TR-273

線画の奥行き知覚のモデル

岡 夏樹

June, 1987

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191～5
Telex ICOT J32964

Institute for New Generation Computer Technology

情報処理学会
知識工学と人工知能研究会
53-4
1987.7.10

線画の奥行き知覚のモデル

岡 夏樹

新世代コンピュータ技術開発機構
oka@icot.junet

我々は，一枚の線画から物体の３次元形状をほぼ定量的に思い浮かべてしまう。いろいろな物体についてのこうした解釈を統一的に説明できる「尤度モデル」を提案する。本モデルでは，ある線画に対して幾何学的に可能な物体のうち，一番もっともらしい物体が知覚される。尤度は，「物体の尤度」と「見え方の尤度」の積で表わされる。「見え方の尤度」という新しい概念を取り入れることにより，規則的な形を含まない物体の定量的解釈ができるようになった。「見え方の尤度」の近似的な尺度として「投影面積の尤度」を提案し，また，それらと「投影面積」との相関関係を示した。３次元形状の伝達を意図した線画についても考察した。

Cognitive Model for Quasi-quantitative Interpretation of Line Drawings

Natsuki OKA

Institute for New Generation Computer Technology
4-28, Mita 1-chome, Minato-ku, Tokyo 108 JAPAN
oka%icot.jp@relay.cs.net
oka@icot.jp

The three-dimensional shape of an object can be perceived quasi-quantitatively from a single two-dimensional line drawing. The new inclusive cognitive model proposed in this paper explains this type of interpretation for various kind of object. It interprets a drawing as one of many geometrically possible objects which is most likely to produce the drawing. The likelihood of producing the drawing depends on both object likelihood and view likelihood. The new concept, view likelihood, enables the quasi-quantitative interpretation of irregular-shaped objects. The likelihood of the projected area is proposed as an approximation to view likelihood. This paper also examines line drawings which communicate the three-dimensional shape of objects.

1．はじめに

本論文では，平行投影された線画のうち，投射線と投影面が垂直なものを扱う。図１に示すような１枚の線画に幾何学的に対応する３次元物体は，図２にいくつかの例を示したように無限にある。しかし，我々は図１のような線画を見ると，（思い浮かべる物体に多少個人差はあるだろうが）ある一つの３次元物体を思い浮かべる。本論文は，人間のこのような線画解釈の一つのモデルを提案する。

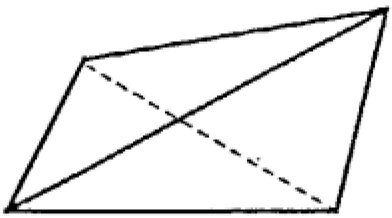


図１

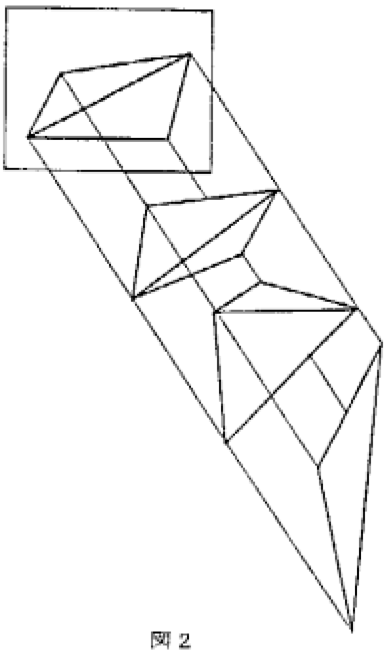


図２

2．関連する研究

線画解釈の研究としては，Huffman〔Huffman〕，Clowes〔Clowes〕，Waltz〔Waltz〕等によるラベル付けがよく知られている。図３に示すように，線画中の線を凸の稜(+)，凹の稜(-)等のように"定性的に"解釈する。Mackworth〔Mackworth〕は，gradient spaceを利用した解釈法を提案したが，それによる解釈もラベル付けと同様定性的である。

