Neuromorphic vision sensors and preprocessors in system applications

Jörg Kramer and Giacomo Indiveri

Institute of Neuroinformatics, Swiss Federal Institute of Technology/University of Zurich
Gloriastrasse 32, 8006 Zurich, Switzerland

ABSTRACT

A partial review of neuromorphic vision sensors that are suitable for use in autonomous systems is presented. Interfaces are being developed to multiplex the high-dimensional output signals of arrays of such sensors and to communicate them in standard formats to off-chip devices for higher-level processing, actuation, storage and display. Alternatively, on-chip processing stages may be implemented to extract sparse image parameters, thereby obviating the need for multiplexing. Autonomous robots are used to test neuromorphic vision chips in real-world environments and to explore the possibilities of data fusion from different sensing modalities. Examples of autonomous mobile systems that use neuromorphic vision chips for line tracking and optical flow matching are described.

Keywords: vision, motion, optical flow, focal-plane array, retina, neuromorphic, analog VLSI, tracking

1. INTRODUCTION

Neuromorphic vision sensors and preprocessors are increasingly being used to implement the first steps of visual processing in artificial systems. A typical application domain are autonomous mobile systems, for which size, power consumption and speed are important parameters. While until recently research in neuromorphic vision concentrated on designing and optimizing individual circuits, part of the effort is now shifting towards multichip neuromorphic systems and hybrid systems interfacing neuromorphic circuits with other types of processors.

Early visual processing in biological systems relies mainly on neurons arranged in a retinotopic manner with local interconnections. Such structures lend themselves nicely to implementation with integrated circuits whose essentially two-dimensional structure forbids high-density long-range connectivities. Analog circuits fabricated with CMOS technology and operated in the subthreshold domain have proven to be a suitable medium for efficient emulation of neural circuits,[1–3] because of similar computational primitives, a close relation of physical structure and function, and low power consumption, thus allowing for high computational density and massive parallelism. Another important common feature is the asynchronous operation of the circuits, which is preferable to clocked operation for sensory processing, given the continuous nature of sensory signals. Clocked systems introduce temporal aliasing artifacts that can significantly compromise the time-sensitive computations performed in real-time sensory processing systems. The limited precision and matching of biological and subthreshold analog components is adequate for processing noisy sensory data and sufficient for performing a variety of perceptual tasks requiring collective computation of neural ensembles rather than calculation of sparse values with high precision.

2. VISION SENSORS

2.1. Retinal Image Sensors

Unlike conventional image sensors neuromorphic image sensors do not merely compute a linear function of image brightness at the location of each sensing element (pixel), but their processing typically involves a nonlinear mapping function, local spatiotemporal filtering and adaptation.

A very useful pixel for a variety of neuromorphic vision circuits is the adaptive photoreceptor by Delbrück and Mead[4] shown in Fig. 1(a). It consists of a photodiode in series with a transistor in source-follower configuration and a feedback loop with a high-gain amplifier and an adaptive gain stage. The photoreceptor logarithmically maps

Other author information:
WWW: http://www.ini.unizh.ch/; Fax: ++41-1-6344983
J.K.(correspondence): E-mail: kramer@ini.phys.ethz.ch; Telephone: ++41-1-6342663
G.I.: E-mail: giacomo@ini.phys.ethz.ch; Telephone: ++41-1-6342680

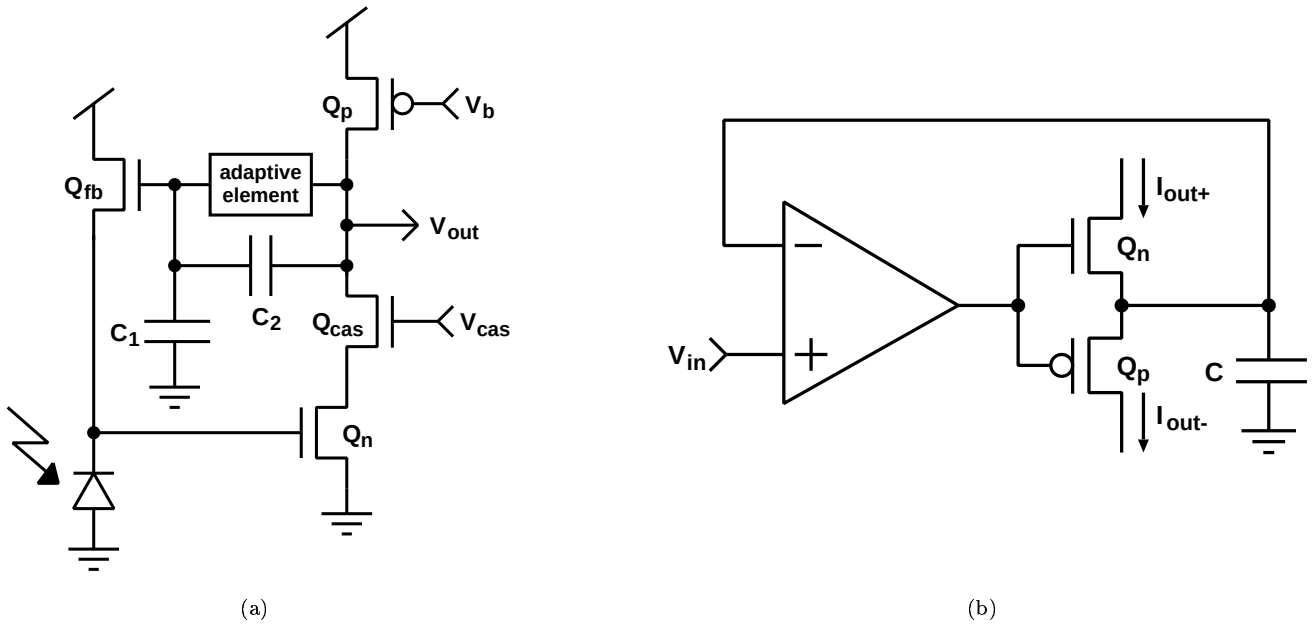


Figure 1. (a) Adaptive photoreceptor circuit with a large dynamic range for contrast encoding. (b) Temporal differentiator circuit with thresholding and rectification.

irradiance onto its output voltage. This compressive transfer function results in an output voltage change of about 80 mV for each decade in background illumination change and in a dynamic operation range of more than six orders of magnitude of background illumination without any saturation effects. Furthermore, the logarithmic characteristic encodes the logarithm of contrasts in output voltage differences, which are thus independent of background illumination. The transient response of the photoreceptor is linearly amplified with respect to the DC response and the output voltage adapts slowly towards its DC value. The transient response shows a dynamic range of one or two decades, depending on the chosen gain. The adaptive photoreceptor exhibits the characteristics of a temporal bandpass filter, where the high-frequency cut-off either depends on the irradiance or on a bias voltage (labeled V_b in Fig. 1(a)) and the low frequency cut-off is determined by the adaptation rate.

