

# Fast Neural Decision System Based on DSPs and Parallel Processing

J.M. Seixas, L.P. Caloba, A.R. Anjos, A.C.H. Dantas, R. Linhares  
COPPE/EE/UFRJ, C.P. 68504, Rio de Janeiro 21945-970, Brazil  
Tel: 55-21-2807393, Fax: 55-21-5900788, E-mail: seixas@lacc.ufrj.br

## Abstract

A prototype of an online event validation system is developed for application in a high-energy collider experiment. The system mainly uses neural networks for extracting rare events with physics significance from a huge background noise. It is based on processing the information collected from different detectors placed around the collision point. Combining a feature extraction phase for each detector with a global decision phase for final decision on discarding or not a given event, the system acts on events previously selected by a first-level analysis that reduces the event rate to 100 kHz. To cope with this input frequency, the proposed system is being emulated in a 16 node transputer based parallel machine that has a fast digital signal processor running as a coprocessor for each node.

## 1 Introduction

In the field of experimental high-energy physics, a new collider machine is being developed at CERN (Switzerland) for investigating the fundamental structure of the matter. This is to be achieved by colliding high-energy particles and analyzing the reaction products by means of a set of detectors with electronic readout. This collider, LHC (Large Hadron Collider), is planned to be operational by 2005 and will produce an enormous amount of data, as the period between collisions is expected to be as low as 25 nanoseconds.

At LHC, the interesting events will be extremely rare. Events with physics significance will be stored in a rate lower than 100 Hz, so that such events have to be extracted from the huge background generated by the colliding machine. As event selectivity for LHC is mandatory, a multi-level triggering system is being designed for iden-

tifying the relevant information whereas background events are discarded. The first level system focuses on processing speed and discards background events that may be considered relatively easy to be identified. At the output of such level, the event rate is reduced to 100 kHz. The second-level triggering (LVL2) system acts on events that passed the conditions of the previous level, so that it has more time (10 microseconds) for achieving a decision. Thus, a more sophisticated technique can be implemented in order to reduce further the event rate by a factor of 100. Finally, the surviving events are to be analyzed by a third level trigger, which implements an even more elaborated algorithm.

One possible approach to address the LVL2 system design is to split the data processing into two phases. In the first, variables with physics significance (or features) are extracted from preprocessed data related to restricted areas of each sub-detector, the Regions of Interest (ROIs), which are identified by the first-level trigger. Next, the global decision phase operates on features related to all ROIs and subdetectors for a given event. The global decision phase correlates the information generated by the different detectors in order to achieve a final decision on keeping or discarding the particular event under analysis at this trigger level.

Both feature extraction and global decision phases have been suggested to be performed by neural processing [1, 2]. In fact, implementations on a fast (40 MHz) digital signal processor (ADSP-21060) pointed out that the neural level-two trigger is feasible to be realized on this flexible technology. Due to the sequential-parallel architecture of the LVL2 system and the natural parallelism of neural networks, the capability to support parallel processing of this DSP was considered quite interesting in such studies.

In this paper the implementation of the neural LVL2 system is moved further towards parallel data processing by using a transputer based parallel machine, TN-310 system from Telmat Multinode (France). This machine is a multiple instruction multiple data parallel computer with a distributed memory architecture. 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 uses a 25 MHz digital signal processor (DSP) as a coprocessor to the T9000 transputer for signal processing applications. This DSP (ADSP-21020) is, in fact, the core processor of the ADSP-21060, which allows standing alone C codes previously developed for the DSP standing alone implementation to be ported to this new environment.

Data used in this paper are obtained from Monte Carlo simulations for the four detectors involved in the LVL2 system. Only events that passed the conditions of a simulated first-level trigger algorithm were used in the analysis. The global decision phase and feature extraction for calorimeters (detectors for energy measurement) were addressed by neural processing and involved the generation of a large sample of QCD jets and single electrons [3], and the simulation of important reactions [4]. For the other three detectors, which include tracking systems and muon detection, a simulation of current classical algorithms for feature extraction was used. For this, C codes simulated module interfaces and performed main functionalities, not necessarily handling all constraints.

