

エネルギー最小化原理による画像の3次元解釈

角所 考\* 淡 誠一郎\* 安部 憲広\* 北橋 忠宏\* 三宅 誠\*\*

\* 大阪大学産業科学研究所

\*\* NHK放送技術研究所

あらまし 単眼画像から3次元構造を復元する定量的画像解釈の問題において、利用する拘束条件を仮説として取り扱い、解釈結果に応じた修正を許容することにより、多様な世界に柔軟に対応しようとするアプローチが最近見られはじめた。しかしこのような手法では、画像中の誤差などをどのように考慮して適用仮説を満たしかつ整合のとれた解釈を生成するか、またどのように適用仮説の修正、変更等を行うかが必然的に問題となる。本論文ではこれらの問題を解決するために、適用仮説に基づく最適な解釈をニューラルネットワークのエネルギー最小化によって求める手法を提案し、さらに適切な仮説の適用をも同時に行うニューラルネットワークの構築を試みる。

Three Dimensional Interpretation of an Image  
based on Energy Minimization

Koh KAKUSHO\* Seiichiro DAN\* Norihiro ABE\* Tadahiro KITAHASHI\* Sei MIYAKE\*\*

\* The Institute of Scientific and Industrial Research, Osaka University,  
Ibaraki-shi, 567 Japan

\*\* NHK Science and Technical Research Laboratories,  
Tokyo, 157 Japan

**Abstract** This paper describes a specific neural network for 3D-shape reconstruction from 2D-shape under some hypothetical constraints about the geometrical structures of objects. The constraints define the network's energy which is minimized when they are satisfied at once. This network minimizes the energy and recovers the 3D-structures most consistent with all constraints. Furthermore, this network automatically controls the application of hypothetical constraints according to the state of recover process and finds the most likely geometrical structures consistent with the input image. These two processes are interactively and concurrently performed based on the same principle, the energy minimization by neural networks.

## 1. まえがき

単眼の2次元画像から面の形状や傾きなどの3次元構造を求める定量的画像解釈は不良設定問題であり、対象に関する何らかの拘束条件が必要である。従来の手法には利用する拘束条件の絶対不変性を前提としたものが多かったが、最近、この拘束条件を常に成り立つとは限らないとする立場から「仮説」と呼び、この適用を対象に応じて変えることにより、多様な世界に柔軟に対応しようとするアプローチが見られるようになった<sup>[1]-[4]</sup>。このようなアプローチによる定量的画像解釈をここでは「仮説に基づく画像解釈」と総称する。

このような方法では必然的に、適切な仮説の適用とそれに基づく3次元解釈の生成という2つの処理を実現する必要がある。しかし解釈生成の際には、正しい仮説が適用されていても画像中の誤差や雑音のために個々の仮説に基づく部分解釈間に整合がとれることは希であり、適用仮説の充足と部分解釈間の整合性の維持とをどのように両立させるかが問題となる。また仮説相互の依存関係にもこの部分解釈間の整合性に基づく間接的な関係が含まれ、これは入力画像の条件によって変化するため、仮説の適用、修正の際にこの依存関係をどのように考慮に入れるかが問題となる。

そこで本稿では最適化手法やライン過程の考え方を導入し、さらにニューラルネットワークを利用してこれらの問題を解決することを考える。最適化手法は近年初期視覚の問題に盛んに利用されており<sup>[5][6]</sup>、ニューラルネットワークのエネルギー最小化等によって並列的に実現できる点で注目されている。またGemanらのライン過程を用いた画像復元の手法<sup>[7]</sup>は、統計学的仮説検定理論による正しい仮説の自動生成と、それに基づく画像の復元とを同時に行う方法であると見ることができ、Kochらは同様の処理がニューラルネットワークを用いて効率的に実現できることを示している<sup>[8]</sup>。これらは主に初期視覚の問題単独での成果であり、仮説に基づく画像解釈のような高次視覚と初期視覚との相互作用の問題に導入した例はまだ少ない。

