

ニューラルネットワークによる濃淡対象物の識別

3M-6

坂口 嘉之

美濃 導彦

池田 克夫

東洋紡績総合研究所

京都大学工学部
高度情報開発実験施設

京都大学工学部
情報工学教室

1 はじめに

圧力エネルギーを持つSNAKEにより、濃淡画像から対象物の輪郭を抽出し、閉曲線からなる輪郭の特徴を不変情報に変換した後、ニューラルネットワークによる識別を行った。

2 システムの構成

図1に、システムの全体図を示す。入力は、ビデオカメラからの濃淡画像(8bit 512x400)であり、最終段の出力数は、識別対象物の数 N である。システムは、機能的に三つのブロックに分けられ、それぞれを輪郭抽出部、特徴抽出部、認識部と呼ぶことにする。

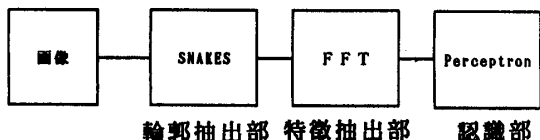


図1 システムの全体図

2-1 輪郭抽出部

[圧力エネルギーを持つSNAKE]

ここで、濃淡画像から対象物の輪郭抽出を行う。輪郭抽出には、Kassらにより提案された Active Contour Models(SNAKES)^[1]を発展させた、圧力エネルギーを持つSNAKEを用いた。圧力エネルギーを持つSNAKEは、画像の輝度変化と、輪郭線を対象物に押しつけようとする圧力を拘束条件に持つスプライン関数である^[2]。閉曲線からなるSNAKEは、画像のエネルギーにより初期位置から徐々に変形を開始し、対象物の輪郭で収束する。この圧力エネルギーを持つSNAKEは、画像中にエネルギー勾配がない場合にも凹形状に収束でき、閉じた曲線で正確な輪郭抽出ができる。圧力エネルギーを持つSNAKEのエネルギーを以下に示す。

$$E_{snake} = \sum_{i=1}^n (E_{int}(v_i) + E_{pressure}(v_i) + E_{image}(v_i))$$

$$E_{int}(v_i) = \alpha_i |v_i - v_{i-1}|^2 + \beta_i |v_{i-1} - 2v_i + v_{i+1}|^2$$

$$E_{pressure}(v_i) = w_{press} E_{press}(v_i)$$

$$E_{image}(v_i) = w_{line} E_{line}(v_i)$$

[均等長割り付け]

SNAKEの各要素間の距離を等しくするために、均等長割り付けを行なう。そのアルゴリズムを以下に示す。

1. 輪郭の全長(L)を計算し、平均要素間長(L_{mean})を計算する。
2. 始点よりL_{mean}の距離にある輪郭上の次の点をニュートン法により求める。
3. この時の誤差が、0.01L_{mean}/L以下になったときにその点を確定し、次の点の座標を計算する。
4. 以後この計算(1~3)を終点まで繰り返す。

均等長割り付けは、対象物の輪郭を n 角形で近似することに相当する。

2-2 特徴抽出部

抽出された輪郭は、n 個の二次元座標 (n=64) からなる点列である。この点列の重心(x_s, y_s)を求め、重心から各要素点までの距離を値とする関数をフーリエ変換する。得られたフーリエ・スペクトルの振幅項は、画像の形状と大きさを表し、位相項は、画像の回転を表している。振幅項を振幅項の最大値で正規化し、位相項を無視することにより、画像の収縮・拡大、回転に無関係な特徴ベクトルに変換できる。図2に、画像の回転によるスペクトルの変動の一例を示す。この図は、ボトルを20度毎に一回転させて、18個の正規化振幅の最大値と最小値、及び、平均値を示したものである。均等長割付を行うことで、スペクトルの変動を小さくすることができた。

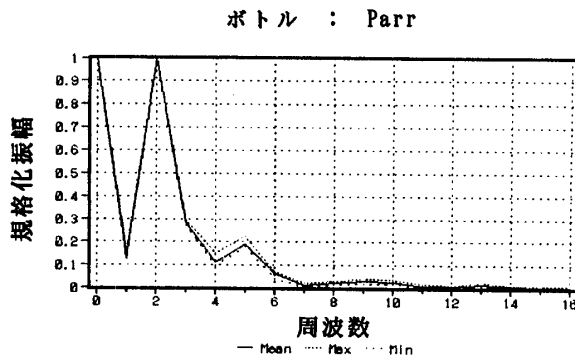


図2 回転によるスペクトルの変動

CLASSIFICATION OF INTENSITY IMAGE OBJECT
USING NEURAL NETWORK.