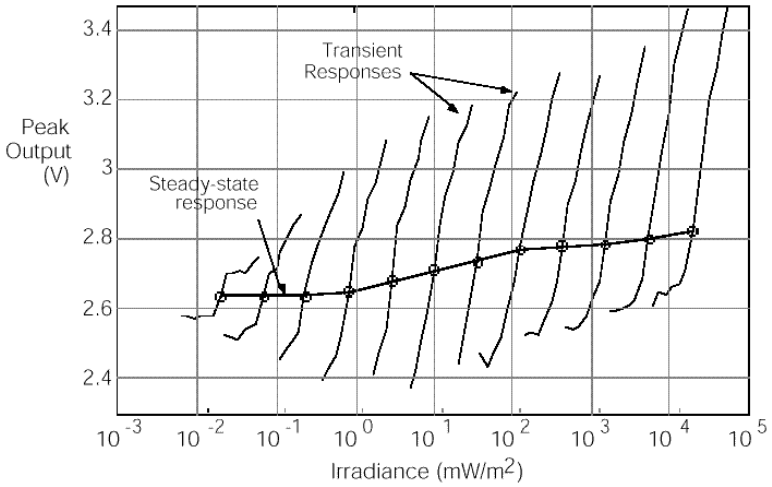


Figure 2. Transfer characteristics of the adaptive photoreceptor circuit of Fig. 1(a) (from Delbrück & Mead[5]).

The transfer characteristics of an implementation of the adaptive photoreceptor are presented in Fig. 2. They resemble those of certain biological photoreceptors (*e.g.* turtle cones). The strong response to local changes in the image and the invariance against the DC light level make the adaptive photoreceptor a well-suited transducer for systems concerned with perceptual tasks in natural environments, where the signal is generated by the different reflectivities of the imaged object surfaces and varies only over a relatively small range, but its scaling due to global lighting conditions ranging from starlight to sunlight varies over a much larger range, but does not contain a large amount of information.

For applications requiring either quantitative information on temporal changes or high-resolution timing response, such as motion sensing, a temporal differentiator or highpass filter may be added to the adaptive photoreceptor to further enhance its response to transient features. Fig. 1(b) shows a temporal differentiator circuit that computes the rectified first temporal derivative of a voltage signal and encodes the result in two currents (I_{out+} resp. I_{out-}) for the two signs of the derivative.[6] This circuit also exhibits a built-in threshold that may be useful for noise suppression in sensory processing.

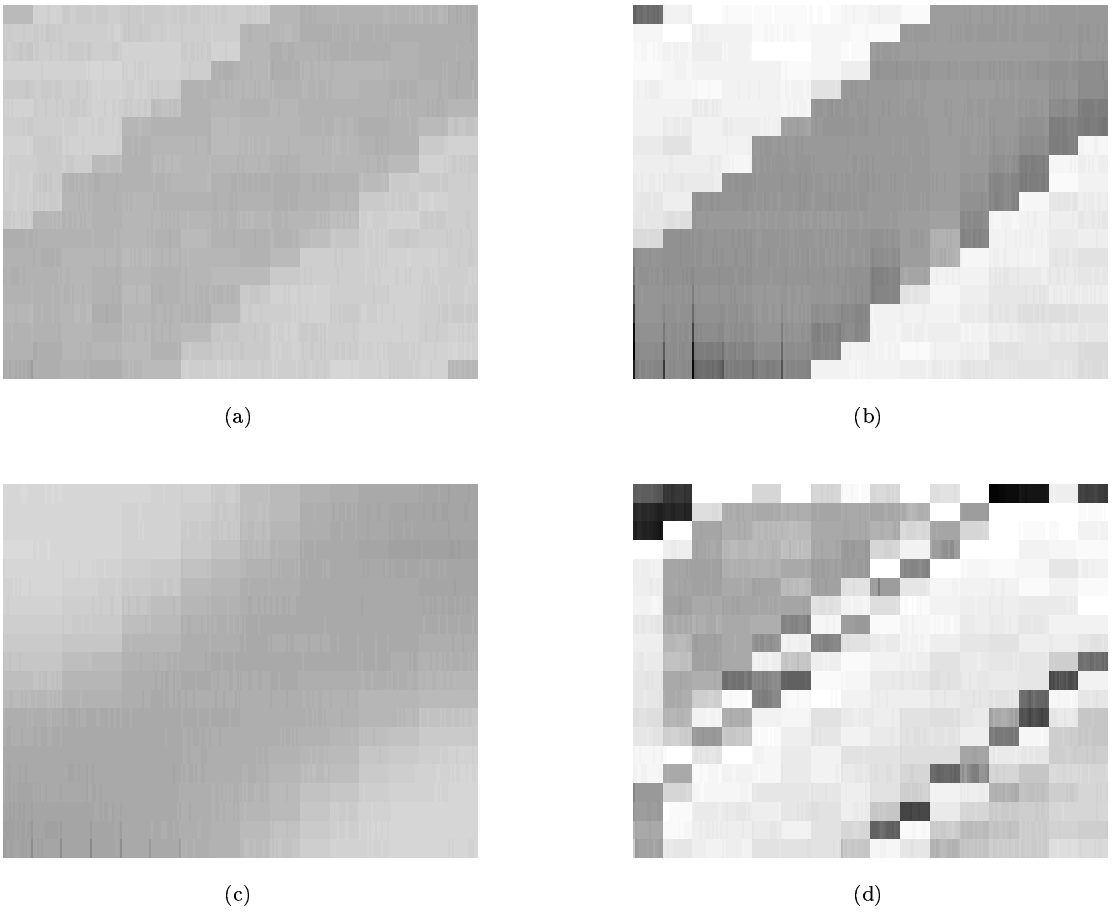


Figure 3. Response of an integrated artificial retina to the image of a black stripe on a white background. (a) Adapted response to stationary image without spatial interaction. (b) Transient response to moving image without spatial interaction. (c) Transient response to moving image for excitatory coupling. (d) Transient response to moving image for inhibitory coupling.

Local spatial filtering in imaging arrays may be obtained by interconnecting the individual pixels with resistive networks mediating excitatory and inhibitory interactions. Excitatory interaction results in spatial lowpass-filtering (smoothing) of the signal, which is a useful operation for suppression of noise and spatial aliasing effects. Inhibitory

interaction implements spatial highpass-filtering that approximates a second spatial derivative and thus enhances edges. This is a very useful function, given that edge extraction is the first step in most image processing systems and is computationally expensive if performed on a serial processor. In Sect. 5.1 we shall see an example of how the capabilities of a fully parallel, real-time, low-power edge extractor based on a neuromorphic spatial-filtering array may be exploited in autonomous mobile systems. The combination of excitatory and inhibitory interactions implements a spatial bandpass filter, which together with the temporal bandpass characteristics of an adaptive photoreceptor provides the functionality of laminar cells in the invertebrate retina and of the outer plexiform layer of the vertebrate retina.

