

Spiking Elementary Motion Detector in Neuromorphic Systems

M. B. Milde

mmilde@ini.uzh.ch

*Institute of Neuroinformatics, University of Zurich, and ETH Zurich,
8057 Zurich, Switzerland*

O. J. N. Bertrand

olivier.bertrand@uni-bielefeld.de

*Neurobiology, Faculty of Biology, Bielefeld University, 33615 Bielefeld,
and Cognitive Interaction Technology, Center of Excellence,
Bielefeld University, 33501 Bielefeld, Germany*

H. Ramachandran

rharshawardhan@techfak.uni-bielefeld.de

*Faculty of Technology, Bielefeld University, 33615 Bielefeld, and Cognitive Interaction
Technology, Center of Excellence, Bielefeld University, 33501 Bielefeld, Germany*

M. Egelhaaf

martin.egelhaaf@uni-bielefeld.de

*Neurobiology, Faculty of Biology, Bielefeld University, 33615 Bielefeld,
and Cognitive Interaction Technology, Center of Excellence, Bielefeld
University, 33501 Bielefeld, Germany*

E. Chicca

chicca@cit-ec.uni-bielefeld.de

*Faculty of Technology, Bielefeld University, 33615 Bielefeld, Germany,
and Cognitive Interaction Technology, Center of Excellence, Bielefeld University,
33501 Bielefeld, Germany*

Apparent motion of the surroundings on an agent's retina can be used to navigate through cluttered environments, avoid collisions with obstacles, or track targets of interest. The pattern of apparent motion of objects, (i.e., the optic flow), contains spatial information about the surrounding environment. For a small, fast-moving agent, as used in search and rescue missions, it is crucial to estimate the distance to close-by objects to avoid collisions quickly. This estimation cannot be done by conventional methods, such as frame-based optic flow estimation, given the size, power, and latency constraints of the necessary hardware. A practical alternative makes use of event-based vision sensors. Contrary to the frame-based

approach, they produce so-called events only when there are changes in the visual scene.

We propose a novel asynchronous circuit, the spiking elementary motion detector (sEMD), composed of a single silicon neuron and synapse, to detect elementary motion from an event-based vision sensor. The sEMD encodes the time an object's image needs to travel across the retina into a burst of spikes. The number of spikes within the burst is proportional to the speed of events across the retina. A fast but imprecise estimate of the time-to-travel can already be obtained from the first two spikes of a burst and refined by subsequent interspike intervals. The latter encoding scheme is possible due to an adaptive nonlinear synaptic efficacy scaling.

We show that the sEMD can be used to compute a collision avoidance direction in the context of robotic navigation in a cluttered outdoor environment and compared the collision avoidance direction to a frame-based algorithm. The proposed computational principle constitutes a generic spiking temporal correlation detector that can be applied to other sensory modalities (e.g., sound localization), and it provides a novel perspective to gating information in spiking neural networks.

1 Introduction

Both animals and humans move most of the time while interacting with the world. This self-induced motion (i.e., ego motion), generates continuous change in the retinal image of the animal or human. While the encoding of motion in insects (e.g., flies and bees) is done with graded potential (modeled with the classical Reichardt detector, Hassentstein & Reichardt, 1956; Borst & Helmstaedter, 2015), dedicated structures in the mammalian brain encode motion information using action potentials (i.e., spikes). However, the precise mechanisms and circuitry to encode motion in cortical structures remain elusive and are subject to ongoing research (Foster, Gaska, Nagler, & Pollen, 1985; Priebe, Lisberger, & Movshon, 2006; Perrone & Thiele, 2002; Roksztin et al., 2010). We know that spikes are dominant mode of communication of information in the vertebrate nervous system. Hence, elementary motion is expected, in principle, to be encoded with spikes at some processing stage. The classical elementary motion detector (EMD) model (Hassentstein & Reichardt, 1956) does not account for spike-based motion estimation.

First-order motion (i.e., elementary motion) can be described by the characteristic pattern of changes of brightness induced by the motion of objects in the visual scene. This pattern is called optic flow (Gibson, 1950, 1979). During translational movements, a nearby object appears to move faster than its background. This apparent motion of objects during translational ego motion provides spatial information about the environment, which can be exploited to construct a map based on the relative distances (Bertrand,

Lindemann, & Egelhaaf, 2015; Faessler et al., 2016) and avoid collisions with obstacles while navigating through an environment (Milde, Bertrand, Benosman, Egelhaaf, & Chicca, 2015; Clady et al., 2014; Kramer, Sarpeshkar, & Koch, 1997; Serres & Ruffier, 2017; Bertrand et al., 2015; Mueller, Bertrand, Lindemann, & Egelhaaf, 2018; Zingg, Scaramuzza, Weiss, & Siegwart, 2010). Furthermore, optic flow has been used in conventional frame-based machine learning applications to segment images (Weinzaepfel, Revaud, Harchaoui, & Schmid, 2013; Chen, Papandreou, Kokkinos, Murphy, & Yuille, 2016), classify objects in videos (Rahtu, Kannala, Salo, & Heikkilä, 2010), and perform tracking (Manen, Kwon, Guillaumin, & Van Gool, 2014).