以下では仮説に基づく画像解釈の問題への上記諸成果の応用として、まずニューラルネット

ワークのエネルギー最小化によって、与えられた仮説に対する最適な解釈を並列に求める手法について提案する。次にこれを改良し、仮説の適用とそれに基づく解釈の生成とを同時に行うニューラルネットワークの構築を試みる。またこれらを用いて実際に行った簡単な線画の定量的解釈の実験結果についても併せて報告する。

## 2. 幾何学的拘束に基づく画像解釈

仮説に基づく画像解釈において何を仮説として利用するかは問題や処理のレベルによって様々である。本稿では画像中の部分特徴に対し、その3次元構造に関する幾何学的拘束を与えるものを仮説として扱い、これを対象世界に関する知識として利用し、解釈を求める。対象世界は室内空間のような人工物から構成されるものを想定している。ただし本稿では仮説に基づく画像解釈におけるニューラルネットワークの応用可能性及び有効性に関して知見を得ることを主眼としているため、画像解釈問題及び利用する仮説自体は議論の本質を失わない程度に単純化した設定を考える。

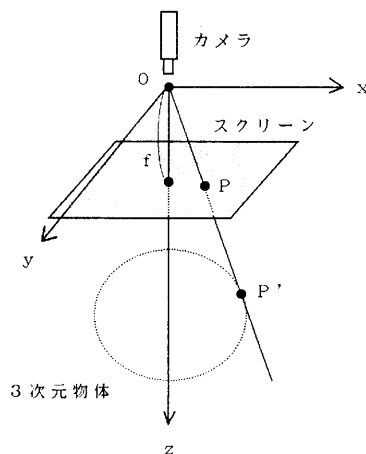


図1 カメラ中心座標系

### 2.1 画像解釈問題の定式化

カメラモデルは中心投影とし、カメラパラメータは既知とする。カメラ位置を原点Oとし、スクリーンに平行にx, y軸、カメラの光軸方向にz軸をとったカメラ中心座標系(図1)を

考え、画像上の点P ( $i_P, j_P, f$ ) の実際の3次元位置をP' ( $x_P, y_P, z_P$ ) として、点Pにおける奥行きパラメータ  $k(i_P, j_P)$  を次のように定義する。

$$x_P = i_P k(i_P, j_P)$$

$$y_P = j_P k(i_P, j_P)$$

$$z_P = f k(i_P, j_P)$$

( $f$  : カメラの焦点距離) (1)

今、カメラの焦点距離とカメラ中心からスクリーンに下ろした垂線の足の位置が与えられたとき、P' の位置はこの奥行きパラメータのみによって一意に決定される。そこで定量的解釈の問題を画像中のすべての特徴点の奥行きパラメータを決定する問題として形式化する。

## 2.2 入力画像

人工物の多くは多面体として近似でき、しかも四角形平面から構成される場合が多い。そこでここでは対象を四角形平面のみからなる多面体に限定する。この限定は対象世界をひどく制限するように思われるが、多面体近似できるような人工物はたとえそれが四角形以外の面を含んでいても、四角形面のみの復元によって全体が復元できることが多い。その場合には画像中の四角形領域のみを解釈の対象にすると考えてもよい。

入力画像は物体の各面に対応した四角形領域に分割可能であるものとし、オクルージョンを含まないものとする。この画像はあらかじめボトムアップ処理により領域分割、多角形近似などの特徴抽出を終えた後、領域の頂点のスクリーン座標及び領域間の接続関係のデータとして与えられるとする。オクルージョンを含まないという条件は議論を簡単にするためのもので、以下に述べる手法自体の本質とは無関係である。

## 2.3 領域解釈に対する仮説

画像領域の3次元解釈に対する仮説を次の3つに限定し、このいずれかが満たされるものとする。室内空間に自然に置かれた人工物からなる世界では、このいずれかを満たす平面により構成される物体が多く、我々人間も外界の認知においてこのような仮説を利用していると考えられる。

長方形仮説：「対象領域は3次元で長方形である。」

水平仮説：「対象領域は3次元で水平面である。」

鉛直仮説：「対象領域は3次元で鉛直面である。」

画像中のある四角形領域に長方形仮説や水平仮説を適用すれば、それぞれその領域の3次元の方向や形状が求められるが<sup>[2]19]</sup>、これらをつなぎ合わせて全体の解釈を構成しようとしても、画像の誤差などのためにそれらに整合がとれないことが多い。正しい仮説として図2の領域A, C, Dに長方形仮説を、Bに水平仮説を独立に適用し、解釈結果を図1の座標系の  $x + y - z = 0$  平面,  $xz$ ,  $yz$ ,  $xy$  平面上に平行投影したものを図3 (a)~(d) に示す。

