Attentive motion sensor for mobile robotic applications

Chiara Bartolozzi
and Neeraj K. Mandloi
Italian Institute of Technology
Genova, Italy
e-mail: chiara.bartolozzi@iit.it

Giacomo Indiveri
University of Zurich and ETH Zurich
Institute of Neuroinformatics
Zurich, Switzerland
e-mail: giacomo@ini.phys.ethz.ch

Abstract—We present a compact vision sensor comprising a one-dimensional array of adaptive photoreceptors, spatio-temporal feature extraction circuits, feature normalization circuits, and an attentional readout circuit that selects the most salient region in the feature map. The sensor comprises also digital input and output circuits for directly interfacing it to digital processing units, making it an ideal device for mobile robotic applications. We describe the sensor architecture and present experimental results measured from the fabricated chip. As we identified unexpected results from one of the computational stages, we compare the measured responses to circuit simulations and propose improvements for new revisions of the chip.

I. INTRODUCTION

Processing detailed sensory information in real-time is a computationally demanding task for both natural and artificial sensory systems. This is especially true for vision, where the amount of information provided by the sensors typically exceeds the parallel processing capabilities of the system. An effective strategy to cope with large amounts of input data in face of limited computational resources is to use selective attention, a process that allows the system to select sub-regions of the input and process them serially, shifting from one sub-region to another, in a sequential fashion [1].

In active vision and robotic applications these strategies can be used to decide which salient sub-regions of the sensory input space to process, dramatically reducing the bandwidth requirements for information transfer, and the system’s overall computational load [2]. With this goal in mind, we designed a compact one-dimensional neuromorphic vision sensor that implements an on-chip model of selective attention, to select and track features of interest as they move in the environment. We will refer to this device as the Tracker-Motion Sensor (TMS). In the TMS, “saliency” is defined as a combination of spatial and temporal contrast features, as well as stimulus velocity. All these features are computed in parallel on the sensor focal-plane, weighted by adjustable biases, and summed into the input of a winner-take-all (WTA) circuit, which selects the most salient input and produces in output the address of the winning pixel, using both analog and digital representations.

The low-power vision processing capabilities of the TMS, combined with the parallel/distributed computation approach, allow both a low weight budget and a redundant fault-tolerant architecture. Furthermore, the current-mode neuromorphic circuits used in the TMS for designing the WTA network are ideal for hardware models of selective attention systems [3]. Several neuromorphic attention systems of this kind have been proposed in the past [4]–[9]. These systems typically contain photo-sensing elements and processing elements on the same focal plane, apply the competitive selection process to visual stimuli sensed and processed by the focal plane processor itself, and perform visual tracking operations. The TMS extends their functionality by implementing a more complex saliency map computation stage, which comprises a flexible way of combining the different features. In addition, the TMS has digital input address decoders for selecting the pixel to read-out from. These, combined with the digital output address encoders, provide the possibility to selectively read-out the sensor’s photo-receptors (e.g. to measure visual features) at the location of the WTA output, or of a small window of interest around the winning pixel (e.g. by means of a microcontroller). The fast (real-time) processing characteristics of TMS, combined with these flexible attention-based read-out techniques, make it an ideal device for on-line active-vision tasks and/or mobile robotic applications.

In this paper we present the TMS architecture, describe...
its main analog processing blocks, and present preliminary experimental results measured from the fabricated chip. We point out unexpected behaviors in the current implementation, also by comparing SPICE simulations to measured data. We conclude by proposing circuit modifications that can improve the system’s performance and by describing application domains that are optimally suited for the TMS.

II. THE TRACKER MOTION SENSOR ARCHITECTURE

Figure 1 shows the block diagram of the TMS, each pixel comprises a photoreceptor, a spatial derivative circuit, a temporal derivative circuit and circuits for the computation of velocity and direction of motion. A feature-map circuit combines these signals at each pixel location to encode the pixel saliency, and provides this signal in input to a global winner-take-all (WTA) circuit, which implements the attentional selection mechanism.