In the next two sections, the neural processing for feature extraction and global decision phases are described. Next, the TN-310 system is detailed and the parallelism of actions is explained. The last section addresses conclusions.

## 2 Feature Extraction

For calorimetry, the ROI description was based on a 11 x 11 matrix of deposited energy in the 121 cells of a fine-grained calorimeter. Feature extraction was addressed by searching preprocessing methods on ROI data capable to reduce the high dimensionability of data input space, and perform electron/jet separation efficiently. In this way, processing speed and discrimination efficiency could be optimally combined for LVL2 system design.

The two main lines of feature extraction in-

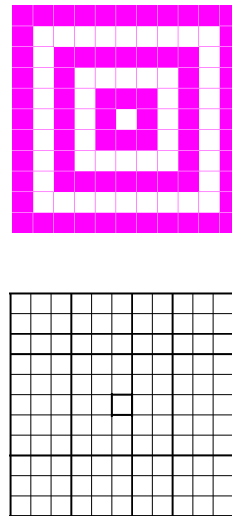


Figure 1: Ring (top) and grouped sum (bottom) structures. Thicker lines define the border of each sector/ring.

involved topological information and principal component analysis. In terms of topological mapping, grouping cells in a way that the outermost cells can have their discriminant information boosted by a weighting profile revealed to be effective. In this way, ring sums and grouped sum structures (Figure 1) were obtained. Rings [1] are built concentrically around the cell of maximum energy deposition in a ROI, so that this particular cell becomes the first ring. Outer rings are obtained by adding up the energy of all cells that belong to a given ring. Cells can also be grouped differently, forming ROI sectors. In this case, cells belonging to a sector have their energies added up to form group sums [6]. For both methods, the grouped information was fed into the input nodes of a neural classifier. Performance of such classifiers was similar and for a 97% electron efficiency, less than 7.3% of jets were misclassified as electrons. Networks were trained by backpropagation and used hyperbolic tangent as the activation function. The number of nodes in the hidden layer of this fully-connected network were 3 (rings) and 5 (grouped sums).

Another possibility is to perform principal discriminating analysis [6]. In this method, the principal components that are capable to discriminate electrons from jets are extracted by training a scalable neural network (Figure 2). The network is built by increasing the number of discriminating components up to the point where the discrimi-

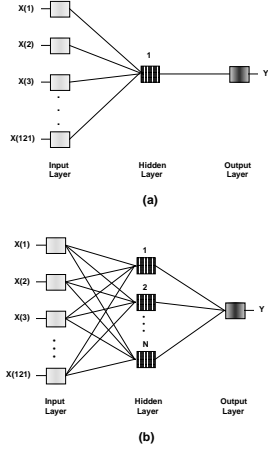


Figure 2: Extracting the first (a) and the N-th (b) discriminating components.

nation error for jets can not be significantly decreased for a fixed high electron efficiency value, say 97%. For this single hidden layer network, each time a new component is added, in the form of adding a new neuron in the hidden layer, the hidden layer of the network is trained only for the weights which connect the input nodes to this new neuron. The other weights in this layer are kept frozen, as they represent discriminating components previously extracted. The weight vector that connects the hidden layer to the output neuron is retrained each time a new component is added, so that the network can find the optimal way to combine the contribution of the components that become available. Only three components were required to achieve a similar performance as obtained from topological methods.

Table 1 summarizes the processing time obtained for each feature extraction method when an Assembly implementation on the ADSP-21020 was tested. For implementation issues grouped sum technique was selected.

Feature	Processing Time ( $\mu s$ )
Rings	4.15
Grouping	6.0
PCA	14

Table 1: Processing times for the various features.

