International Journal of Embedded and Real-Time Communication Systems
Volume 11 • Issue 1 • January-March 2020

An Event-Based Neural Network Architecture with Content Addressable Memory

Sivaganesan S, Karpagam Academy of Higher Education, Tamil Nadu, India
Maria Antony S, KIT-Kalaignarkarunanidhi Institute of Technology, Tamil Nadu, India
Udayakumar E, KIT-Kalaignarkarunanidhi Institute of Technology, Tamil Nadu, India

ABSTRACT

A hybrid analog/digital very large-scale integration (VLSI) implementation of a spiking neural network with programmable synaptic weights was designed. The synaptic weight values are stored in an asynchronous module, which is interfaced to a fast current-mode event-driven DAC for producing synaptic currents with the appropriate amplitude values. It acts as a transceiver, receiving asynchronous events for input, performing neural computations with hybrid analog/digital circuits on the input spikes, and eventually producing digital asynchronous events in output. Input, output, and synaptic weight values are transmitted to/from the chip using a common communication protocol based on the address event representation (AER). Using this representation, it is possible to interface the device to a workstation or a microcontroller and explore the effect of different types of spike-timing dependent plasticity (STDP) learning algorithms for updating the synaptic weights values in the CAM module.

KEYWORDS

Artificial Neural Networks, Complementary Metal-Oxide Semiconductor, Content Addressable Memory Cell, VLSI

1. INTRODUCTION

1.1 Biological Neuron

A neuron is a specialized type of cell found in the bodies of all eumetozoans. Only sponges and a few other simpler animals lack neurons. The features that define a neuron are electrical excitability and the presence of synapses, which are complex membrane junctions that transmit signals to other cells. The body's neurons, plus the glial cells that give them structural and metabolic support, together constitute the nervous system. In vertebrates, the majority of neurons belong to the central nervous system, but some reside in peripheral ganglia, and many sensory neurons are situated in sensory organs such as the retina and cochlea (Myers et al., 2007).

The soma is usually compact; the axon and dendrites are filaments that extrude from it. Dendrites typically branch profusely, getting thinner with each branching, and extending their farthest branches a few hundred micrometers from the soma. The axon leaves the soma at a swelling called the axon hillock, and can extend for great distances, giving rise to hundreds of branches. Unlike dendrites, an axon usually maintains the same diameter as it extends. The soma may give rise to numerous dendrites, but never to more than one axon (Liu et al., 2010). Synaptic signals from other neurons are received

DOI: 10.4018/IJERTCS.2020010102

by the soma and dendrites; signals to other neurons are transmitted by the axon. A typical synapse, then, is a contact between the axon of one neuron and a dendrite or soma of another. Synaptic signals may be excitatory or inhibitory. If the net excitation received by a neuron over a short period of time is large enough, the neuron generates a brief pulse called an action potential, which originates at the soma and propagates rapidly along the axon, activating synapses onto other neurons as it goes.

1.2 Artificial Neural Networks

Artificial neural networks are parallel computational models, comprising densely interconnected adaptive processing units. These networks are composed of many but simple processors (relative, say, to a PC, which generally has a single, powerful processor) acting in parallel to model nonlinear static or dynamic systems, where a complex relationship exists between an input and its corresponding output (Hasler et al., 2010). Artificial neural networks are viable models for a wide variety of problems, including pattern classification, speech synthesis and recognition, adaptive interfaces between humans and complex physical systems, function approximation, image compression, forecasting and prediction, and nonlinear system modeling.

These networks are "neural" in the sense that they may have been inspired by the brain and neuroscience, but not necessarily because they are faithful models of biological, neural or cognitive phenomena. In fact, many artificial neural networks are more closely related to traditional mathematical and/or statistical models, such as nonparametric pattern classifiers, clustering algorithms, nonlinear filters and statistical regression models, than they are to neurobiological models (Sontag et al., 2000; Indiveri et al., 2007).

1.3 VLSI- Neural Networks

Artificial neural networks (ANN) are parallel algorithms. Their inherent parallelism makes them particularly suited to parallel VLSI implementations, i.e. the development of dedicated circuits or architectures that are able to perform many operations in parallel. As ANN involve similar operations to be performed in parallel, it is particularly easy to develop dedicated parallel architectures: most of them are based on the repetition of identical devices, each of them performing one of the operations. Parallel architectures for ANN are justified by the fact that most ANN learning algorithms involve a high computational load. Since the early beginning of the neural networks research, one has naturally been tempted to develop specialized machines, in order to lighten the computational load.

Another nice feature of ANN is that most algorithms mainly involve simple operations. Indeed, those algorithms with biological inspiration naturally use the same type of operations as found in biological cells. Of course, there are exceptions: it is not because pseudo-inverse of matrices may be found in "neural" algorithms that we could expect our brain to perform large matrix inversions. Inherent parallelism, simplicity of operations and high computational load are the reasons why VLSI implementations of ANN became so popular. Since the mid-80s, many attempts have been made towards the realization of VLSI circuits, or larger machines, suited to the simulation of ANN (Yu et al., 2010). The development of VLSI circuits implementing large numbers of operations in parallel is interesting on the point of view of the circuit designer: as large number of cells have to be integrated, one must carefully optimize them. Furthermore, connectivity, transmission of information, and maximization of the use of resources are interesting problems generated by such circuits.

2. EXISTING SYSTEM

2.1 Chip Architecture

The chip was fabricated using AMS 0.35 Complementary Metal-Oxide Semiconductor (CMOS) technology. It comprises five main blocks: the asynchronous controller, the SRAM block, the neural core, the bias generator and the AER Input/output (IO) interfaces (Scholze et al., 2011). The