

幾何学的拘束に基づく物体認識への
コネクショニスト・アプローチ

2E-5

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

1. はじめに

渡辺ら^{[1][2]}は画像解釈の新しい手法として、「隣接領域は3次元的に接続している」や、「四角形領域の3次元形状は長方形である」など対象世界の一般的な幾何学的拘束に関する様々な仮説を用意し、画像に対してこれを選択、適用することにより解釈を求める方法を提案した。この処理は、

- I. 互いに矛盾しない最適の仮説の組合せを選択する。
- II. 選択した仮説を画像に適用し、これを満足する解釈を求める。

の2つからなる。しかし利用される仮説間には複雑な依存関係が存在し、上記I, IIの処理においてこれを常に考慮に入れる必要があるが、一様な逐次解釈アルゴリズムだけを用いてこのような処理を実現するのは困難である。我々はまずこのうちのII.の処理を相互結合型コネクショニスト・モデルのエネルギー最小化によって実現することを考え、簡単な画像解釈を行った。本稿ではその内容について報告する。

2. コネクショニスト・モデルによる画像解釈

今回解釈に利用するのは次のような幾何学的拘束に関する仮説である。

- 仮説A: 「画像中のすべての領域の内部は、3次元的にはなめらかな面である」
- 仮説B: 「画像中のすべての隣接領域は、境界線において3次元的に接続している」
- 仮説C: 「(画像中のある角に対し、)角は3次元的には直角である」

これらの仮説を画像に適用した場合の解釈を求めるため、それぞれの仮説の示す幾何学的拘束をエネルギー関数として表現し、相互結合型コネクショニスト・モデルによってこれを最小化することにより、適用される仮説をすべて満足する解釈を求める。

2.1 システムへの入力

入力された画像は、線画化、細線化されたものとする。また背景と判断される部分は除外されており、物体を表す領域については、線分を共有する領域対が抽

出されているものとする。さらに各領域の頂点は認識され、座標で表されているものとする。細線化され、線分近似された線画は2値データとして上記の座標データとともにコネクショニスト・モデルの入力とされる。

2.2 ユニット

ユニットは画像の各ピクセルに対し1つ用意する。いま、カメラの位置および向きを既知として図. 1のようなカメラ中心座標系を考えれば、スクリーン上の点R(x, y, f) (f: カメラの焦点距離) の実際の3次元位置は、奥行きパラメータkが決まれば、点R'(k·x, k·y, k·f)として一意に定まる。このkを各ユニットの状態値とする。

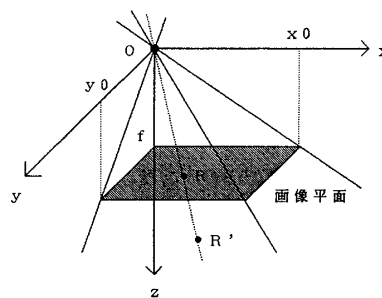


図. 1 カメラ中心座標系

2.3 エネルギー関数による幾何学的拘束の表現

(1) 仮説Aの示す幾何学的拘束の表現

つぎのようなエネルギー関数E_Aにより表現する。

$$E_A = \sum \left\{ k(x_2, y_2) - \frac{2k(x_1, y_1)k(x_3, y_3)}{k(x_1, y_1) + k(x_3, y_3)} \right\}^2$$

(ただし、 \sum は、

画像中において一直線上でピクセルとして隣接し、

かつ、

同時に境界線上にある、もしくは点(x₂, y₂, f)が境界線上にない

ような画像内のすべての3点の組

{(x₁, y₁, f), (x₂, y₂, f), (x₃, y₃, f)} についての和を表わす。)

A Connectionist Approach to Object Recognition Based on Geometric Constraints

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

* The Institute of Scientific and Industrial Research, Osaka University

** NHK Science and Technical Research Laboratories

このエネルギー関数は、一直線上で隣合う画像内のすべての3つのピクセルに対応する点の3次元位置が、図. 2のようにエッジを横切らない範囲において一直線上に並ぶ場合に最小となる。

