

# A software-hardware selective attention system

L. Carota, G. Indiveri, and V. Dante

## Abstract

Selective attention is a mechanisms used to serially select and process salient subregions of the input space, while suppressing inputs arriving from non-salient regions. This mechanism is found in many biological systems and can be a useful engineering tool for developing artificial systems that need to process in real time sensory data. In this paper we present a hardware model of a selective attention system, implemented using a neuromorphic VLSI chip interfaced to a workstation, via a custom PCI board. The chip makes use of an *address event* (spike based) representation for receiving input signals, selecting salient inputs and sequentially shifting from one salient input to the other, and for transmitting output signals in the form of spikes. The PCI board processes the spikes (address events) received from the chip and acts as an interface between the VLSI chip and an algorithm that generates saliency maps. We describe the characteristics of the system and present experimental data showing the system's response to saliency maps generated from natural scenes.

## 1 Introduction

Selective attention is a mechanism used by a wide variety of biological systems to optimize their limited parallel-processing resources by identifying relevant subregions of the sensory input space and processing them in a serial fashion, shifting sequentially from one subregion to the other. This mechanism acts as a dynamic filter that allows the system to determine what information is relevant for the task at hand, and to process it, while suppressing the irrelevant information that the system is not able to analyze simultaneously. It can be a very effective engineering tool for designing artificial systems that need to process in real-time sensory information and that have limited computational resources. In biology selective attention mechanisms are modulated by stimulus-driven and goal-driven factors to facilitate the emergence of a “winner” from several potential targets [11]. The selective attention system we present in this paper implements a real-time model of the the stimulus-driven form of selective attention, based on the *saliency map* concept, originally put forth by Koch and Ullman [9]. The system processes images on a workstation, generates corresponding saliency maps, and produces focus of attention (FOA) scanpaths. We generate saliency maps using the visual attention software model developed by Itti et al.[8]. The program generates bi-dimensional (gray-level) saliency map images, in which the brightness of the pixel corresponds to its saliency. We then translate the saliency map into a series of spike trains (events), the brightest pixel generating the spike train with highest firing rate. Each pixel in the saliency map image is assigned an address. We then send the *address-events* to a selective attention chip [7] via a custom digital PCI-AER board [2]. The communication protocol used to transmit address-event from the PCI board to the selective attention chip (and back) is based on the *Address-Event Representation* (AER) [10, 3]. The selective attention chip contains a competitive network of (silicon) integrate and fire neurons that selects the input with highest frequency spikes, and implements inhibition of return dynamics (a key feature of many selective attention systems) [5, 12].

In the next Section we briefly describe the saliency map generation process, the characteristics of the PCI-AER board, and the behavior of the selective attention chip. In Section 3 we present