### 3 Global Decision

In terms of the global decision phase, data are analysed for each ROI of a given event. In average 5 ROIs are produced per event. The feature space that is being used comprises 12 features from the four different detectors that contribute for the decision taken of the LVL2 system. To correlate such diverse information, a fully-connected network with 6 nodes in the hidden layer is used. The four output nodes of the network provide particle probabilities for electrons, muons, pions and jets. These probabilities are computed for each ROI and are used to identify events representative of the physics channels of interest. Table 2 shows the performance of the neural classifier for particle identification.

Particle	Identified as		
	electron	jet	muon
electron	97.7%	2.3%	0
jet	5.6%	94.4%	0
muon	0%	0.1%	99.9%

Table 2: Performance of the particle discriminator.

### 4 TN-310 System

The basic architecture of the TN-310 system is shown in Figure 3. The main components are the basic processors (transputers and their DSPs) and the C104 based switching network that allows each node to access data held anywhere in the system. All nodes comply to the HTRAM (High performance TRAnsputer Modules) standard. In particular, the system in consideration houses two interconnected boards of eight DSP HTRAMs each. In terms of memory, each node comprises 256 kbytes SRAM used as shared memory, to transfer data to and from the DSP, and 8 Mbytes of T9000 private DRAM. The ADSP 21020 can be programmed from the T9000 through C runtime library calls.

The T9000 (see Figure 4) is a 32-bit microprocessor that exhibits multiprocessing capabilities. Its architecture supports the creation and scheduling of any number of concurrent processes. Communication between processes on different processors takes place over virtual channels. Multiple virtual channels are multiplexed onto each physical link by the virtual channel processor.

Each message to be transmitted is split into a sequence of packets, and packets from different mes-

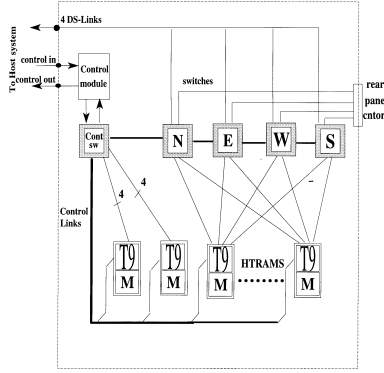


Figure 3: The basic architecture of TN 310 system.

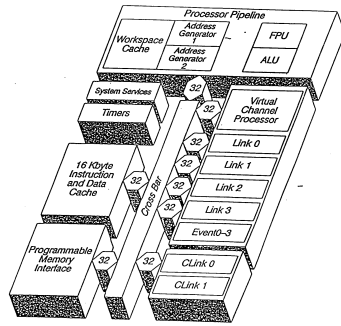


Figure 4: T9000 transputer.

sages may be interleaved over each physical link. As a consequence of interleaving packets from different messages, processes communicate simultaneously via each physical link.

Each node of the system has access to the communication network through four high speed (100 Mbits/s) serial links (DS-links). Four C104 chips are used per board, each one connected to a corresponding link of the T9000s. The C104 chips enable messages to be routed from any of its links to any other link.

The ADSP-21020 (Figure 5) is a 32/40-bit floating point processor exhibiting a 40 ns instruction cycle time. The basic architecture of this device includes three independent computational units: ALU, multiplier with fixed-point accumulator and shifter. The units are connected in parallel and can use up to 16 internal registers at any moment. Every instruction is executed in a single cycle, so

that multiplication with accumulation and search for the next operands can be performed in one cycle.

The TN 310 system can be accessed through a host machine, in our case a PC running MS-DOS and Windows. For developing applications, a C toolset software layer allows to describe processes and the communication among them in a relatively flexible way, so that ultimate processing speed can be achieved.

## 5 Validation System Implementation

The prototype being implemented in the TN-310 environment processes ROI based data. The basic scheme can be seen in Figure 6. A host node is responsible to interface with the outside world and reads from disk event data and neural network's parameters, the latter being stored in the memory of neural processors during the initialisation procedure of the system. This node continuously feeds the feature extractor nodes with ROI data according to the subdetector assigned to the respective node.