A. Analog core:

1) Photoreceptor circuit: The input of each pixel is the logarithmic adaptive photoreceptor proposed in [10]. It responds with low gain to static or slowly changing stimuli and with high gain to fast changing stimuli centered around the adaptation point. To minimize power-consumption, we implemented a self-biasing option [11] for the photoreceptor amplifier, such that the photoreceptor bias current is proportional to the total amount of photocurrent generated in the whole array. An optimal amount of power is used at each illumination level by scaling the amplifier bias current and bandwidth proportional to the input bandwidth.

2) Spatial derivative circuit: The spatial derivative is implemented by a “bump-antibump” circuit [12], that compares the photoreceptor outputs of two neighboring pixels and produces an output current with a bell shaped profile, which reaches its minimum value in correspondence of equal inputs, and increases for increasing difference (in absolute value).

3) Temporal derivative and edge detection circuit: The temporal circuits are those originally proposed in [13]: they comprise a hysteretic differentiating amplifier which converts rapid changes in the input voltage (the photoreceptor output) into current pulses with amplitudes proportional to the input signal’s temporal contrast (see Fig. 2).

4) Velocity circuits: To compute the velocity of detected contrast changes, we use a modified version of the “Facilitate and sample” (FS) circuit [13]. The temporal derivative pulse produced by a moving edge triggers the generation of both a sharp “fast” and a slowly decaying “slow” voltage pulse. The fast pulse of one pixel samples the slow decaying pulses of its neighbors. As the sampled voltage is proportional to the time the edge takes to travel across pixels, it is directly proportional to its velocity.

There are two main differences between the circuit used in the TMS and the original FS circuit. The first lies in a thresholding and signal restoration stage that we introduced between the temporal derivative output and the fast pulse generation circuit: a two-node winner-take-all (WTA) circuit [14] compares the temporal derivative output to a set threshold and eventually produces a normalized pulse which is optimally suited for driving the fast pulse generation circuit. The second improvement lies in the slow pulse generation circuit: to generate this pulse we are using a current-mode log-domain filter with tunable time-constant (the Differential Pair Integrator [15]) which produces a purely exponentially decaying signal. In its original conception, the FS produced a non-linear decay signal, using a diode-capacitor circuit [13], which had a very steep slope at short time intervals and a very shallow slope for longer intervals. This resulted in inhomogeneous sampling resolution for different speeds, while the exponentially decaying signal provides a uniform sampling resolution for a wide range of (tunable) speeds.

Each pixel computes left and right velocity, if an edge is traveling in one direction, one of the two will be significantly higher than the other, determining the estimated direction of motion. The comparison between the two sampled values is implemented by a three inputs current-mode WTA; the third input sets a minimum threshold used to drive the circuit when no moving edges are detected.

5) Feature maps and selective attention: The saliency value of each pixel is computed as a weighted sum of the currents corresponding to the measured temporal and spatial derivatives and velocity. It is performed by three current normalizer circuits [16], that set the diverse currents to a common range and allow to differently weight each contribution to the final saliency map. The attentional selection is implemented by a current-mode WTA and a self-inhibition circuit [15], needed for scanning the array in order of decreasing saliency [17]. Depending on the time constant of the self-inhibition an increasing number of pixels can be selected as attentional target.

B. Asynchronous logic I/O and analog readout

The position of the winning location is reported off-chip using both analog position-to-voltage circuits, as well as digital encoders. The selection of a winner in the competition stage is signaled off-chip via a digital address and two control signals that signal an error exception in case more than one pixel is active or no pixel is active at all, respectively.

The analog currents corresponding to the spatial and temporal derivatives and to the velocity measure are transformed into voltages by on-chip current-voltage converters. A digital asynchronous decoder is used to address the pixel from which...
we want to read the measurements. The addressing can be driven by an external microcontroller that can be programmed either to perform a sequential scanning, or to perform the attentional readout, by selecting as reading outputs the pixels selected by the WTA.