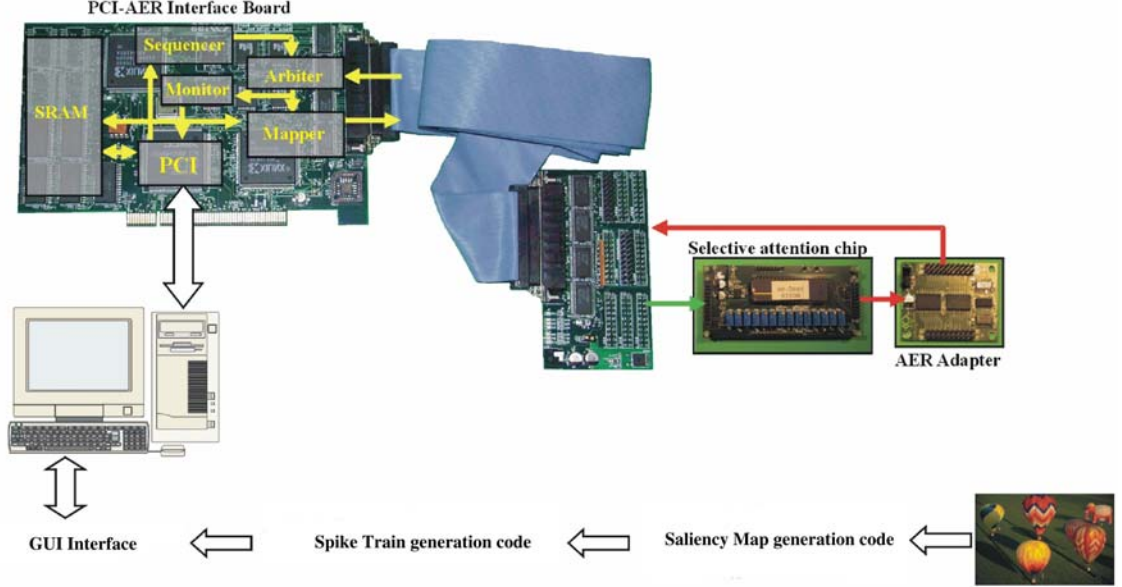


Figure 1: The experimental setup. We generate saliency maps from images, convert them to sequences of spike trains, and send them to the selective attention chip via the PCI-AER board. The selective attention chip is interfaced to the PCI board via a smaller adapter board. The PCI-AER board also reads spikes generated by the selective attention chip and sends them back to the PC for data logging.

experimental results measured from the system and in Section 4 we draw the concluding remarks.

## 2 The selective attention system components

The selective attention system we present has three main components (see Fig. 1): the software algorithm for generating saliency maps from arbitrary images, the PCI-AER board, for interfacing the algorithm to neuromorphic AER chips, and the selective attention chip, for implementing the WTA competitive stage, and the inhibition of return dynamics.

### 2.1 The Saliency Map generation code

The Saliency Map (SM) code simulates a two-dimensional layer of leaky integrate-and-fire neurons that receive the input from different "conspicuity" maps. These maps code for image brightness, color, orientation, *etc.* After being extracted from color images of  $N_i \times M_i$  pixels, they are normalized, re-sampled, and combined to generate an  $N_s \times M_s$  gray-level image that represents the final saliency map (with  $N_s < N_i$  and  $M_s < M_i$ ). In our case we extract three conspicuity maps (for intensity, color and orientation) from  $800 \times 800$  pixel images and generate  $8 \times 8$  saliency maps. Next to generating saliency map images, we generate, for each image pixel, a spike train using a Poisson distribution with mean rate proportional to the pixel's gray level. We then send sequences of spike trains to the selective attention chip, via the PCI-AER board. In the experiments presented we

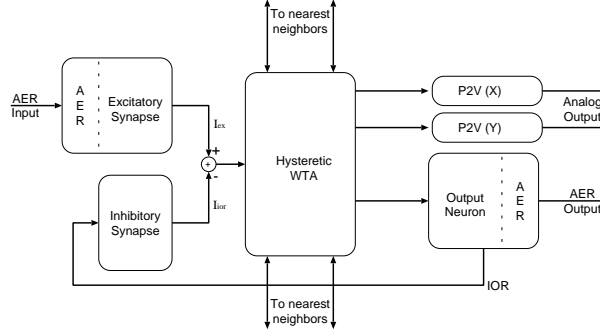


Figure 2: Block diagram of a basic cell of the  $8 \times 8$  selective attention architecture.

used a one-to-one mapping between the pixels in the saliency map and the neurons in the selective attention chip.

## 2.2 The PCI-AER board

The PCI-AER board is designed to provide a programmable interface between the AER bus and a standard PCI bus and allow a number of AER-compliant neuromorphic devices to communicate with each other. The present version of the PCI-AER board [2] implements the following functions:

- The MAPPER, which implements the programmable connectivity pattern between up to four sender chips, and up to four receiver chips through the storage of the connection matrix on a look-up table;
- The MONITOR, which allows to tap the transactions on the AER bus, to attach time information to them and to forward the joint information to a PC via PCI;
- The SEQUENCER, that can generate events on the AER bus, emulating a virtual neural chip and/or a flux of external spikes to the neural chips.

The board is managed on the PCI side via low-level Linux drivers and a high-level GUI interface.

In this project we exploited the MAPPER function to arrange the transmission of the Poisson spike trains to the right neuron-addresses of the selective attention chip, the MONITOR to collect and store on file the time-stamped spike trains coming from the chip and the SEQUENCER to send the software generated spikes to the chip.

## 2.3 The selective attention chip

The selective attention chip contains an array of  $8 \times 8$  cells. Each cell comprises an excitatory synapse, an inhibitory synapse, a hysteretic winner-take-all (WTA) cell [7], a local inhibitory output neuron [6], and two position-to-voltage (P2V) circuits [4] (see Fig. 2).

The P2V circuits produce two analog output voltages encoding the  $x$  coordinate and the  $y$  coordinate of the winning cell. The excitatory synapse is a current-mirror integrator [1] interfaced to the input AER circuitry. It receives off-chip address events, and integrates them into an excitatory current  $I_{ex}$ . The inhibitory synapse is a similar circuit that 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 of the WTA cell (see Fig. 2).

