A Biomimetic Focal Plane Speed Computation Architecture

Vivek Pant[1] and Charles M. Higgins[1,2]
[1]*Department of Electrical and Computer Engineering,* [2]*ARL Division of Neurobiology, The University of Arizona, Tucson, AZ.*
vivek@email.arizona.edu, higgins@ece.arizona.edu

Abstract: A sensor was designed to compute speed at the image focal plane for robotic navigation. It employs an array of parallel sensing and computing blocks, and outputs a signal that varies linearly with image speed.

1. Introduction

Robotic visual navigation is a complex task that requires data to be acquired from an image sensor and processed in real time. Conventional computer architectures may be used for autonomous navigation in many cases. However, they require fast analog-to-digital converters to communicate the data from a high-resolution sensor (such as a CCD array) to a serial digital processor. They further demand a fast algorithm to process this data in order to perform real-time decisions. In a system that is required to operate on a limited power budget, the cost of high-speed data transfer and processing is often a bottleneck. The insect visual system presents an elegant alternative approach to both sensory data acquisition and processing. Since this system has survived and evolved over hundreds of millions of years, it provides us with highly optimized neuronal algorithms to design architectures for autonomous navigation. Such neuromorphic architectures utilize an array of sensor circuits to transduce and parallel process the visual information at the image focal plane itself. This bypasses the problems of data communication and analog-to-digital conversion and makes the neuromorphic system fast. Also, since these systems utilize analog circuits to process the data, their output may be used directly as a control signal for navigation.

2. Algorithm

It has been shown that insects utilize image speed for performing complicated navigational tasks such as steering through obstacles, measuring flight distances, and smooth landing [1,2]. Image speed may be computed in many ways, for example by using a feature-based scheme that computes speed as a ratio of the distance between features in two successive frames to the inter-frame interval ($\Delta x/\Delta t$). However, in a complex motion scenario this computation may lead to multiple velocity vectors oriented such that determining the net speed is difficult. Spatio-temporally tuned motion-energy models are able to handle such complex scenarios better and may be used to predict the direction of motion unambiguously [3]. However, these motion energy models, when used to extract visual image speed, suffer from dependence on spatial frequency. A non-directional motion unit that has spatio-temporal characteristics and also low dependence on spatial frequency was recently proposed [4,5]. A single such unit (see Figure 1a) receives input from three adjacent photodetectors (PD). This signal is then temporally high-pass filtered (HPF). The high-pass filtered outputs from the two lateral photodetectors are then low-pass filtered (LPF) and added. This signal is then added with the central photodetector response to get an output whose mean response is roughly proportional to image speed. The time constant of both the low-pass and the high-pass filter was chosen to be τ. For a sinusoidal input grating $I = C/2 \cdot \sin(\omega_x x + \omega_t t)$, where C is the contrast, and ω_x and ω_t are the spatial and temporal frequencies of the grating, the output of this model is given as:

$$O_{speedS} = \frac{C}{2}\frac{\omega_t \tau}{\left(1+(\omega_t \tau)^2\right)}\sqrt{4\left(\cos(\omega_x)^2 + \cos(\omega_x)\right) + (\omega_t \tau)^2 + 1} \quad . \qquad (1)$$

The solid lines in Figure 1b show the response of a single unit at different spatial frequencies (ranging between 0.1 to 0.25 cycles per sensor) as the image speed is varied. For this range of image speeds, the sensor response is seen to be roughly independent of the spatial frequency.

3. Chip Architecture

We have developed a VLSI sensor architecture based on this non-directional summation (ND-S) model that may be used to estimate image speed in a scene. The architecture comprises an array of pixels, each computing non-directional motion with respect to its adjacent pixels (Figure 2a). An adaptive photoreceptor transduces light incident on a pixel into an electrical signal. This circuit can operate at over five orders of magnitude of light