Yoshiyuki SAKAGUCHI* Michihiko MINOH**

Katsuo IKEDA**

* TOYOBO, Ltd.

** KYOTO University.

2-3 認識部

ここは、誤差逆伝播法を用いた、入力層(16点)と隠れ層(中間層)(20点)、出力層(5点)から成る多層パーセプトロンである^[4]。学習データには、5種類のボトル(回転なし)の正規化スペクトル(1次~16次)を用いた。

3 実験

5種類の任意の角度で回転しているボトルの識別実験を行った。図3に実験に用いた5種類のボトルの画像とSNAKEの収束結果を示す。SNAKEの初期位置は、画像枠一杯に設定し、原画像を20度毎に回転させた。その結果の一つを図4に示す。対象物の回転によらず、唯一の識別細胞が興奮しており、出力の変動も小さく、対象物の識別ができていたことが確認された。また、他の四つのボトルも同様の結果が得られた。

また、SNAKEは、逐次収束していくので、収束過程での識別が可能である。図5に、ズームアップして撮像したボトル画像の認識過程を示す。これにより、識別対象物の特徴的な輪郭が得られた時点で正しい識別ができることが分かった。また、画像中の対象物の大きさに関わらず識別が可能なが示された。



図3 識別に用いたボトルの画像

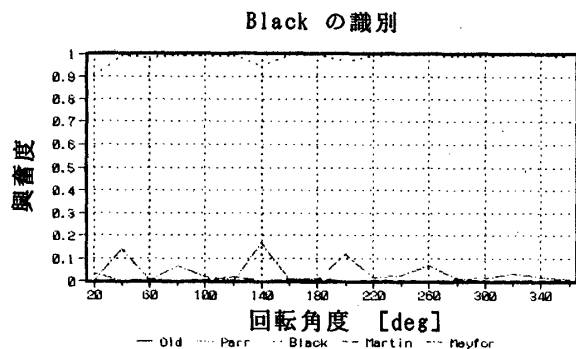


図4 回転に対する識別の変化

Meyfor の識別

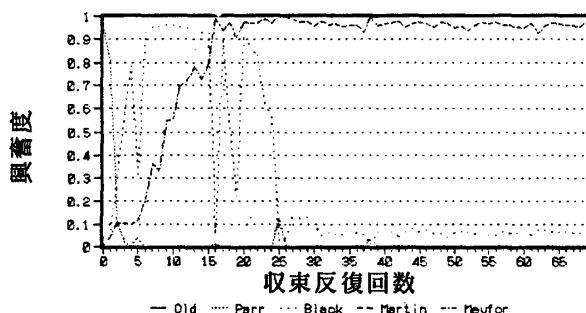


図5 収束過程での出力の変化

4 考察

SNAKEとニューラルネットワークを組み合わせることにより、画像中の対象物の回転角度、大きさ、位置に関わらず対象物の正確な識別ができた。また、学習の完了したニューラルネットワークは、各画像の輪郭のスペクトルの特徴を見つけ、その変動分を許容するものであった。今回は、5種類の対象物の識別を行ったが、実用的にはより多くの対象物を識別する必要がある。その場合には、大規模なネットワークを構築するより、このような小規模のネットワークを並列に用いて統合するほうが学習も速く、修正も容易であり、実用的であると考えられる。

5 おわりに

今回は、画像中の対象物が一つの場合であったが、複数の対象物があった場合でも、SNAKEを分裂型にすることで対処できる^[3]。この場合、分裂したSNAKEに対して、後のニューラルネットワークに、さらなるネットワークを層状に組み合わせることによって、より高度な識別が可能となろう。これは、ニューラルネットワークの学習方法と共に、現在研究を進めている。また、SNAKEは、初期位置が輪郭の近くであった場合、非常に高速に収束するので、ここで提案した手法を実用的な問題に適用していく。

参考文献

[1] KASS, WITKIN: Snakes: Active Contour Models, International Journal of Computer Vision, 321-331(1988).
 [2]坂口,美濃,池田: 圧力エネルギーを持つSNAKES, 電子情報通信学会1990年春期全国大会予稿集.
 [3]坂口,美濃,池田: SNAKEパラメータの設定についての検討, 信学技報, PRU90-21, 1990.
 [4]D.E.Rumelhart and J.L. McClelland: Learning Internal Representation by Error Propagation, Parallel Distributed Processing, Vol.1., MIT Press, 1986.