Neural Classifiers Implemented in a Transputer Based Parallel Machine

J. M. Seixas[1], A. R. Anjos[1], L. P. Calôba[1], C. B. Prado[1], J. C. R. Aguiar[2] and A. C. H. Dantas[1]

[1] COPPE/EE/UFRJ, C.P. 68504, Rio de Janeiro 21945-970, Brazil
[2] CEPEL, Rio de Janeiro, Brazil

Abstract. A transputer based parallel machine is used as a development platform for fast neural signal processing applications in physics and electricity. The 16 node machine houses 32-bit floating point digital signal processors running as coprocessor for the transputers, so that signal processing can be optimized. The application in physics consists in a prototype of an online validation system for a high event rate collider experiment, which is implemented using neural networks for physics process identification. In electricity, a nonintrusive load monitoring system for household appliances is developed using a neural discriminator to identify seven groups of equipment.
keywords: parallel processing, neural networks, classifiers, principal component analysis, real-time systems.

1 Introduction

Neural networks find a vast area of applications in signal processing domain. In particular, as classifiers, neural networks have been extensively used due to their ability in combining high classification efficiency and processing speed. As inner products are the main mathematical operations required by the neural processing during the production phase, neural classifiers can be efficiently implemented in commercial devices, such as digital signal processors (DSPs). Ultimate limits in processing speed can be reached if the natural parallelism of neural networks is explored. Therefore, when both performance and speed are of concern in a project, neural networks may be considered an efficient solution.

In this paper we describe two applications of neural processing in a parallel computing environment. In the first one, an online validation system is designed for a high-event rate collider experiment in high-energy physics, which is being developed at CERN (Switzerland). In this experiment (LHC), bunches of particles will collide in periods of 25 nanoseconds, so that a large amount of experimental data will be produced. However, events with physics significance will be extremely rare. Thus, the incoming data flow needs to be reduced by a highly sophisticated online validation system for deciding whether a given event should be discarded or stored by the data acquisition system. The second application involves the design of a nonintrusive electrical load monitoring system for household appliances. As household appliances respond for a significant fraction of the total demand in power consumption, the knowledge of the consumption profile of this segment is valuable for energy conservation and alleviation of the electrical system in peaking periods.

Both applications described in this paper are developed for the TN-310 system, a multiple instructions multiple data parallel computer with a distributed memory architecture [1]. The system in consideration houses 16 nodes based on T9000 transputers, which communicate with each other by means of a network of C104 chips. Each node has access to the communication network through four high speed (100 Mbits/s) serial links (DS-links). For optimizing signal processing applications, the system includes fast DSPs (ADSP-21020) running as coprocessors for the transputers.

The application development makes use of the C toolset environment, in order to achieve ultimate speed. This software layer features hardware configuration and a set of libraries and development tools to support ANSI C programming.

In our case, the TN-310 system is accessed through a PC running MS-DOS and MS-Windows.

2 The Applications

For both applications to be described, fully-connected feedforward neural classifiers were designed. These classifiers were trained on preprocessed data by means of backpropagation method [2]. The preprocessing phase was introduced in order to reduce the dimensionability of data input space, so that more compact classifiers could be developed.

2.1 Online Validation System

The LHC (Large Hadron Collider) will produce event rates up to 100 MHz but only very rare events will have physics significance to the experiment. In order to remove such deep background noise that hides potentially interesting events, an online event validation system is being designed as a multi-level triggering system. In the first level, a very fast algorithm will be capable to reduce the event rate to 100 kHz. The second-level triggering (LVL2) system will act on events that passed the conditions of the previous level. Not all regions of the detectors contain valuable information for a given event, so that only restricted areas in the detector (known as Regions of Interest - ROI) will be moved by the first-level system to the LVL2 system. This will alleviate bandwidth requirements on the LVL2 system, which is expected to achieve a further reduction factor of 100 in the event rate. The surviving events will be analyzed by a third level trigger.

The prototype being implemented in the TN-310 system concerns the LVL2 operation. The architecture used splits the LVL2 operation into two phases. In the first, raw detector data is translated into features capable to efficiently identify the relevant physics processes. This feature extraction works on ROI information provided by the detectors that participate in the LVL2 decision:

calorimeters (for energy measurement), trackers (for image display of an interaction) and muon chambers (for muon detection). Next, the global decision phase correlates detector information for analysis refinment. Features are combined to compute the probability of a particle to be found in a given ROI, so that physics processes can be identified.