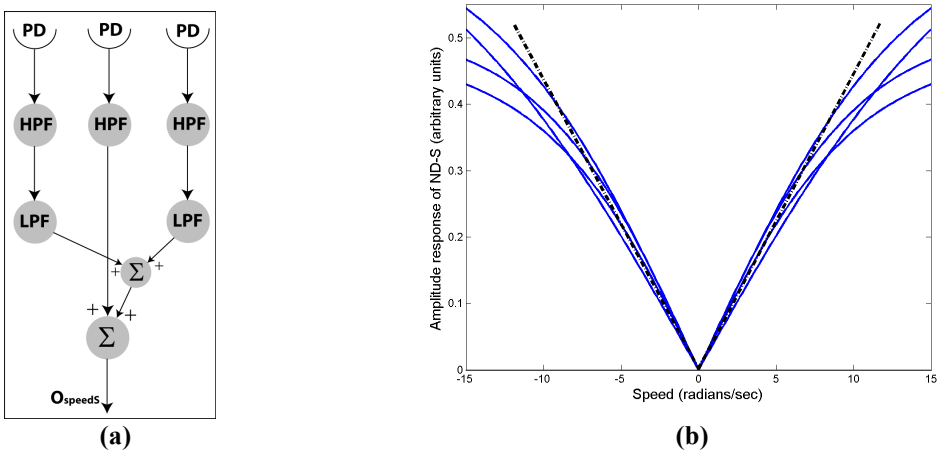


Figure 1: (a) A non-directional summation unit (ND-S). (b) Response of a ND-S unit at different spatial frequencies (solid line curves). The dotted line shows linear speed response.

intensity and has a band-pass characteristic [6]. A g_m-C type first-order low-pass filter with tunable time constant implements the next computational block. The DC offset of the photoreceptor and filter outputs is subtracted prior to the next stage to get pure AC signals. These AC signals from the two low-pass filters and the central photoreceptor are then added using a current-mode circuit. The amplitude response is finally computed by summing the outputs of current-mode absolute value circuits in each pixel. A single transistor gate (S) in each pixel is used to combine outputs from pixels and is controlled by peripheral digital circuitry.

A chip incorporating this sensor has been fabricated in 0.18 μm CMOS process through MOSIS on a 4×4 mm^2 die. The chip has a grid of 70×70 ND-S type pixels. Two different layouts were fabricated for each pixel, differing in fill factor (the percentage area of a pixel devoted to phototransduction). One pixel type was designed with a high fill factor (pixel type 1), and the other with bigger capacitors for better control over the circuit time constants (pixel type 2) but lower fill factor. In this publication we present results from ND-S type 2 pixel. The photodiode was fabricated as a well-type diffusion and was surrounded by a guard-ring to minimize substrate currents. The fill factor for the pixel (13%) was carefully designed so that the entire signal processing circuitry may be packed without sacrificing the optical information. The chip has peripheral digital scanner circuitry to access the pixel core. Scanners are addressable flip-flop circuits and act as an indexing scheme to route a signal to and from a specific pixel (addressed by its x- and y-position) to a bond pad on the chip. A photomicrograph of the chip is shown in Figure 2b.

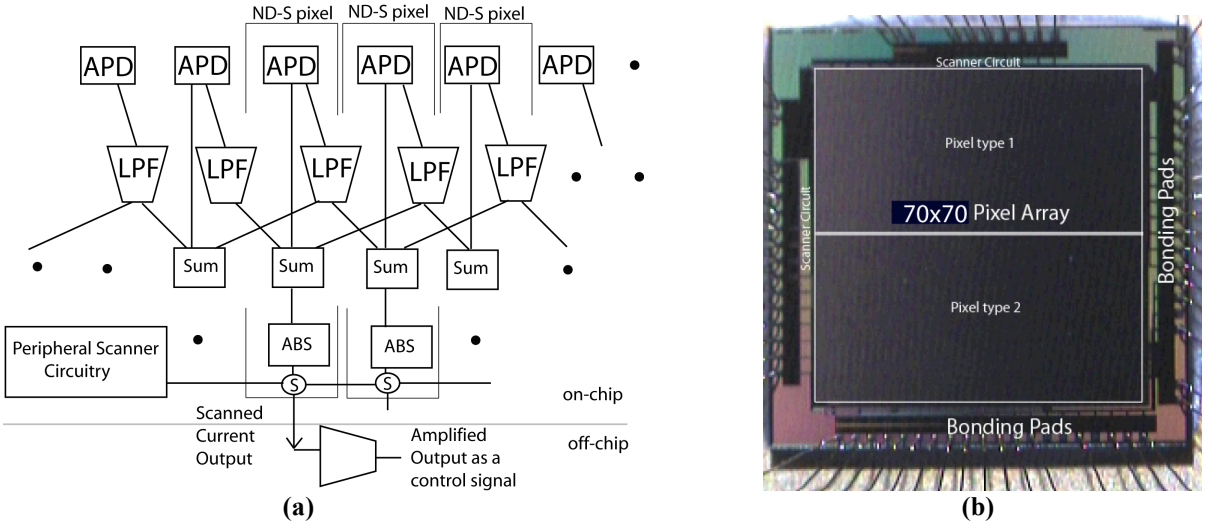
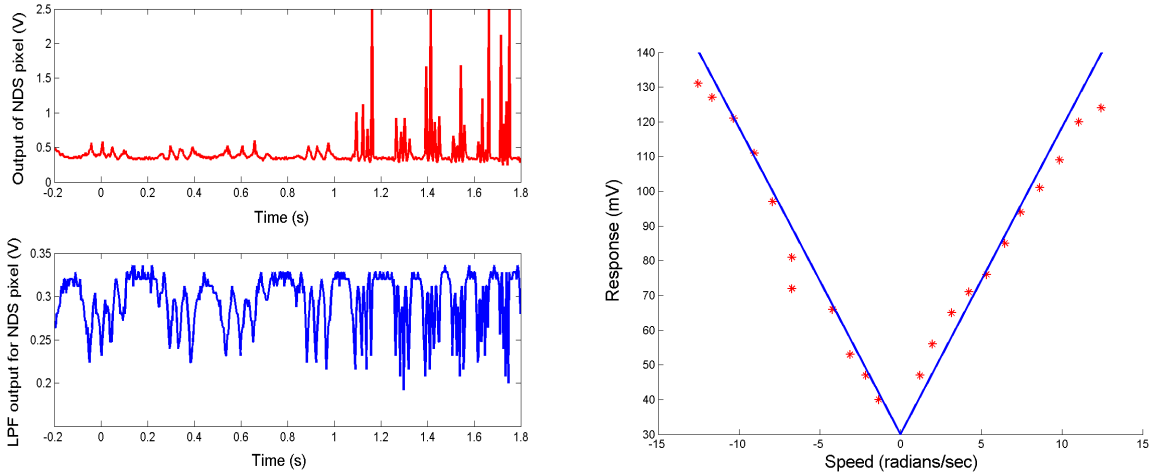