In the selective attention chip, each hysteretic WTA cell is connected to its four nearest neighbors. If lateral excitation is enabled, the system tends to select new winners in the immediate neighborhood

of the currently selected cell. The winning cell supplies a current to the position-to-voltage row and column circuits. It also sources a DC current into a neuron connected to it. Each action potential generated by this neuron produces an address event. The amplitude of the injection current (and hence the frequency of the address events) is independent of the WTA’s cell input. Next to transmitting the pixel’s address off chip, the output neuron is instrumental for implementing the IOR mechanism: The spikes generated by the winning cell’s output neuron are integrated by the cell’s inhibitory synapse. As the integrated inhibitory post-synaptic current  $I_{ior}$  increases, the cell’s net input current  $I_{ex} - I_{ior}$  decreases. As soon as this net input current decreases below the value of a net input current exciting a different cell, the WTA network switches state and selects the new cell as the winner. When the old winning cell is de-selected, its corresponding local output neuron stops firing and its inhibitory synapse recovers, decreasing the inhibitory current  $I_{ior}$  back to zero. Depending on the time constants and strength of the excitatory and inhibitory synapses, on the input stimuli and on the frequency of the output neuron, the WTA network will 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* [13].

### 3 Results

We measured the spikes (address-events) generated by the selective attention chip, in response to different saliency maps, created from images of natural scenes. With these experiments we point out the non-linear filtering properties of the selective attention mechanism (by analyzing the *mean rates* of the spiking activity) and its dynamics (by analyzing the *raster plots* of the spikes of the winning neurons).

In Fig. 3 we show the mean firing rates of the output neurons of the selective attention chip (bottom row), measured via the PCI board, in response to software saliency maps (middle row) generated from two images of “natural” scenes (top row). Note how neuron (8,8) of the selective attention chip has non-null activity, on the bottom left image, even if it is not stimulated with input spikes. This is most probably due to border effects on the chip layout and is being taken into account for the next generation of selective attention chips.

In Fig. 4 we show the input image subdivided into the  $8 \times 8$  regions that form the saliency map, together with the mean firing rates of the four winning neurons of the selective attention, next to the raster plot of the spikes generated by the same neurons. The raster plot shows how the selective attention chip changes the position of the focus of attention (FOA) over time, attending the four corresponding regions of the input image with different delays. The FOA dynamics are determined by some of the chip’s parameters that can be set by changing external bias voltages and that control properties such as synaptic weights, time constants, neuron’s output spike frequency, refractory period, *etc.*

### 4 Conclusion

We demonstrated a mixed software-hardware model of saliency-based selective attention, based on the emerging AER communication protocol, and custom neuromorphic VLSI devices. We presented experimental results showing how the various components of the system interact and how the system produces behaviors that are in accordance with physiological and psychophysical data.

This work was mainly intended as a demonstration of the capabilities of the PCI-AER board, designed to implement the communication between software/hardware neuromorphic components. Now that we verified the correct behavior of all the individual components of the system, we plan to use the system presented as a real-time research tool for investigating properties of selective