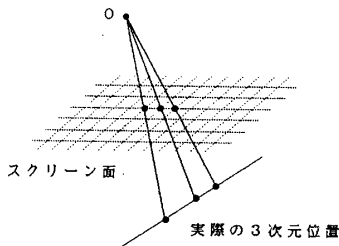


図. 2 エネルギー関数EAの意味

(2) 仮説Bの示す幾何学的拘束の表現

2.1, 2.2により、得られる解釈は仮説Bの示す幾何学的拘束を満足する。

(3) 仮説Cの示す幾何学的拘束の表現

仮説Cを適用する画像中の角Lを $\angle R_1 R_2 R_3$ (R_1, R_2, R_3 は領域の頂点)で表すと、この角に対する仮説Cの幾何学的拘束は次のエネルギー関数で表現される。

$$E_{c^L} = \sum_P \cos^2(\angle R_1' R_2' P') + \sum_Q \cos^2(\angle Q' R_2' R_3')$$

(画像中の点Rの3次元位置をR'で表す。)

ただし、P, Qはそれぞれ線分 $R_2 R_3, R_1 R_2$ 上の任意の点。

2.4 ユニットの状態変化

時刻tにおける第iユニットの状態値を $k_i(t)$ とする。ユニットの状態変化は、

$$E = a \cdot E_a + b \cdot \sum_L E_{c^L}$$

$$\frac{d k_i}{d t} = -\alpha \cdot \frac{\partial E}{\partial k_i}$$

によって行う。

ただし \sum_L は仮説Cが適用される画像中のすべての角についての和を示す。

なおここで得られる解釈は各部分の相対的な位置関係のみを表現する。絶対的な距離情報は、任意の1点の距離データを与えることにより容易に得られる。

3. 実験

上で述べたコネクショニスト・モデルを用いて、図. 3で示す簡単な図形に対し、仮説A, Bおよびすべて

の内角に仮説Cを適用した場合の解釈を求めた。図. 4は上式における比例定数をそれぞれ、

$$a = 8.33 \cdot 10^{-3}, \quad b = 4.00 \cdot 10^{-6}, \quad \alpha = 1.00$$

として各ユニット10000回の状態変化を同期的に行った結果である。スクリーン面上に並んでいた点が、次第に立体的に浮き上がり、3次元解釈が得られた。入力画像では抽出できていなかった線も稜としてほぼ復元できている。

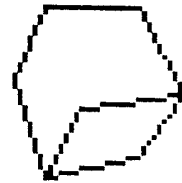


図. 3 入力図形

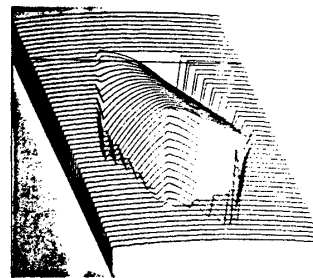


図. 4 出力結果

4. むすび

仮説の示す幾何学的拘束をエネルギー関数として表現し、これをコネクショニスト・モデルによって最小化することにより簡単な画像の解釈を試みた。なお今回定義したエネルギー関数は対応する幾何学的拘束を表現するものとして必ずしも最適なものは考えておらず、今後修正していく予定である。

謝辞

日頃、有益な御助言、ご援助を戴く、(株)ATR視聴覚機構研究所社長淀川英司氏ならびに主幹研究員、乾敏郎、川人光男両氏に深く感謝致します。また御討論を戴いた(株)日本IBM大和研究所の渡辺佳映氏に感謝致します。

参考文献

- [1] 渡辺, 安部, 北橋: "思い込み優先による二次元領域の三次元解釈", SICIE関西支部シンポジウム「あいまい情報処理と知的システム制御」講演論文集(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).