Several attempts at building artificial retinas with electronic hardware have been made within the past thirty years, showing a development from implementations with discrete components[7] over first integrated versions[8] to high-density arrays[9] with resolutions of up to 48,000 pixels.[10] A more recent approach[11] couples an array of the adaptive photoreceptors described above directly with excitatory and inhibitory networks. This implementation makes use of the large transient gain of the adaptive photoreceptor to enhance temporal features and to improve the ratio of signal to fixed-pattern noise. We built a modified version of this retina with pixels arranged on a hexagonal grid and each pixel forming an excitatory and an inhibitory connection with each of its six nearest neighbors. The connection strengths, temporal adaptation rate and rise time of the photoreceptor may be set by biases, allowing us to control the passband of the spatiotemporal filter. The circuit was implemented with 1.2 μm CMOS technology as a test array of 21 $\times$ 16 pixels with a pixel size of 120 μm $\times$ 150 μm. Fig. 3 shows the response of the retina to a black bar on a white background for different adaptation states and bias conditions. In Fig. 3(d), which shows the effect of the inhibitory network, the positive and negative overshoots emphasizing sharp edges can clearly be seen. Such overshoots are also observed in human psychophysics and are known there as *Mach bands*. Another psychophysical effect that can readily be observed on adaptive retinal circuits is an afterimage or contrast-reversal appearing after adaptation to a certain stimulus and sudden removal of that stimulus. Interposition of neutral density filters between the stimulus and the imaging lens showed that the retina responds with a high transient gain across more than seven decades of background illumination, corresponding to a relative illumination range from starlight to sunlight. The response is substantially invariant to background illumination in this range, except that towards lower illumination levels the background gray level shifts to slightly smaller values and the adaptation time increases. The response time constant also increases due to lowpass filtering of the photoreceptor, whose bandwidth is proportional to the photocurrent, *i.e.* to image brightness. A larger version of this retina with 125 $\times$ 94 pixels has been submitted for fabrication. Using state-of-the-art silicon processing technology artificial retinas with resolutions approaching video standards could be implemented on 1/2" imagers.

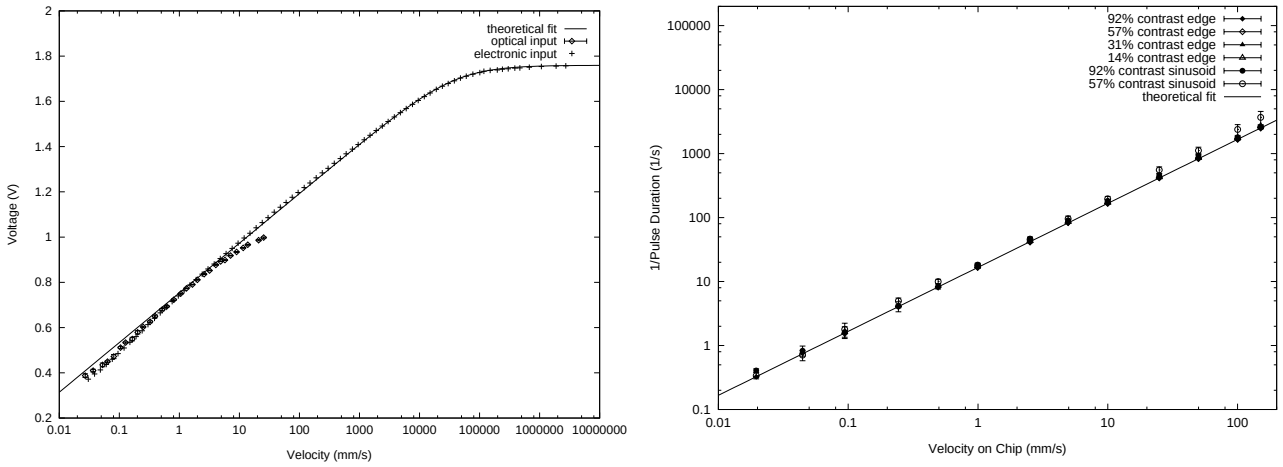
Orientation-selectivity can easily be implemented at the retinal level without increasing the pixel size significantly, by using different connection strengths for the excitatory and inhibitory networks along different directions on the retina.

2.2. Motion Sensors

A variety of insect species have been shown to perform local motion processing very early in their visual pathway, and directionally selective cells have been found in the retina of some higher animals. Especially for autonomous systems moving in a predominantly stationary environment navigation tasks may be significantly facilitated by estimating *ego*-motion from relative motion measurements. In Sect. 5.2 it will be demonstrated how a robot can navigate through its environment purely based on a low number of relative velocity estimates.

We developed different compact elementary velocity sensors for estimating local direction of motion and velocity at each sensor location. The estimates are based on time-of-travel measurements of extracted features in the focal plane. As feature extractors nonlinear temporal or spatial edge detectors as described in Sect. 2.1 have been used. The response characteristics of two types of such velocity sensors, called *FS sensor*[12] and *FTI sensor*[13] respectively, are shown in Fig. 4. Both sensors have separate output terminals for the different directions of each motion component. The FS sensor encodes velocity logarithmically in its output voltage. As for the adaptive photoreceptor described in Sect. 2.1 the logarithmic compression allows to encode a large dynamic range within a limited output voltage range. The mapping encodes the logarithm of *velocity contrast, i.e.* velocity ratios, in output voltage differences. This property and has been exploited in a system application presented in Sect. 5.2.

The FTI sensor encodes velocity inversely in the duration of asynchronous binary output pulses. This representation is suitable for implementing pulse-based communication and processing, which is encountered in most parts



(a) (b)

Figure 4. Response characteristics of motion-sensing elements. (a) FS motion-sensing element. Response to optical and electrical stimuli. (b) FTI motion-sensing element. Response to optical stimuli of different types and contrasts.

of any neural systems. Both sensors encode velocity over more than three orders of magnitude and are invariant to contrast over a significant contrast range.

