

要 旨

SOM の並列学習と そのデータ駆動型プロセッサへの 効率的な実装法

福永 諭

自己組織化マップ (Self-Organizing Map:SOM) はデータ分析の分野においてよく知られている手法である。しかし，ユニットの数が増えるにつれて，計算量が膨大になるため，並列処理による高速化が求められる。一般的な並列手法としては競合層分割法が挙げられるが，これは更新時の負荷が不均一になるという欠点を抱える。この問題を解決したものが，入力層分割法である。しかし，入力層の数だけ PE を準備する必要のあるこの手法は，実装時にプログラムの規模が大規模になってしまう。本研究では既に入力層分割法を発展させ，プログラム規模が大きくなり過ぎない一般化入力層分割法 (Generalized Input layer Parallel Model:GIPM) を開発している。本論文では，この GIPM を採用した SOM (以降 GIPM-SOM) と，同様にして競合層を一般化した，一般化競合層分割法 (以降 GCPM-SOM) をデータ駆動型プロセッサである DDMP (Data Driven Media Processor) 上に実装し比較した。その結果，GIPM-SOM の方が，処理時間及び並列処理性能において，高性能であることを明らかにした。

キーワード 自己組織化マップ，SOM，データ駆動型プロセッサ，並列処理，一般化入力層分割法

Abstract

Parallel Implementation of Self-Organizing Map on a Data-Driven Processor

Satoshi Fukunaga

The Self-Organizing Map(SOM) is a popular tool for data analysis. However, the execution time of the training phase is long for many application domains that either deal with very large data sets or require real-time responses. In order to resolve these problems, many researchers have developed the parallel SOM algorithm. Popular parallel SOM algorithm is used Competitive layer Parallel Model(CPM). But CPM algorithm make load imbalance among PEs occurs. To solve this problem the Input layer Parallel Model(IPM) was proposed. Then we present a novel scheme consisting of a new parallel training algorithm Generalized Input layer Parallel Model(GIPM). In this thesis, GIPM algorithm and Generalized Competitive layer Parallel Model(GCPM) algorithm implementation on a data-driven multiprocessor system, and compare its performance. As a result, it was shown that GIPM is high-speed and high performance more than GCPM.

key words SOM , Parallel Algorithms , Data Driven Processor, Generalized Input layer Parallel Model