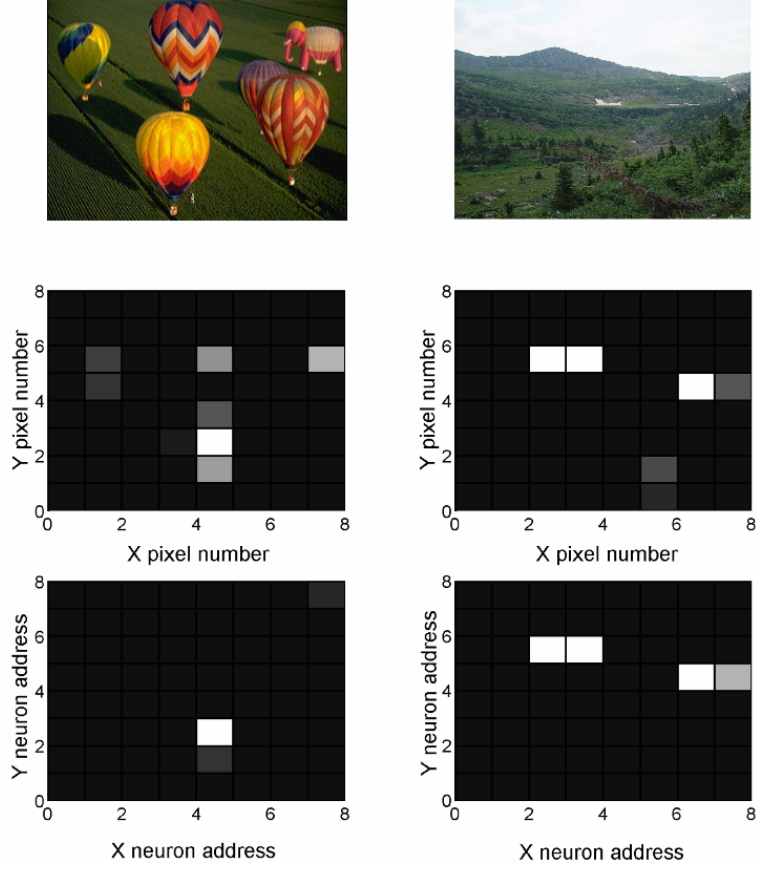


Figure 3: Original image, saliency maps and (gray-level) average firing rate of the selective attention chip neurons, for two natural scenes.

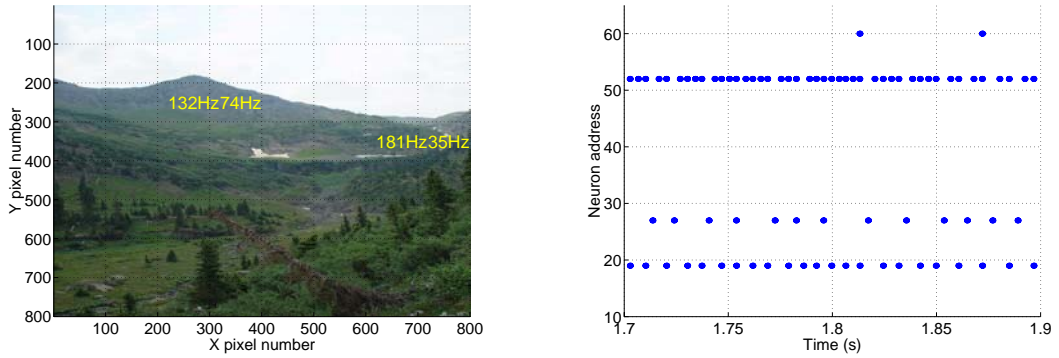


Figure 4: Original image (on the left) and raster plot of the selective attention chip output neurons (on the right). The firing rate of the corresponding regions of the saliency maps (averaged over 10s) are superimposed on the image. The raster plot of the  $8 \times 8$  neurons has the neuron address on the vertical axis and time on the horizontal axis.

attention systems and eventually identifying the key software/hardware components necessary for designing selective attention systems useful also for engineering applications.

## Acknowledgments

This work is based on experiments that were carried out at the 2002 Workshop on Neuromorphic Engineering (<http://www.ini.unizh.ch/telluride>). We are grateful to Dirk Walther and Stefano Fusi for their help with the software on the saliency map generation and spike train generation respectively, and to Rodney Douglas and Paolo Del Giudice for their support.

## 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. Dante and P. Del Giudice. The pci-aer interface board. In A. Cohen, R. Douglas, T. Horiuchi, G. Indiveri, C. Koch, T. Sejnowski, and S. Shamma, editors, *2001 Telluride Workshop on Neuromorphic Engineering Report*, pages 99–103, 2001. <http://www.ini.unizh.ch/telluride/previous/report01.pdf>.
- [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] B. Gibson and H. Egeth. Inhibition of return to object-based and environment-based locations. *Percept. Psychophys.*, 55:323–339, 1994.
- [6] G. Indiveri. Modeling selective attention using a neuromorphic analog VLSI device. *Neural Computation*, 12(12):2857–2880, December 2000.
- [7] G. Indiveri. A neuromorphic VLSI device for implementing 2-D selective attention systems. *IEEE Trans. on Neural Networks*, 12(6):1455–1463, November 2001.
- [8] L. Itti, E. Niebur, and C. Koch. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.
- [9] C. Koch and S. Ullman. Shifts in selective visual-attention – towards the underlying neural circuitry. *Human Neurobiology*, 4(4):219–227, 1985.
- [10] 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.
- [11] E. Niebur and C. Koch. Computational architectures for attention. In R. Parasuraman, editor, *The Attentive Brain*, pages 163–186. MIT Press, 1998.
- [12] Y. Tanaka and S. Shimojo. Location vs feature: Reaction time reveals dissociation between two visual functions. *Vision Research*, 36(14):2125–2140, July 1996.
- [13] A. L. Yarbus. *Eye movements and vision*. Plenum Press, 1967.