Conjunction Search Using a 1-D, Analog VLSI-based, Attentional Search/Tracking Chip

Timothy Horiuchi and Ernst Niebur
The Zanvyl Krieger Mind/Brain Institute
Johns Hopkins University

September 7, 1998

Abstract

The ability of animals to select a limited region of sensory space for scrutiny is an important factor in dealing with cluttered or complex sensory environments. Such an "attentional" system in the visual domain is believed to be involved in both the perception of objects and the control of eye movements in primates. While we can intentionally guide our attention to perform a specific task, it is also reflexively drawn to "salient" features in our sensory input space. Understanding how high-level task information and low-level stimulus information can combine to control our sensory processing is of great interest to both neuroscience and engineering. Towards this end, we have designed and fabricated a one-dimensional, analog VLSI vision chip that models covert attentional search and tracking. We extend previous analog VLSI work (Morris and DeWeerth, 1997) on delayed inhibition in a winner-take-all network to use extracted image edges as input to the attentional saliency map and to perform serial search on a particular feature conjunction (spatial derivative and direction-of-motion). We further demonstrate the ability to modify the circuit's parameters "on-the-fly" to switch between a search mode and a tracking mode.

1 Introduction

Raw visual information enters our system in a staggeringly high-dimensional space. While the information needed to perform a particular visual task might be present in this sea of data, it is only a tiny fraction of what is captured by our retinas. The difficulty lies in identifying and extracting the information which is relevant to a given situation; this is where biological visual systems excel.

Studies of visual cortex have revealed a fine-grained parallel architecture transforming the raw image into multiple feature maps at many scales. Close to the signal transducers (e.g., the photoreceptors in the retina or on a chip) parallel processing of sensory information clearly is an efficient strategy. In vision, only a small number of elementary features are typically used, making it possible to provide dedicated processors for each such feature at every location. These simple features, however, carry only a limited amount of useful information which, in general, is insufficient for the needs of higher organisms or complex machinery. Because of the astronomical number of possible combinations of features, it is practically impossible to provide specialized detectors for each behaviorally-important stimulus at every position (Tsotsos, 1990). More importantly, the processing power required to to process all of the available information in parallel far exceeds what the primate brain could possibly provide, given reasonable size and power constraints.

It has been found, indeed, that access to higher processing areas and visual awareness is severely limited by a selection process, commonly called "selective attention" (Rensink, O'Regan, & Clark, 1997). Perception in higher animals is found to be a hybrid parallel-sequential process in which the extraction of elementary features is performed in parallel for all locations of the sensory surface while

1

more advanced processing is a sequential process (see, however Palmer & McLean, 1995; Eckstein, 1998, for alternative accounts of human data in terms of parallel processes). Although attention has been most heavily studied as a spatial effect (in the head-based and environment-based coordinate frames), there is growing support for an object-based attention as well (see Kanwisher & Driver, 1992, for a review of object-based attention).

One of the standard tools for studying attention is visual search. In this psychophysical paradigm, human observers look through a set of items for a designated target and report whether the target is or is not present. The observers' reaction time (RT), starting from when the display appears, typically depends on the visual characteristics of the target and non-targets. One of the striking results of this paradigm is that RTs in experiments where a target differs from the non-targets by a single feature depend very little on the number of non-targets, while RTs in experiments where a target is defined by a conjunction (i.e., combination) of features is proportional to the number of non-targets. One of the most influential theories put forward to explain these results is the Feature Integration Theory (Treisman & Gelade, 1980; Treisman, 1988). It explains this difference by proposing that elementary features are processed in parallel by dedicated feature detectors while conjunctive targets require access to more sophisticated circuitry which is not available in parallel and therefore requires serial scrutiny. More recent research, however, has called this elegant theory into question since many conjunction searches were found to be more efficient than would be predicted by a strictly serial search (e.g. Cohen, 1993; Dehaene, 1989; Egeth, Virzi, & Garbart, 1984; McLeod, Driver, & Crip, 1988; Treisman & Sato, 1990; Wolfe, 1992).

While the necessity of an explicit attentional mechanism (separate from the main stream of information flow) is still hotly debated, several neurally-plausible models have been proposed to guide experimental neuroscience. One influential model of stimulus-driven (or "bottom-up") attention (which we have incorporated into our work) proposes the creation of a map of image saliency which is followed by a winner-take-all mechanism (Koch and Ullman 1985). The location of the highest saliency becomes the attended location. While the saliency map is thought to be constructed in parallel from generic image primitives (e.g., orientation, texture, color-discontinuity), the relative weighting of feature types is thought to be task-dependent (i.e., "top-down").

Compatible with this model is the phenomenon of "inhibition-of-return" (IOR) (Posner & Cohen, 1984; Rafal, Calabresi, Brennan, & Scioltio, 1989; Shimojo, Tanaka, Hikosaka, & Miyauchi, 1995). This effect results in an increase of RTs for detecting target stimuli at the location of attentive fixation after approximately 300 msec and lasting up to 1000 msec. The effect is particularly strong once attention has been drawn away from the location. While there is evidence that part of this "inhibition" is related to object identity (Tipper, Driver, & Weaver, 1991), of greater interest here is its spatial component. In particular, Kwak and Egeth (1992) showed directly that at least a part of the inhibition-of-return effect is spatially-defined, i.e. the return is inhibited at the *location* of the last attended item, not to its other properties. Gibson and Egeth (1994) generalized earlier approaches and showed that both object-based and environment-based descriptions can influence location-based inhibition-of-return.

A straightforward way to take the observed inhibition of return phenomenon into account computationally and which is consistent with the winner-take-all model consists of applying slowly-decaying inhibition to the attended location's input once the attention has shifted away. Alternatively, if a large enough inhibition is applied *during* attentional fixation, the winner-take-all system can be made to automatically shift to a new location and leave this slowly-decaying inhibition at the attended location, generating internal dynamics which scan different locations in order of decreasing saliency.

