

ニューラルネットワークによるパターンの注視と認識

1 J-2

—行動決定ネットとパターン認識ネットの自己組織的様相—

堀内誠致 横井浩史 嘉数侑昇

北海道大学工学研究科

1. はじめに

未知環境における自律移動ロボットを考える場合、何がロボットにとって必要な情報であるかを事前に教えることはできないため、その行動を通して自律的に獲得する必要がある。

本研究では、画像中のパターンを自己組織的に獲得し、得られた各パターンと行動をマッピングさせるための学習モデルを提案する。ここで、得られたパターンに拡大・縮小、移動、回転といった変換が施されている場合やパターンの一部分が遮蔽されている場合にも同一パターンと判定する機構が必要になる。

従来、画像処理のテンプレートマッチングの分野においては一般化 Hough 変換[1]と呼ばれる手法がパターンの変形や部分遮蔽に非常にロバストであることが知られている。一般化 Hough 変換の特徴はテンプレートとそのポーズ（変換）の2つの情報に分割し、マッチングを行う点である。実際、2種類の情報に分けることは人間の場合にも当てはまることが実験心理学的に確かめられている[2]。

以下の節ではこれらを踏まえたモデルを提案する。

2. 学習モデル

本研究で用いる学習モデルを図 1 に示す。

このモデルは2つに大きく分けられる。自己組織化マップ（SOM）を用いて画像からテンプレートを作製し、そのテンプレートが画像中に存在すればパラメータ空間に位置、拡大率、回転角度といった情報を書き込む部分と、そのようにして得られたパラメータ空間と行動をニューラルネットワークによって学習する部分である。

ラルネットワークによって学習する部分である。

2.1. テンプレートとパラメータ空間の生成

(1) 特徴抽出

最初に、環境から得られた画像(256×256)を2値化し重心位置を求め、疎画像化(64×64)する。次に求めた重心位置を中心とする極座標系を部分的に展開する。この極座標系は円周方向に8等分(45度)、半径方向に4等分、計32の小領域からなる。この各小領域に画像の輝度値(256階調)が割り当てられる。したがって、SOMへのある時刻 t 、回転角度 θ 、拡大率 s での入力データ $x(t, \theta, s)$ は

$$x(t, \theta, s) = [x_1, x_2, \dots, x_i, \dots, x_{32}]$$

となる。ただし、 $0 \leq x_i \leq 255$ である。

(2) テンプレートの生成

ここで用いる SOM は入力層ニューロン 32 個、

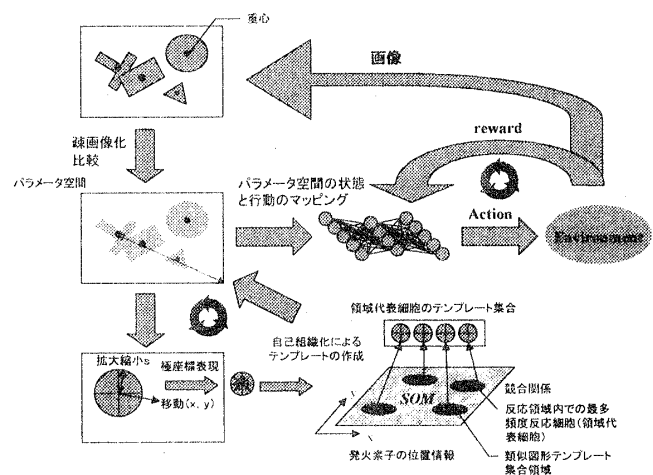


図 1: 学習モデル

競合層ニューロン 100 個で構成される。

競合層への入力データ $x(t, \theta, s)$ の, θ は極座標系が円周方向に 8 等分されているので 8 通り, s は半径方向に 4 分割され画像の解像度が 64×64 であるので最小値 4, 最大値 64 (重心の位置に依存する) となり, ある時刻 t での入力データセットは最大 240 個となる。

競合層の勝者ニューロン C_{j^*} は式(1)により入力 $x(t, \theta, s)$ と j 番目の競合層ニューロン C_j への結合荷重ベクトル $w_j(t)$ との距離 d_j が最小になるものとする。ここで, θ, s は入力データセットの中で距離 d_j を最小にする組み合わせとなる。

$$d_j = \min_{\theta, s} \sqrt{\|x(t, \theta, s) - w_j(t)\|^2} \quad (1)$$

特に最小となる距離 d_j が,

$$d_j \leq 16 \quad (2)$$

を満たすときそれをテンプレートとみなす。

(3) 結合荷重の更新

勝者ニューロン C_{j^*} とその近傍領域 S_{j^*} に含まれるニューロンの結合荷重 $w_j(t)$ を式(3)によって更新する。

$$w_j(t+1) = w_j(t) + \eta(x(t) - w_j(t)), j \in S_{j^*} \quad (3)$$

ここで, 近傍領域 S_{j^*} は勝者ニューロンを中心とする 9 個のニューロン, 学習係数 η は 2 である。

(4) パラメータ空間の生成

(2)テンプレートの生成において勝者ニューロン C_{j^*} が式(2)を満たす場合, パラメータ空間 P に極座標系の中心座標 (x, y) , 回転角度 θ , 拡大率 s , 勝者ニューロン番号 j を追加する。

2.2. パラメータ空間と行動のマッピング

パラメータ空間と行動のマッピングにはバックプロパゲーション (BP) 型ニューラルネットワークを用いる。ここで, パラメータ空間は 5 次元であるためすべての場合を入力とすると非常に多くの入力層ニューロンが必要になる。したがって, 実験内容に即してここでは拡大率 s , 回転角度 θ の情報は用いない。

3. 実験内容と結果

使用する画像は図 2 のように 3 領域に分割し各領域に 3 つの図形 (○, △, ×) のうちいずれか 1 つが入るようにする。ただし, 各図形は移動・回転・拡大縮小といった変換を受けている。BP

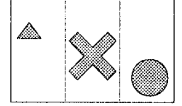


図 2: 図形の例

型ニューラルネットワークの構成は入力層 300, 中間層 100, 出力層 4 (左, 中, 右, 無) とする。学習内容は○のある部分を示すように学習を行う。

実験の結果, 学習は以下のように進行した。最初, テンプレートが生成されていないため, パラメータ空間に有効な値が追加されず適切な学習ができない状態であった。しかし, テンプレートが自己組織的に生成されはじめると, パラメータ空間とのマッピングが可能となった。

自己組織化されたテンプレートは上記 3 つの図形以外にも生成されており, 図形が小さいときに重心位置が微妙にずれることが原因と思われる。

4. まとめ

本研究では, 画像中の未知のパターンからテンプレートを自己組織的に生成し, テンプレートとそのポーズ情報に分離することでパターンを効率良く表現し, 行動とマッピングできることを示した。

現在, 画像から重心を求め, それを中心にテンプレートやパラメータ空間を生成しているが, 重心の微妙なずれによって正しく生成されない場合がある。また, パラメータ空間は 5 次元空間であるため, 行動とのマッチングを行う場合, 組み合わせ問題が生じる。これらは今後の課題としたい。

参考文献

- [1] 和田俊和, 松山隆司: Hough 変換に基づく図形検出法の新展開, 情報処理, Vol.36, No.3, pp.253-263, Mar. 1995.
杉尾武志, 乾 敏郎: 図形認知における変換情報の表現, 日本心理学会第 61 回大会論文集, pp.615, 1997.