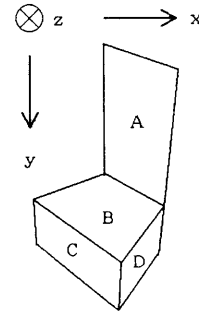


図2 入力図形

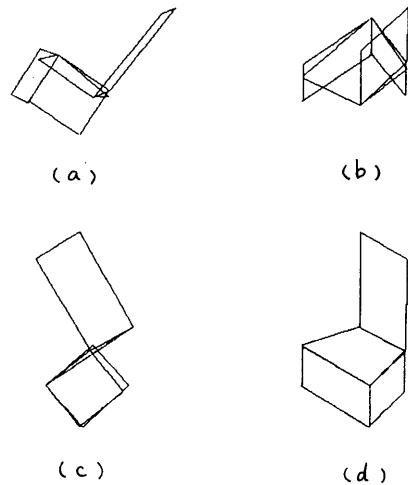


図3 画像の誤差による解釈の不整合

この解釈結果は、デジタル画像の量子化誤差のために接合部の整合がとれていない。また3次元空間において鉛直方向の面は無数に存在するため、鉛直仮説は単一領域だけでは一意に解釈を決定できず、隣接領域の解釈との整合性を用いて解釈を絞り込む必要がある。

これらのことから解釈の生成においては、仮説の示す拘束条件の充足と部分解釈間の整合性の維持とを同時に行うのが望ましいといえる。画像解釈におけるこのような問題には最適化手法が有効であることが知られている<sup>[6]</sup>。そこでニューラルネットワークを用いて、部分解釈間の整合性を保ちながら、同時に適用されたすべての仮説をできるだけ満足する解釈を並列に求めることが考えられる。その方法について次に述べる。

### 3. ニューラルネットワークによる解釈の生成

#### 3.1 ニューラルネットワークの構成

ニューラルネットワークのユニットは各々が画像中の各頂点对応し、その状態値は正の連続値で、対応する頂点の奥行きパラメータを表現するものとする。したがって、ユニットの状態値分布は入力画像の定量的解釈の解空間を表現する。

#### 3.2 仮説の示す拘束条件のエネルギー表現

解釈結果が平面で構成されるための平面性条件のエネルギーおよび2.3で定義した3つの仮説のエネルギーを以下のように定義する。ただし、点PからQへ向かうベクトルを $\overrightarrow{PQ}$ で表わし、ベクトルのスカラー積およびベクトル積を記号 $\cdot$ および $\times$ で表す。すなわち、

$$\begin{aligned} \overrightarrow{PQ} &= (x_Q - x_P, y_Q - y_P, z_Q - z_P) \\ \overrightarrow{P_1P_2} \cdot \overrightarrow{Q_1Q_2} &= (x_{P_2} - x_{P_1})(x_{Q_2} - x_{Q_1}) \\ &\quad + (y_{P_2} - y_{P_1})(y_{Q_2} - y_{Q_1}) \\ &\quad + (z_{P_2} - z_{P_1})(z_{Q_2} - z_{Q_1}) \\ \overrightarrow{P_1P_2} \times \overrightarrow{Q_1Q_2} &= \\ &= ((y_{P_2} - y_{P_1})(z_{Q_2} - z_{Q_1}) - (z_{P_2} - z_{P_1})(y_{Q_2} - y_{Q_1}), \\ &\quad (z_{P_2} - z_{P_1})(x_{Q_2} - x_{Q_1}) - (x_{P_2} - x_{P_1})(z_{Q_2} - z_{Q_1}), \\ &\quad (x_{P_2} - x_{P_1})(y_{Q_2} - y_{Q_1}) - (y_{P_2} - y_{P_1})(x_{Q_2} - x_{Q_1})) \end{aligned} \quad (2)$$