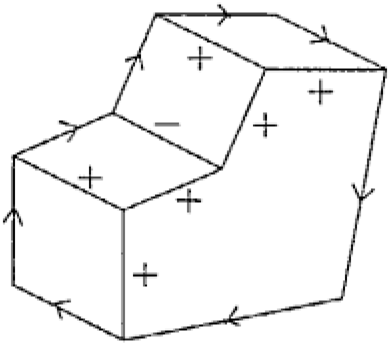


図３

一方人間は，より定量的な解釈を行なっている。すなわち，稜の凹凸にとどまらず，かなり定量的に（例えば，各面がどのくらいの角度で交わっているか）３次元形状を知覚している。＊　本論文で提案するモデルは，このようなより定量的な解釈をめざしたものである。

＊　解釈される線画により，どの程度はっきり３次元形状が知覚されるかということや，どのくらい解釈に個人差があるかということは，変動するだろう。例えば，直方体のような規則的な形状の物体の線画であれば，３次元形状が明確に思い浮かべられ，その解釈には，個人差が少ないだろう。

Roberts のシステム〔Roberts〕は，次の仮定のもとに線画中の物体の位置と大きさを定量的に解釈した。一つは，線画中の物体は，あらかじめ与えてあるmodel（３次元座標値が与えられている）と相似であること，もう一つは，物体は何かにより支えられていることである。

Roberts の方法でいろいろな物体を解釈するためには，多数のmodel が必要になる。また，model を増やすとともに線画が多数のmodel とマッチする可能性が出てきて，それらからどれを選ぶかという問題が生じる。本論文での主な興味は，このような幾何学的に可能な多数の物体のうちでどれを知覚するかというところにある。

Liardet らのシステム〔Liardet〕は，図４のように座標軸の投影が指定された線画を扱った。画像上で座標軸と平行な線分は，実際のシーンにおいても平行であると仮定して定量的に解釈する。３次元座標値の決まらない頂点があれば，ユーザにconstruction line の入力を促す。construction line はz軸と平行な線分である。例えば図４では，ユーザは点線で示すようにconstruction line を指定しなければならない。

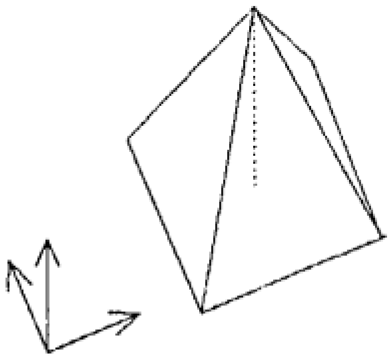


図４

広谷らのシステム〔広谷〕も，座標軸の投影が指定された線画を扱った。画像上で座標軸と平行な線分は，シーンにおいても平行であると仮定し，さらに物体はxy平面上に置かれていると仮定して定量的に解釈する（図５）。この仮定だけでは，角錐や斜角柱の解釈は定まらない。

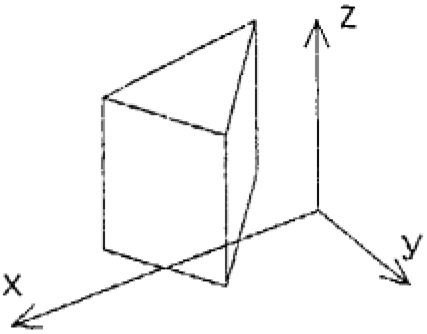
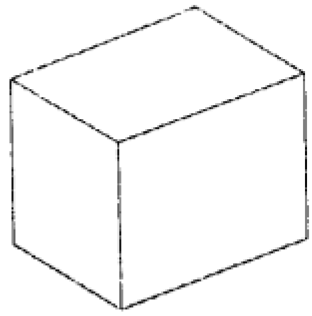
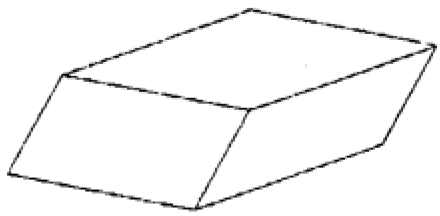