Feature extraction is performed in parallel for each subdetector using the concept of data parallelism. Each feature extractor node assigned to a given subdetector runs the same C code on input data distributed by the host node. Features are gathered by a local network, which is, in fact, split into two nodes. The first network receives the features extracted in the first layer of data processing and labels features according to event, ROI and subdetector they belong. It also passes labeled data to the second local network, which feeds the global decision processors running in parallel in the second layer of data processing. This second local network is also responsible to collect global decision results in terms of particle probabilities and send them to the outside world through the host node.

Tests on this implementation are currently being performed in terms of processing speed. The complete system is expected to include data pre-processing, where raw data from the ROIs are collected from a number of Readout Buffers (ROBs) and transformed into valuable detector data for feature extraction. Neural feature extraction for the tracking systems is also under development. On top of this, more updated simulation data are being considered, which include longitudinal segmentation of the calorimetry system and more de-



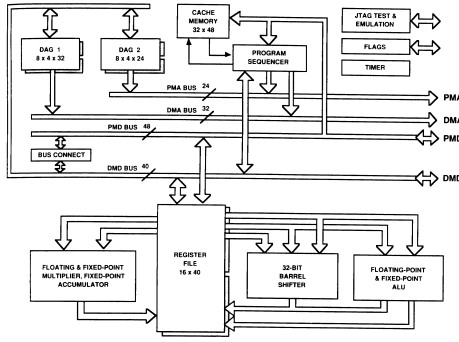


Figure 5: ADSP-21020 block diagram.

tailed information on the tracking systems.

## 6 Conclusions

A prototype of an online event validation system for a high event rate experiment in the field of high energy physics has been described. The system is being implemented in a transputer based parallel machine that uses a fast DSP to optimize signal processing applications.

The validation system performs local data processing on data readout from different subdetectors for translating raw data into features with valuable physics information. Then, global processing follows, so that detector features can be correlated to achieve a more precise decision rule. Events that meet the requirements of the validation system proceed to a further level of analysis. The rest of events get discarded in order to reduce the background of the experiment.

The implementation of such system is being carried on a 16 node machine and parallelism is explored in both feature extraction and global decision phases of data processing. Final tests on such prototype are currently being performed using simulated detector data.

## Acknowledgements

We are thankful to the support provided by CNPq, FUIB (Brazil) and the European Commission. We also thank the RD11 and ATLAS/Trigger-DAQ collaborations at CERN for providing the data sets and a framework that has been discussing different technologies which address efficient LVL2 design.

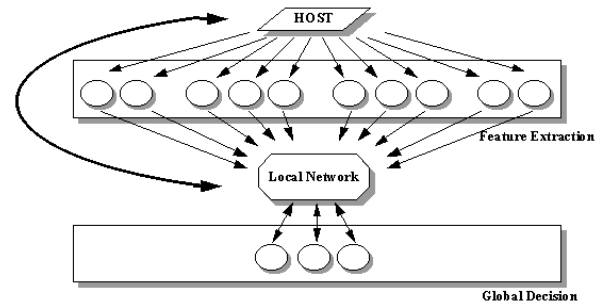


Figure 6: Simplified scheme for the implementation of the validation system.

## References

- [1] J. M. Seixas et al. Neural Second-Level Trigger System Based on Calorimetry. Computer Physics Communications, 95 (1996) 143;
- [2] J. M. Seixas et al. Implementing a Neural Second-Level Trigger System on Digital Signal Processor Technology: The Global Decision Problem. First Int. Workshop on Electronics for LHC Experiments, Lisbon, Portugal (1995) 324.
- [3] G. Klyuchnikov et al. "A Second-level Trigger Based on Calorimetry Only", CERN-ATLAS/DAQ-007, 1992.
- [4] Test Data for the Global Second-Level Trigger. R.K. Bock et al. CERN/EAST 93-01 (1993).
- [5] Neural Feature Extraction for Calorimeters Based on Optimum Weighting Procedures. J.M. Seixas et al. AIHENP, Lausanne, Switzerland (1996).
- [6] Neural Discriminating Analysis for a Second-Level Trigger System. L.P. Calôba et al. CHEP, Rio de Janeiro, Brazil (1995).