さらにワールド座標系における鉛直方向ベクトルを $\vec{n}$ とする。

[定義1]  $\vec{v}_1, \vec{v}_2$ を3次元空間ベクトルと

し、 $\vec{v}_1, \vec{v}_2$ の成す角を $\theta(\vec{v}_1, \vec{v}_2)$ で表す。ここで $\text{rect}(\vec{v}_1, \vec{v}_2)$ 、 $\text{para}(\vec{v}_1, \vec{v}_2)$ を次式で定義する。

$$\begin{aligned} \text{rect}(\vec{v}_1, \vec{v}_2) &= \cos^2(\theta(\vec{v}_1, \vec{v}_2)) \\ &= \frac{(\vec{v}_1 \cdot \vec{v}_2)^2}{|\vec{v}_1|^2 |\vec{v}_2|^2} \\ \text{para}(\vec{v}_1, \vec{v}_2) &= \sin^2(\theta(\vec{v}_1, \vec{v}_2)) \\ &= 1 - \cos^2(\theta(\vec{v}_1, \vec{v}_2)) \end{aligned} \quad (3)$$

[定義2] 画像中の領域Aにおいて、i番目の頂点 $P_i$ に隣接する2頂点を $P_{i1}, P_{i2}$ とする。この領域Aに対する長方形仮説のエネルギー $E_{RA}$ 、水平仮説のエネルギー $E_{HA}$ 、鉛直仮説のエネルギー $E_{VA}$ を次式で定義する。

$$\begin{aligned} E_{RA} &= \sum_{i=1}^4 \text{rect}(\overrightarrow{P_iP_{i1}}, \overrightarrow{P_iP_{i2}}) \\ E_{HA} &= \sum_{i=1}^4 \text{para}(\overrightarrow{P_iP_{i1}} \times \overrightarrow{P_iP_{i2}}, \vec{n}) \\ E_{VA} &= \sum_{i=1}^4 \text{rect}(\overrightarrow{P_iP_{i1}} \times \overrightarrow{P_iP_{i2}}, \vec{n}) \end{aligned} \quad (4)$$

[定義3] 平面性条件のエネルギー $E_F$ を次式で定義する。

$$\begin{aligned} E_{FA} &= \sum \text{para}(\overrightarrow{P_2P_1} \times \overrightarrow{P_2P_3}, \\ &\quad \overrightarrow{P_3P_2} \times \overrightarrow{P_3P_4}) \\ E_F &= \sum E_{FA} \end{aligned} \quad (5)$$

ただし第1式の $\sum$ は領域Aにおけるすべての連続する4頂点 $P_1, P_2, P_3, P_4$  ( $P_1, P_2, P_3, P_4$ はこの順に並んでいるものとする)についての和、第2式の $\sum$ は画像中のすべての領域についての和である。

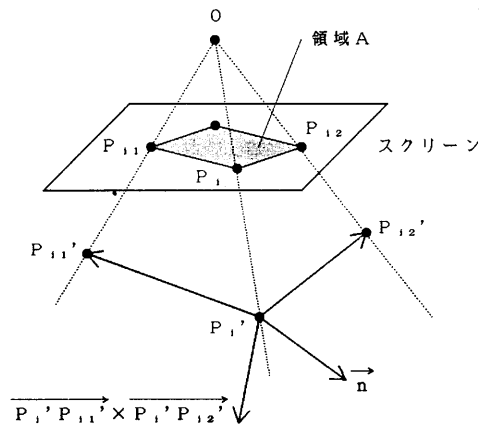


図4 エネルギーの意味(1)

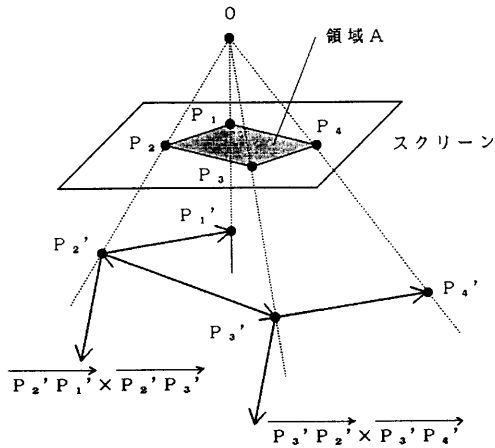


図5 エネルギーの意味(2)