Most algorithms processing optic flow-based information rely on frames acquired from conventional imaging sensors. However, successive images in a video do not change at every pixel location. Thus, these algorithms perform unnecessary computation due to the redundancy in the data. The computational cost can be lowered by using event-based vision sensors (Posch, Matolin, & Wohlgenannt, 2010; Lichtsteiner, Posch, & Delbruck, 2008; Brandli, Berner, Yang, Liu, & Delbruck, 2014; Posch, Serrano-Gotarredona, Linares-Barranco, & Delbruck, 2014; Son et al., 2017). These sensors have the advantage that only changes in temporal contrast, encoded as events, trigger the generation of data, thereby providing a sparse coding of the visual input. Optic flow estimation can take advantage of this sparse representation, as already demonstrated by several studies (Benosman, Clercq, Lagorce, Ieng, & Bartolozzi, 2014; Benosman, Ieng, Clercq, Bartolozzi, & Srinivasan, 2011; Liu & Delbruck, 2017; Conradt, 2015; Rueckauer, & Delbruck, 2016), especially when the asynchronous encoding scheme is maintained, hence preserving precise timing (i.e., very low latencies), and “*pseudo-simultaneity*” (Camunas-Mesa et al., 2012).

Unlike synchronous processing, which is performed in a serial manner (as in a CPU or microcontroller), asynchronous computing, similar to neural computing in the biological brain, naturally preserves temporal information without the need to explicitly encode time and offers the advantages provided by distributed and fully parallel computation.

Neuromorphic processing systems, especially mixed-signal (analog/digital) ones, combine all the aforementioned properties and are thus well suited to operate on event-based data. Further advantages are provided by neuromorphic processors operating in the sub-threshold regime (Mead, 1989). The current flowing across a transistor that is operated in the subthreshold range depends exponentially on the voltage supplied to the gate of the transistor. This exponential relationship is needed to model transfer characteristics found in biological neurons and synapses. In analog subthreshold neuromorphic processors, this exponential transfer characteristic can be implemented with a single transistor in contrast to digital superthreshold systems, in which a high computational load is required (Partzsch et al., 2017). Neuromorphic sensory-processing systems are perfectly suited to be incorporated in autonomous agents in order to

estimate optic flow from visual input. Such an agent could be a flying, walking, or wheeled robot, which should, depending on the field of application, be capable of navigating autonomously in any given environment. This kind of agent is required especially in the context of search and rescue missions, where not only size represents a crucial constraint but also the payload (Calisi, Farinelli, Iocchi, & Nardi, 2007; Ko & Lau, 2009). Especially the latter directly affects the agent's operation time. Asynchronous neuromorphic sensory-processing systems provide solutions that have been shown to scale with power consumption, in contrast to conventional synchronous GPU-based solutions (tens of mW versus hundreds of W). Furthermore, their inherent parallel computing architecture makes them even better suited for scaling up due to their distributed computation.

Our work shows how a spike-time-dependent gain modulation of the well-known differential pair integrator (DPI) synapse (Bartolozzi & Indiveri, 2007) adaptively scales the synaptic efficacy. This adaptive synaptic efficacy (ASE) scaling enables a downstream neuron to encode the time-to-travel between two spatially adjacent inputs (e.g., pixels) into an instantaneous burst of spikes. The ASE circuit in combination with a DPI synapse and an adaptive, exponential leaky integrate-and-fire (LIF) neuron (Indiveri, Chicca, & Douglas, 2006) describe the spiking elementary motion detector (sEMD) presented here. The time-to-travel is inversely proportional to the amplitude of optic flow. The size and duration of the burst produced by the LIF neuron directly reflect the temporal correlation of two spatially adjacent inputs: the closer in time the two spikes arrive relative to each other, the more the neuron spikes. We will show in detail how the sEMD can be modeled in software and how the principle can be further abstracted and emulated in mixed-signal neuromorphic circuits (hardware). Further, we will demonstrate that the sEMD can be used to extract a collision avoidance direction of a moving robotic agent in an outdoor cluttered environment. The proposed circuitry might constitute a possible connectivity scheme of how biological synaptic structures are organized in order to estimate temporal correlation from discrete action potentials.

