

Selective Attention のための特徴抽出機構に関する研究

4 L - 5

中野 格 横井 浩史 嘉数 侑昇
北海道大学工学部

1. はじめに

自律システムの必要な要件の一つとして、自身の環境情報の獲得がある。この際、センサの物理特性、メモリの有限性、処理速度など、一般に様々な制約により、正確かつ広範囲な環境マップの作成は困難である。そこで、自身の目的に応じ、必要な環境情報を効率的に獲得するための選択的な注視機構が必要である。

ここで必要な情報はタスク、システム特性などに応じ、その個々において異なることから、システム自身の観点から決定する必要がある。これは Situatedness[1]として知られるところである。本稿では CCD カメラにより得た実画像から探索対象を認識する上で有効な特徴を抽出する手法を提案する。

2. 処理の概要

本システムでは入力画像からエッジを抽出し、ここから局所的な特徴を抽出する。これらを自己組織的に統合することによって多様な抽象度にある情報を生成する。システムは探索目標に応じ、その中で重要な(探索する上で注目すべき)情報をその目標の特徴として BP によって強化する(Fig.1)。ここでは多様な抽象度にある情報を自己組織的に生成するユニットとしてネオコグニトロン[2]を用いた。次にネオコグニトロンについて簡単に説明する。

3. ネオコグニトロン

ネオコグニトロンは福島邦彦によって開発された階層型ニューラルネットワークである。入力層 U_0 には光受容細胞が平面状にならんでいる。入力パターンが投影されると、光受容細胞は受光した光の強さに応じた出力を出す。入力層 U_0 の後ろには S 細胞と名付けられた細胞の層 U_1 と C 細胞と名付けら

れた細胞の層 U_2 が交互に配置されている。S 細胞は特徴抽出機能を持つ細胞で局所的な特徴の一つに選択的な反応を示す。それに対して C 細胞は特定の細胞面内にある自身の受容野内の S 細胞が一つでも出力を出せば出力を出す。

従って、C 細胞はその入力側の S 細胞と同じ特徴に反応するが、入力側の個々の S 細胞に比べ特徴の呈示位置のずれにはあまり影響を受けない。このような処理を何層にもわたって行うことにより、特徴の抽出と統合を行う。

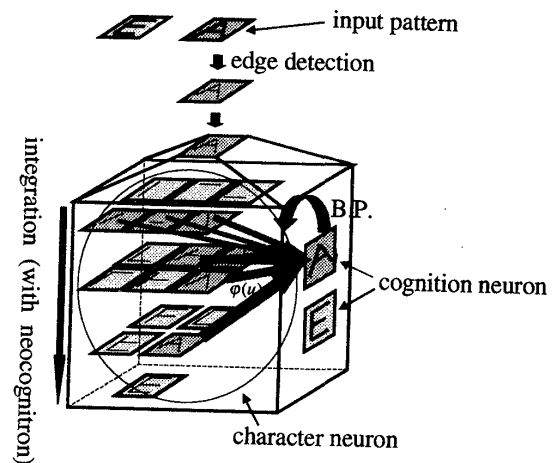


Fig.1 concept view of this system

4. システムの構造

- (1) CCD から 480(H)×480(W), 256 階調の画像を入力し、平滑化を行ない 95(H)×95(W)の低解像度に落とす。
- (2) S 細胞により、(1)の中心 33(H)×33(W)の領域を特徴抽出し、これを C 細胞で処理することによりノイズの影響を抑え、かつエッジの欠落を補完する(Fig.2)。

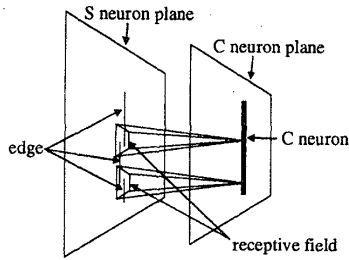


Fig.2 interpolation of edge

- (3) ネオコグニトロン自己組織化機能を用い、多様な抽象度にある特徴にそれぞれ特化した反応特性を持つ細胞を生成し、これから特徴細胞群を生成する。特徴細胞はそれが表す特徴の抽象度に関係なく認識細胞が等価に扱うよう2値化する。

$$p(u) = \begin{cases} 1 & \exists u > \text{threshold} \\ 0 & \forall u \leq \text{threshold} \end{cases}$$

$$u \in p$$

$p(u)$: 特徴細胞 p の出力

u : ネオコグニトロンの細胞の出力

- (4) 認識細胞の出力に対し、その正誤のみを教師信号(T)として与え、BPにより認識細胞と正しい認識に貢献した特徴細胞との結合を強化する。

$$O_j = f\left(\sum_i w_{ij} \cdot p(u)_i + B_j\right)$$

O_j : 認識細胞の出力

w_{ij} : 特徴細胞 i から認識細胞 j への結合荷重

B_j : j 番目の認識細胞が持つ閾値

$$f : f(x) = \frac{1}{1 + e^{-x}}$$

学習は最も大きな出力を出す認識細胞に関してのみ以下のようにBPを用いて行なう。

$$w_{i_{\max}} = w_{i_{\max}} \cdot \eta \cdot \delta \cdot O_{\max}$$

$$B_{\max} = B_{\max} \cdot \eta \cdot \delta$$

O_{\max} : 最大値を出力する認識細胞の出力値

$w_{i_{\max}}$: 特徴細胞 i から最大値を出力する認識細胞への結合荷重

B_{\max} : 最大値を出力する認識細胞の閾値

η : 学習効率

δ : 出力の誤差

ここで、 δ は以下の様に導く。

$$\delta = (T - O_{\max}) O_{\max} (1 - O_{\max})$$

$$T = \begin{cases} 1 & (\text{if } O_{\max} \text{ is right}) \\ 0 & (\text{if } O_{\max} \text{ is wrong}) \end{cases}$$

5. 実験・結果

文字 B を描いた壁面の探索を想定し、CCD カメラをランダムに動かすことによって得た 10 種類の画像を 200 回呈示した。その後、B を視野の中心に捕らえた際の認識細胞の発火状態を Fig.3 に示す。

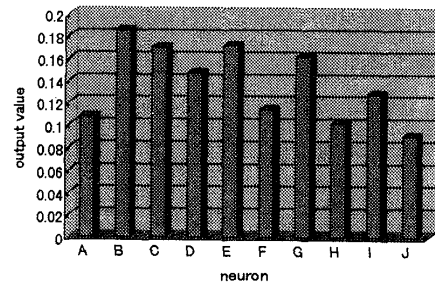


Fig.3 outputs of neurons

Fig.3 より B に対応する認識細胞が最大値を出力していることがわかるが、他の細胞に対して特化した出力特性を有しているとは言えない。

6. おわりに

本稿では Selective Attention のための特徴抽出機構を提案したが、その認識能力は十分なものとは言えない。より特化した特徴細胞を生成すべく、その自己組織化機能に対して改善が必要である。

7. 参考文献

[1]Pfeifer : "Symbols, Patterns, and behavior : towards a new understanding of intelligence", 第 10 回人工知能学会全国大会, 1996
 [2]Fukushima : "Neocognitron : A self-organization Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position", Biological Cybernetics, 36, pp.193-202, 1980