これらのエネルギーの意味について説明する。まず  $E_{RA}$  は、領域 A の各頂点に対し、その点で交わる 2 辺の方向ベクトルが直交するとき 0 となる。また  $E_{HA}$ 、 $E_{UA}$ 、 $E_{FA}$  では領域 A の各頂点に対し、その点で交わる 2 辺の方向ベクトルのベクトル積を考え、 $E_{HA}$ 、 $E_{UA}$  はこれらがすべて  $\mathbf{n}$  と平行または垂直となるとき (図 4)、 $E_{FA}$  はこれらが領域 A 中の隣接する頂点同士で互いに平行となると (図 5) にそれぞれ 0 となる。なお  $E_{FA}$  は必ずしも領域 A に対する平面性条件の厳密な表現とはなっておらず、3 次元的に折り畳まれた状態 (図 6) でも 0 となる。これは隣接するユニット間の局所的な関係のみによってエネルギーを定義することを優先したためである。エネルギーの局所性によりニューラルネットワークの動作は局所並列となり、四角形以上の多角形に拡張した場合にも対象画像の複雑さからの影響を受けにくいものとなる。このニューラルネットワークでは、ユニットの初期状態はすべて 1.0 で入力画像そのものに対応しており、平面性条件のエネルギーの山を飛び越えるような急激なユニットの状態変化を許さない限り図 6 のような状態に陥ることはない。したがってユニットの状態変化を比較的小刻みにすることや平面性条件のエネルギーの重みを大きくすることによりこの問題は回避できる。ただし今回定義したエネルギーはいずれも、対応する拘束条件の表現として必ずしも最適なもの

のであるとは考えておらず、今後改良する必要があると考えている。

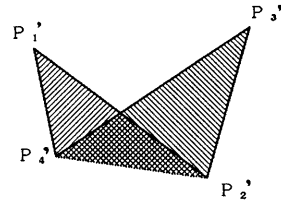


図6 エネルギー  $E_F$  の問題点

### 3.3 ユニットの状態変化

3.2 で定義したエネルギーを用いて、画像解釈に用いるニューラルネットワークのエネルギー  $E_1$  を定義する。

$$E_1 = c_0 E_F + c_1 \sum E_{RA} + c_2 \sum E_{HA} + c_3 \sum E_{UA} \geq 0 \quad (6)$$

ただし  $\sum$  はそれぞれの仮説の適用領域についての和、 $c_0 \sim c_3$  は平面性条件および 3 つの仮説の重みを決定する正定数である。このエネルギーは式 (1) ~ (5) より奥行きパラメータ、すなわちユニットの状態値の関数となる。画像中の頂点 P に対応するユニットの時刻 t における状態値を  $k_P(t)$  で表わし、各ユニットを式 (7) で与えられる  $E_1$  の最急降下方向に状態変化させると、式 (8) よりエネルギーは時間的に必ず減少する。

$$\frac{d k_P(t)}{d t} = -\alpha \cdot \frac{\partial E_1}{\partial k_P(t)} \quad (\alpha : \text{正定数}) \quad (7)$$

$$\begin{aligned} \frac{d E_1}{d t} &= \frac{\partial E_1}{\partial k_P(t)} \cdot \frac{d k_P(t)}{d t} \\ &= -\alpha \cdot \left( \frac{\partial E_1}{\partial k_P(t)} \right)^2 \leq 0 \end{aligned} \quad (8)$$

### 3.4 実験結果

3.1 ~ 3.3 で定義したニューラルネットワークを用いて正しい仮説が適用されている場合の解釈を求めた。図 2 に対し 2.3 と同じ仮説を与えた場合の解釈結果を図 7 に示す。図 3 の場合とは異なり、整合性のとれた解釈が得られている。前にも述べたようにユニットの初期状態値はすべて 1.0 (対象物体の全頂点がスクリーン上に存