While some of the models developed in either neurobiology or psychology have been implemented in simulation studies (Niebur and Koch (1996); Itti, Niebur and Koch (in press); Ahmad, 1992; see Niebur and Koch, 1998 for recent reviews and references), little hardware development for system-level simulation has been done so far in this area. To date, most neuromorphic analog VLSI systems, part of whose goal is to develop real-time hardware models of neural systems at the organism-level, have remained at the early sensory processing or simple motor control level (Douglas, Mahowald,

& Mead, 1995). While several examples of reflexive sensorimotor loops have been developed (e.g. DeWeerth, Nielsen, Mead, & Astrom, 1991; Horiuchi, 1995; Etienne-Cummings, Van der Spiegel, & Mueller, 1996), little work has been done on how these systems can be integrated to process information in more complex situations. This report describes one contribution towards this end.

In the hardware modeling domain, Morris and DeWeerth (1997) (see also Wilson, Morris, & DeWeerth, 1998) designed a one-dimensional analog VLSI chip demonstrating the use of a winner-take-all circuit with delayed inhibition to produce attentional shifting. This chip also incorporated the use of fast positive feedback to provide hysteresis, or persistence, in an attentional fixation. In previous work (Horiuchi et al., 1997), winner-take-all circuits were also used to model visual attention to track moving edges in the presence of non-targets as well as to extract target motion for controlling saccadic and smooth pursuit eye movements. While this system incorporated several forms of positive feedback which improved tracking ability, one problem with this system was the inability to consider a target on the basis of a conjunction of features or to disengage attention from an unintended target.

In this paper we describe a one-dimensional chip that uses extracted image features as input to the saliency map, performs serial search in the image using the winner-take-all mechanism with delayed inhibition, and determines the location of a conjunction-defined target using a detector applied to the attentionally-gated feature stream. We further demonstrate the ability to modify the circuit's parameters "on-the-fly" to switch between a search mode and a tracking mode.

2   Architecture

Figure 1 shows a block diagram of the chip's different processing elements and the flow of information. The first stage of processing consists of image transduction and feature extraction. Adaptive photoreceptor circuits (labeled P in Figure 1) transduce the incoming pattern of light into an array of voltages. Temporal (TD) and spatial derivatives (SD) are computed from these voltages and are then used to compute the direction of image motion (DM) at each pixel.

The second stage of processing is the selection of the most salient location in the image. The spatial derivative alone is used as the input to the saliency map to drive the winner-take-all (WTA) circuit; thus the WTA selects the "strongest" edge. The activated WTA circuit then gates the local image features onto the chip's global output lines. In this way, the spatial derivative, the direction-of-motion, and the pixel position (using the "position-to-voltage," P2V, circuit) at the selected, or attended, location is known. There are two additional currents on the input to the WTA which provide the circuit with some interesting dynamics: a fast positive feedback and a delayed negative feedback. The fast positive feedback serves the role of producing hysteresis in the response to a particular stimulus. When a pixel is selected by the WTA circuit, a small additional current is added to its input node, allowing it to win by a larger margin. In order for a different pixel to become the winner, it must provide an input that is greater than the current winner's input **plus** the hysteresis input. This reduces quick oscillations between locations with comparable inputs. In addition to the positive feedback, the winning status also triggers the growth of an inhibitory current (IOR) on the input node. This produces a "fatigue" behavior, effectively reducing the saliency at that location once a pixel has been chosen. Once the inhibition is large enough to counteract both the hysteresis and pre-existing differences in inputs, the winning status jumps to the next most salient location, producing a scanning behavior.

The third stage of processing for conjunction search occurs off-chip, where feature conjunctions are detected and, in a feedback loop with the chip, circuit parameters are modified to facilitate tracking behavior. This will be discussed in subsection 4.1 where we show an example.
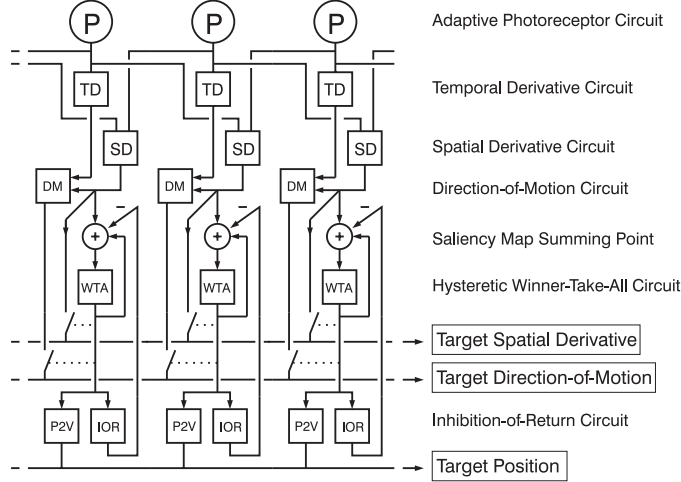
3

P   P   P   Adaptive Photoreceptor Circuit

TD TD TD Temporal Derivative Circuit

SD SD SD Spatial Derivative Circuit

DM DM DM Direction-of-Motion Circuit

+   +   +   Saliency Map Summing Point

WTA WTA WTA Hysteretic Winner-Take-All Circuit

Target Spatial Derivative

Target Direction-of-Motion

P2V IOR P2V IOR P2V IOR Inhibition-of-Return Circuit

Target Position

Figure 1: System block diagram of the tracking chip with the inhibition-of-return circuits. P = adaptive photoreceptor circuit, TD = temporal derivative circuit, SD = spatial derivative circuit, DM = direction-of-motion circuit, WTA = Winner-Take-All circuit, P2V = position-to-voltage circuit, IOR = inhibition-of-return circuit. This 2.0 μm, double-poly, CMOS chip contains a one-dimensional array of 23 pixels. The temporal derivative (TD), the spatial derivative (SD), and the direction-of-motion (DM) are computed at each location in the image which is focused onto the chip surface. The SD circuit outputs are used as the input to the saliency map and a "winner" is chosen by the WTA circuits. A limited amount of positive feedback provides hysteresis which prevents rapid switching between competing candidate locations. The winning pixel location gates information about the direction-of-motion, spatial derivative, and the pixel's position onto the global output lines.