Another implementation[12,14] of a velocity-sensing element showing an inverse relationship between velocity and durations of output pulses compares a pulse duration representing the measured velocity to a pulse duration representing an adapted average velocity, in a similar way as the adaptive photoreceptor of Sect. 2.1 compares the measured irradiance value to an adapted value. An optical illusion observed in psychophysics and also exhibited by this sensor, which corresponds to the afterimage effect in the adaptive photoreceptor is the *motion aftereffect* or "waterfall illusion", where after adaptation to a certain velocity a stationary image is perceived to move at the same speed in the opposite direction. Motion adaptation can be useful, for example, to report changes in *ego*-motion of systems moving in stationary environments.

Local motion-sensing elements can only provide reasonable velocity estimates for motion in one dimension. For tasks involving relative motion estimation of objects moving on a plane, 1-D motion-sensing is often sufficient. In order to estimate image motion in the 2-D focal plane, local velocity sensors may be used to compute motion parameters for each dimension. Combining these motion parameters is not sufficient to make reasonable estimates of object motion if the brightness gradient is approximately constant within the receptive field of the motion sensors. The best estimate that can be done in the absence of any further information and more global connectivity is that of the velocity component along the brightness gradient, called *normal optical flow*. The normal optical flow gives a lower bound for the speed, but no upper bound, and its direction may deviate from the direction of motion by up to $\pm 90^o$. At least two measurements of the motion parameters for the two dimensions from regions with different directions of the brightness gradient, but moving at the same velocity, are required to uniquely determine the velocity. A motion-sensing scheme that integrates measurements of motion parameters across an entire array has been implemented to estimate 2-D global velocity.[15] Such a scheme may be useful for estimation of *ego*-motion, but does not provide any spatial resolution. A modified version of this algorithm is being developed at our institute.[16] The measured motion parameters are only integrated across small image regions, such that estimates of 2-D velocities with spatial resolution are possible on a single chip.

3. COMMUNICATION INTERFACES FOR 1-D OR 2-D ARRAYS

The vision sensors described in Sect. 2 are compact enough to be monolithically integrated as 1-D or 2-D arrays in the focal plane. But especially elements performing higher-level processing, such as motion sensing, are relatively large and thus have a low density and a small fill factor of the photosensitive area if arranged in 2-D arrays. This results in low resolution and susceptibility to spatial aliasing. It is thus desirable to integrate different processing layers on separate substrates. For example, feature-based motion-sensing arrays may have the motion-computation stage separated from the feature-extraction stage.

But this confronts us with the general technological problem of how to communicate the large data flow from integrated parallel processing arrays to different devices. The best solution would be to interconnect with a wire each position of the array with every position on every other device it may want to communicate with, as is typically done for intrachip communication. Unfortunately, a conventional integrated circuit only provides a limited number of communication wires along its periphery, which are easily outnumbered even by the elements of medium-density arrays. In addition, there is a scaling problem for two-dimensional arrays, since the periphery is essentially one-dimensional.

Attempts at solving this problem include three-dimensional circuit integration techniques and communication via optical interconnects. Unfortunately, neither of these approaches has yet been developed for mass production and they are thus still expensive. Furthermore, it will be difficult to achieve a substantial non-local fanout with such schemes.

The most successful strategy to deal with the communication problem is temporal multiplexing, where different signal channels share a communication wire. Since such a wire has to transmit the combined data flow from the different channels its bandwidth has to be of the order of the sum of the signal bandwidths of these channels. Implementations of neuromorphic systems with standard silicon processing technology allow multiplexing of roughly 10^4 signal channels on a single wire, assuming a desired temporal resolution of about 1 ms and a maximum bandwidth of 10 MHz for the communication system.

The commercially best-established multiplexers are scanners that read values from the different channels in a fixed regular sequence. The main advantages of scanners are their simplicity and that they are used as interfaces to most standard devices for storage and display. Neuromorphic arrays often contain free-running on-chip scanners using a single clock input and generate a signal that can be directly displayed on a multisync monitor using one or two additional transistors for impedance conversion and amplification.[17] However, most display and storage devices use the PAL or NTSC video standards. Such devices include video monitors, video cassette recorders and most frame grabbers. For generating a video signal, the same on-chip scanners can be used to read out the signal, but timing and signal levels are much more critical than for multisync monitors. The conversion of the signal levels from the chip into video format can be done with off-the-shelf components, but they usually require an input signal with proper horizontal and vertical timing and a composite video synchronization signal. The latter may be easily generated on chip from the horizontal and vertical synchronization signals provided by the scanners, but the tolerance on the video timing is very low and the matching of horizontal and vertical timing is critical and thus not easily achieved with free-running on-chip scanners. In order to get around this difficulty we are developing an interface that generates a standard video signal from the scanned output of neuromorphic arrays of arbitrary sizes. The interface will consist of a custom-built timing-controller chip generating the horizontal and vertical timing for the on-chip scanners and a commercially available RGB-to-video encoder. On-chip scanner logic providing the necessary interface to the timing-controller chip has already been developed and is implemented on our retina chips. The logic also provides an input control bit selecting between the timing-controlled mode and the free-running mode.

For arrays with unevenly distributed activity patterns scanning is an inefficient way of multiplexing, since bandwidth is wasted reading inactive pixels. In traditional imaging arrays most pixels are active most of the time, but retinomorphic vision circuits that report only changes in the image use coding schemes that are much more efficient for images of predominantly static environments. For such circuits it is worth investigating other multiplexing schemes. Most standard protocols for communication between neuromorphic chips are asynchronous and event-based, *i.e.* no clocks are required and no bandwidth is wasted on inactive pixels. The events from different locations in an array are sent as source or destination addresses via a common bus. The main difference between such event-address protocols is the way they handle synchronous signals from different locations. The address bus may either be controlled by an arbiter that buffers events if necessary to avoid collisions[18] or colliding events may just get discarded.[19] Which of these schemes is preferable depends on the nature of the signals and on the application. For systems with multiple

senders and receivers some form of arbitration seems desirable. In our institute arbitrated neuromorphic multichip systems are being developed, as well as two-chip sender-receiver systems *e.g.* for feature-based motion computation.

An approach to avoid the communication problem altogether is on-chip computation of sparse higher-level image features that can be represented by a low number of signals not exceeding the number of available output pins. In the past, we have built several such systems for computing focus of expansion,[20] time to contact[21] and motion discontinuities.[22] A new system that uses such an on-chip data reduction strategy is described in the following Section.

4. INTEGRATED OPTICAL TRACKING SYSTEM

Selecting the spatial position of a feature of interest from highly dimensional input signals such as visual images is a nonlinear operation that, in one shot, eliminates the problem of communicating high-dimensional signals to further processing stages and provides a basis for tracking operations. Tracking features as they move in the environment is a relatively complex task that researchers are still trying to cope with, both in the engineering community[23] and in the biological one.[24] Within this framework, we have designed a compact single-chip system able to measure the horizontal position of high-contrast spatial edges in real time.