Both phases may be performed by neural processing [3]. For the calorimeter subsystem, feature extraction was performed by grouping cells of deposited energy in a ROI in the manner shown in Figure 1. Thicker lines define the border of each region, and cells belonging to a region have their energies added up to form group sums. For this implementation, the neural network comprised 3 hidden nodes and a single output neuron, so that electron/jet (of particles) discrimination could be efficiently performed. For the other three detectors involved in the LVL2 decision scheme, a simulation of current classical algorithms was implemented.

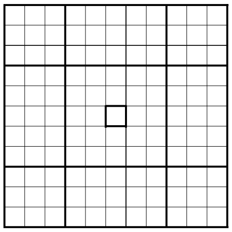

Fig. 1. The grouped sum structure.

For the global decision phase, it was explored the capability of neural networks in correlating information in a multidimensional feature space. A set of twelve features was fed into the neural classifier which comprised six neurons in the hidden layer and four output nodes, so that electrons, pions, jets and muons could be detected.

Simulated data sets for the second-level trigger operation at LHC conditions were used for training the networks.

2.2 Electrical Load Monitoring System

Both transient and steady-state information were used to characterize data acquired from seven groups of equipment: refrigerating, resistive heating, universal motor, ventilating, consume electronics, incandescent lamp and fluorescent lamp.

Starting from the time each appliance had been switched on, current signals were sampled from the AC line by means of a digital storage oscilloscope for a 2 second acquisition window. Steady-state analysis provided current and phase angle information and measurements were made after a minimum of 2 seconds operation.

A principal discriminating analysis was performed on samples from the transient response. This analysis aimed at finding the most discriminating components on data input space, so that classification can be achieved with minimal number of hidden nodes in the network. Only two neurons in the hidden-layer were needed and the output layer comprised seven nodes. Each output node was assigned to a group of equipments and maximum probability was used to detect the winning class for a given input.

3 Implementations

The system implementation for both applications made use of a master node to perform communication with the outside world and to supervise the continuous distribution of data through slave processors. The slave processors act as feature extractors or global decision units for the physics application (figure 2), and perform preprocessing and neural classification for the load monitoring system. For the validation system, a local network is used to label and group data from the feature extraction phase and to transmit features to the global decision layer.

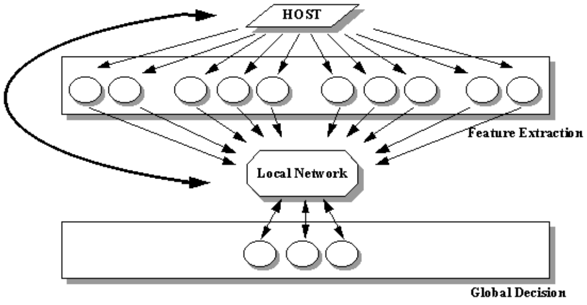

Fig. 2. Simplified scheme for the implementation of the validation system.

The data parallelism approach that is being used intends to minimize the communication overheads of the TN-310 system. In fact, the minimum time required for packet (32 bytes) transmission was measured to be $\sim 7\ \mu$s. Therefore, as neural networks for both applications are relatively compact (and fast), data communication time can be considered significant to the overall processing speed. Consequently, data parallelism will minimize dependencies among nodes.

As communication time represents the bottleneck of applications in this environment, it may be useless to develop the application using all resources of the machine. For instance, when data are distributed to, say slave $\# M$ in the processing chain, the first slave may have finished its processing and become free. Thus, further nodes added to the chain will not improve processing speed. This fact was explored to develop application using minimal machine resources.

The activation function (hyperbolic tangent) of the neural networks was implemented by means of a lookup table, in order to achieve shorter computation times. The sampling resolution for building such table and which provided the full reproduction of the simulated network operation with minimum memory requirements was achieved to be 0.01. Saturation of the activation function was considered to be reached at 7.

The prototype of the validation system for physics was implemented using 13 nodes and was capable to cope with a 2.6 kHz input frequency. It can be considered to emulate a vertical slice of the actual second-level triggering system, which will be running in practice close to this speed. Particle identification above 94% was achieved by the system.

For the load monitoring system, principal component analysis and classification was capable to run in less than 200 μs. Here, as data input vectors comprised 200 components, data transmission required barely as much time as data processing. Therefore fewer nodes were needed to implement the system. Over 100 different pieces of equipment studied, the system was able to classify correctly more than 84% of the sample.

References

1. TN-310 System Manual and Training Set. Telmat Multinode, France (1995).
2. Neural Networks. A Comprehensive Foundation. S. Haykin. Macmillan (1994).
3. Neural Second-Level Trigger System Based on Calorimetry. J.M. Seixas, L.P. Caloba, M.N. Souza, A.L. Braga and A.P. Rodrigues **Computer Physics Communications**, Volume. 95, pp 143-157 (1996).