1.1 Event-Based Optic Flow Estimation. Event-based algorithms use temporal contrast changes to estimate optic flow. These changes are detected asynchronously by so-called event-based vision sensors, which send an event whenever the light intensity changes by a sufficient amount (see section 2.1). Approaches to estimating event-based optic flow developed during the last three decades range from the gradient-based method using the Lukas-Kanade (Lucas & Kanade, 1981) algorithm (Benosman et al., 2011), local-plane fitting (Brosch, Tschechne, & Neumann, 2015; Milde et al., 2015), or relational networks (Martel, Chau, Dudek, & Cook, 2015) to correlation-based methods based on either delay lines (Horiuchi, Lazzaro, Moore, & Koch, 1991; Horiuchi, Bair, Bishofberger, Lazzaro, & Koch, 1992),

block matching of event frames (Liu & Delbruck, 2017), or the time-to-travel algorithm (Kramer, Sarpeshkar, & Koch, 1995).

This work proposes a novel correlation-based motion detection scheme for analog very large scale integration (aVLSI) systems inspired by the time-to-travel algorithm (Kramer, Sarpeshkar, & Koch, 1995). As stated earlier, the time-to-travel of events across the retina is inversely proportional to the relative velocity. Kramer and colleagues (Kramer, Sarpeshkar, & Koch, 1995, 1996, 1997) converted a fast brightness change into a single current pulse using a temporal edge detector circuit. In Kramer's temporal edge detector circuit, a current pulse was fed to a pulse-shaping circuit, which produced a slow, monotonic, decaying voltage signal. The respective pulse produced by a pixel x acts as a so-called *facilitation pulse*, while a pulse of the neighboring pixel $x + 1$ triggers the measurement (i.e., the *trigger pulse*). The time to travel is directly encoded in the absolute output voltage of the circuit. The voltage is set by the relative time of the facilitation to the trigger pulse and stored using a standard sample-and-hold circuit (Kramer et al., 1997). As soon as the measurement is triggered, the trigger pulse causes positive feedback, which ultimately resets the circuit to its resting state.

Conradt (2015) used the time-to-travel algorithm implemented on a microcontroller to extract optic flow directly from a dynamic vision sensor (DVS). Events produced by the DVS are time-stamped and used to compute the Δt between adjacent pixels in order to extract the time-to-travel. This approach has the advantage that motion estimation is possible within a wide dynamic range of velocities. However, it has the drawback that the time-to-travel is encoded as a fixed-point number, which cannot directly be used by a neuromorphic processor.

Giulioni, Lagorce, Galluppi, and Benosman (2016) used an event-based vision sensor connected to a mixed signal analog/digital neuromorphic processor in order to estimate optic flow in a more biologically plausible manner with low-power and low-latency requirements. They used the same time-to-travel idea as originally proposed by Kramer et al. (1995). Their circuit motif can be summarized by feedforward inhibitory connections to direction-selective neurons as identified by Barlow and Levick (1965). The time-to-travel is expressed by the number of spikes a so-called elementary motion unit (EMU) emits. Four EMUs share one facilitation neuron, which is connected with excitatory synapses to four direction-selective cells (one for each cardinal direction). Each direction-selective unit has one additional trigger neuron connected to an inhibitory synapse. The time for which one of the four direction-selective neurons is active with respect to the facilitation neuron's activity encodes the relative velocity of a stimulus.

We propose a novel motion extraction mechanism, which we call the spiking elementary motion detector (sEMD). Similar to Conradt (2015) and Giulioni et al. (2016), we decouple the motion estimation from the sensor, in contrast to Kramer et al. (1995). This enables us to use various event-based vision sensors and identify the best-suited one, given the constraints of our

task. Conradt (2015) used a synchronous processor (i.e., a microcontroller), to calculate the time-to-travel. As argued earlier, synchronous processors in principle can operate on event-based data, but the processor does not have the intrinsic distributed and parallel nature that would be needed to optimally exploit the sparsity and asynchronicity of event-based data. Like Giulioni et al. (2016), we emulated the motion detector in mixed-signal analog/digital neuromorphic hardware and processed events coming from an event-based vision sensor using artificial neurons and synapses. By doing so, we make the motion estimate easily available to a downstream network of spiking neurons. While Giulioni et al. (2016) needed 9 neurons and 8 synapses to realize a full motion detection unit (i.e., a 4-way motion detector), our implementation consists only of 4 neurons, 4 synapses, 12 additional transistors, and 4 capacitors. Thus, the sEMD requires less silicon area for its implementation, while encoding the time-to-travel in the number of spikes of the neuron, similar to Giulioni et al. (2016). The 12 transistors and 4 capacitors constitute 4 so-called adaptive synaptic efficacy (ASE) circuits (see section 2.3 for more details). This circuit adaptively scales the synaptic efficacy dependent on the relative timing of the trigger and the facilitation pulses. The essence of the computation of the proposed circuitry is to realize a temporal correlation detector, which is encoded by an instantaneous synaptic efficacy modulation. The synaptic efficacy modulation not only determines the absolute number of spikes produced by the neuron but also affects the interspike interval (ISI) distribution within a burst. In contrast to Giulioni et al. (2016), where the ISIs within a burst were kept constant, the ISIs within a burst in the sEMD response exponentially increase over time. This exponential increase enables the circuit to encode information about the time-to-travel already by the first two spikes of a burst. Thus, the sEMD can perform a fast but imprecise motion estimate with two spikes and provides a more accurate estimate within few milliseconds.