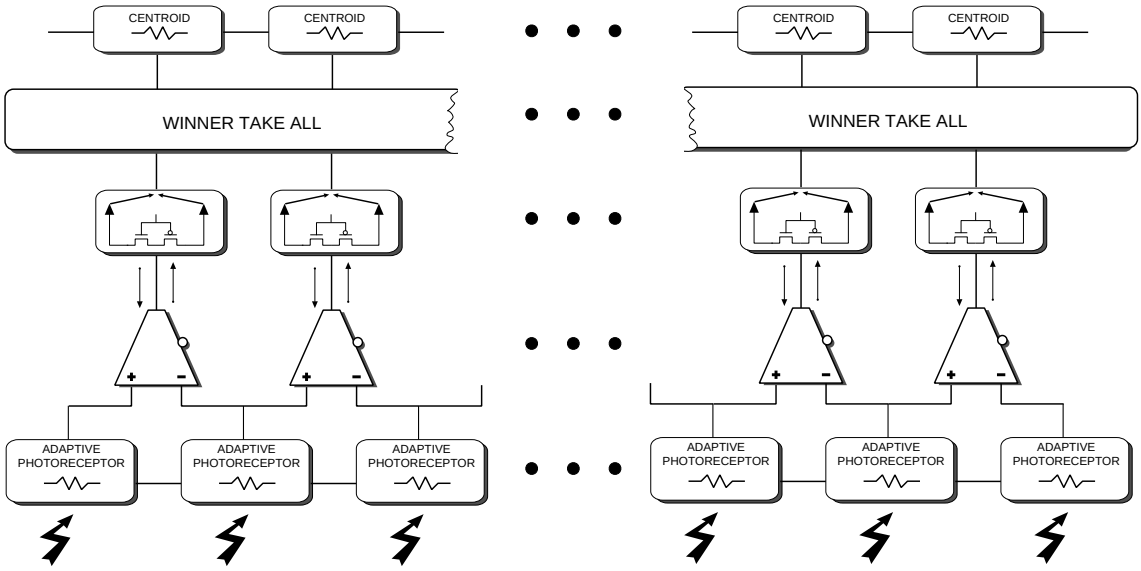


Figure 5. Block diagram of single-chip tracking system. Spatial edges are detected at the first computational stages by adaptive photoreceptors connected to transconductance amplifiers. The edge with strongest contrast is selected by a winner-take-all network and its position is encoded with a single continuous analog voltage by a *centroid* circuit (see text for details).

The chip is structured in a hierarchical way, in which image brightness data coming from the bottom input stage is processed in parallel. A block diagram of the system is depicted in Fig. 5. The first stage is a one-dimensional array of adaptive photoreceptors with lateral excitatory coupling, to implement spatial smoothing (see also Sect. 2.1). The second stage is composed of an array of simple transconductance amplifiers[25] which receive input from neighboring photoreceptors and implement a spatial derivative operation. The output current of the transconductance amplifier is in fact linearly proportional to the difference of the input voltages for small differences and saturates smoothly for large signal differences. By looking at the sign of this output current we also have information on the polarity of the detected edge. If the output current is positive it corresponds to an ON edge (dark-to-bright transition), and if it is negative it corresponds to an OFF edge (bright-to-dark transition). Furthermore, modulating the control parameter which determines the amount of lateral excitatory coupling between adaptive photoreceptors, it is possible to electronically tune the sensor's spatial-frequency selectivity.

At the third computational stage the current polarity of the transconductance amplifiers is measured.[23] This current can be gated so that the sensor is able to selectively respond only to ON edges, only to OFF edges or to both types. A schematic diagram of the circuit which characterizes this (and the preceding) computational stage is shown in Fig. 6.

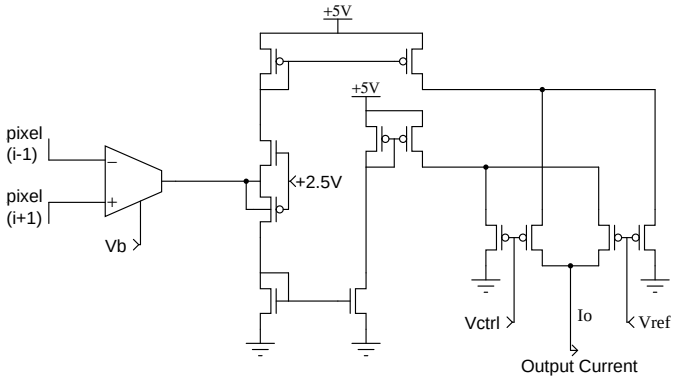


Figure 6. Circuit diagram of current-polarity detector. The transconductance amplifier (in the left part of the image) provides in output a current which is a sigmoidal function of its differential input voltage. Positive currents being "sourced" from the amplifier are gated to the n-type current mirror (in the bottom part of the image). Negative currents being "sunk" by the amplifier are obtained from the p-type current mirror in the very top part of the figure. Both types of currents are then conveyed to the output node Io. Depending on the values of the control voltage signals $Vctrl$ and $Vref$, both current polarities are allowed to flow or only one of the two polarities is allowed to flow (or, in a not very useful case, neither of the two types of current are allowed to flow).

Figure 7 shows the output of the array of adaptive photoreceptors together with the output of the polarity-selective spatial derivative circuit. The two lower traces representing the spatial derivative of the input stimulus were taken at different instants in time, with control values such that the sensor was selective only to ON edges at first and only to OFF edges later.

