A 2D Neuromorphic VLSI architecture for modeling selective attention

Giacomo Indiveri
Institute of Neuroinformatics
Swiss Federal Institute of Technology/University of Zurich

December 15, 1999

Abstract

Selective attention is a mechanisms used to sequentially select the spatial locations of salient regions in the sensor's field of view. This mechanism overcomes the problem of flooding limited processing capacity systems with sensory information. It is found in many biological sensory systems and can be a useful engineering tool for artificial visual systems. We present a hardware model of a selective attention mechanism implemented on a VLSI chip, using analog neuromorphic circuits. The chip makes use of a spike based representation for receiving input signals, transmitting output signals and for shifting the selection of the attended input stimulus over time. The chip can be interfaced to neuromorphic sensors and actuators, for implementing multi-chip selective attention systems. We describe the characteristics of the circuits used in the architecture, and present experimental data measured from the system.

1 Introduction

Selective attention is a mechanism used by a wide variety of biological systems to determine which of several inputs will be analyzed by the system's (limited) parallel-processing resources. Specifically, selective attention allows the system to pick out relevant subregions of the input and process them, sequentially shifting from one subregion to the other, while suppressing the irrelevant information that the system cannot analyze simultaneously.

Recent theories suggest that the selection mechanism, which facilitates the emergence of a "winner" from several potential targets, can be modulated by stimulus-driven and goal-driven factors [13]. Stimulus-driven selective attention appears to be a rapid, bottom-up, task-independent mechanism, whereas goal-driven selective attention appears to act in a slower, top-down, volition controlled manner. In this paper we present an analog VLSI 2D architecture which implements a real-time model of the the stimulus-driven form of selective attention, based on the concept of a *saliency map* [10]. The architecture proposed receives inputs from synaptic circuits and projects

0-7695-0619-4/00 $10.00 © 2000 IEEE

its outputs to local inhibitory neurons. The synaptic input circuits and the output inhibitory neurons are part of the I/O interfacing circuits of the chip, that use an *Address-Event Representation* (AER) to communicate with other chips and devices [11, 3]. In this representation input and output signals are transmitted as asynchronous streams of binary words which encode the address of the sending node and carry its analog information in their temporal structure, very much like natural spike trains do in biological systems.

Several VLSI systems for implementing *visual* selective attention mechanisms have already been presented [2, 5, 7, 12]. These systems contain photo-sensing elements and processing elements on the same focal plane, and typically apply the competitive selection process to *visual* stimuli sensed and processed by the focal plane processor itself. Unlike these systems, the device here proposed is able to receive input signals from *any* type of AER device. Therefore input signals need not arrive only from visual sensors, but could represent a wide variety of sensory stimuli obtained from different sources. In general, decoupling the sensing stage from the processing stage and using the address-event representation to receive and transmit signals has several advantages: a multi-chip AER attention system could use multiple sensors to construct a saliency map, the input of the selective competition processing stage; visual sensors used for generating the saliency map could be relatively high-resolution silicon retinas and would not have the small fill factors that single-chip 2D attention systems are troubled with; top-down modulating signals could be fused with the bottom-up generated saliency map to bias the selection process; multiple instances of the same selective attention chip could be used to construct hierarchical selective attention architectures.

2 The Selective Attention Chip

The chip contains of an array of 8×8 cells, layed out on a square grid, with local (nearest neighbor) excitatory and inhibitory connections. Each cell comprises an excitatory synapse, an inhibitory synapse, a hysteretic WTA cell, a local inhibitory output neuron, and two output (analog) position to voltage (P2V) circuits [4]. All of these circuits, except for the P2V circuits, have been used and fully characterized in a similar 1D selective attention chip [6].

Fig. 1(a) shows the circuit diagram of a cell's input excitatory synapse. The circuit is a current-mirror integrator interfaced to the input AER circuitry [1]. It receives off-chip pulse-frequency modulated digital pulses, and integrates them into an analog excitatory current I_{ex} (see Fig. 1(a)). Similarly, the inhibitory synapse of Fig. 1(b) integrates the on-chip spikes of the same cell's output neuron into an inhibitory current I_{ior}. The synaptic currents I_{ex} and I_{ior} are subtracted and sourced into the input node V_{in} of the WTA cell (see Fig. 2(a)).