2 Methods

2.1 Neuromorphic Silicon Retina. Event-based vision sensors are fundamentally different from traditional cameras. Contrary to the frame-based approach, they produce data only when there are changes in the visual scene, which makes them inherently efficient. The event-driven nature of the sensors guarantees high temporal resolution (more than 1 million events per second Lichtsteiner et al., 2008) compared to standard cameras (24–60 frames per second, that is, about 0.4 to 1 million events per second with 128×128 pixel resolution).

The dynamic vision sensor (see Figure 1) used in this study has a 128×128 pixel resolution, an intrascene dynamic range of 2 lux to 100 klux, and a power consumption of 23 mW. The DVS pixel produces an event if a change in luminance bigger than about 10% occurs; ON-events encode positive changes while OFF-events are triggered by negative changes (for

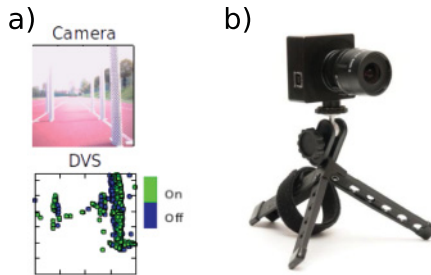


Figure 1: The dynamic vision sensor. (a) Comparison of the same scene as seen by a conventional frame-based camera (top) and an event-based DVS camera (bottom). Note that the DVS only sees objects at locations where the temporal contrast changes (e.g., edges). To generate the DVS picture, events within a 35 ms time window were accumulated while the agent was moving translationally at 0.7 m/s. (b) The dynamic vision sensor (taken from DVS, 2009).

more details, see Lichtsteiner et al., 2008). Upon occurrence of an event, the corresponding pixel's location (the x and y coordinates), the polarity of the event (pol , i.e., “on” or “off”), and a time stamp (t), are transmitted in the form of an address event representation (AER) event (see equation 2.1) (AER-Caltech-Memo, 1993; Mahowald, 1992, 1994). The AER protocol is an asynchronous handshaking protocol, where each event is written onto a common transmission bus, which is shared by all pixels on the chip:

$$e = (x, y, t, pol). \quad (2.1)$$

A DVS mounted on a mobile robotic platform emits events triggered by contrast changes produced by the boundaries of objects present in the environment. These events can be used to extract the relative motion of the objects due to the translational ego motion of the camera (see section 2.2).

2.2 Spiking Elementary Motion Detector in Software. We designed a spiking neural network in Brian2 (Goodman & Brette, 2008) for implementing spiking elementary motion detectors compatible with event-based vision sensors. The network in its simplest form processes the sensor's events in order to filter out noise and extract relative motion from the stream of events. Each neuron in the network is described by the differential equation of a simple linear LIF neuron (Indiveri et al., 2006),

$$\frac{dV}{dt} = \frac{1}{C_m} \times (I_e - (I_{lk} + I_i) \times (1 - e^{-\frac{V}{U_r}})), \quad (2.2)$$

where V is the membrane voltage of the neuron, t is time, C_m is the membrane capacitance, U_t is the thermal voltage, I_e is the excitatory current, I_i is the inhibitory current, and I_{lk} is the leak current. Following each presynaptic spike, the corresponding current is updated with a given synaptic weight w :

$$I_{syn} = I_{syn} + w. \quad (2.3)$$

All synaptic currents are described by a differential equation:

$$\frac{dI_{syn}}{dt} = -\frac{I_{syn}}{\tau}, \quad (2.4)$$

where I_{syn} is the synaptic current, t is time, and τ is the time constant of the exponential decay of the excitatory post-synaptic current (EPSC).

2.2.1 Spatiotemporal Filtering. The DVS response is noisy and thus needs to be filtered. This noise is, on the one hand, due to shot noise, which triggers events without a cause (Lichtsteiner et al., 2008; Yang, Liu, & Delbruck, 2015). On the other hand, the DVS is, like any other neuromorphic hardware, prone to device mismatch¹ (Pelgrom, Duinmaijer, & Welbers, 1989) and temperature sensitivity (Nozaki & Delbruck, 2017). The sources of noise result in noise events and so-called hot pixels—those that continually emit events. To prevent the noise from altering the motion detection, the first layer of the network implements a spatiotemporal filter for removing events that are isolated in time and space. A spatial neighborhood of 3×3 pixels of the DVS is selected. All pixels in one neighborhood are connected to a single spatiotemporal filtering neuron, which generates a spike only if at least six local contrast changes within the neighbourhood are detected within 35 ms (see Figure 2a). The neuron's parameters are set to guarantee that even when all pixels within the neighborhood are active within 35 ms, only one output spike is produced (see supplementary information Table 2).