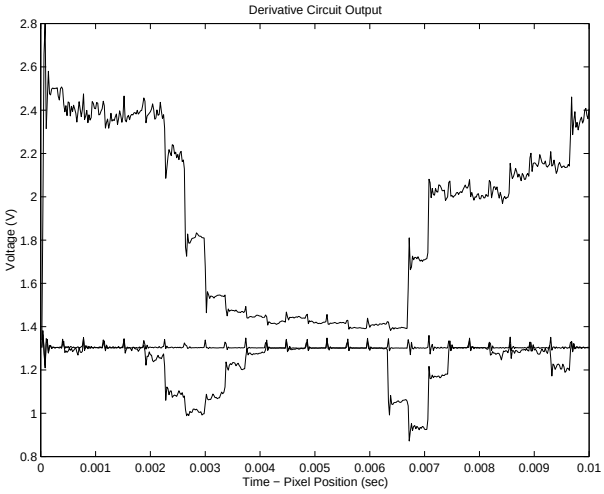
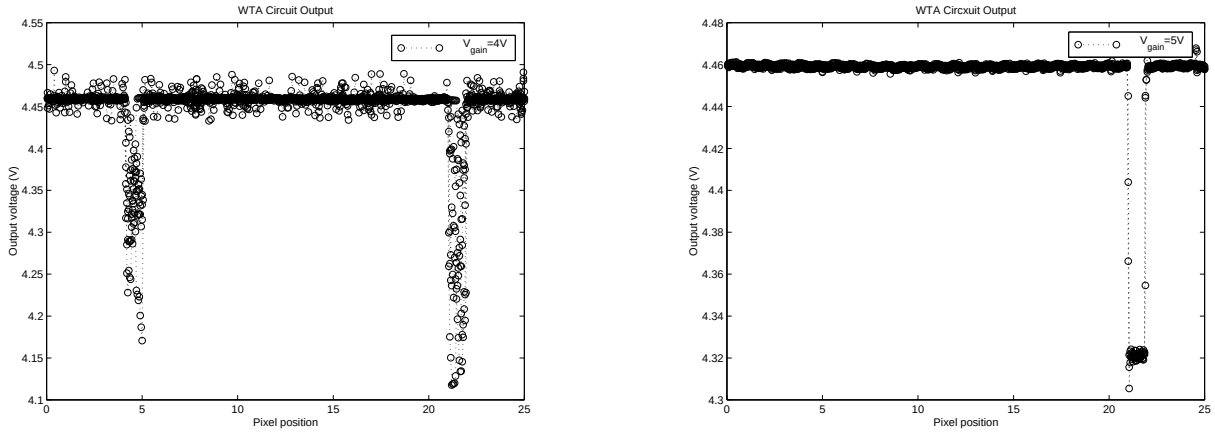


Figure 7. Response of the array of adaptive photoreceptors to a black bar on a white background (upper trace) and output traces of the spatial derivative circuit(lower traces).

The fourth computational stage uses a recently proposed type of winner-take-all (WTA) network[26] to implement the highly nonlinear operation of feature selection. This circuit contains as many elements as there are pixels. The outputs of all of the WTA elements are null, except for the element receiving the strongest input. The interesting

feature of this circuit lies in its recursive interactions with its input stage: a positive feedback loop reinforces the choice of the winning node by adding a set bias current to the input of the winning node and by spreading it to the winner's neighbors. This operation ensures that the circuit does not flicker between edges with similar strength in the presence of noise (see Fig. 8).



(a) (b)

Figure 8. (a) Traces of multiple responses of the WTA network to a black bar on a white background, with the positive feedback loop deactivated. As the ON and OFF edges have very similar contrast, the circuit oscillates in selecting both edge polarities; (b) Traces of multiple responses of the WTA network with active positive feedback loop to the same black bar. Once a decision is made on which WTA element encodes the winner, the output does not oscillate and the network is able to faithfully track the position of the selected edge.

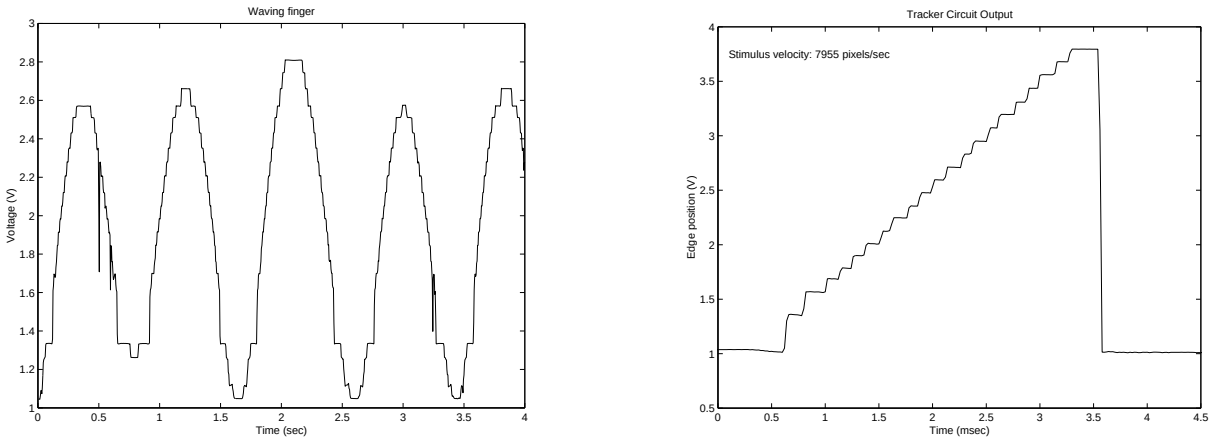
Once the circuit selects a winner, it will tend to lock onto it (due to the hysteresis induced by the positive feedback loop). As the hysteretic component of the winner's current, summed through the positive feedback pathway, is also spread laterally, nearest neighbors of the winning element will be favored in the selection of the new winner, whereas elements in the periphery will be inhibited. As a consequence, the system will track reliably moving edges, while being indifferent to possible "distractors" that may be present in the periphery of the spatial stimulus position.

Finally, at the last computational stage there is a resistive layer connected to a network of "centroid circuits", as proposed by DeWeerth,[27] that allows the system to encode the position of the WTA network with a single output voltage.

The overall system is able to detect edges present in the visual scene in a wide variety of illumination conditions. The adaptive photoreceptors at the input stage allow the system to perform in conditions that range from dim artificial room illumination to bright natural sun light. Figure 9 shows the output of the system for a low-contrast real-world type of stimulus and for a high-contrast one.

5. AUTONOMOUS MOBILE SYSTEMS

To exploit the advantages of neuromorphic vision sensors, we chose to use them in conjunction with a mobile robot for solving vehicle-guidance and autonomous navigation tasks. These types of tasks in fact require compact and power-efficient computing devices. Such devices should moreover be robust to noise, tolerant to adverse conditions induced by the motion of the system (*e.g.* to jitter and camera calibration problems) and possibly able to adapt to the highly variable properties of the world. Although traditional digital technology provides a fine infrastructure for such systems, it is not suited to the computation density and speed particularly encountered in sensory and motor processing.



(a) (b)

Figure 9. (a) Output of the system in response to a waving finger moving in front of the chip; (b) Output of the system in response to a fast-moving dark bar on a stationary light background.

The mobile robot we used in our experiments is *Koala* (K-Team, Lausanne). It measures 32 cm in length, 31 cm in width and is 11 cm high. It has an on-board Motorola 68331 processor, 1 MByte of RAM, and two to three hours of autonomous operation from its battery.