Each hysteretic WTA cell is connected to its four nearest neighbors, both with lateral excitatory connections and lateral inhibitory connections (see Fig. 2(a)). The inhibitory connections are modulated by the bias voltage V_{inh}, and control the spatial extent over which competition takes place. If lateral inhibition is maximally turned on, all WTA cells of the architecture are connected together (global inhibition) and only one winner can be selected at a time. Lateral excitatory connections are modulated

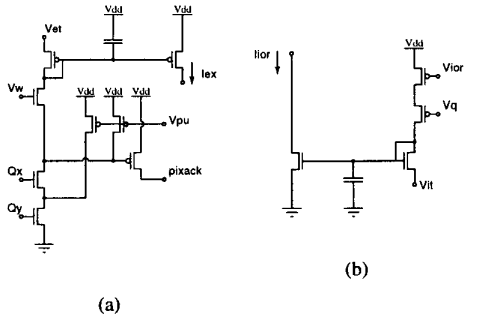


Figure 1: Synaptic circuits. (a) Input excitatory synapse; (b) Inhibitory synapse

by the bias voltage V_{ex} and control the amount of lateral facilitatory coupling between cells. If lateral coupling is enabled, the system tends to select new winners in the immediate neighborhood of the currently selected cell. When the cell in consideration is selected as a winner, its output transistors source DC currents into the two position-to-voltage row and column circuits. The winning WTA cell also sources a DC current I_{inj} into the local inhibitory neuron connected to it. The amplitude of the injection current I_{inj} is independent of the input, but depends on the bias voltage V_{wta} and on the control voltage V_{inj}. This current, integrated onto the neuron's capacitor C_m of Fig. 2(b), allows the neuron's membrane voltage V_{mem} to increase linearly with time. As soon as V_{mem} reaches the threshold voltage V_{thr}, the neuron generates an action potential: the comparator and the inverters of Fig. 2(b) drives V_{out} to the positive power supply rail. This activates row and column request signals (R_x and R_y), which produce an address event. The output AER circuit's acknowledge signals (A_x and A_y) reset the pulse by allowing the neuron's membrane capacitance to discharge at a rate controlled by V_{pw}.

The output neuron in each pixel, besides transmitting the pixel's address off chip, also implements the inhibition of return (IOR) mechanism (a key feature of many selective attention systems). While a cell is winning, its output neuron generates spikes at a constant rate (the frequency of which is proportional to I_{inj}). These spikes are integrated onto the cell's inhibitory synaptic circuit (see node V_{ior} in Fig. 2(b) and in Fig. 1(b)). As the integrated inhibitory post-synaptic current I_{ior} starts to cancel out the cell's input excitatory synaptic current I_{ex}, a different cell eventually receives the largest net input current and is selected by the WTA network. When the previous winning cell is de-selected (and its corresponding local output neuron stops firing), its inhibitory synapse recovers, decreasing the inhibitory current I_{ior} to zero. pending on the dynamics of the IOR mechanism, the WTA network will continuously switch the selection of the winner between the largest input and the next-largest, or between the largest and more inputs of successively decreasing strength, thus generating focus of attention *scanpaths* [14]. The dynamics of the IOR mechanism depend on the time