This layer not only cancels out noise but also reduces the amount of data to be processed by downsampling the spatial resolution. The spatiotemporal filtering neurons help to ensure that the spikes fed into the sEMD neuron are triggered by a common cause, thus ensuring the presence of a change in contrast often due to a moving edge. The output of two neighboring spatiotemporal filtering neurons is used to estimate the time-to-travel.

2.2.2 Measuring Time-to-Travel with Spiking Neurons. We propose a novel mechanism for coding time-to-travel with spiking neurons. To illustrate the

¹Device mismatch is the response variability introduced to the electronic circuits due to imperfections in the manufacturing process. This leads to variability across all circuit blocks in an array.

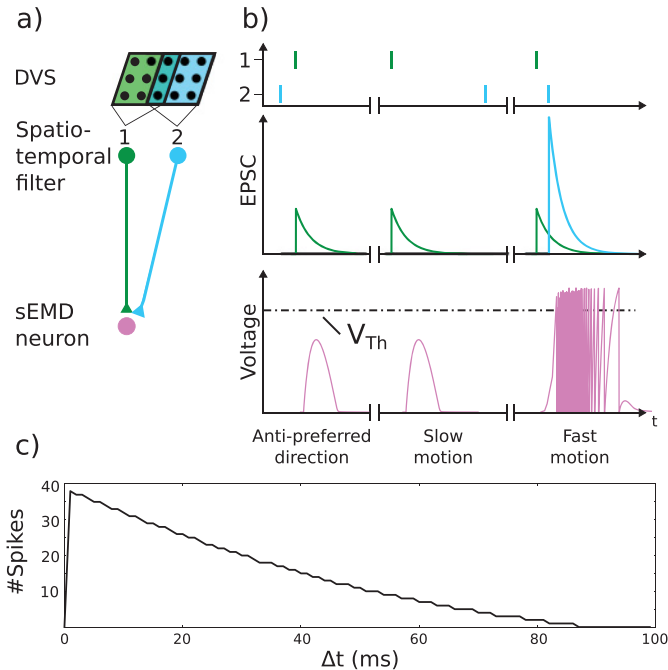


Figure 2: The spiking elementary motion detector (sEMD) model in Brian2. (a) Nine DVS pixels are connected to one spatiotemporal filtering neuron. The spatiotemporal filter neuron emits a single spike only if six or more DVS pixels are active within 35 ms. Only a single sEMD neuron is depicted. Both spatiotemporal filter neurons—facilitation (green) and trigger (blue)—are connected via a normal synapse to the sEMD neuron. (b) The direction selectivity of the sEMD neuron is installed by modulating the synaptic efficacy of the trigger synapse (blue) by the EPSC of the facilitation synapse (green) (see equation 2.5). Only if neuron 1 spiked first and shortly before neuron 2 is the EPSC produced by neuron 2 strong enough to elicit spikes in the sEMD neuron. (c) Tuning curve of the sEMD. The number of spikes of the sEMD neuron distinguishes different Δt between facilitation and trigger spikes.

principle, let us assume that we have two spatiotemporal filtering neurons and one sEMD neuron (see Figure 2a). The synaptic efficacy of the trigger synapse (blue) is determined by the EPSC of the facilitation synapse (green). The trigger synapse can only inject current into the neuron (i.e., produce an EPSC), if the synaptic current of the facilitation synapse is nonzero (see Figure 2b; compare the anti-preferred direction versus the fast motion). Otherwise, the spike, which is propagated along the trigger synapse, has no effect on the postsynaptic membrane potential (see Figure 2b, slow motion and anti-preferred). Furthermore, the synaptic weight of the facilitation synapse

is set to guarantee that if there is no EPSC propagated along the trigger synapse, the sEMD neuron will not produce any output spikes (see Figure 2b, anti-preferred direction and slow motion). If the facilitating spatiotemporal filtering neuron spikes at t_0 , the membrane potential rises. Further, if the trigger neuron emits a spike at $1 \text{ ms} < t_1 - t_0 < 80 \text{ ms}$, the EPSC of the facilitation synapse is nonzero (compare Figure 2b, antipreferred motion and fast motion). As a consequence, the sEMD neuron emits a burst of spikes where the number of spikes encode the time-to-travel. The synaptic weight of the trigger synapse (w_{e2}) is multiplied by the synaptic current of the facilitation synapse (I_{e1}) and added to the actual synaptic current of the trigger synapse (I_{e2}):