在する)とし、全ユニットを同期的に1000回状態変化させた。なおパラメータは、

$c_0 = c_1 = c_2 = c_3 = 1.0$ ,  $\alpha = 2.0 \times 10^{-4}$  に設定している。

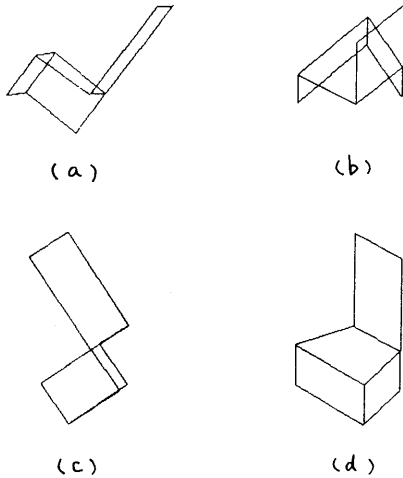


図7 ニューラルネットワークによる解釈結果

#### 4. 仮説の自動決定

3. のようなニューラルネットワークにより、与えられた仮説に基づく最適な解釈が得られるとしても、いかにして適切な仮説を適用するかという問題は依然として残される。そこで次に Geman, Kochらの成果を応用し、適切な仮説の適用とそれに基づく解釈の生成とを同時に行うニューラルネットワークの構築を試みる。

##### 4.1 仮説ユニット

3. で構成したニューラルネットワークのユニットを今後は「解釈ユニット」と呼ぶ。そしてこれに加えて、画像中の領域ごとに新たにユニットを付加し、これを「仮説ユニット」と呼ぶ。仮説ユニットには、長方形仮説ユニット、水平仮説ユニット、鉛直仮説ユニットの3種類があり、対応する領域に特定の仮説を適用するか否か(これを「適用判断」と呼ぶ)を0(適用しない)から1(適用する)までの連続的な状態値で表現する。仮説ユニットにはそれぞれ隠れユニットが付属しており、隠れユニットは

シグモイド関数を通して仮説ユニットの状態値を決める。

##### 4.2 仮説の適用に関する評価基準

仮説の適用判断のためには、何らかの評価基準(あるいは仮説の適用のための制約条件)が必要である。ここでは次のような評価基準を考え、これをできるだけ満たすように仮説の適用判断を行うことにする。

① 仮説と解釈の一致 ある仮説が適用されている領域に、その仮説の示す拘束条件を満たす部分解釈が構成できること。

② 仮説の適用判断の明確さ 中途半端な判断が行われないこと(判断結果が0もしくは1のどちらかに収束すること)。

③ 適用内容の合理性 水平仮説と鉛直仮説が同じ領域に同時に適用されないこと。

④ 適用される仮説の数 各領域において少なくとも1つの仮説が適用されること。

⑤ 2次元特徴との一致 隣接領域の境界線は3次的にルーフエッジであると考え、隣接領域に同時に水平仮説が適用されないこと。

##### 4.3 評価基準のエネルギー表現

3.2で定義した4つのエネルギーに加え、4.2の評価基準をエネルギーとして表現し、ニューラルネットワーク全体のエネルギー $E_2$ を次式で定義する。なお画像中の領域Aに対応する長方形仮説ユニット、水平仮説ユニット、鉛直仮説ユニットの状態値をそれぞれ $r_A$ ,  $h_A$ ,  $v_A$ で表す。

$$\begin{aligned}
 E_2 = & c_4 E_F \\
 & + \sum \{ c_5 (E_{RA} \cdot r_A + c_{R'}(1-r_A)) \\
 & \quad + c_5 (E_{HA} \cdot h_A + c_{H'}(1-h_A)) \\
 & \quad + c_5 (E_{VA} \cdot v_A + c_{V'}(1-v_A)) \\
 & \quad + c_6 r_A(1-r_A) \\
 & \quad + c_6 h_A(1-h_A) \\
 & \quad + c_6 v_A(1-v_A) \\
 & \quad + c_7 h_A \cdot v_A \\
 & \quad + c_8 (1-r_A)(1-h_A)(1-v_A) \} \\
 & + \sum c_9 h_A h_{P2} \\
 \geq & 0 \quad (9)
 \end{aligned}$$

