

5H-9

超並列計算機上のニューラルネットワークを用いた 顔画像認識に関する研究

山森一人 阿部亨 堀口進

北陸先端科学技術大学院大学情報科学研究科

1 はじめに

近年、人工知能処理の研究はますます盛んになってきている。この研究分野の1つである、ニューラルネットワークを用いた情報処理は高度並列分散型の情報処理と言ってもよい。しかし、現状では実用規模の応用問題をニューラルネットワークで解くには計算機の能力がまだまだ不足している。そのため、超並列計算機上にニューラルネットワークを実装して処理の高速化を図る研究が行われるようになってきている。

本研究では、超並列計算機nCUBE2上に並列ニューラルネットワークを実装し、その性能評価を行なった。さらに、並列ニューラルネットワークの応用例として顔画像認識を取り上げ、モザイク化した画像による認識と、エッジ抽出を行なった画像による認識法を実装し、その評価を行なう。

2 並列ニューラルネットワーク

バックプロパゲーションアルゴリズム（以下BP法と呼ぶ）はD.E.Rumelhartらによって提唱された、誤差を最小化する最急降下法の一つである。これは一般的にニューラルネットワークを学習させる場合に多く用いられるものであるが、処理時間が長いため様々な高速化手法が試みられている [1]。

通常BP法は層状の構造を持ったネットワーク上に実装される。そのネットワークはinput(入力)、hidden(隠れ)、output(出力)の3層から成る場合が多い。各層はそれぞれ幾つかのユニットを持ち、各ユニットはそれぞれ活性化関数により、入力に応じた出力を持つ。層間のユニットは重み付きのリンクで結ばれているが、同じ層のユニット間は結合していない。このことから、BP法を用いたニューラルネットワークを並列化するに当たり、3種類の手法が用いられている [2]。

1. ユニット並列
2. 学習セット並列

Face Image Recognition on Parallel Neural Network
Kunihito Yamamori, Toru Abe, Susumu Horiguchi
Japan Advanced Institute of Science and Technology
15. Asahidai, Tatsunokuchi, Ishikawa 923-12, Japan

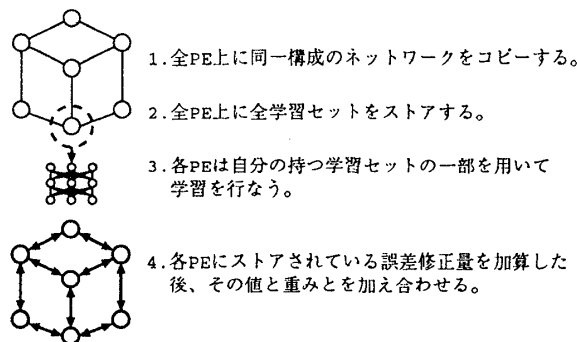


図2: シミュレータのアルゴリズム

図1: 学習セット並列モデルのアルゴリズム

3. パス並列

ユニット並列モデルは、ニューラルネットワーク中の各ユニットがそれぞれ並列に動作可能である点を利用したものである。このモデルは最も分かり易いものではあるが、大規模なシミュレーションを行なう場合、膨大な数のプロセッシングエレメント（以下PEとする）を用意せねばならず実用的ではない。また、PE間通信が多く発生するという欠点を持つ。

学習セット並列モデルは、BP法における学習が各学習パターンによる重み修正量の線型和であることを利用している。この線型性により、全ての学習パターンが独立に並列処理できる。

パス並列モデルは、BP法で行なわれるフォワードパス、バックワードパスをそれぞれ並列に行なうものである。すなわち、各学習パターンによる誤差ベクトルが出力層から隠れ層へ伝えられている間（バックワードパス）、入力層は隠れ層へ次の学習パターンを伝えることができる（フォワードパス）。このモデル単独では大きな速度向上は望めないため、他の2つのモデルと組み合わせる必要がある。

本研究においては通信量が小さく、最も高速化が期待できる学習セット並列モデルを採用する。実装したアルゴリズムの概要を図1に示す。図2に、学習セット並列モデルにおいてEncode/Decode問題を使用しPE

数と学習セット数を変化させたときの学習速度高速化率を示す。

3 顔画像認識への応用

並列ニューラルネットワークの応用例として顔画像認識をとりあげる。本研究では2種類のモデルを用いて認識を行なう。1つは顔画像を複数のブロックに分割し、各ブロック内の平均濃度をニューラルネットワークに入力して認識を行なう[3]。もう1つのモデルは各画像からエッジを抽出したあと複数ブロックに分割し、各ブロック内にある線分の大きさをニューラルネットワークに入力して認識を行なう。画像のサイズは512×512ピクセル、各ピクセルは256階調の濃度情報を持つ。今回は32人分の顔画像を用いて認識実験を行なった。