図５

Kanade〔Kanade〕は，平行線のヒューリスティクス（画像上で平行な２本の線分はシーンにおいても平行である）と斜対称のヒューリスティクス（画像上の斜対称な形はシーンにおいては対称な形である）を提案した。斜対称とは，対称軸に必ずしも垂直でない線に沿った対称を言う。これらのヒューリスティクスにより，図６(a) のような線対称な形を含む物体の線画を定量的に（この例では，直方体であると）解釈できる。しかし，図６(b) のように線対称な形を含まない物体の線画の定量的な解釈はできず，直方体ではないことが言えるだけである。



(a)



(b)
図６

人間は，construction line のような余分な指定なしに，線対称な形を含まない物体の線画（例えば図１や図６(b)）でもある程度定量的に奥行きを知覚している。本論文では，このような知覚のモデル化を試みる。

３．尤度モデル

これまでの研究では，見慣れた物体のとき（例えば，Roberts のmodel ）や規則的な形の物体のとき（例えば，Kanadeのヒューリスティクス）の解釈の方法が提案されてきた。一方，人間は見慣れた物体や規則的な物体でないときも線画をほぼ定量的に解釈できる。本論文で提案する尤度モデルは，これら

1）見慣れた物体のとき
2）規則的な形の物体のとき
3）その他

の場合の解釈を統一的に説明できる。

尤度モデルでは，以下に定義する尤度Lにしたがって線画を解釈する。すなわち，ある線画に対して一番もっともらしい（尤度が高い）物体が，通常は知覚され，尤度の低い物体は意識的に思い浮かべることはできるものの，通常は知覚されない。なお，この考え方は，統計学において，最尤推定法として知られているものと対応する。

本モデルでは，尤度Ｌｔを次のように定義する。

Ｌｔ（線画Ｄ；物体Ｏ，視点Ｖ）
＝Ｌｇ（線画Ｄ；物体Ｏ，視点Ｖ）
×Ｌｏ（物体Ｏ）
×Ｌｖ（視点Ｖ｜物体Ｏ）

すなわち，Ｌｔ（total likelihood＝全体の尤度）（与えられた"線画Ｄ"に対する，"視点Ｖ"から見た"物体Ｏ"の尤度）は，以下に（詳しくは次節以後で）説明するＬｇ，Ｌｏ，Ｌｖの３項の積で表わされる。単に尤度と言ったときには，このＬｔを指すことにする。Ｌｇ（geometrical likelihood＝幾何学的尤度）は，与えられた"線画Ｄ"に対する，"視点Ｖ"から見た"物体Ｏ"の幾何学的な対応（投影・逆投影の関係）のもっともらしさを表わす。Ｌｏ（object likelihood ＝物体の尤度）は，"物体Ｏ"の存在のもっともらしさを表わす。Ｌｖ（view likelihood ＝見え方の尤度）は，"物体Ｏ"における"視点Ｖ"からの見え方のもっともらしさを表わす。なお，式中で，；の左側は観測値，右側は推測値である。｜は条件付き確率を表わす。

尤度関数には，個人的な経験（どのような＜画像，立体＞の対を見てきたか）と期待（その場面で何が見えることを予期するか）が反映される。これは，「これまでに経験した頻度が多い物体に見えやすい」という自然な考え方に基づくが，これに対して，「人間は初めて見る線画でも奥行きを知覚できる」という反論が予想される。しかし，「初めて」見る線画が存在し，人間がそれを解釈できることは，本論文に述べるような確率的なモデルの妥当性を否定するものではない。なぜならば，人間の情報処理の特質の一つとして，「おおよそのマッチングをとることができる」ことが挙げられるからである。すなわち，「初めて」見る線画でも，それと似た線画を見た経験はあると思われる。なお，本論文においては，Ｌｖを物体の投影面積の尤度で近似することを後述するが，これは「おおよそのマッチング」の一つの実現法になっている。