式の意味について説明する。まず第1項は式(5)で定義した平面性条件に対応する。また最初の $\sum$ は画像中のすべての領域についての和である。この中の第1~3項は評価基準①に対応

し、 $E_{RA}$ ,  $E_{HA}$ ,  $E_{UA}$ がそれぞれ $c_R$ ,  $c_H$ ,  $c_U$ より小さい場合には、 $r_A$ ,  $h_A$ ,  $v_A$ を1に、逆の場合には0とした方が項の値はより小さくなる。特に $r_A$ ,  $h_A$ ,  $v_A$ が1であるときにはこれらの項は $E_{RA}$ ,  $E_{HA}$ ,  $E_{UA}$ そのものとなる。第4～6項は評価基準②に対応し、 $r_A$ ,  $h_A$ ,  $v_A$ が0もしくは1のときに最小となる。第7項は評価基準③に対応し、 $h_A$ ,  $v_A$ の少なくとも一方が0のとき最小となる。第8項は評価基準④に対応し、 $r_A$ ,  $h_A$ ,  $v_A$ のうち少なくとも1つが1のとき最小となる。さらに後の $\Sigma$ は画像中のすべての隣接領域対A, Bについての和で評価基準⑤に対応し、 $h_A$ ,  $h_B$ の少なくとも一方が0のとき最小となる。 $c_R$ ,  $c_H$ ,  $c_U$ は評価基準①の厳しさを表す正定数、 $c_4 \sim c_9$ は平面性条件および5つの評価基準の重みを決める正定数である。

#### 4.4 状態変化

各ユニットは $E_2$ の最急降下方向に状態変化させる。すなわち解釈ユニットの状態変化は式(7)と同じ(ただし $E_1$ の代わりに $E_2$ を用いる)であり、仮説ユニットの状態変化は次式とする。簡単のため長方形仮説ユニットについてのみ示すが、他も同様である。

$$r_A(t) = \frac{1}{1 + \exp(-u_{RA}(t))}$$

$$\frac{d u_{RA}(t)}{d t} = -\beta \cdot \frac{\partial E_2}{\partial r_A(t)}$$

( $\beta$ : 正定数) (10)

ただし時刻 $t$ における長方形仮説ユニットの状態値を $r_A(t)$ 、付属の隠れユニットの状態値を $u_{RA}(t)$ で表す。

#### 4.5 実験結果

図8に示す画像を与え、解釈ユニットの初期値をすべて1.0、仮説ユニットの初期値をすべて0.5(隠れユニットの初期値を0.0)とし、性急な適用判断を避けるために仮説ユニットの状態変化1回につき解釈ユニットの状態変化5回の割合でニューラルネットワークを動作させた。これにより、解釈ユニットは仮説ユニットの示す仮説の適用判断に基づく最適な解釈の生成を試み、一方仮説ユニットは解釈の状況に応じて仮説の適用判断の内容を更新していく(表1)。このようなプロセスを繰り返し、仮説ユニット

の状態変化400回、解釈ユニットの状態変化2000回のは4.2の評価基準①～⑤に基づく適切な仮説が自動的に適用され、同時にこれに基づく最適な解釈が図9のように生成された。なお本実験ではパラメータを、