2.1   Circuits

The transduction from light intensity to electrical signals is done with an adaptive photoreceptor circuit (Delbrück and Mead, 1989). This circuit is a high-gain amplifier in a low-pass, negative-feedback loop with a photodiode (Figure 2, left). In the feedback loop is a diode-type non-linearity which creates an intensity-dependent time-constant which spans many orders of magnitude. This basically acts as an auto-gain control circuit at the pixel level.

The temporal derivative circuit (Figure 2, middle) consists of mirroring the current required to charge a 1 pF capacitor in an op-amp follower circuit. Using this circuit, the output is split into a positive and negative signal stream ("on" and "off" channels). The two voltage references, **nsource** and **psource** are used to amplify the current.

The spatial derivative circuit (Figure 2, middle) uses a source-degenerated transconductance amplifier to measure the differences in neighboring photoreceptor voltages and this bi-directional current is then rectified into a positive and negative signal stream. Amplification in the spatial derivative signal is controlled by the SD bias parameter. A copy of the bi-directional spatial derivative is also created for later use.

The direction-of-motion circuit (Figure 2, right) computes a normalized product of the spatial and temporal derivatives ($\frac{TD \cdot SD}{|SD| + |TD|}$) which does not give the pure velocity, but the velocity modulated by the spatial derivative. This type of signal effectively reports the sign of the velocity with larger signals for larger spatial derivatives, producing a type of confidence measure.

The input to the saliency map is constructed from the absolute-value of the SD current. Due to the rectification in the SD circuit, this is a simple summation of currents, however, there is an
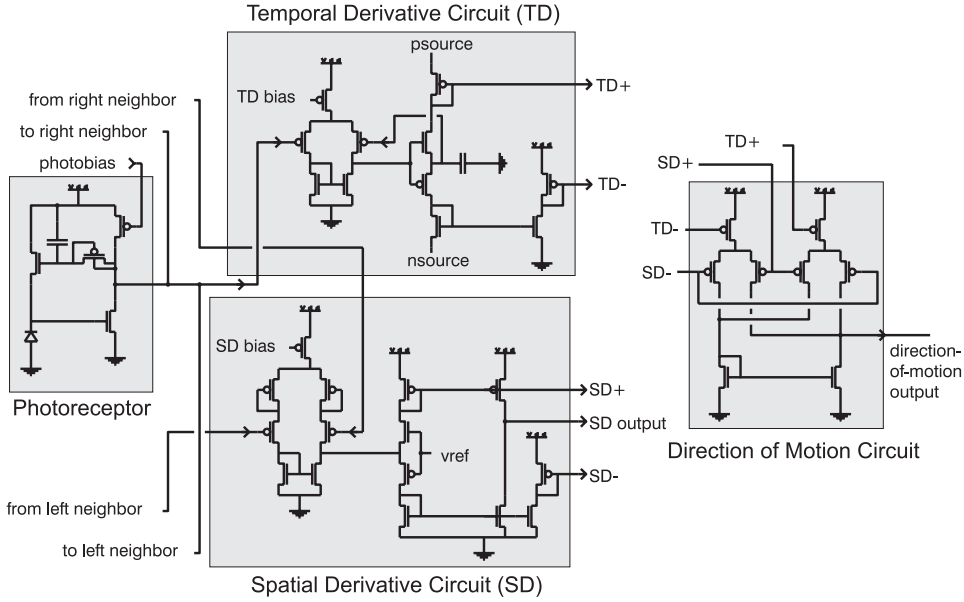
4

Figure 2: Circuit diagram for the photoreceptor (P), spatial derivative (SD), temporal derivative (TD), and direction-of-motion (DM) circuits.
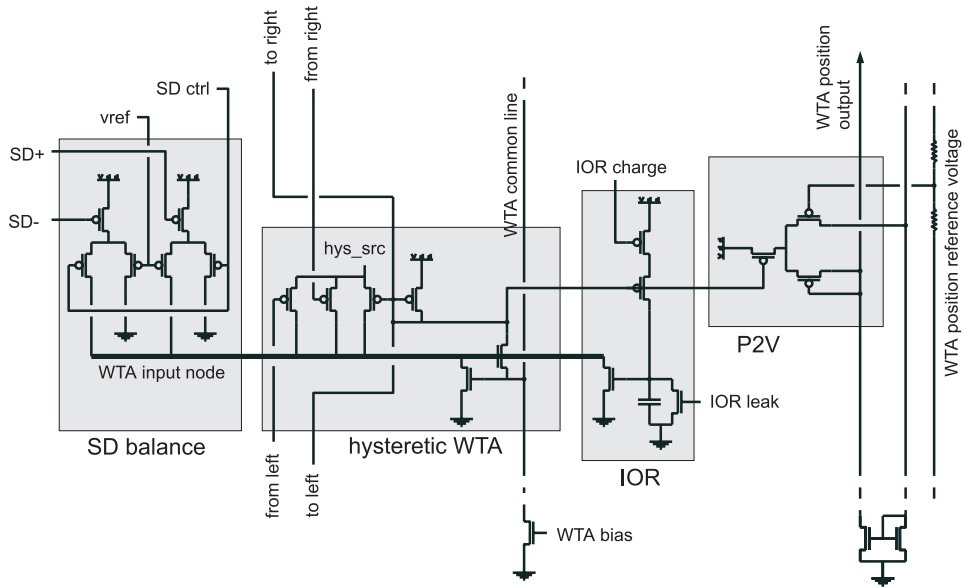

Figure 3: Circuit diagram for the SD balance, winner-take-all (WTA), inhibition-of-return (IOR), and position-to-voltage (P2V) circuits.

additional stage (SD balance circuit) of current steering where we provide the ability to bias the saliency map either towards positive or negative spatial derivatives (Figure 3, left). This balance is controlled by the signals **vref** (usually fixed) and **SD ctrl**. This circuit can be used to pre-bias the

5

chip in anticipation of particular target, or it can be used in a "feature-hysteresis" mode where a target is first chosen with no bias, then the bias is set to match whichever type of target was chosen.

