

可変長符号化に基づく新たな画像2値化法

2D-2

陳涛 高木幹雄

東京大学生産技術研究所

1 はじめに

ニューラルネットの画像処理への応用は最近注目されつつあり、様々な画像処理応用方面への提案がされている[1]。本論文では、RLC (Run Length Coding) に基づく新たな2値化法について、修正したニューラルネット構造を中心に実験的検討を加える。改良した逆伝播学習法を利用し、未知画像の任意のヒストグラムを入力すると、視覚的に適切な閾値出力するものである。画像の前処理・変換とヒストグラムの分割・グループ等が本システムに含まれている。これから、その出力結果を後継ニューラルネットの入力として画像認識・パターン識別を行なう。

2 ニューラルネット

階層型ニューラルネットにより、入力パターンから出力パターンへの任意の変換が可能であることはよく知られている。その機能を利用した、画像の輪郭抽出・ボケ除去フィルタ、2値化、画像圧縮・画像再構成や、又その特殊な場合であるパターン分類・パターン識別、波形認識と復元等の研究論文がある[1, 2, 4]。

2.1 修正した階層型ニューラルモデル

出力層の内部変数と出力はシグモイド関数 $F(\alpha)$ で関係付ける。

$$F(\alpha) = \frac{1}{1 + e^{-2\lambda\alpha}} \quad (1)$$

ここで、 $\lambda(1/T)$ はシグモイド関数の傾きを決定する温度 (T) パラメータである。出力層のユニット Y の出力誤差を次式で定義する。

$$E_p^{bp} = \frac{1}{2} \sum_{i=0}^{M-1} (O_{pi} - Y_{pi})^2 \quad (2)$$

p は第 p 種類の入出力パターンを表す。ここで、誤差関数に対して E^{bp} を最小化する [2, 3]。適当な正係数 η 、慣性係数 α 、学習係数 η を導入して、次式になる。

$$\Delta w_{ij}^{(n+1)} = \eta \delta_j x_i' + \alpha \Delta w_{ij}^{(n)} \quad (3)$$

δ_j は

$$\delta_j = \begin{cases} Y_j(1 - Y_j)(O_j - Y_j) & \text{出力ノード} \\ x_j(1 - x_j) \sum_k \delta_k w_{jk} & \text{その他} \end{cases} \quad (4)$$

2.2 2値・多値化に用いられるモデル

ヒストグラム $H(GL)$ ($0 \leq GL \leq M$) を各グループに分け、最適なグループ分割点を求められ、先ず次の様に正規化する。

$$h(j) = H(j) / \sum_{i=0}^{M-1} H(i) \quad (0 \leq j \leq M) \quad (5)$$

ヒストグラムをグループに分ける時、ヒストグラムの形状を考えなければならない。グループの数を G と、グループ分割距離を DST と設定すると、分割点 $D \subseteq [d_0, d_2, \dots, d_{G-1}]$ は次の式で決定される。

$$\sum_{j=d_g}^{d_{g+1}} h(j) = \frac{1}{G} \quad (0 \leq g \leq G-2) \quad (6)$$

入力層のユニット X_k の入力信号 $h^*(k)$ は

$$h^*(k) = \begin{cases} ON & \text{if } d_k - d_{k-1} \leq DST \\ OFF & \text{その他} \end{cases} \quad (7)$$

3 画像2値化

画像 ($N_1 \times N_2$) の濃淡レベルが正の整数集合 $GL \subseteq [0, M]$ (レベル数 M; レベル M が最も濃いものとする) であるとき、画像2値化では、2値画像を得る為の閾値 $T \subseteq GL$ を適切に選択することが目標となる。

ここで、 $f(x, y)$ は座標 (x, y) のピクセルの輝度を表す。閾値を $T \subseteq GL$ とし、閾値処理により得られる2値化画像の輝度を $B \subseteq (0, 1)$ とすると、閾値を T として画像 $f(x, y)$ を閾値処理して得られる2値化画像 $f'(N_1 \times N_2) \rightarrow B(0, 1)$ は、

$$f'(x, y) = \begin{cases} 1 & \text{if } GL > T \\ 0 & \text{その他} \end{cases} \quad (8)$$