The TMS behavior can be tuned by means of 36 biases controlled by an on-chip programmable bias generator [18]. The programmable, precise and temperature independent properties of this on-chip bias generator are critical for application of the TMS in uncontrolled environments (e.g., such as outdoors) and for the use on compact autonomous systems.

III. TMS TEST SETUP AND CHARACTERIZATION

To test the TMS with real-world stimuli we mounted a lens onto it and projected a moving dark bar on a white background from a rotating drum with well-controlled rotational speed. We verified that the adaptive photoreceptor circuits behave as expected and measured the response of the following computational stages.

1) Spatial derivative circuits: Figure 3 shows the response of the adaptive photo-receptors (top) and the spatial derivative circuits (bottom), measured from seven neighboring pixels. During the time the bar takes to travel from one pixel to its neighbor, the spatial derivative is at its maximum (see negative peak in Fig. 3). Mismatch effects have been minimized in this stage, by careful layout; to quantify them we measured the response of seven pixels over ten repetitions of the same experiment. In the top panel of Fig. 4 the values of both minimum peak and steady state outputs of each pixel are shown; the two values are significantly different in all pixels; their difference, plotted in the lower panel of Fig. 4, is a measure of the input stimulus spatial derivative.

Figure 5 shows the spatial derivative measure of one pixel as a function of different speeds of the rotating drum; clockwise and counter-clockwise rotations can be clearly distinguished thanks to the asymmetric nature of the bump response.

2) Temporal derivative circuits: Figure 6 shows both photo-receptor response (top) and temporal derivative circuit response (bottom), for different values of photo-receptor adaptation bias. The photoreceptor data confirms the expected behaviour: a transient peak is followed by slow adaptation, the peak amplitude and the time course of adaptation can be tuned by means of the $V_{adap}$ bias. Conversely, the response of the (ON) temporal derivative (TmpDiff) circuit is not behaving as expected: the output peak of the temporal differentiator has an unexpected delay from the onset of the photoreceptor step. To investigate this issue we performed SPICE circuit simulations using as input waveforms the same ones measured from the photoreceptors. Figure 7 shows the measured and simulated output of the circuit in response to the measured photoreceptor waveform for $V_{adap} = 0V$ superimposed. As shown, the SPICE simulations don’t explain the delay in the TmpDiff response. We carried out further measurements using a wide range of bias settings: the TmpDiff circuit produces many spurious peaks mainly in response to noise and flicker. We found that these can be removed by increasing the circuit unity gain frequency. This can be achieved either by biasing the amplifier in sub-threshold or by decreasing the absolute value of its capacitors [13]. As changing bias settings in the TMS did not produce the desired outcome, we are investigating the effect of using smaller capacitor in the layout, and are considering changing design completely, with a radically different temporal derivative circuit (e.g., using the one proposed in [19]).
3) Velocity, feature-map, and selective attention circuits:
The unexpected output of the temporal difference circuit strongly affects the following processing stages. In light of this, we are currently in the process of testing the velocity and subsequent circuits. We are furthermore characterizing the analog and digital outputs, in order to determine what specifications are needed by conventional digital processing modules (e.g., micro-controllers) and develop a mixed analog/digital system optimally suited for mobile robotic applications.

IV. Conclusion

We presented a novel analog VLSI vision sensor which extends the capabilities of previously proposed designs by combining in one single chip spatial derivative measurements, temporal derivative measurements, velocity measurements, as well as selective attention processing features. We endowed the device with digital inputs and outputs in order to directly interface it to micro-controllers for mobile robotic applications. The first processing stages of the chip work as expected. However, an unusual behavior measured in the temporal derivative stages has delayed the application of this device to the mobile robotic tasks that it was designed for. We are currently in the process of “debugging” the analog VLSI circuits, and are planning a second revision of this device. We are confident however that the architectural and functional solutions adopted for this sensor are ideal for a wide range of application scenarios, irrespective of the details (and issues) present in the individual processing stages.

ACKNOWLEDGMENT

This work was supported by the EU eMorph grant #ICT-231467.

REFERENCES