The winner-take-all circuit (Lazzaro et al., 1989) is a two-transistor (per pixel) circuit that allows input currents to compete via a common voltage reference line that generates inhibition equal to the maximum input current. The "winning" pixel's input node voltage is non-zero and the global bias current generated by **WTA bias** is used to generate a feedback current. This feedback gain is controlled by the voltage **hys_src**. This current is copied to the winning pixel's neighbors to increase the likelihood of "attention" shifting locally (upon movement of the image) instead of jumping across the array.

The inhibition-of-return circuit (Figure 3, right) is activated by a copy of the feedback current which is limited by the bias voltage **IOR charge**. This current linearly charges a 260 fF capacitor which in turn generates an exponentially-rising inhibitory current on the input node. While the leak transistor controlled by **IOR leak** sets the decay rate of the inhibition, the combination of **IOR leak** and **IOR charge** set the onset rate of the inhibition.

The P2V circuit (Figure 3, right) is also activated by the feedback current. This circuit basically functions as a voltage follower, driving the local WTA-position-reference voltage onto the global line (DeWeerth, 1992). This reference voltage is provided by a polysilicon resistor ladder set to different voltages on the two ends of the array.

The WTA input node voltage is also used to generate a logical gating signal (not shown) to gate both the SD output current and the DM output current onto global lines. Also not shown are the various circuits used to monitor internal signals.

This chip was fabricated using the Orbit 2.0 μm double-poly CMOS process on a 2220 x 2220 μm^2 die. The array consists of 23 pixels. Each pixel strip is 64 μm wide by 1370 μm long. This length includes all of the processing described above, as well as the necessary circuitry for accessing intermediate signals for both monitoring and circuit tuning.

3 Performance

To demonstrate the operation of the chip, we focus both static and dynamic test images onto the surface of the chip and measure the various output signals. Our stimuli consist of medium-contrast, black-and-white patterns printed onto paper mounted on a rotating mechanism to produce either a static picture or a one-dimensional "movie."

3.1 Static Images

Figure 4 shows outputs from the chip for a *static* scene consisting of two dark bars. Each bar produces two regions of activity on the spatial derivative map (Figure 4, left), one from the light-to-dark transition on the left edge and one from the dark-to-light transition on the right edge. The four bumps of activity will be the dominant features of the saliency map that drive the WTA circuit.

The larger panel on the right shows the position of the WTA-selected pixel and the SD current at that position as a function of time. As described above, the IOR circuits produce inhibition that cause the WTA circuits to select a new pixel after a short time. In this example, the rate of decay for the IOR circuit (once the WTA has selected another pixel) is high so that the strong input locations recover after a short time and only the eight most salient locations are visited. By creating a scatter plot of spatial derivative vs. WTA position (Figure 5), a selective reconstruction of the original saliency map can be made. Only the attended pixel locations, however, will be "painted" in.
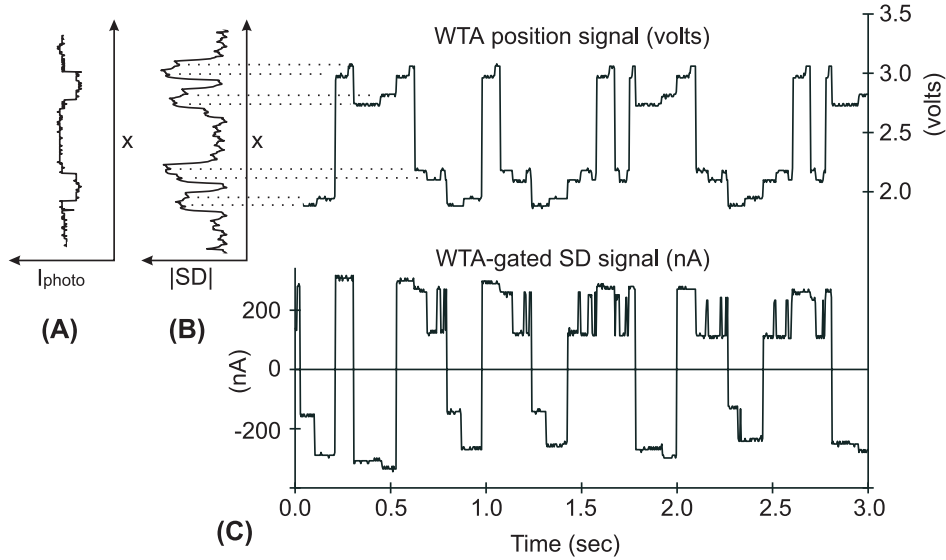
Figure 4: Winner-take-all trajectories for a static image of two dark bars. The image of two dark bars is focused onto the photoreceptor array (A) which produces four bumps of activity in the spatial derivative circuits (B). (Note that internally, the chip preserves the polarity of the spatial derivative, but only reports an unsigned SD signal for scan out). The winner-take-all circuits select pixels sequentially as the delayed, negative-feedback inhibition keeps the winning-status unstable. In panel (c), the upper trace shows the position of the winning pixel as a function of time and the lower trace shows the spatial derivative at the selected pixel. The parameters have been chosen for this particular image such that only about eight locations are ever visited.
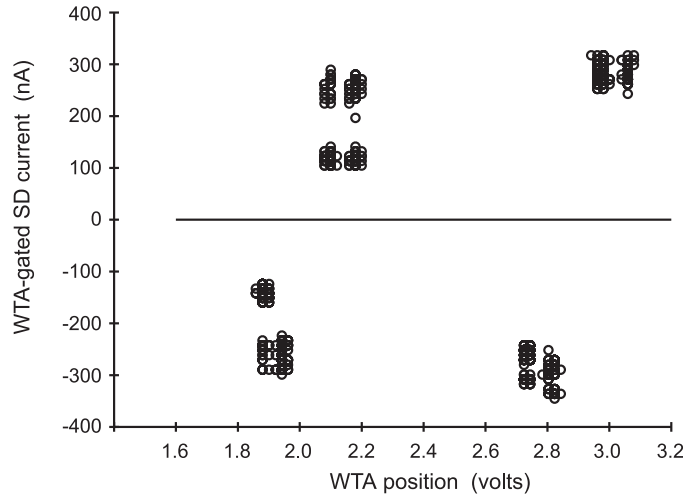