と表される。

3.1 従来の統計法による2値化

閾値を定める際には、ヒストグラムがよく利用される。ヒストグラムの形状・分布を利用する方法では、ヒストグラムの形状や仮定される分布や画像集合の事前情報から、閾値を決定する。例えばヒストグラム上のピークを正規分布に従うと仮定して検出し、それを利用してある評価関数を最大化すると閾値を得たり [5]、或はヒストグラムを正規化し、それを分布関数として取り扱い、定式化された近似誤差を最小化することによって閾値を求める [6]。しかしながら、その場合ヒストグラムと得られた閾値との関係は明瞭ではなく、その上、大量の画像集合に対して的確な事前情報を得ることが一般に困難となる。そう言う統計法の結果は必ずしも人間の視覚に一致するとは言えない。

3.2 新たな2値化モデル

2値・多値化問題をヒストグラムから閾値への変換であると捉え

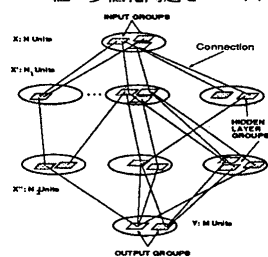


図1: 2値・多値化に用いられるモデル



図2: 原画像 LENA

New Binarization Method Based On Run Length Coding

Tao Chen Mikio Takagi
University of Tokyo

ると、ヒストグラムと閾値の間での適応な変換器を設定することが、上述問題を解決する一方策と成り得る。そこで我々は、ニューラルネットのパターン変換器としての機能に着目し、ヒストグラムと適切な閾値との関係を学習させたニューラルネットの画像2値・多値化への応用を提案する [12]。

2値・多値化基準： 画像内に埋没しているオブジェクト部分と背景部分を分析判別法と最小誤差法で良好に分離出来る範囲内、RLC可変長符号化法で符号化された2値化・多値化画像のデータ量を最大化するような閾値。

2値・多値化した画像にある各オブジェクトの数が多ければ多い程画像のRLC量が大きくなる。統計的に、画像ピクセルの濃淡レベルが一樣乱数の場合、RLCのデータ量が一番多い。単純な画像の場合はRLCデータ量が少ない。原画像が持つエッジや不連続成分など高域の周波数成分が出来るだけ2値・多値化画像に残される為に、RLCと言う補助基準を利用した。シミュレーションによると、視覚的に良い2値・多値化画像を得られる範囲内で、2値・多値化した画像のRLCデータ量が多い方が、原画像の細部を良く保存されることが分かった。良い教師信号の範囲内、RLCデータ量の多い2値・多値化画像に対応した閾値を最終閾値にする。この方式で、原画像の不連続性を良く保持した2値・多値化画像からの濃淡画像復元に有効であることが分かった。RLCデータを得る為に、2値化画像に対して、上から下へスキャンする。スキャン線と第kオブジェクト輪郭線の交差点のY座標を y_{ki} 、X座標を $x_{ki(2j)}$ と $x_{ki(2j+1)}$ とすると、RLCのデータ量は $E^{rl} = \sum_{k=1}^K O_k \sum_{i=0}^{j_k-1} y_{ki} \sum_{j=0}^{j_k-1} x_{ki(2j)} x_{ki(2j+1)}$ 。ここで、Kは2値化画像内のオブジェクト数、 y_{k0} と $y_{k(j_k-1)}$ は各々第kオブジェクトの最小と最大Y座標である。

RLC符号の要素を考慮すれば、最終的な学習アルゴリズムは修正した逆伝播学習法になる。これは式 $E = \beta_{bp} E^{bp} - \beta_{rl} E^{rl}$ で与えられるEを最小化することである。 β_{bp} の項は出力と教師信号との誤差最小化する保証項である。 β_{rl} の項は得られた2値化画像のRLCデータ量最大化する保証項である。

