

H-006

## リアルタイム学習に向けた誤差逆伝搬網のハードウェア化 Hardware Implementation of Back Propagation Neural Network for Real Time Learning

間所 洋和<sup>†</sup>  
Hirokazu Madokoro

佐藤 和人<sup>†</sup>  
Kazuhiro Sato

石井 雅樹<sup>†</sup>  
Masaki Ishii

### 1. はじめに

誤差逆伝搬網 (Back Propagation Neural Network, 以下BPと記す) [1] は, ニューラルネットワークの中でも優れた学習能力と写像特性を示すことから, 様々な分野に応用されている. 我々は, 人と人とのコミュニケーションを積極的に支援する情報機器の開発を目的に, そのプラットフォームとなる評価用FPGA (Field Programmable Gate Array) ボードの試作開発と, FPGA 上への学習機構の実装を目指している.

BP のハードウェア化に関しては, 様々なモデルが提案されているが, その評価は2値図形パターンの認識などに留まっている [2, 3]. 本研究では, 画像情報をリアルタイムに学習するBPのハードウェアモデルを提案し, FPGA への実装を想定したシミュレーション実験から, 本提案モデルの識別能力と表現能力について評価する.

### 2. 提案モデル

BP の特徴は, 次式で示すシグモイド関数の特性を利用している点にある.

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

ノイマン型計算機上でC言語など高級言語を用いてBPを記述した場合は, 浮動小数点演算や $e^x$ 関数などの算術演算関数を用いることができる. しかしながら, ハードウェアロジックとして実装した場合, ゲート規模や回路合成の可否などから, 浮動小数点表現や算術関数を簡略化する必要がある. そこで本研究では, 浮動小数点演算を固定小数点表現に, シグモイド関数を階段関数に置き換える. なお, モーメント項は考慮していない.

固定小数点表現は8ビットで表現し, 最上位ビットを符号ビット, 符号ビットに続く上位4ビットを整数部, 残りの下位3ビットを小数部とした. よって, 本モデルで表現できる値の範囲は,  $-15.875$  から  $+15.875$ , 分解能は  $0.125$  である.

シグモイド関数は図1で示す階段関数に近似した. また, 学習で用いるシグモイド関数の微分関数は図2で示すピラミッド型の階段関数に近似した. 教師信号は0と1の2値とし, 0の場合は負の方向に誤差を修正するため, 図2の階段関数を $x$ 軸に対称変換した関数として用いる. なお, 本稿では, 通常のBPは浮動小数点モデルと表現する.

### 3. 評価実験

#### 3.1 XOR

本モデルにおける基礎的な能力を評価するために, XOR問題のシミュレーション実験を行なった. ネット

<sup>†</sup>秋田県工業技術センター

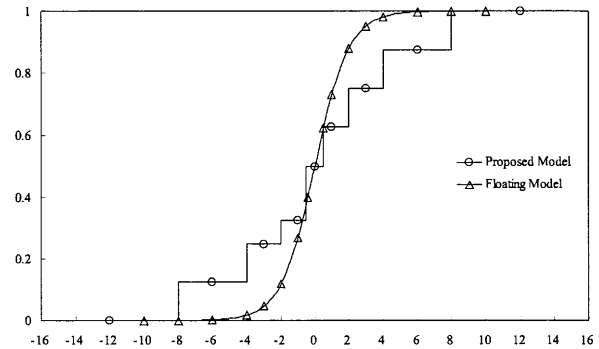


図1: シグモイド関数

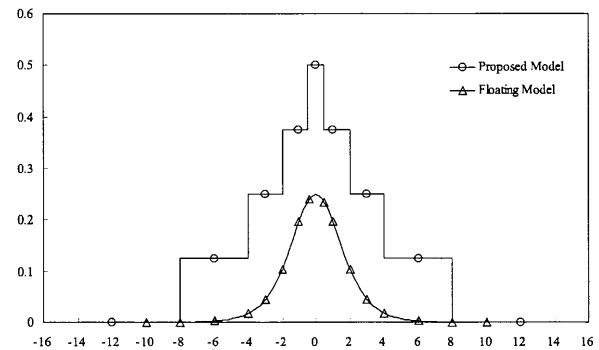


図2: シグモイド関数微分形

ワークは, 入力層2ユニット, 中間層4ユニット, 出力層1ユニットとした.

学習におけるエラーレートの推移を図3に示す. 提案モデルでは, 約1,000回程度の学習でエラーレートが0.1以下まで収束した. 一方, 浮動小数点モデルでは, 2,500回以上の繰り返しが必要となった. 提案モデルの学習回数が半分以下になったのは, 図1に示すシグモイド関数の微分値を大きく設定したためと考える. 微分値を大きく設定したのは, 固定小数点表現の分解能の制約からであり, 結果として学習の収束速度を高める要因になった. 一方, 本モデルは誤差の修正が大きいと, 浮動小数点モデルと比較してエラーレートの振幅が大きく揺らぎが発生している.