Figure 5: Scatter plot of the winning pixel location and its spatial derivative when selected by the WTA mechanism.

4 Dynamic Stimuli

Figure 6 shows the scatter plots of four different *dynamic* stimulus conditions which were designed to stimulate the four quadrants of the SD-motion stimulus space. The stimuli were single edges
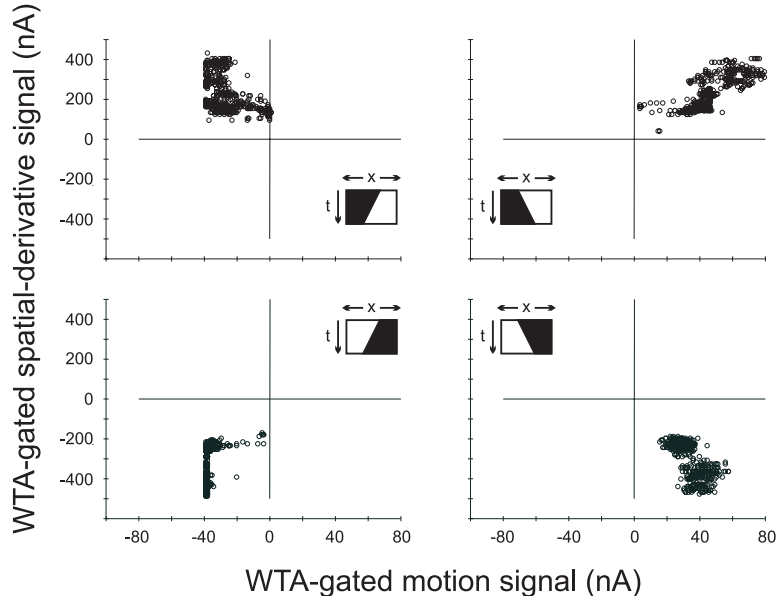
Figure 6: Feature conjunction display (SD vs. direction). Four different stimulus conditions are shown: a white-black edge moving left and right (top two panels), and a black-white edge moving left and right (bottom two panels). The vertical axis in all plots is the spatial derivative and the horizontal axis in all plots is the direction-of-motion. The chip selects the one strong edge in each stimulus condition and reports its characteristics. The four stimulus conditions produce outputs in the four different quadrants. From these outputs it is possible to make a "quadrant detector" to signal the existence of a target with the desired feature conjunction.

(positive and negative derivative) that were slowly moved either leftward or rightward. These single targets were tracked by the WTA circuit and their feature characteristics were reported off-chip.

Next, a stimulus was used which consists of multiple targets that lie in all four quadrants of the SD-direction stimulus space (Figure 7(A)). With the inhibition-of-return circuits operating, the WTA circuit shifts from one pixel to another, reporting the feature characteristics for each pixel visited. By constructing an off-chip quadrant detector (two zero-crossing detectors and a logical AND) we can detect when the WTA has selected a particular target of interest. In this example, we have chosen to detect the pixels of positive derivative and leftward moving edges. Figure 7(B) shows the output of the WTA position and the quadrant detector as a function of time.

Figure 8 shows the spatial derivative and direction-of-motion outputs for the same experiment as in Figure 7. By producing the scatter plot of spatial derivative vs. direction, we see that the WTA circuit has visited all four quadrants (i.e. all four moving targets).

4.1 Switching From Search to Tracking

Along with the ability to search for a predetermined target, it is also important to have the control to stop and scrutinize, or track, a target. While the inhibition-of-return acts as a destabilizing influence, producing repetitive shifts, the position-based hysteresis current and the feature-based hysteresis system both act to stabilize the WTA, producing tracking behavior. The modes of searching (destabilized WTA) and tracking (stabilized WTA) have different optimal parameter settings.