3.1 モザイク化画像による認識

まず顔の部分だけの情報を取り出すため、256×256ピクセルの大きさで顔の中心部分を切りだす。その後8×8個のブロックに画像を分割し、各ブロック内のピクセルが持つ平均濃度をブロックの代表値としてニューラルネットワークに入力する。ニューラルネットワークの出力ユニットは人数分用意し、ある画像に対して特定の出力ユニットが反応することで認識を行なう。

3.2 エッジ化画像による認識

モザイク画像を用いた場合と同じく、256×256のサイズで顔の中心部分を切りだす。その後平滑化処理、二値化処理を行なった後、ガウシアン-ラプラシアンフィルタによりエッジ情報を抽出し、Hilditchのアルゴリズムを用いて細線化処理を行なう。これらの処理が終わったあと、画像を4×4個のブロックに分割する。各ブロック内において、左下から右上に向かい、図3に示すようなパターンを検索し、その正規化した個数をニューラルネットワークへの入力とした。この手法により、ブ

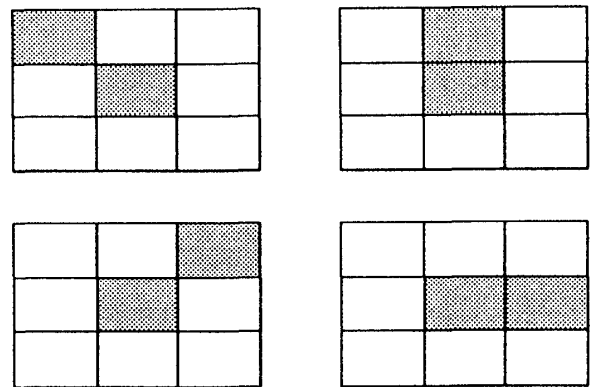


図3: 検索パターン

ロック内の線分の大きさに関する情報が得られる。また、線分の位置に関する情報は詳細には得られないが、各ブロックの位置により、おおまかではあるが抽出することができる。

4 おわりに

超並列計算機上に実装した3種類のニューラルネットワークの性能を比較・評価を行なった。超並列計算機でアプリケーションを実装する場合、PE間通信がその速度向上のボトルネックになる。このため本研究では、並列BPモデルの中では最も通信量が少ない学習セット並列モデルについて検討を行なった。その結果、256のPEを使用したとき、単一PE使用時に比べ約100倍の高速化が可能であることを示した。

さらに並列ニューラルネットワークの応用例として、顔画像認識を取りあげ、モザイク画像とエッジ画像を用いてそれぞれ個人識別が可能であることを述べた。

今後の課題としては、より通信量の少ない学習アルゴリズムをさらに検討する必要がある。また顔画像認識で用いた人数をさらに増やした場合や、元画像にノイズがある場合などについての評価が必要である。

参考文献

- [1] Scott E. Fahlman, "An Empirical Study of Learning Speed in Back-Propagation Network", CMU-CS-88-162, Sep. 1988.
- [2] Alexander Singer, "Implementation of Artificial Neural Network on the Connection Machine", Thinking Machines RL-90(1990).
- [3] 小杉 信, "モザイクとニューラルネットを用いた顔画像の認識", 電子情報通信学会論文誌, Vol. J76-D-II, No. 6, pp. 1132-1139, 1993.

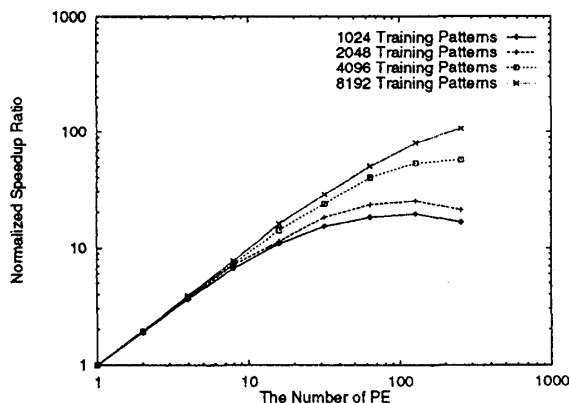


図2: 学習セット並列モデルの高速化率