5.1. Line Tracking

As a first experiment, we interfaced the integrated optical tracking device to Koala for solving line-following tasks in realistic environments (as opposed to very well controlled laboratory environments). The chip was mounted onto a wire-wrap board together with a 4 mm lens with an f-number of 1.2, and it was attached to the front of Koala with the lens tilted towards ground at an angle of approximately 60°, in a way to image onto the retinal plane the features present on the floor approximately 10 cm ahead (see Fig. 10(a)). The output voltage of the tracking chip was directly applied to one of the analog inputs of the robot and digitized at 8 bit resolution. To implement the line-following task Koala used a very simple control algorithm with the following structure:

1. Read the tracking chip's output.
2. Shift and rescale the measured value so that the variable encoding the edge position is zero when the measured edge is in the center of the chip's visual field.
3. Set the forward component of the speed to a value which is weighted by a Gaussian function of the edge position: the speed will be maximum when the edge is in the center of the visual field and it will decay to a minimum value as the edge shifts to the sides.
4. Execute motor command sending forward and rotational components of speed to the robot's motors.

The algorithm basically applies the (rescaled) output of the sensor directly to the motors. We make almost no use of the on-board CPU's processing power (leaving it free for other processes, such as learning). The computationally expensive part of the processing (involving visual preprocessing and target selection) is done in real-time by the neuromorphic sensor.

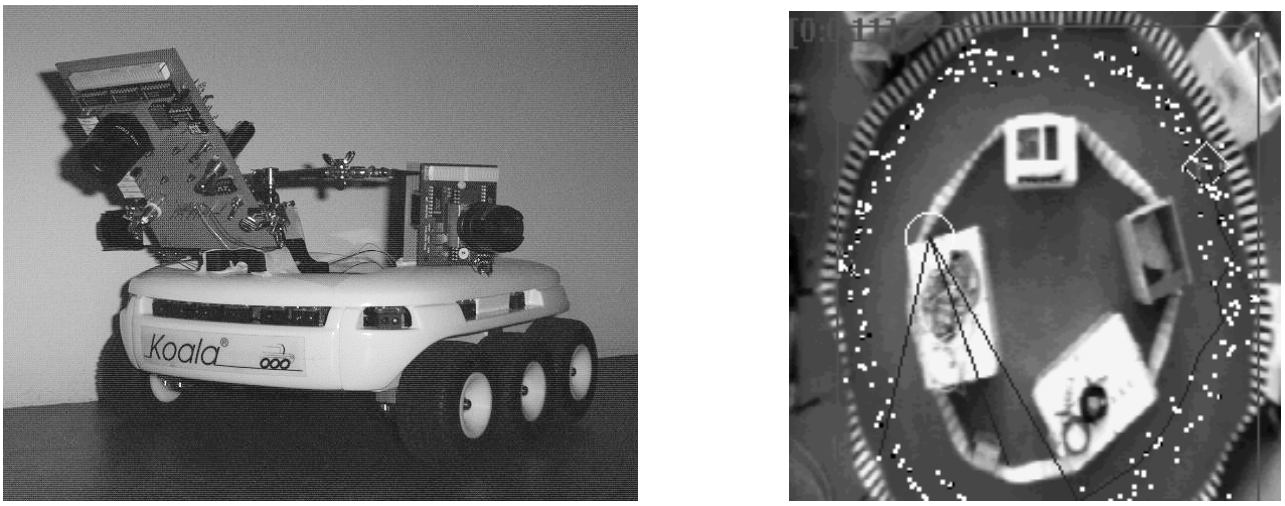
Using this simple control algorithm, the robot is able to reliably track black cables randomly layed out on the floor, for a wide variety of conditions (*e.g.* floors with different texture, cables of different sizes, extreme illumination conditions, etc.) For a quantitative analysis of the tracking performance of a similar system using a current-mode silicon retina as a visual preprocessor see Indiveri and Verschure, 1997.[28]

5.2. Optical Flow Matching

A second task was based on an experiment designed to study motion processing in honeybees.[29] The outcome of the experiment suggested that honeybees are able to fly along the center of a tunnel by balancing the velocities of the images of the two side walls in their eyes. Using two arrays of FS motion-sensing elements (*cf.* Sect. 2.2) with the outputs averaged across each array, that were mounted on the Koala robot, we implemented a simple optical flow matching algorithm to enable the robot to navigate through a corridor. The two FS motion-sensing arrays were placed onto either side of the robot with their optical axes set at equal and opposite angles with respect to the robot's forward direction, as seen in Fig. 10(a).

Given the logarithmic mapping of measured velocity to output voltage of the FS sensor, the difference between the output voltages of the two sensors is proportional to the logarithm of the ratio of the measured velocities, which in turn is inversely proportional to the ratio of the distances from the corridor boundaries and independent of speed and corridor width. We used this difference to correct the robot's trajectory. For boundaries with high-contrast features and low curvatures, a linear function of the difference turned out to be an appropriate choice. Lower-contrast and higher-curvature boundaries required superlinear correction functions.

Fig. 10(b) shows the trajectory of the robot in a circular corridor with high-contrast gratings printed onto the corridor walls. The correction function used in this particular experiment was quadratic. The robot covered several laps in either direction with the same parameter settings.



(a) (b)

Figure 10. (a) Koala robot equipped with neuromorphic vision sensors for line tracking and optical flow matching. The sensors pointing sideways are used for optical flow matching experiments, while the sensor in front is used for line tracking tasks. (b) Optical flow matching experiment: Trajectory of the robot recorded for several laps in both directions using a camera suspended from the ceiling. The bright dots indicate robot positions recorded at regular time intervals.

The system was able to navigate through corridors with boundaries exhibiting dense high-contrast features, using an extremely simple control algorithm. As for the line tracking task, only a small fraction of the robot's computing resources were used. The robotic platform has built-in proximity sensors, the capability of receiving several other sensory inputs and the possibility of executing processes in parallel. For example, we simultaneously implemented the line tracking and optical flow matching algorithms in a competitive fashion. This represents a first step towards investigation of data-fusion strategies which, together with learning, will be necessary to perform more sophisticated tasks in less well-controlled environments.

6. CONCLUSIONS

We described a variety of neuromorphic sensors suitable for transducing and preprocessing visual information in system applications. Different strategies for communicating high-dimensional signals from such sensors to other neuromorphic circuits, to machine-vision systems and to standard devices for data storage and display were outlined. We presented a single-chip neuromorphic system with additional on-chip processing stages to extract higher-level features and thus reduce the dimensionality of the output signal. Advantages of neuromorphic vision systems were evidenced in two application examples involving a mobile robot: a line tracking system and an optical flow matching system.

The advantages of neuromorphic vision sensors, namely low power consumption, compactness and real-time response, allowed us to integrate different sensory modalities on a small mobile platform and increase the complexity of its behavioral responses in its interactions with the environment.

ACKNOWLEDGMENTS

Fuyuki Okamoto designed and implemented the scanner logic for the timing-controller interface. Paul Verschure helped to set up the optical flow matching experiment. Participants in the NSF Neuromorphic Workshop at Telluride provided many helpful discussions. This work was supported by the Swiss National Science Foundation SPP Grant and the U.S. Office of Naval Research. Fabrication of the integrated circuits was provided by MOSIS.