#### 3.2 実画像応用

実画像への応用として, 一般的な環境下で撮影したシーン画像からヒトの肌領域を抽出する実験を行なった. 学習で用いた画像を図5(a)に示す. 学習データは, マウスクリックを通じて肌領域を代表する座標と肌領域以外の座標を5地点ずつ設定する. これらの各座標を中心として第2近傍までの領域からRGBの各成分値を取得す

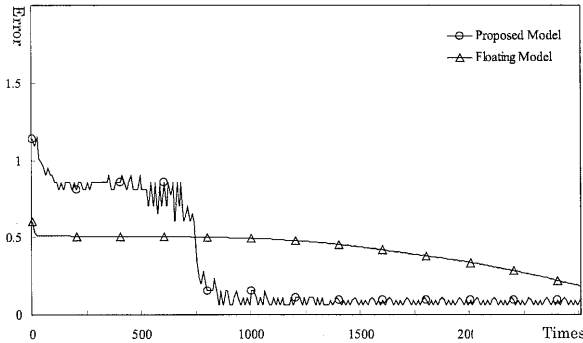


図 3: XOR 学習結果

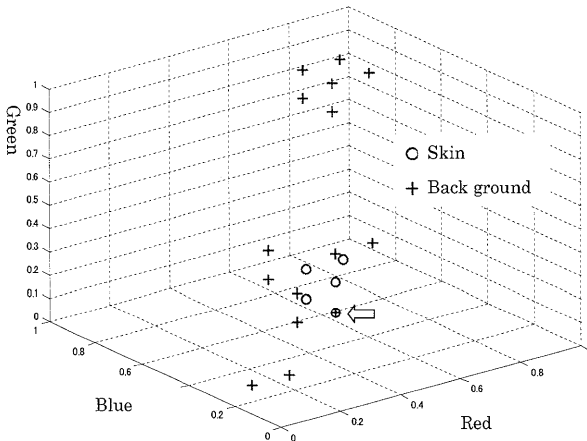


図 4: 学習データ

る。学習データの空間分布を図 4 に示す。入力層の各ユニットには上記学習データの各成分を入力し、出力層には肌領域又は肌領域以外を示す教師信号を与える。したがって、BP のネットワーク構成は、入力層 3 ユニット、出力層 2 ユニットとした。中間層については、これまでの予備実験に基づいて 3 ユニットとした。

浮動小数点モデルの肌領域抽出結果を図 5(b) に、本モデルの抽出結果を図 5(c) に示す。浮動小数点モデルでは背景の一部で誤抽出が認められるものの、顔領域を中心とした肌領域が安定して抽出できている。一方、本モデルでは、肌領域内において欠損領域が多数見受けられるものの、全体的には顔領域を中心とする肌領域が抽出できている。

### 3.3 考察

本モデルの最大の課題は分解能にある。現状、本モデルのデータ表現は、 $-15.875$  から  $+15.875$  までの  $0.125$  刻みとなっている。結合荷重の表現は中間層を増やすことによって補間できるが、入力層に提示される入力データは  $0$  から  $1$  の間で正規化するため、9 段階の精度でしか表現できない。サンプリング精度が粗い場合は画像上におけるノイズ低減に寄与するが、その反面、特徴量が打ち消される。また、サンプリングの影響で図 4 の矢印で示すような矛盾する学習データが発生している。同じ入力値に対して異なる教師信号が与えられた場合、ネットワークの内部に矛盾が生じ、学習の収束を阻害する。

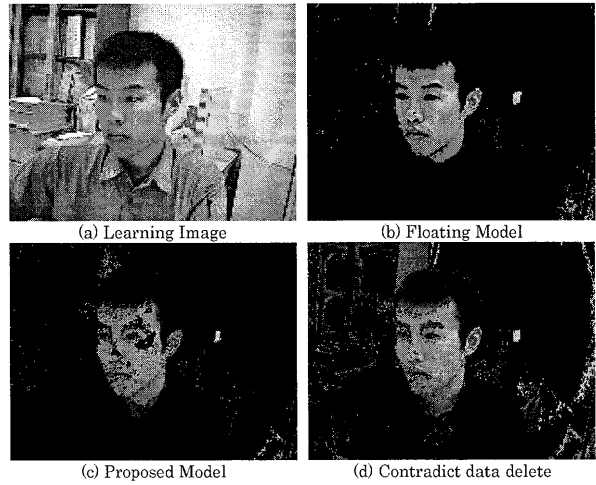


図 5: 肌領域抽出結果

そこで、学習データ中から矛盾するデータを取り除いて学習を行った結果、エラーレートは、取り除く前が  $29.91$  であったのに対して、取り除いた後は  $20.72$  まで収束した。しかしながら、肌領域の抽出結果 (図 5(d)) は、背景領域において誤抽出が増加している。この結果は、取り除いた学習データが BP が形成する識別境界付近に位置していたことを意味する。すなわち、矛盾するデータを取り除くことによってネットワークは特化した方向に収束し、汎化能力が低下したことになる。以上のことから、今後は固定小数点の分解能を増やすためにビット幅を拡張する必要がある。

## 4. まとめ

本研究では、リアルタイム学習に向けた BP のハードウェアモデルを提案し、肌領域の抽出実験に適用した。その結果、本提案モデルは一般的なシーン画像に対しても十分な学習能力と分類能力を有することを示した。

今後は、ハードウェア記述言語を用いて本モデルを FPGA に実装し、表情認識や手話認識などへの応用と、これら機能の IP (Intellectual Property) 化を目指す予定である。

## 参考文献

- [1] D. E. Rumelhart, G. E. Hilton, and R. J. Williams, "Learning representations by back-propagation errors," *Nature*, vol. 323-9, pp.533-536, Oct. 1986.
- [2] H. Hikawa, "A digital hardware pulse-mode neuron with piecewise linear activation function," *IEEE Trans. Neural Networks*, vol.14, no.5, pp.1028-1037, Sept. 2003.
- [3] Y. C. Kim and M. A. Shanblatt, "Architecture and statistical model of a pulse-mode digital multilayer neural network," *IEEE Trans. Neural Networks*, vol.6, no.5, pp.1109-1118, Sept. 1995.