$$c_R = 0.005, \quad c_H = c_U = 0.01, \quad c_4 = 1.0,$$

$$c_5 = 2.0, \quad c_6 = 2.5, \quad c_7 = c_9 = 3.0,$$

$$c_8 = 1.5, \quad \alpha = 2.0 \times 10^{-4}, \quad \beta = 0.01$$

に設定している。

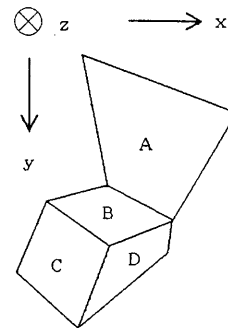


図8 入力図形

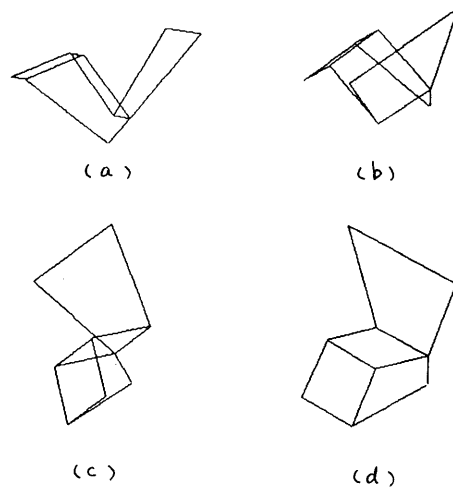


図9 解釈結果

表1 仮説ユニットの状態変化

仮説ユニット の状態変化 (回)	領域	仮説ユニットの状態値		
		長方形仮説 ユニット	水平仮説 ユニット	鉛直仮説 ユニット
100	A	0.43	0.00	0.44
	B	0.75	0.04	0.00
	C	0.80	0.00	0.38
	D	0.42	0.00	0.65
200	A	0.34	0.00	0.72
	B	0.99	0.00	0.00
	C	0.99	0.00	0.11
	D	0.22	0.00	0.96
400	A	0.00	0.00	1.00
	B	1.00	0.00	0.00
	C	1.00	0.00	0.00
	D	0.00	0.00	1.00

## 5. むすび

仮説に基づく画像解釈にニューラルネットワークを導入し、与えられた仮説に基づく最適な解釈を生成するニューラルネットワークを構築した。さらにこれを改良し、適切な仮説の適用をも同時に行うニューラルネットワークを構築した。

仮説の適用制御の問題は仮説推論や制約推論を扱う人工知能研究においても問題となり、様々な手法が提案されているが、画像などのパターン情報でかつ誤差や雑音を含むような対象を扱う適当な方法はまだ確立されていない。本稿はこのような問題へのニューラルネットワークの応用可能性を示唆するものでもある。ただし今後、仮説に基づく画像解釈の手法自体を仮説推論等の枠組みを用いて整理し、この応用可能性を理論的な立場からも検討する必要があると考えている。

また今回提案したニューラルネットワークは、ローカルミニマの回避やエネルギーのパラメータ設定などの問題に対する抜本的な解決策をまだ構想していない。これは入力画像が単純であるため事実上問題が生じていないことによるが、今後入力画像の複雑化および利用する仮説の増加などによりこの問題が深刻化することが予想

され、焼き鈍し<sup>[7]</sup>やパラメータ学習<sup>[10]</sup>などの導入について検討する必要があると考えている。

**謝 辞** 日頃御援助を賜るATR視聴覚機構研究所淀川英司社長、並びに有益な御助言を戴く同研究所主幹研究員、乾敏郎、川人光男両氏に深く感謝致します。

## 文 献

- [1] 石塚：“画像の認知と仮説推論”，信学技報PRU87-123 (1988)。
- [2] 渡辺，安部，北橋：“優先順序付き仮説による画像領域の三次元解釈”，信学技報PRU88-148 (1989)。
- [3] 田中，川島，金谷：“平行性仮説による多面体の3次元形状復元”，信学論 (DII)，J72-DII，4 (1989)。
- [4] 前原，川島，金谷：“直交性仮説による多面体の3次元形状復元”，信学論 (DII)，J72-DII，6 (1989)。
- [5] 坂上，横矢：“弛緩法と正則化”，情報処理，30，9 (1989)。
- [6] 杉原：“線画の代数構造を利用した単眼画像からの立体復元法”信学論 (D)，J66-D，5，(1983)。
- [7] S.Geman and D.Geman：“Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images”，IEEE Trans., PAMI-6, 6 (1984)。
- [8] C.Koch, J.Marroquin and A.Yuille：“Analog “neuronal” networks in early vision”, Proc. Natl. Acad. Sci. USA, 83 (1986)。
- [9] R.Horand：“New Methods for Matching 3-D Objects with Single Perspective Views”，IEEE Trans., PAMI-9, 3 (1987)。
- [10] 川人，池田，三宅：“神経回路網の学習と視覚情報処理”，テレビジョン学会誌，42, 9 (1988)。
- [11] 角所，淡，安部，北橋，三宅：“仮説に基づく画像の3次元解釈へのコネクショニスト・モデルの利用”，人工知能学会全大 (1989)。