REFERENCES

1. C. Mead, "Neuromorphic electronic systems," *Proc. IEEE* **78**, pp. 1629–1636, 1990.
2. X. Arreguit and E. A. Vittoz, "Perception systems implemented in analog VLSI for real-time applications," in *PerAc'94 Conference: From Perception to Action, Lausanne, Switzerland*, 1994.
3. R. Douglas, M. Mahowald, and C. Mead, "Neuromorphic analogue VLSI," *Ann. Rev. Neurosci.* **18**, pp. 255–281, 1995.
4. T. Delbruck and C. A. Mead, "Photoreceptor circuit with wide dynamic range," in *Proc. 1994 IEEE Int. Symp. Circuits and Systems, London*, IEEE, (Piscataway, New Jersey), 1994.
5. T. Delbruck and C. A. Mead, "Analog VLSI phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits," tech. rep., California Institute of Technology, 1996. CNS Memo No. 30.
6. J. Kramer, R. Sarpeshkar, and C. Koch, "An analog VLSI velocity sensor," in *Proc. 1995 IEEE Int. Symp. Circuits and Systems, Seattle, Washington*, pp. 413–416, IEEE, (Piscataway, New Jersey), 1995.
7. K. Fukushima, Y. Yamaguchi, M. Yasuda, and S. Nagata, "An electronic model of the retina," *Proc. IEEE* **58**, pp. 1950–1951, 1970.
8. C. A. Mead and M. A. Mahowald, "A silicon model of early visual processing," *Neural Networks* **1**, pp. 91–97, 1988.
9. K. A. Boahen and A. G. Andreou, "A contrast sensitive silicon retina with reciprocal synapses," in *Advances in Neural Information Processing Systems*, J. E. Moody, S. J. Hanson, and R. P. Lippmann, eds., vol. 4, pp. 764–772, Morgan Kaufmann, (San Mateo, California), 1992.
10. A. Andreou and K. Boahen, "A 48,000 pixel, 590,000 transistor silicon retina in current-mode subthreshold cmos," in *Proc. 37th Midwest Symposium on Circuits and Systems, Lafayette, Louisiana*, pp. 97–102, 1994.
11. S. Liu and K. Boahen, "Adaptive retina with center-surround receptive field," in *Advances in Neural Information Processing Systems*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, eds., vol. 8, pp. 678–684, The MIT Press, (Cambridge, Massachusetts), 1996.
12. J. Kramer, R. Sarpeshkar, and C. Koch, "Pulse-based analog VLSI velocity sensors," *IEEE Trans. Circuits and Systems II* **44**, pp. 86–101, 1997.
13. J. Kramer, "Compact integrated motion sensor with three-pixel interaction," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**, pp. 455–460, 1996.
14. J. Kramer, G. Indiveri, and C. Koch, "Analog VLSI motion projects at Caltech," in *Advanced Focal Plane Arrays and Electronic Cameras*, T. M. Bernard, ed., *Proc. SPIE* **2950**, pp. 50–63, SPIE, (Bellingham, Washington), 1996.
15. J. Tanner and C. Mead, "An integrated analog optical motion sensor," in *VLSI Signal Processing, II*, S.-Y. Kung, R. E. Owen, and J. G. Nash, eds., pp. 59–76, IEEE, New York, 1986.

16. A. Stocker, "Smooth optical flow computation by a constraint solving circuit," *INNS/ENNS/JNNS Newsletter* (21), pp. 5–6, 1998. Appearing with Vol. 11, No. 1 of Neural Networks.
17. C. A. Mead and T. Delbrück, "Scanners for visualizing activity of analog VLSI circuitry," *Analog Integrated Circuits and Signal Processing* **1**, pp. 93–106, 1991.
18. J. Lazzaro, J. Wawrzynek, M. Mahowald, M. Sivilotti, and D. Gillespie, "Silicon auditory processors as computer peripherals," *IEEE Trans. Neural Networks* **4**, pp. 523–528, 1993.
19. A. Mortara, E. A. Vittoz, and P. Venier, "A communication scheme for analog VLSI perceptive systems," *IEEE J. Solid-State Circuits* **30**, pp. 660–669, 1995.
20. G. Indiveri, J. Kramer, and C. Koch, "System implementations of analog VLSI velocity sensors," *IEEE Micro* **16**, pp. 40–49, 1996.
21. G. Indiveri, J. Kramer, and C. Koch, "Parallel analog VLSI architectures for computation of heading direction and time-to-contact," in *Advances in Neural Information Processing Systems*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, eds., vol. 8, pp. 720–726, The MIT Press, (Cambridge, Massachusetts), 1996.
22. J. Kramer, R. Sarpeshkar, and C. Koch, "Analog VLSI motion discontinuity detectors for image segmentation," in *Proc. 1996 IEEE Int. Symp. Circuits and Systems, Atlanta, Georgia*, pp. 620–623, IEEE, (Piscataway, New Jersey), 1996.
23. T. Horiuchi, T. G. Morris, C. Koch, and S. P. DeWeerth, "Analog VLSI circuits for attention-based, visual tracking," in *Advances in Neural Information Processing Systems*, M. C. Mozer, M. I. Jordan, and T. Petsche, eds., vol. 9, pp. 706–719, The MIT Press, (Cambridge, Massachusetts), 1997.
24. E. Niebur and C. Koch, "Control of selective visual attention: modeling the"where" pathway," in *Advances in Neural Information Processing Systems*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, eds., vol. 8, pp. 802–808, The MIT Press, (Cambridge, Massachusetts), 1996.
25. C. A. Mead, *Analog VLSI and Neural Systems*, Addison-Wesley, Reading, Massachusetts, 1989.
26. G. Indiveri, "Winner-take-all networks with lateral excitation," *J. Analog Integrated Circuits and Signal Processing* **13**, pp. 185–193, 1997.
27. S. P. DeWeerth, "Analog VLSI circuits for stimulus localization and centroid computation," *Int. J. of Comp. Vision* **8**, pp. 191–202, 1992.
28. G. Indiveri and P. Verschure, "Autonomous vehicle guidance using analog VLSI neuromorphic sensors," in *Artificial Neural Networks - ICANN'97, Lausanne, Switzerland*, vol. 1327 of *Lecture Notes in Computer Science*, Springer Verlag, 1997.
29. M. V. Srinivasan, M. Lehrer, W. Kirchner, and S. W. Zhang, "Range perception through apparent image speed in freely-flying honeybees," *Vis. Neurosci.* **6**, pp. 519–535, 1991.