Figure 2: (a) Block diagram of the computational architecture. The output from the chip is measured by using a current-sense amplifier circuit external to the chip. (b) Photomicrograph of the integrated circuit.

4. Results

The chip was tested for artificial and real world stimuli. Figure 3a shows the raw output from the chip when a real-world stimulus moves in its visual field. It was seen that the response of the chip to an initial slow moving stimulus was much smaller than its response to the same stimulus moving fast. To characterize this further we used a setup incorporating a drum stimulus covered with a square-wave grating. The speed of this rotating stimulus was varied from -15 to 15 radians/second. A typical output current response from a single pixel was of the order of 0.1μA. The output from an ND-S pixel was converted to a voltage signal using a current-sense amplifier with a gain of 7.5×10^5. Each data point was recorded by averaging over five cycles. The pixel response largely follows a linear fit (see Figure 3b). Further characterization of this chip is currently in progress. All the data reported in this paper was recorded from a single ND-S pixel (for tabulated data, see Table 1). The chip was tested with a supply voltage of 2.2V and consumed 19.4mW of electrical power. All the circuits in a pixel were operated in the subthreshold regime of the MOS transistor to minimize power consumption. The power consumed by a single pixel was less than 1.1μW.



(a) **(b)**

Figure 3: (a) Raw data from the ND-S type pixel to a real-world stimulus. A hand was moved back and forth in front of the chip, slowly before 0.8 seconds, and then more quickly afterwards. The lower trace shows the response of the low-pass filter while the upper trace shows the corresponding net response of an ND-S pixel. (b) Response characteristics of a single ND-S type pixel. Data (*) plotted above was recorded by taking an average of the pixel's response over five stimulus presentations. The solid line is the response of a linear model.

Table 1: Measured chip performance

Process resolution	0.18μm (through MOSIS)
Chip size	70 × 70 array of ND-S type pixels on a 16 mm^2 die
Pixel size	$23.2\times23.6\mu m^2$
Pixel element count	1 photodiode, 23 transistors, and 3 capacitors
Fill factor	13% for pixel type 1 17.4% for pixel type 2
Typical output current	100 nA

5. References

[1] M.V. Srinivasan, S. W. Zhang, and N. J. Bidwell, "Visually mediated odometry in honeybees," *J. of Experimental Biology* 200, pp. 2513-2522, 1997

[2] A. Barron and M.V. Srinivasan, "Visual regulation of ground speed and headwind compensation in freely flying honey bees (*Apis mellifera* L.)," *J. of Experimental Biology* 209, pp. 978-984, February 2006

[3] E.H. Adelson and J.R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Optical Society of America A*, vol. 2, no. 2, pp. 284-299, 1985

[4] C. M. Higgins, "Non-directional motion may underlie insect behavioural use of image speed," *Biological Cybernetics*, vol. 91, no. 5, pp. 326-332, November 2004

[5] Z. Rivera-Alvidrez, "Computational modeling of neurons involved in fly motion detection," *MS thesis*, Department of Electrical and Computer Engineering, The University of Arizona, August 2005

[6] T. Delbrück, "Silicon retina with correlation-based, velocity-tuned pixels," *IEEE Trans. Neural Networks*, vol. 4, pp. 529-541, May 1993