$$I_{e2} = I_{e2} + (w_{e2} \times I_{e1}). \quad (2.5)$$

This leads to a nonlinear scaling of the EPSC (see equation 2.4) depending on the time-to-travel; in addition to providing motion perception, it also gives us a direction-selective mechanism for free. The sEMD neuron encodes the time-to-travel, which is represented by the Δt in the firing of the pre-synaptic spatiotemporal filtering neurons. This Δt is converted into an instantaneous burst of spikes (see Figure 2c). In our example, a stimulus moving from right to left does not produce any output spikes; the membrane voltage stays below the spiking threshold V_{Th} (see Figure 2b, antipreferred). This is due to the multiplication of the synaptic weight w_{e2} with the synaptic current I_{e1} , which is zero in this case.

We showed that the sEMD implemented in software is able to encode the time-to-travel of events produced by an edge moving in the visual field of an event-based sensor (see Figure 2).

2.3 Spiking Elementary Motion Detector in Hardware. The sEMD features a direction-selective nonlinear scaling mechanism of synaptic currents. The application of this motion detection scheme to robotic sensory-motor tasks poses strong constraints on latency, power consumption, and, in some cases, size and weight (e.g., in mini- and flying robots). We address these constraints by using subthreshold neuromorphic solutions not only at the sensor but also at the computation level. The subthreshold operation of the circuit produces very small currents (on the order of nano-amperes), resulting in low-power consumption. In this regime, the transfer function of the transistor is exponential, providing a powerful tool for the implementation of biological models at low “cost” (power consumption and silicon area), given the otherwise high computational cost of the exponential function (Partzsch et al., 2017). Therefore, we designed a circuit implementing the sEMD model described in section 2.2 and fabricated it (see Figure 3, left) using the standard Austria microsystems (AMS) 180 nm CMOS technology. The resulting chip comprises an array of eight sEMD blocks (see

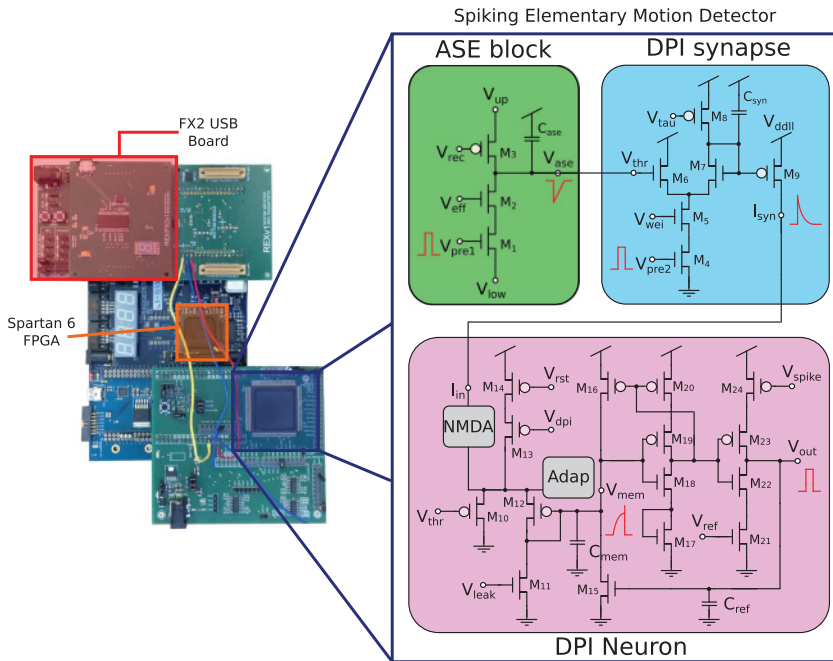


Figure 3: (Left) Overview of the spiking elementary motion detector (sEMD) test chip. The sEMD chip is interfaced by a Raggedstone 2 (Raggedstone2, 2017) board, containing a Spartan 6 FPGA. The FPGA is interfaced by the FX2 USB board (Fasnacht, Whatley, & Indiveri, 2008) in order to set biases and stimulate and record activity. (Right) Detailed schematics of the sEMD. The three computational blocks of the sEMD are shown in green (ASE), blue (DPI synapse), and pink (DPI neuron). The ASE block receives the facilitation pulse (red) as input at the gate of M_1 and outputs the voltage V_{ase} to the gate of transistor M_6 (threshold) of the differential pair integrator (DPI) synapse. The voltages V_{eff} and V_{rec} determine the amplitude and recovery rate of the V_{ase} voltage. The DPI synapse is turned on by receiving the trigger pulse at the gate of the M_4 transistor. The resulting output is the EPSC I_{syn} whose amplitude is set by the voltages V_{wei} and V_{ase} , and V_{tau} controls the time constant of recovery. The output current I_{syn} is injected into the neuron (I_{in}). The neuron integrates the incoming current and starts spiking once the membrane potential rises above its threshold set by the voltage V_{thr} . The leak and the refractory period of the neuron can be set by the corresponding voltages V_{leak} and V_{ref} . The NMDA-block and the adaptation block are turned off and not used in this study. For details, see Indiveri et al. (2006).

