Fast Optoelectronic Neural Network for Vision Applications

Marta Ruiz-Llata and Horacio Lamela

Universidad Carlos III de Madrid,
28911 Leganés, Madrid, Spain
http://gdo.uc3m.es/gotl/

Abstract. This paper reports the recent steps to the attainment of a compact high-speed optoelectronic neuroprocessor based on an optical broadcast architecture that is used as the processing core of a vision system. The optical broadcast architecture is composed of a set of electronic processing elements that work in parallel and whose input is introduced by means of an optical sequential broadcast interconnection. Because of the special characteristics of the architecture, that exploits electronics for computing and optics for communicating, it is readily scalable in number of neurons and speed, thus improving the performance of the vision system. This paper focuses on the improvement of the optoelectronic system and electronic neuron design to increase operation speed with respect to previous designs.

1 Introduction

Neural networks are computational approach that is nowadays demonstrating applicability to many computational cost problems [1].On the contrary of traditional processors, an artificial neural network can be seen as a set of very simple processing elements with a high degree of interconnectivity between them. The computation strength of a simple processing element is very small, but the massive cooperation between lots of them results in a powerful machine for tasks as pattern recognition and classification [2].

For vision there have been a wide range of neural systems with application ranging from early to high level vision operations. Examples are the hardware implementation of the bio inspired artificial retina [3] or – more recent – pulse or spiking neural networks (PCNN) [4] for low or medium level tasks. For high level operation such image classification, neural networks are emulated via software [1]. Software based implementations of neural networks are often not fast enough to meet the real time requirements of some vision application but the main disadvantage is that their development require high computer literacy and math background; which contrast with one of the most exciting properties of neural networks that is its ability of learning.

There have been also hardware implementations of neural network processors. The idea was the development of parallel processing architectures to optimize neural network algorithms; examples were the CNAPS processor [5] and the analogue Intel ETANN [6] among others [7]. Recent trends are the implementation of neural proc-

J. Cabestany, A. Prieto, and D.F. Sandoval (Eds.): IWANN 2005, LNCS 3512, pp. 502 – 509, 2005.

essing systems instead of parallel processing architectures. As example we can find the VindAx processor from Axeon Ltd. [8] implementing a self-organizing map and the IBM zero instruction set computer (ZISC) implementing a radial basis function (RBF) neural network [9]. The ZISC has been tested and compared with other neural network hardware implementation for a quite complicated vision task: face tracking and identification [10]. Although higher operation speed, the performance of this system is worse than the other proposed implementations (RBF on FPGA and DSP); the performance difference was linked to the limit in the size of input pattern to 64 elements.

Our work is focused on the implementation of a neural processing system for smart vision applications. The performance demanding for such a device are fast operation speed and large size. Large size means that the hardware architecture must not impose a (small) limit in the size of input patterns nor to the number of neural nodes and interconnections between them.

It is also our understanding that optics may help in the implementation of neural processing architectures, but with a different approach that is used in optical computing architectures – such optical correlators [11] and optical neural networks [12]. In those systems optics is used for computing, basically for spatial filtering, in optical correlators, and for multiplication and addition, in optical vector-matrix multipliers. Some problems encountered in the realization of large-scale optoelectronic systems are: optical alignment, interconnection weight reconfiguration and assignment by spatial light modulators (SLM), and the construction of the associated opto-mechanical subsystems with reduced dimensions. Our point of view is that optics must be used only for interconnects, while electronics must be used for computing [13].

In this paper we describe in section 2 the basic principles of the optoelectronic architecture for neural networks we have proposed [14] and that has been used as the processing core of a vision system [15][16]. In section 3 we describe the implementation of a new fast prototype based on a new neuron design. In section 4 we describe the vision system that will be benefited by the new fast optoelectronic neural network design and its preliminary results. We finish with conclusions in section 5.

2 Broadcast Neural Network Processor

As represented in equation (1), the output of a single neuron y_j in a neural network is a non-linear function of a weighted sum of inputs:

$$y_j = f\left(\sum_{i=1}^{N} w_{ji} x_i\right) \tag{1}$$

where N is the size of the input pattern, x_i is one element of the input and w_{ji} is the interconnection weight between input i and neuron j. Hardware implementations of neural processors, especially optoelectronic implementations [11][12], usually try to construct an architecture where basic operations are executed simultaneously for all neurons. The hardware architecture we propose simplifies the optical interconnection scheme and avoids the use of SLM but takes advantage of the high spatial and temporal communication bandwidth of optical beams. As represented in figure 1, it is composed of a set of K cells comprising of M optoelectronic neurons that share the same