３．１．幾何学的尤度（Ｌｇ）

線画と物体の幾何学的な対応関係（投影・逆投影の関係）に対する人間の判断は，厳密なものではない。したがってＬｇは，厳密に幾何学的に対応する物体で極大値をとり，その周辺でなめらかに減少する。

しかし，このような曖昧さを扱うことは本論文の主題ではないので，以後は，簡単のため幾何学的に可能か不可能かのどちらかであるとする。また，幾何学的な対応関係を判定するメカニズムについては，本論文の範囲外とし，以下では，「幾何学的に可能な物体のうちで我々はどれを知覚するか」という問題に絞って論じる。

３．２．物体の尤度（Ｌｏ）

Ｌｏは物体の存在のもっともらしさを表わす。Ｌｏには，個人的な経験（どのような物体を見てきたか）と期待（その場面で何が見えることを予期するか）が反映される。

線画中の物体は与えられたmodel と相似であるとする仮定[Roberts]や，対称形や平行であることの仮定[Kanade]は，次のようなＬｏを考えたことに相当する。

$$
\mathrm{Lo}(\mathrm{O}) = \begin{cases} k\ (\text{定数}) & \text{Oは仮定された物体} \\ 0 & \text{その他} \end{cases}
$$

このような方法（２値のＬｏだけによる制限）によって解釈が一意に決まるのは，仮定された物体の中でちょうど一つが幾何学的に可能なときだけである。これはかなり限られた場合である。本論文では，以下に述べるように，Ｌｖと多値のＬｏの組み合わせにより，より一般の場合（例えば図１や図６(ｂ)）に解釈が一意に決まることを自然に説明できるモデルを提案する。

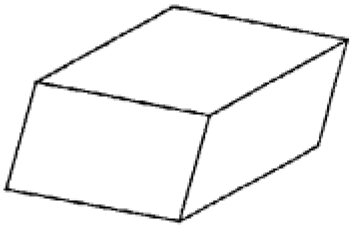
３．３．見え方の尤度（Ｌｖ）

Ｌｖは，ある物体についての見え方のもっともらしさを表わす。もっともらしい見え方とは，「大体そのように見える」ことが多い見え方である。あるいは，少し視点をずらしても見え方があまり変わらない視点からの見え方である。ある視点から見た（のと大体同じ様に見えた）経験が多いと，その視点からの見え方の尤度は高い。

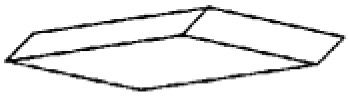
なんらかの原因で，特定の視点から見ることが多い物体では，その視点からの見え方の尤度が高くなる。例えば，車は路上にあるものを水平よりやや上方から見る事が多いから，そのような視点からの見え方の尤度が高くなる。普段机の上に置いてあるものは，斜め上から見るような見え方の尤度が高くなる。Roberts［Roberts］や広谷ら［広谷］の「物体は支えられている」という仮定に基づく解釈は，このような特定の視点からの見え方の尤度と関係がある。

上述のような特定の片寄りのない，一般的な場合について以下に考察する。なお，Liardetら［Liardet］，広谷ら［広谷］やKanade［Kanade］の「画像上で平行なものは，シーン上でも平行である」という仮定は，特殊な視点からの見え方（実際は平行でないものが，たまたま平行に見える）の尤度が低いことに対応する。また，「他に制約がないなら，斜対称のヒューリスティクスを満足する中で，最小の傾きの面が一番もっともらしい［Kanade］」という主張は，ある一つの面についての見え方の尤度を述べていると見なせる。

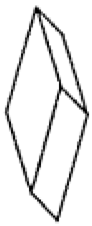
まず，一般的な場合の見え方の尤度に対する直感的な理解を助ける例を示す。図７は，あるひとつの平行六面体をいくつかの視点で描いた線画である。同様に，図８は，あるひとつの三角錐をいくつかの視点で描いた線画，図９は，あるひとつの直方体をいくつかの視点で描いた線画である。



(a)



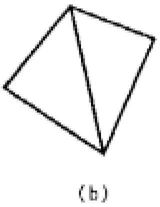
(b)



