An Architecture of Large Scale Neural Networks

Eiko Sugawara
School of Information Science,
Japan Advanced Institute of Science and Technology

July 2, 2004

Abstract

In today's advanced VLSI technology, large scale neural networks (NNs) can be implemented into a single chip; consequently, high-speed, low-power, NNs can be realized in an extremely small area. Such integrated NNs are expected to be used as embedded systems in practical applications. However, a major issue in realizing such embedded systems is efficient self-reconfiguration strategy such that a system can automatically avoid faults in the run time without any help of host computer. A self-reconfiguration of NN is the task to adapt an NN to a given application automatically, changing the number of neurons and their links, and acquiring the optimal weights for their links. This paper concentrates on multi-layer NNs, which is the most popular architecture used in image and signal processing, and proposes a self-reconfigurable multi-layer NN.

First, a bus-connected multi-layer NN and simple fault compensation scheme called function shifting are proposed. The problem of implementing a multi-layer NN is the huge wiring area for weighted links between adjacent layers. The proposed architecture reduces the number of links and facilitates the incorporation of fault tolerance mechanism. The proposed NN with the function shifting is implemented using FPGA and the hardware overhead of the function shifting is shown to be quite small.

Second, weight training scheme is considered. The drawback of available weight training schemes, e.g. back propagation, is the large computational cost and additional complex hardware for each neuron, which in turn increases the probability of their being faulty. We employ genetic algorithms (GAs) for weight training. The advantage of GAs is fast computation and independent hardware. A GA processor is implemented in hardware and the hardware overhead is also discussed.

Finally, we propose a self-reconfigurable multi-layer NN employing the function shifting and GAs. The proposed architecture offers the combined advantages of low hardware overhead for adding spare neurons and fast weight training time. In this architecture, a certain number of faulty neurons are replaced by the function shifting and the weights of remaining faulty neurons are trained by GA. The proposed system is implemented in hardware to show the possibility of self-reconfiguration.

Key Words: **Multi-layer neural networks, Hardware implementation, Fault tolerant, Self-reconfiguration**