Figure 3, right). In order to stimulate and characterize the sEMD test chip, we used the pyNCS framework (Stefanini, Neftci, Sheik, & Indiveri, 2014). The differential pair integrator (DPI) silicon synapse (Bartolozzi & Indiveri,

2007), the short-term adaptation circuit proposed in Ramachandran, Weber, Aamir, and Chicca (2014), and the DPI adaptive, exponential LIF neuron (Indiveri et al., 2006) are used as building blocks for the implementation of the sEMD model in hardware (see Figure 3).

2.3.1 Adaptive Synaptic Efficacy Circuit. Short-term plasticity, influencing the strength of the synapse (Markram, Pikus, Gupta, & Tsodyks, 1998), can be classified by polarity: short-term depression reduces the synaptic strength, and short-term facilitation increases the synaptic strength. Circuits modeling short-term plasticity are normally connected to the gate of the weight transistor of a synapse. We, however, used a simple neuromorphic short-term depression circuit presented in Ramachandran et al. (2014) to alter the synaptic efficacy independent of the weight. Hence, we call this block adaptive synaptic efficacy (ASE) circuit hereafter. In response to the incoming spike (event/pulse), the ASE circuit alters the synaptic efficacy and offers an independent control over the recovery rate of the efficacy. In our architecture, the ASE circuit receives the first input event and thus facilitates the motion estimate. The ASE circuit features a weight, w_{ase} , as well as a time constant, τ_{ase} , set by V_{eff} and V_{rec} , respectively (see Figure 3). Both biases affect the output voltage of the ASE circuit in response to a pre-synaptic spike. The output voltage of this circuit is provided as a gate voltage to the threshold transistor of the DPI synapse as shown in the block diagram (see Figure 3).

2.3.2 Synapse Circuit. The DPI synapse presented in Bartolozzi and Indiveri (2007) is one of the most used synapses in subthreshold neuromorphic chips. In response to the presynaptic input spikes (events/pulses), the DPI synapse gives an exponential decaying EPSC as its output depending on the parameter setting (see Figure 3). The weight, w_{dpi} , which is set by V_{wei} of the synapse, determines the amplitude of the EPSC evoked during a pre-synaptic spike, whereas the time constant τ_{dpi} , which is set by V_{tau} , dictates the temporal evolution of the EPSC in between the pre-synaptic spikes. The DPI synapse has a threshold bias. The transistor that sets the synapse's threshold is usually employed to implement any global computation that affects the EPSC of the synapse, such as the homeostasis mechanism. In fact, a voltage supplied to the gate of the threshold transistor scales the amplitude of the resulting EPSC along with the w_{dpi} associated with the synapse circuit (see Figure 3). To realize a quarter-way motion detector, we used one DPI synapse that acts as a trigger pulse generator. To facilitate a motion estimate, we use an ASE block connected to the threshold of the corresponding synapse.

2.3.3 Neuron Circuit. We used the adaptive exponential LIF neuron circuit presented in Indiveri et al. (2006). The neuron integrates the incoming synaptic current, which charges its membrane capacitor C_{mem} (see Figure 3).

Furthermore, the silicon neuron offers a tunable leakage current, as well as a spiking threshold. All biases to tune the neuron's behavior can be set externally and are stored on chip once they are loaded. If the membrane voltage surpasses the spiking threshold value, a positive feedback loop is activated, during which C_{mem} is quickly charged. Consequently, the neuron consumes very little power during the spike generation (Indiveri et al., 2006). Right after the spike generation, the membrane potential is reset to zero by enabling the reset transistor (M_{15}), and then the refractory period kicks in. The length of the refractory period is also determined by an external bias (V_{ref}), which sets the gate voltage of the refractory transistor M_{21} . The current through the refractory transistor discharges C_{ref} and slowly turns off the reset transistor M_{15} , thus preventing the neuron from eliciting another spike. In our implementation, the refractory period is set to approximately 1 ms (see supplementary information Table 3 for a detailed list of parameters used in this study). We tuned the parameters of the neuron so that no spike is elicited in response to small EPSCs; however, the neuron spikes for large-input currents.

The ASE circuit determines the amplitude of the EPSC generated by the DPI synapse, which causes spiking output of the neuron. Therefore, the firing rate of the neuron is determined by the threshold value of the synapse, which in turn depends on the time-to-travel of the event across the focal plane. It should be noted that the DPI synapse receives events independent of the ASE circuit. In this way, the sEMD circuit obtains a direction-selective response and encodes the time-to-travel in its spiking output.