To demonstrate the circuit's ability to switch between these two important behavioral modes of an attentional system, a small external circuit was constructed to use the quadrant detector signal

(A)

X

t

"4-quadrant stimulus"
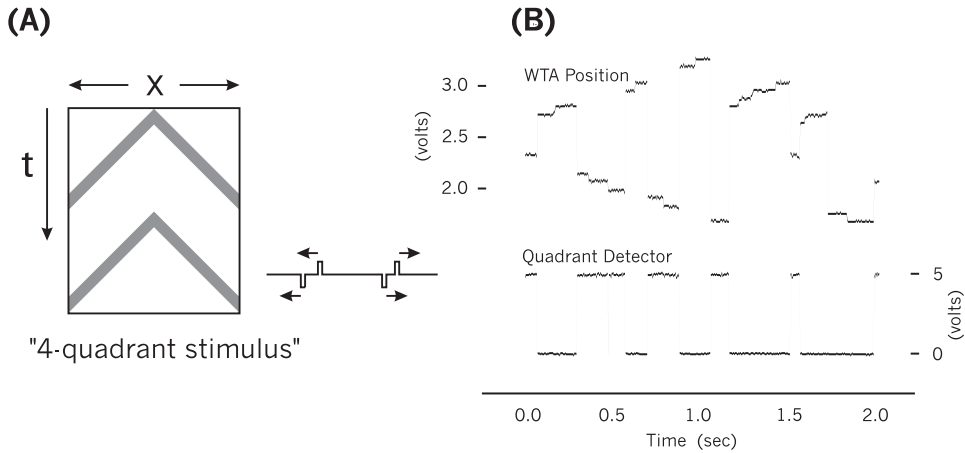
(B)

WTA Position

Quadrant Detector

Figure 7: Example of a SD-direction conjunction search using the inhibition-of-return. In this example, a moving stimulus which contains stimuli that fit into all four quadrants (of the SD-direction space) is used **(A)**. An external detector circuit was constructed to indicate the presence of a positive spatial derivative edge moving to the right. Panel **(B)** shows the position of the WTA-selected pixel (upper trace) and the output of the external quadrant detector circuit (lower trace) as a function of time. Using the quadrant-detector circuit output, it is possible to disable the inhibitory circuits and shift into a tracking-only mode.
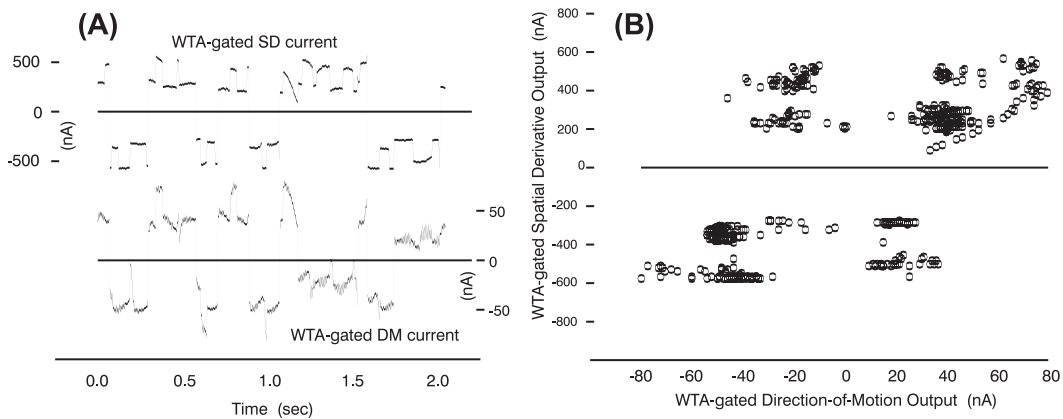

(A) WTA-gated SD current

WTA-gated DM current

(B)

WTA-gated Spatial Derivative Output (nA)

WTA-gated Direction-of-Motion Output (nA)

Figure 8: From the same experiment as in Figure 7, the spatial derivative and motion outputs are considered. **(A)** Spatial derivative (upper trace) and direction-of-motion (lower trace) as a function of time. **(B)** Scatter plot of the spatial derivative vs. the direction-of-motion. This plot shows that all four possible conjunction targets were visited

to trigger a change in three user-definable parameters. In this demonstration, as in the previous subsection, a stimulus is presented that consists of four possible feature-conjunction targets. The task of the circuit is to locate and then track the target which consists of the conjunction of positive spatial derivative and leftward motion.

The three parameters that were modified when switching from the search mode to tracking mode were: 1) the inhibitory onset rate was changed to zero to prevent inhibition from building up, 2) the WTA-based positive feedback to both the local pixel and its two nearest neighbors was increased
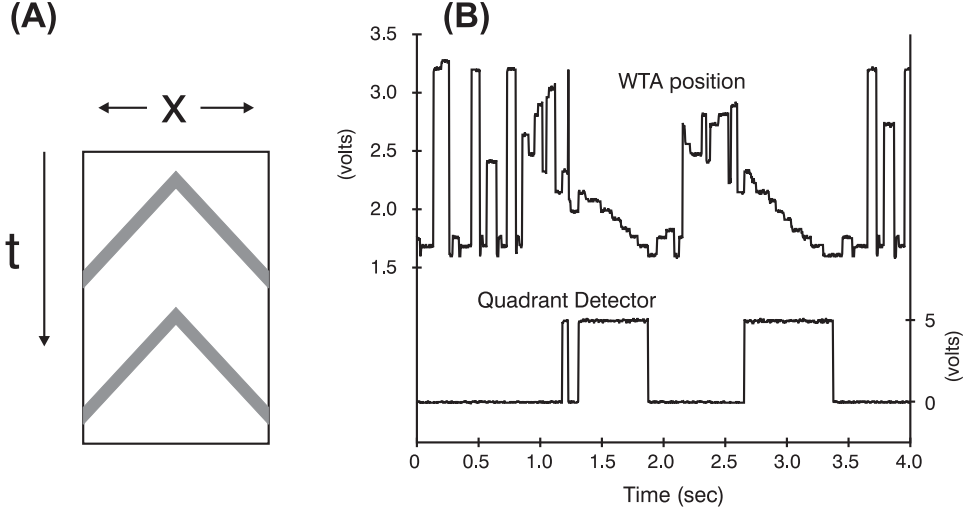
Figure 9: **(A)** The stimulus used for the search experiment contains four edges which have different spatial derivatives) and different directions of motion. This is the same stimulus as in Figure 7. **(B)** The WTA position and the quadrant detector output are plotted as a function of time. In this experiment, the quadrant detector output is used to switch the parameters on the chip from values that promote switching behavior, as in Figure 7, to those that promote tracking behavior. In this particular example, the target edge (positive spatial derivative and leftward motion) is initially being tracked, but as the edge leaves the array and a new group of edges are available, the system begins to search the different possible edges until the target edge is found and tracking is resumed.

to improve tracking, and 3) the feature-based hysteresis was activated to strongly bias the saliency map to the type of edge being tracked. The effect of these parameter changes are visible in Figure 9 where the scanning behavior gives way to tracking when the quadrant detector signal is activated. When the tracked target leaves the field of view, the chip begins scanning different edges until the target edge is found again.