4 シミュレーション結果

実験対象とした画像はNOAA画像とSIDBA標準画像である。画像のサイズは512x480と512x512、濃淡レベル数は256である。

4.1 2値化結果

パラメータDSTとGは実験的に定める。使用するニューラルネットは、入力層のユニット数を32、中間層のユニット数を64と256、出力層のユニット数は入力層のユニット数に等しく32と設定する。所望の閾値の範囲は出力信号が $Y_i = \max(Y(0), \dots, Y(G))$ であるユニット Y_i から決定出来る。これから、RLC補助基準を用いて、視覚的に最適な閾値を決定する。学習環境は、学習係数0.2、慣性係数は0.8、接続重みとオフセットの初期値は-0.5から0.5の一樣乱数である。本階層型ニューラルネットの学習法は、教師付き学習法であるので、教師信号 $e(i)$ ($i=0, \dots, G$) を設定する必要がある。適切と思われる閾値に対応するユニットとその近傍ユニットの教師信号の値を1とし、それ以外は0とする。テストサンプルに対する適切な閾値範囲は、分析判別法と最小誤差法により決定する。教師信号が設定出来れば、ニューラルネットワークの学習は、入力と教師信号の組、すなわちヒストグラムと適切な閾値の組を順次提示することにより進行する。トレーニングを繰り返しながら、ノイズを加える。

5 終りに

ニューラルネットワークを用いて画像2値・多値化処理とエッジ・輪郭抽出について述べ、NOAA画像と人物画像に応用した例について報告した。今後の方針として、画像2値・多値化を行ったり、エッジ・輪郭を抽出したものを画像の特徴としてデータベースに書き込み、この特徴に基づく分類・認識を行なう。

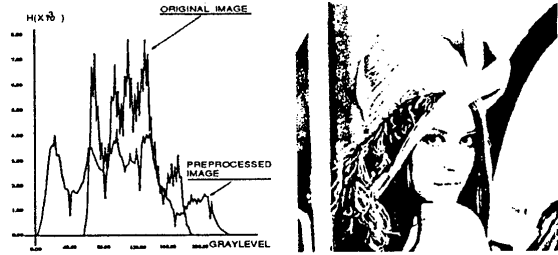


図3: 原画像と線形処理後画像のヒストグラム

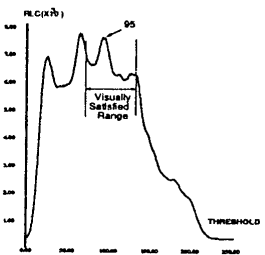


図5: RLCにより得られた閾値



図4: 判別分析法に基づいた2値化画像



図6: 出力2値化画像

参考文献

- [1] D.E.Rumelhart and J.L.McClelland, Parallel Distributed Processing, MIT Press, Vol.1,2, 1989
- [2] R.P.Lippmann, An Introduction to Computing With Neural Nets, IEEE ASSP Magazine, April 1987
- [3] S.I.Amari, Proceedings of IEEE, Vol 78, No.9, p1443-1463, September 1990
- [4] Proceedings of IEEE, Vol 78, No.9-10, p1443-1463, September 1990
- [5] J.O.Eklundh and A.Rosenfeld, Peak Detection Using Difference Operators, IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-1, 3, pp.317-325, July 1979
- [6] J.Kittler and J. Illingworth, Minimum Error Thresholding, Pattern Recognition, Vol.19, No.1, pp.41-47, 1986
- [7] J.J.Hopfield, Neurons With Graded Response Have Collective Computational Properties Like Those of Two-State Neurons, Proc. Natl. Acad. Sci. USA, Vol.81, pp.3088-3092, May 1984
- [8] Medical Imaging Technology, Vol.9, No.4, September 1991
- [9] A.Rosenfeld and A.C.Kak, Digital Picture Processing, New York Academic Press, 1976
- [10] H. Shvaytser, Learnable and Nonlearnable Visual Concept, IEEE Trans. on PAMI, vol.12, No.5, May 1990
- [11] R.A.Kirsch, Computer determination of the Constituent Structure of Biological Image, Computers and Biomedical Research, V4.3, pp.123-134, June 1971
- [12] T.Chen and M.Takagi, Image Binarization By Back Propagation Algorithm, Proceedings of 17th International ISPRS Congress, Washington D.C., August 1992