(c)

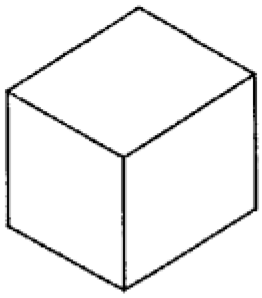
図７

(a)

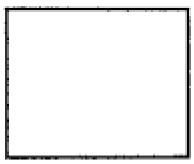


(b)

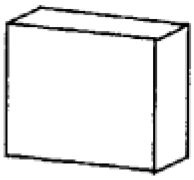
図８



(a)



(b)



(c)

図９

一般に，図８のような細長い物体（極端な場合は線分）では，その軸と垂直な方向からの見え方の尤度が高く，図７のような平たい物体（極端な場合は平面）では，その平たい広がりと垂直な方向からの見え方の尤度が高い。図９のような等方的な物体（極端な場合は球）では，細長い物体や平たい物体と比べて，見え方の尤度は大きく，その視点による変化は小さい。

ここでは，このような一般的な場合の「見え方の尤度」を近似し，計算可能な尺度を導入する。それは，「投影面積の尤度」である。すなわち，見え方の尤度を投影面積がその視点での値に近い値である確率で近似する。この尺度によると，個々の頂点，辺，面の様子が無視されてしまう欠点があるが，全体的な様子はよく捉えているし，ある意味で「おおよそのマッチング」を表現していると言える。

また，ある物体における「投影面積の尤度」と「投影面積」には正の相関関係がある。投影面積が最大になるところで視点による投影面積の変化はゆるやかで，投影面積の尤度は大きい。結局，「見え方の尤度」と「投影面積」には正の相関関係があることになる。なお，例えば直方体の場合，最大の面を正面から見た図（図９(b)）より，斜めから見た図（図９(a)）の方が投影面積が大きいことをここで注意しておく。

岡が提案した「投影面積のヒューリスティクス」〔岡〕は，尤度モデルの考え方では，見え方の尤度（投影面積で近似したもの）に関連した解釈法であると位置付けることができる。以下このヒューリスティクスを簡単に紹介する。

<投影面積のヒューリスティクス>

（他に特に手掛かりが無ければ）物体は投影面積が最大になる方向から投影されていると仮定して解釈せよ。

なお，後にも述べるが，このヒューリスティクスによる解釈は，３次元形状の伝達を目的に描かれた線画を対象とした場合には，正しく解釈できる可能性が上がる。このヒューリスティクスの適用例を２つ挙げる〔岡〕。