It should be noted that, when taking the data shown in Figure 8, we chose to make the chip find the correct edge by searching all possible edges. Alternatively, it would have been possible to pre-bias the saliency map to only look for one type of edge and scan for the correct direction-of-motion.

It is relatively simple to tune the parameters for a particular behavior once an image has been selected, such as number of targets visited, the duration of an attentional fixation, or speed of shifting, however, these characteristics are not independently controlled. The shifting behavior of the attentional system will depend critically on the mean input strengths, the relative input strengths, the pattern of input movement, the level of hysteresis, and the rate of onset and decay of the inhibition-of-return circuit.

4.2 Power Consumption

The fabricated chip draws approximately 1.08 mA at 5 volts, (or 5.4 mW) which includes the power to the padframe. The image processing and winner-take-all selection circuits alone draw approximately 120 μA at 5 volts (or 600 μW).

10

5 Summary and Considerations for the Future

In this paper we have described a one-dimensional analog VLSI-based model of covert attentional search using an augmented winner-take-all (WTA) circuit and WTA-gated features to detect conjunction-defined targets. This work combines and extends the work of Horiuchi et al. (1997) and Morris and DeWeerth (1997) to produce a single chip capable of operating on visual edges as the input saliency map as well as reporting the position, the polarity, and the direction-of-motion of the selected edge.

From a neural modeling perspective there are a number of shortcomings. The first and most obvious is the pixel-based nature of the WTA circuit. Most objects in our world occupy regions of space that vary dramatically in both size and shape, and comprise many different features. To address part of this issue, Morris et al. (1998) demonstrated a switch-based coupling of adjacent pixels at the WTA input node to create "super-pixels" which are then selected as a region. Another limitation of this model is the assignment of the inhibition-of-return function to the retinal coordinate frame. If the attentional system were used to drive an eye movement, the inhibition associated with different objects in the image would be shifted off the correct locations and inappropriately placed onto new locations. It may be quite possible, however, that this problem is only an apparent one since human psychophysics has found little or no integration across saccades (Bridgeman, Hendry, & Stark, 1975; McConkie & Zola, 1979; Irwin, Yantis, & Jonides, 1983; Irwin, 1991; Grimes, 1996; Irwin, 1996). This would remove any difficulty with different coordinate frames since only information from the presently used frame is required. Along the same lines, more extreme results have been recently obtained indicating that visual search may not rely on *any* stored representation of the visual input (Horowitz & Wolfe, 1998). It remains to be seen, however, how these data relate to the vast body of literature showing interactions between subsequent attentional fixations in the form of inhibition-of-return. Another more subtle difficulty in relating this model to neurobiological function is the extreme non-linearity of the winner-take-all concept. There are a few suggestions that we may effectively be able to split our attention into more than a single region across both space and sensory modality (Pashler, 1987; Pylyshyn & Storm, 1988; Wolfe, Cave, & Franzel, 1989). Although this is clearly a minority view at this time, it deserves to be kept in mind.

From an engineering perspective, a higher-resolution, two-dimensional system is clearly desirable for a larger field of view as well as for the use of two-dimensional features (e.g., orientation) on the input. Although this particular design was demonstrated in one-dimensions, scaling the winner-take-all circuit and its gating circuits up to multiple dimensions should not dramatically increase the design, size, or operation of the circuits. While intended as a model for biological attentional search, this circuit and variants of it are well suited to controlled industrial environments where the detection of faults in a manufacturing process often needs to be small, cheap, and fast.

6 Acknowledgements

References

Ahmad, S. (1992). VISIT: A Neural Model of Covert Visual Attention. In Moody, J., Hanson, S., & Lippman, R. (Eds.), *Advances in Neural Information Processing Systems 4*, pp. 420–427. Morgan Kaufman, San Mateo.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacements of the visual world during saccadic eye movements. *Vision Research, 15*, 719–722.

Cohen, A. (1993). Asymmetries in visual search for conjunctive targets. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 775–797.

Dehaene, S. (1989). Discriminability and dimensionality effects in visual search for feature conjunctions: a functional pop-out. *Perception & Psychophysics, 46,* 72–80.

Delbrück, T., & Mead, C. (1989). An Electronic Photoreceptor Sensitive to Small Changes in Intensity. In Touretzky (Ed.), *Neural Information Processing Systems 1*, pp. 720–727. Morgan Kaufmann, San Mateo, CA.

DeWeerth, S. P. (1992). Analog VLSI Circuits for Stimulus Localization and Centroid Computation. *Intl. J. Comp. Vis., 8(3),* 191–202.

DeWeerth, S. P., Nielsen, L., Mead, C. A., & Astrom, K. J. (1991). A Simple Neuron Servo. *IEEE Transactions on Neural Networks, 2*(2), 248–251.

Douglas, R., Mahowald, M., & Mead, C. (1995). Neuromorphic Analogue VLSI. In Cowan, W. M., Shooter, E. M., Stevens, C. F., & Thompson, R. F. (Eds.), *Annual Reviews in Neuroscience, Vol. 18*, pp. 255–281. Annual Reviews Inc., Palo Alto, CA.

