制約に関するさまざまなコストを考慮すれば、最適な視点位置を求めるための戦略が必要となる。ここで、対象物体が凸多面体の場合には、視点を物体に対して対象な位置に移動すれば事足りるが、そうでない場合には、物体形状によって最適な視点位置は異なるため、個々の形状に応じて適切な視点位置を求める必要がある。ところが、3次元空間において、ある面や辺、頂点などが観測可能な視点領域は、それらの面、辺、頂点の向きや位置、範囲などに依存し、不可視部分にはこの情報が存在しないため、適切な視点位置を定めるにはまず不可視部分の形状をある程度予測せざるを得ない。

これらの2つの問題の解決のために不可視部分の予測処理を実現しようとするとき、前者に関しては人間の利用している認知的な制約条件の導入が、また後者に関しては物体形状の存在確率に関する統計的な制約条件の導入が、それぞれ必要不可欠となる。ただし、認知的な制約条件が人間の過去の経験から獲得されたであろうことを考えれば、両者の制約条件にはある程度の関連があると考えられ、本稿では、いずれの観点から見ても妥当であると考えられる最も基本的な制約条件をいくつか導入した。しかし、これらが妥当であ

るとの根拠はわれわれ自身の内省の結果のみであり、今後、上記のいずれかの目的に応じ、利用する制約条件およびその適用順位などに関して、心理学や認知科学における研究成果の導入、もしくは物体の存在確率の評価に関する検討等を行う必要がある。

5.2 エネルギー最小化処理の改良

本稿ではエネルギー最小化処理自体が議論の焦点ではなかったため、これに関しては十分な検討を行っていないが、これに関してもいくつか問題点が残っている。まず、本手法でも行ったように、複数の制約条件を充足する解を得るためのエネルギー関数を、個々の制約条件を表現したエネルギー関数項の線形荷重和によって表現する場合、これらの項の重み付けをどのように決めるかということが問題となる。ここで、利用するそれぞれの制約条件の表現する内容の属性が異なる場合、おのおののエネルギー関数項の値は全く次元の異なるレベルでそれぞれの制約条件の非充足度を表すことになり、重み付け定数の設定問題はさらに複雑なものとなる。本稿では、少なくとも画像との整合性、平面性、直角性の3つの幾何学的制約条件に関しては、すべてベクトルの内積を利用してエネルギー関数を定義し、それらの値の意味する制約条件の非充足度の次元の統一に努めたが、依然この問題の根本的な解決には至っていない。また、処理の並列化を前提として、できるだけ局所的な変数間の関係に基づいてエネルギー関数を定義することを考えたため、提案した個々のエネルギー関数の中には、平面性のエネルギー関数等のように必ずしも対応する制約条件の厳密な表現となっていないものがあり、改良の余地があることは否定できない。

また本手法では、エネルギー最小化のための手法として最も単純な最急降下法を利用したが、これは最小値の探索を保証していない。本手法では、位相情報の決定のための組み合わせ最適化問題と幾何情報の決定のための制約充足問題を共にエネルギー最小化問題として定式化することにより、1つのエネルギー最小化処理によって位相的制約条件と幾何学的制約条件を同時に考慮することが可能となっているが、位相情報の決定のための組み合わせ最適化問題には、図3において示したようにもともと複数の解が存在するため、本手法で定義したエネルギー関数には、通常の可視部分の形状復元のためのエネルギー関数に比べて局所解が多く含まれる。このような局所解の回避のための手法はまだ確立されていないが、処理の複雑化と収束回数

の増加を許容すれば、模擬焼きなましなどの導入などが考えられる。

5.3 構成的推論の導入

エネルギー最小化は与えられた要素に対して適切なラベルを割り当てるといった解釈的推論の一種であり、これをもって、構成要素そのものを新たに生成するような構成的推論を実現することは不可能である¹⁵⁾。本稿で扱った不可視部分の予測処理では、三面頂点の条件等により対象物体を限定したため、位相情報の決定処理が最終的に解釈的推論に帰着され、これをエネルギー最小化の範囲内で扱うことが可能となっている。しかし、位相情報の決定問題は通常複数の解を持ち、これをエネルギー最小化によって解くことは前述のように局所解の問題を生じるため、必ずしも最善の策であるとはいえない。さらに、2.2節でも述べたように、多様な物体を扱うためには、より多くの種類の制約条件を互いの優先順位を動的に変更しながら利用する処理が必要であり、この実現を目指した場合に、位相情報の決定処理が解釈的推論の範囲内に収まらなくなることが予想される。これらのことを考慮すれば、今後、位相情報の決定のために構成的推論等の新たな枠組みを導入することが望ましいが、位相情報と幾何情報との相互依存性を考慮するためには、位相情報決定のための構成的推論と、幾何情報決定のための解釈的推論との相互作用が必要となり、これをどのようにして実現するかが今後の大きな課題となる。

謝辞 本稿に対し、筆者らは査読者の方々より適切なご指摘ならびに貴重なご助言を賜った。また本研究の成果の一部は、文部省科学研究費、稲盛財団助成金ならびに郵政省電気通信フロンティア研究による補助を受けた。ここに記して深謝する。