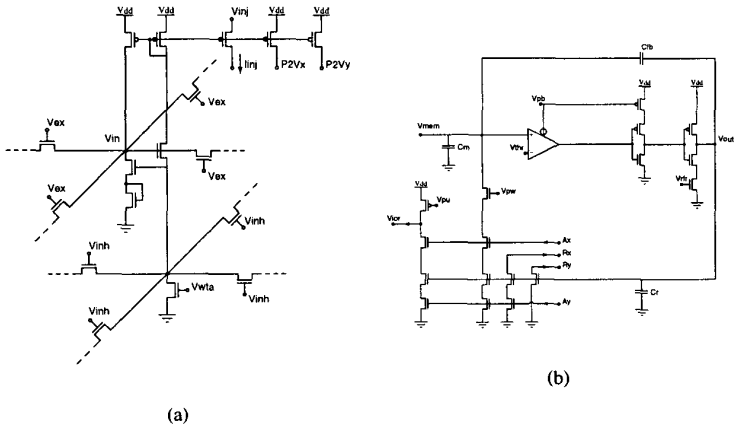
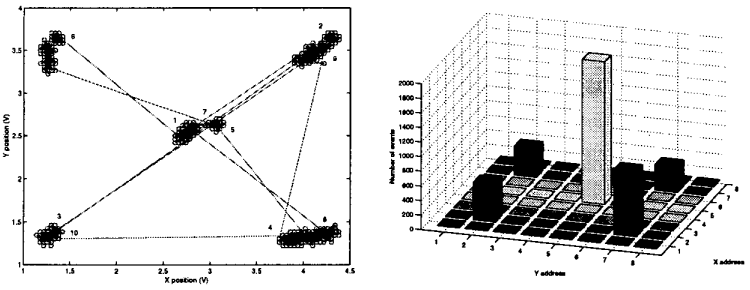


Figure 2: (a) Hysteretic WTA cell; (b) Local output integrate and fire neuron.

constants of the excitatory and inhibitory synapses (set by V_{et} and V_{it} of Fig. 1 respectively), on their synaptic strength (set by V_w and V_q respectively), on the frequency of the input stimuli and on the frequency of the output inhibitory neuron.

3 Experimental Results

In practical applications the chip would be interfaced to analog VLSI neuromorphic sensors that use AER I/O circuitry to provide an appropriate saliency map input [9]. However, to characterize the behavior of selective attention model with well controlled input signals, we interfaced the chip to a workstation. As the stimulation software was able to cycle through all the 64 pixels of the network at a rate of 500Hz and as the input synapses were set to have time constants of the order of milliseconds, they appeared to receive input spikes virtually in parallel. The test input stimulus we used excited the four inner corner pixels (2,2) (2,7) (7,2) and (7,7) with 30Hz spike trains and the central pixel (5,5) with spikes at a rate of 50Hz. Fig. 3(a) shows the analog output of the position-to-voltage circuits in response to 300ms of stimulation with the "static" input map described above. The system selects the central pixel initially and starts to cycle through all the excited cells as the IOR mechanism forces the WTA network to de-select chosen winners. The position-to-voltage circuits are actively driven when the WTA network is selecting a winner (*i.e.* when the output p-type transistors of Fig. 2(a) are sourcing current into the nodes $P2V_x$ and $P2V_y$). At the times in which no pixel is winning, there is no active device driving the position-to-voltage circuits, and their outputs drift towards zero. This is evident in Fig. 3(a), for example, at the pixel position (7,2) in the lower right corner of the figure. When the network selects it as its eighth target, the vertical position-to-voltage circuit outputs approximately 1.3V and



(a) (b)

Figure 3: (a) Output of the position-to-voltage circuits of the selective attention architecture measured over a period of 300ms, in response to a test stimulus exciting four corners of the input array at a rate of 30Hz and a central pixel at a rate of 50Hz; (b) Histogram of the chip's output address-events, captured over a period of 13.42s in response to the same input stimulus.

the horizontal one outputs approximately 4.4V. When the IOR mechanism forces the network to de-select the winner the outputs of the position-to-voltage circuits drift toward zero. As soon as the network selects the pixel (7,7) as the new (ninth) winner, the position to voltage circuits are actively driven again, and their output quickly changes to approximately 4.4V and 3.6V (for the horizontal and vertical circuits respectively).