Eckstein, M. P. (1998). The higher visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 9*(2), 111–118.

Egeth, H., Virzi, R., & Garbart, H. (1984). Searching for conjunctively defined targets. *J. Experimental Psychology, 10*(1), 32–39.

Etienne-Cummings, R., Van der Spiegel, J., & Mueller, P. (1996). A Visual Smooth Pursuit Tracking Chip. In Touretzky, Mozer, & Hasselmo (Eds.), *Advances in Neural Information Processing Systems 8*, pp. 706–712. MIT Press, Cambridge, MA.

Gibson, B., & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics, 55*(3), 323–339.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In Akins, K. (Ed.), *Perception*, Vol. 5 of *Vancouver studies in cognitive science*, pp. 89–109. Oxford University Press, New York.

Horiuchi, T. (1995). An Auditory Localization and Coordinate Transform Chip. In Tesauro, G., Touretzky, D., & Leen, T. (Eds.), *Advances in Neural Information Processing Systems 7*, pp. 787–794. MIT Press, Cambridge, MA.

Horiuchi, T. K., Morris, T. G., Koch, C., & DeWeerth, S. P. (1997). Analog VLSI Circuits for Attention-based Visual Tracking. In Mozer, M., Jordan, M., & Petsche, T. (Eds.), *Advances in Neural Processing Systems 9*, pp. 706–712. MIT Press.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394,* 575–577.

Irwin, D. (1996). Integrating Information Across Saccadic Eye Movements. *Current Directions in Psychological Science, 5*(3), 94–100.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23,* 420–456.

Irwin, D. E., Yantis, S., & Jonides, J. (1983). Evidence against visual integration across saccadic eye movements. *Perception & Psychophysics, 34,* 49–57.

Itti, L., Niebur, E., & Koch, C. (1998). A model of saliency-based fast visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, in press*.

Kanwisher, N., & Driver, J. (1992). Objects, attributes and visual attention: which, what, and where. *Current Directions in psychological science, 1*, 26–31.

Koch, C., & Ullman, S. (1985). Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry. *Human Neurobiology, 4*, 219–227.

Kwak, H., & Egeth, H. (1992). Consequences of allocating attention to locations and to other attributes. *Perception & Psychophysics, 51*(5), 455–464.

Lazzaro, J., Ryckebusch, S., Mahowald, M. A., & Mead, C. (1989). Winner-take-all networks of O(n) complexity. In Touretzky (Ed.), *Advances in Neural Information Processing Systems 1*, pp. 703–711. Morgan Kaufmann, San Mateo, CA.

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading?. *Perception & Psychophysics, 25*, 221–224.

McLeod, P., Driver, J., & Crip, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature, 332*, 154–155.

Morris, T., & DeWeerth, S. P. (1997). Analog VLSI Excitatory Feedback Circuits for Attentional Shifts and Tracking. *Analog Integ. Circ. and Sig. Proc., 13*(1/2), 79–91.

Morris, T. G., Horiuchi, T. K., & DeWeerth, S. P. (1998). Object-based Selection Within an Analog VLSI Visual Attention System. *Circuits and Systems II. in press*.

Niebur, E., & Koch, C. (1996). Control of Selective Visual Attention: Modeling the "Where" Pathway. In Touretzky, D. S., Mozer, M. C., & Hasselmo, M. E. (Eds.), *Advances in Neural Information Processing Systems*, Vol. 8, pp. 802–808. MIT Press, Cambridge, MA.

Niebur, E., & Koch, C. (1998). Computational Architectures For Attention. In Parasuraman, R. (Ed.), *The Attentive Brain*, chap. 9, pp. 163–186. MIT Press, Cambridge, MA.

Palmer, J., & McLean, J. (1995). Imperfect, Unlimited-Capacity, Parallel Search Yields Large Set-Size Effects. Irvine, CA. Society of Mathematical Psycholgy.

Pashler, H. (1987). Detecting conjunctions of color and form: reassessing the serial search hypothesis. *Perception & Psychophysics, 41*, 191–201.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In Bouma, H., & Bouwhuis, D. G. (Eds.), *Attention and Performance X*, pp. 531–556. Hilldale, NJ.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179–197.

Rafal, R., Calabresi, P., Brennan, C., & Scioltio, T. (1989). Saccade Preparation Inhibits Reorienting to Recently Attended Locations. *J. Exp. Psych: Hum. Percep. and Perf, 15*, 673–685.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368–373.

Shimojo, S., Tanaka, Y., Hikosaka, O., & Miyauchi, S. (1995). Vision, Attention, and Action – inhibition and facilitation in sensory motor links revealed by the reaction time and the line-motion.. In Inui, T., & McClelland, J. L. (Eds.), *Attention and Performance XVI*. MIT Press.

13

Tipper, S. P., Driver, J., & Weaver, B. (1991). Short report: object-centered inhibition or return of visual attention. *Quarterly Journal of Exp. Psychology*, *43A*, 289–298.

Treisman, A. (1988). Features and Objects: The Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology*, *40*(2A), 201–237.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Human Perception and Performance*, *16*, 459–478.

Tsotsos, J. (1990). Analyzing vision at the complexity level. *Behav. Brain Sci.*, *13*, 423–469.

Wilson, C. S., Morris, T. G., & DeWeerth, S. P. (1998). Segmentation Coding for Object-Based Attentive Selection Systems. In *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems Conference, Monterey, CA*. IEEE Press, Piscataway, NJ. in press.

Wolfe, J. M. (1992). Effortless texture segmentation and parallel visual search are not the same thing. *Vision Research*, *32*(4), 757–763.

Wolfe, J., Cave, K., & Franzel, S. (1989). Guided search: an alternative to the feature integration model for visual search. *J. Exp. Psychology*, *15*, 419–433.