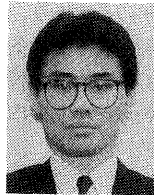
参 考 文 献

- 1) 谷内田正彦：ロボットビジョン，p. 250，昭晃堂，東京（1990）。
- 2) 谷内田正彦，石黒 浩：ロボットの視覚，人工知能学会誌，Vol. 5，No. 6，pp. 720-730（1990）。
- 3) 安田浩之，安藤和久，大西 昇，杉江 昇：物理的に存在しない輪郭線の抽出，電子情報通信学会論文誌（D-II），Vol. J73-D-II，No. 6，pp. 906-913（1990）。
- 4) 島谷 明，鎧沢 勇：重なり線図形における補完を含む解釈候補の自動抽出，電子情報通信学会技術報告，PRU 91-133（1992）。
- 5) 光本浩士，岡崎耕三，田村進一，梶見直樹，副井裕：面对称性に基づく単眼視からの3次元復元，電子情報通信学会論文誌（D），Vol. J71-D，No.

- 10, pp. 2032-2040 (1988).
- 6) 津戸明宏, 八木康史, 谷内田正彦: 移動視により得られたワイヤーフレームモデルからの空間構成の理解, 情報処理学会研究報告, CV 73-9 (1991).
- 7) 杉原厚吉: 多面体線面の代数的解析法, 電子通信学会論文誌 (D), Vol. J 62-D, No. 3, pp. 201-208 (1979).
- 8) 尺長 健, 金子 博: 透視角度変換一角度情報からの形状情報復元一, 電子情報通信学会論文誌 (D), Vol. J 70-D, No. 10, pp. 1931-1940 (1987).
- 9) 前原一夫, 川島孝夫, 金谷健一: 直交性仮説による多面体の3次元形状復元, 電子情報学会論文誌 (D-II), Vol. J 72-D-II, No. 6, pp. 887-895 (1989).
- 10) 角所 考, 淡誠一郎, 安部憲広, 北橋忠宏, 三宅誠: 画像の3次元解釈へのニューラルネットワークの適用, 電子情報通信学会論文誌 (D-II), Vol. J 73-D-II, No. 8, pp. 1297-1302 (1990).
- 11) 横矢直和, 坂上勝彦: 画像理解と最適化原理, 電子情報通信学会誌, Vol. 74, No. 4, pp. 326-334 (1991).
- 12) Marill, T.: Emulating the Human Interpretation of Line Drawings as Three-dimensional Objects, *Int. J. Computer Vision*, Vol. 6, No. 2, pp. 147-161 (1991).
- 13) Kakusho, K., Dan, S., Abe, N. and Kitahashi, T.: Shape Recovery and Error Correction Based on Hypothetical Constraints by Parallel Network for Energy Minimization, *Lecture Notes in Computer Science*, Vol. 654, pp. 159-178 (1992).
- 14) Leclerc, Y. G. and Fishler, M. A.: An Optimization-Based Approach to the Interpretation of Single Line Drawings as 3D Wire Frames, *Int. J. Computer Vision*, Vol. 9, No. 2, pp. 113-136 (1992).
- 15) 松山隆司: 画像理解における推論方式, 人工知能学会誌, Vol. 4, No. 1, pp. 21-29 (1989).

(平成4年12月14日受付)

(平成5年7月8日採録)



角所 考

昭和39年生. 昭和63年名古屋大学工学部電気学科卒業. 平成5年大阪大学大学院工学研究科通信工学専攻博士課程修了. 工学博士. 現在, 日本学術振興会特別研究員, 米国スタンフォード大学ロボット工学研究所客員研究員. 神経回路網および画像理解に関する研究に従事. 第41回情報処理学会全国大会奨励賞受賞. 電子情報通信学会, 人工知能学会各会員.



淡 誠一郎 (正会員)

昭和35年生. 昭和58年大阪大学工学部通信工学科卒業. 昭和63年同大学大学院博士課程修了. 同年, 同大学産業科学研究所助手, 現在に至る. 工学博士. 定理証明, 論理プログラミング, 画像理解などの研究に従事. 高次推論, マルチメディア情報処理などにも興味を持つ. 電子情報通信学会, 人工知能学会, IEEE 各会員.



安部 憲広 (正会員)

昭和21年生. 昭和44年大阪大学基礎工学部電気工学科卒業. 昭和49年同大学大学院博士課程修了. 同年同学部制御工学科に勤務, 同助教授, 同大学産業科学研究所助教授を経て, 平成3年4月より九州工業大学情報工学部機械システム教授. 工学博士. 人工知能, 特に機械システムへの人工知能技術の適用について研究中. 他に言語理解と画像理解の融合, エキスパートシステムについても研究を行っている. 著書「Prolog プログラミング入門」(共立出版), 「Lisp」(培風館), 「Prolog と AI」(近代科学社)など. 電子情報通信学会, 人工知能学会, AAAI 各会員.



北橋 忠宏 (正会員)

昭和14年生. 昭和37年大阪大学工学部通信工学科卒業. 昭和43年同大学大学院博士課程修了. 同年同大学基礎工学部助手, 助教授を経て, 昭和54年豊橋技術科学大学工学部情報工学科教授. 現在, 大阪大学産業科学研究所教授. 3次元物体・3次元運動認識のための視覚システム, 自然言語処理, 学習・推論機構に関する研究に従事. 工学博士. 電子情報通信学会, IEEE, 人工知能学会各会員.