一般の四面体は図10のように表わされるが，この図において投影面積が極大であることより，

z（A）＝z（C）　　かつ

z（B）＝z（D）

であることが言える。ここに投影方向をｚ軸方向とし，ｚ（A）は点Aのｚ座標を示す。また投影面積が最大であることより，

｜z（A）－z（B）｜≦limit

であることが言える。ここにlimit は，線画から計算できる値である。

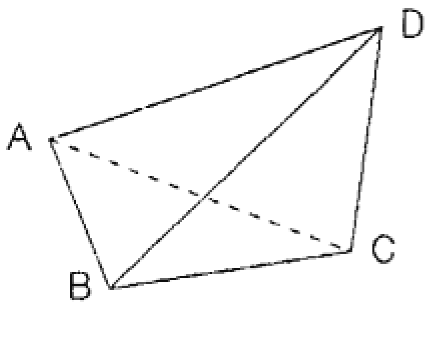


図10

図11は平行六面体であるとすると，投影面積が極大であることより，

z（A）＝z（C）＝z（H）　　かつ

z（B）＝z（E）＝z（G）

であることが言える。また投影面積が最大であることより，

｜z（A）－z（B）｜≦limit

であることが言える。

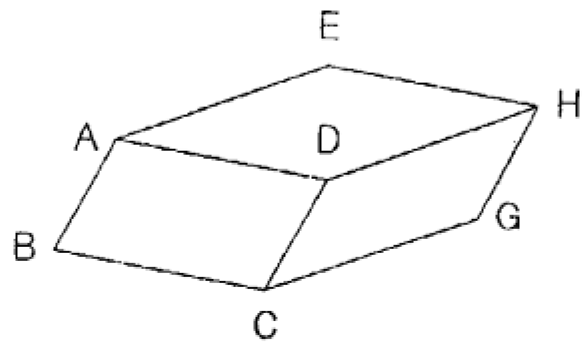


図11

投影面積のヒューリスティクスによる解釈では，このようにまだ曖昧さが残っている（あるひとつの物体であると解釈が決まってはいない）が，次節では尤度モデルにより曖昧さを残さない解釈ができることを示す。

３．４．尤度モデル全体の働き

まず，物体の尤度(Ｌｏ)の働きと見え方の尤度(Ｌｖ)の働きの関係を明らかにする。幾何学的に可能な物体の中にＬｏの値が飛び抜けて高いものが一つだけあれば，Ｌｖの値に関係なく解釈がその物体であると一意に決まる。Ｌｖの値の変動はそれほど大きくないからである。これは，見慣れた物体や規則的な形の物体の解釈に相当し，以下に挙げる<例１>はこの場合である。これまでの線画の定量的解釈の研究では主にこの場合が扱われてきた。

そうでないときは，ＬｏとＬｖの積により，解釈が決まる。<例２>はこの場合である。Ｌｏの値があまり変わらない場合は，解釈においてＬｖの値が支配的になる。<例３>はこの場合である。これまでの研究ではＬｖが考慮されなかったので，<例２>や<例３>のような線画の定量的解釈はできなかった。

A．例１

幾何学的には，図12の物体はいろいろな六面体でありうる。しかし，直方体のＬｏは，他の六面体のＬｏよりずっと高く，幾何学的に可能な直方体は一つだけであるので，Ｌｖには無関係に，解釈はその直方体に決まる。

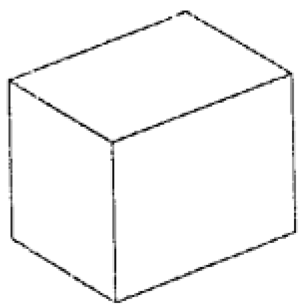


図12

B．例２

幾何学的には，図13の物体はいろいろな六面体でありうるが，直方体ではありえない。平行六面体のＬｏは他の可能な六面体のＬｏより大きいが，可能な平行六面体は無数にあるし，各平行六面体間ではＬｏの値に大きな変化はない。したがって，どの平行六面体が知覚されるかは，主にＬｖで決まる。Ｌｖの値は，

Ｌｖ（ある見え方｜等方的な物体）
＞Ｌｖ（投影面積の大きい方向からの見え方｜等方的でない物体）
＞Ｌｖ（投影面積の小さい方向からの見え方｜等方的でない物体）

大体このような関係があるので，投影面積のヒューリスティクスを満たす平行六面体のうち，奥行きの大きいものほど（等方的な物体ほど）知覚されやすい。

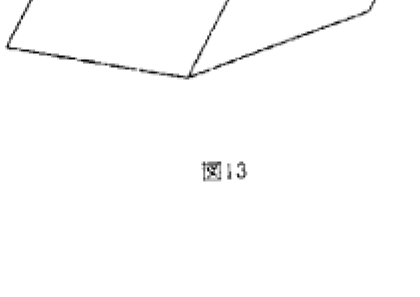


図13

C．例３

図１４に幾何学的に対応する四面体は無数にある。これらの物体のＬｏの値は，互いにあまり変わらない。したがって，解釈においてはＬｖが支配的である。＜例２＞と同様に，投影面積のヒューリスティクスを満たす四面体のうち，奥行きの大きいものが知覚されやすい。