The local inhibitory neurons in the architecture are fundamental circuits both for the output AER interface and for implementing the IOR mechanism. To quantitatively characterize the properties of the chip's output AER interface, we measured the address-events generated by the cells of the 2D architecture, in response to the same test stimulus used for the experiment of Fig. 3(a). In Fig. 3(b) we plotted a histogram of the address events captured by a logic analyzer for a period of approximately 13s. As reflected by the input stimulus, on average the architecture selected the central pixel most frequently, visiting the four corner pixels with a lower, but similar frequency.

4 Conclusions

We presented a VLSI device that implements a 2D model of selective attention for sequentially picking out the spatial locations of the most salient inputs. The device accepts input signals in the form of address-events and transmits its output using the same representation. The chip contains also analog position-to-voltage output circuits that can be used for quickly assessing the position of the winner in the 2D array or for driving actuators, such as DC motors or pan-tilt units. We showed experimental data that validates the functionality of the system, using control inputs generated on a workstation. The possibility to transmit the result of the selective attention competi-

tion to further processing stages using the same representation of input signals allows researchers to design modular multi-chip selective attention systems. An example of an application that uses the device proposed in this article, interfaced to a vision chip, to select salient stimuli and to orient the sensor toward them can be found in [8].

Acknowledgements

Fabrication of the integrated circuits was provided by MOSIS. This work was supported by the Swiss National Science Foundation SPP Grant and the U.S. Office of Naval Research.

References

[1] K.A. Boahen. Communicating neuronal ensembles between neuromorphic chips. In T. S. Lande, editor, *Neuromorphic Systems Engineering*, pages 229–259. Kluwer Academic, Norwell, MA, 1998.

[2] V. Brajovic and T. Kanade. Computational sensor for visual tracking with attention. *IEEE Journal of Solid State Circuits*, 33(8):1199–1207, August 1998.

[3] S. R. Deiss, R. J. Douglas, and A. M. Whatley. A pulse-coded communications infrastructure for neuromorphic systems. In W. Maass and C. M. Bishop, editors, *Pulsed Neural Networks*, chapter 6, pages 157–178. MIT Press, 1998.

[4] S. P. DeWeerth. Analog VLSI circuits for stimulus localization and centroid computation. *Int. J. of Comp. Vision*, 8(3):191–202, 1992.

[5] T.K. Horiuchi and C. Koch. Analog VLSI-based modeling of the primate oculomotor system. *Neural Computation*, 11:243–265, 1999.

[6] G. Indiveri. Modeling selective attention using a neuromorphic analog VLSI device. *Neural Computation*, 1999. Submitted.

[7] G. Indiveri. Neuromorphic analog VLSI sensor for visual tracking: Circuits and application examples. *IEEE Trans. on Circuits and Systems II*, 46(11):1337–1347, November 1999.

[8] G. Indiveri, R. Mürer, and J. Kramer. Active vision using an analog VLSI model of selective attention. *IEEE Trans. on Circuits and Systems II*, 1999. Submitted.

[9] G. Indiveri, A.M. Whatley, and J. Kramer. A reconfigurable neuromorphic VLSI multichip system applied to visual motion computation. In *Proceedings of the Seventh International Conference on Microelectronics for Neural, Fuzzy and Bio-inspired Systems; Microneuro'99*, pages 37–44, Los Alamitos, CA, April 1999. IEEE Computer Society.

[10] C. Koch and S Ullman. Shifts in selective visual-attention – towards the underlying neural circuitry. *Human Neurobiology*, 4(4):219–227, 1985.

[11] J. Lazzaro, J. Wawrzynek, M. Mahowald, M. Sivilotti, and D. Gillespie. Silicon auditory processors as computer peripherals. *IEEE Trans. on Neural Networks*, 4:523–528, 1993.

[12] T. G. Morris, T. K. Horiuchi, and S. P. DeWeerth. Object-based selection within an analog VLSI visual attention system. *IEEE Trans. on Circuits and Systems II*, 45(12):1564–1572, 1998.

[13] E. Niebur and C. Koch. Computational architectures for attention. In R. Parasuraman, editor, *The Attentive Brain*, pages 163–186. MIT Press, 1998.

[14] A. L. Yarbus. *Eye movements and vision*. Plenum Press, 1967.