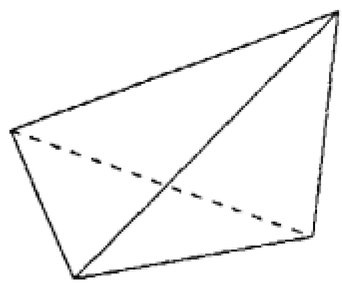


図14

4．　３次元形状の伝達を目的とした線画

Ｌｖに頼った解釈をした場合は，解釈の結果が線画のもとになったシーンと必ずしも一致するわけではない*。対象を３次元形状の伝達を目的とした線画に限定すると，Ｌｖに頼った解釈をした場合でも，最尤モデルによる解釈がもとになったシーンと一致する可能性が大きく増す。これは，なるべく正確に形状を伝達するために，解釈する人間の特性を考慮して線画が作成されるからである。すなわち，３次元形状伝達のための線画は次のように描かれる。

・もっともらしい視点で描く
（見慣れた視点で描く，
偶然規則的な形に見えるようには描かない）
・伝達情報量が大きくなるように描く
（投影面積を大きくする，
奥行きを小さくする）

尤度の高い視点と情報を多く伝える視点が一致していることは興味深い。

＊ただし，これは人間による解釈でも同様であろう。本論文の最尤モデルは，線画のもとになったシーンとどの程度合うかにより評価されるべきではなく，人間による解釈とどの程度合うかにより評価されるべきである。

5. 結 論

本論文では次のことが示された。

- 物体の尤度（Ｌｏ）と見え方の尤度（Ｌｇ）に基づき，人間によるいろいろな物体の線画の準定量的解釈を統一的に説明できる認知モデルを提案した。

- 見え方の尤度（Ｌｇ）という新しい概念により，規則的な形を含まない物体の準定量的解釈ができるようになった。

- 見え方の尤度（Ｌｇ）の近似的な尺度として投影面積の尤度を提案し，また，それらと投影面積との相関関係も示した。

- ３次元形状の伝達を意図した線画についても考察した。

ＣＡＤシステムや，アニメーション・システムの入力において，正確な形状が必要とされない用途においても，現状では大体の形状を入力する手段が無いから，完全な情報を指定しなければならない。本論文で提案した尤度モデルは，このような大まかな３次元形状の入力に応用できる。

今後の主な課題は次の通りである。

- 尤度関数を具体的に決める。

- 尤度関数を＜線画，立体＞の対から学習させる。

［謝　辞］

有益な助言を頂いた群馬大学の松浦勉氏，松下電器産業の野田克彦客員，加藤文之氏，ＩＣＯＴの古川康一次長ならびに世木博久氏に感謝します。

［参考文献］

[Clowes] Clowes, M. B. 1971. On seeing things. Artificial Intelligence 2: 79-116.

[Huffman] Huffman, D. A. 1971. Impossible objects as nonsense sentences. In B. Meltzer and D.Michie (Eds.), Machine Intelligence 6. New York: Elsevier, 295-323.

[Kanade] Kanade, T. 1981. Recovery of the three-dimensional shape of an object from a single view. Artificial Intelligence 17: 409-460.

[Liardet] Liardet, M. et al. 1978. Input to CAD systems: two practical examples. In Latombe (Ed.), Artificial intelligence and pattern recognition in computer aided design. North-Holland, 403-427.

[Mackworth] Mackworth, A. K. 1973. Interpreting pictures of polyhedral scenes. Artificial Intelligence 4: 121-137.

[Roberts] Roberts, L. 1965. Machine perception of three-dimensional solids. In J. Tippett (Ed.),Optical and electro-optical information processing. Cambridge, Mass.: MIT Press, 159-197.

[Waltz] Waltz, D. 1975. Generating semantic descriptions from drawings of scenes with shadows. In P. Winston (Ed.), The psychology of computer vision. McGraw-Hill, New York, 19-92.

[岡] 岡 夏樹. 1986. 暗黙の了解事項に基づく線画の定量的解釈. 情報処理学会第32回全国大会. 1433-1434.

[広谷] 広谷 豊史ら. 1981. ３次元図形の入力の研究. 日本建築学会第３回電子計算機利用シンポジウム. 343-347.