3 Results

To characterize the circuit's response to simple and well-controlled stimuli, we used a single facilitation pulse and (in contrast to typical usage) multiple trigger pulses (as defined in section 1.1) with different timing, as shown in Figure 4. Before the facilitation pulse is provided, the output voltage of the ASE circuit (V_{ase}) is at its resting level (0.9 V; see Figure 4, ASE). As soon as the facilitation pulse is sent to the gate of M_1 , the ASE circuit's capacitor (C_{ase}) is quickly charged, and the ASE circuit's output voltage drops to 0.4 V (see Figure 4, ASE). The size of the voltage drop is set by the gate voltage of the M_2 transistor. As soon as V_{out} deviates from its resting level, the current through the M_3 transistor starts discharging the capacitor again. The time constant of the recovery of V_{out} to its resting voltage is set by the gate voltage V_{rec} of M_3 . The output voltage of the ASE circuit is connected to the gate of the threshold transistor (M_6) of the DPI synapse (see Figure 3), therefore modulating the synaptic efficacy. The trigger pulses are provided to the gate of M_4 of the DPI synapse at eight different times relative to the facilitation pulse (Figure 4; AER input, $\Delta t = 2, 12, 22, 32, 42, 52, 62, 72$ ms). The relative time between the facilitation and the trigger pulse represents the time-to-travel (for more details, see section 1.1). The trigger pulse is used as input

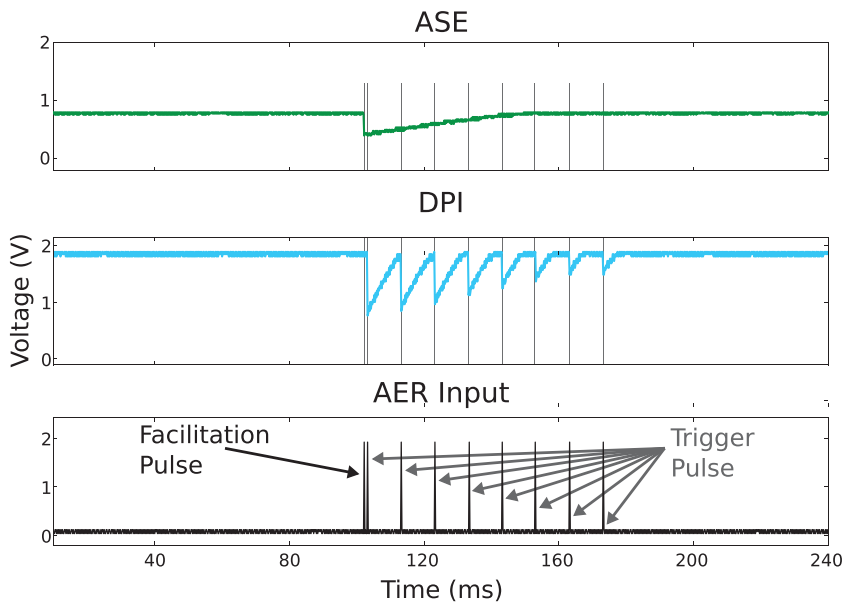


Figure 4: Adaptive synaptic efficacy (ASE) circuit's response and its effect on the DPI's efficacy. Voltage traces of both circuit blocks in response to different Δt . The synaptic weight, w_{dpi} set by the gate voltage of M_5 , is unaltered. The synaptic efficacy is determined by the threshold voltage, which in turn is set by the ASE's output voltage. The biases of the sEMD circuit were set to emphasize its response. The DPI time constant is usually in the range of a few tens of milliseconds.

to the DPI synapse, generating an output current that is a function of the output voltage (V_{ase}). A short time-to-travel results in a low V_{ase} at the time of the trigger pulse, which in turn produces a large change in the DPI's output voltage. Hence, the amplitude of the output EPSC is also large. A longer time-to-travel produces a smaller EPSC amplitude.

The increase in amplitude of the resulting EPSC follows a multiplicative effect on top of the amplitude set by the gate voltage of M_5 (V_{wet}) of the synapse. The time constant of EPSC is determined by the gate voltage of M_8 (V_{lau}) of the synapse and is unaffected by the V_{ase} . Therefore, the downstream neuron can integrate more current and thus elicit more spikes.

To show this effect, we stimulated the circuits with two time-to-travel ($\Delta t = 2, 20$ ms) and observed the neuron's spiking behavior in terms of the number of spikes elicited, as well as the duration of the burst (see Figure 5, Neuron). The gate voltage of M_6 (Figure 5, ASE) is at 0.45 V and 0.65 V when the respective trigger pulses arrive at the DPI synapse. Thus, the resulting EPSC amplitude is larger at $\Delta t = 2$ ms compared to $\Delta t = 20$ ms (see

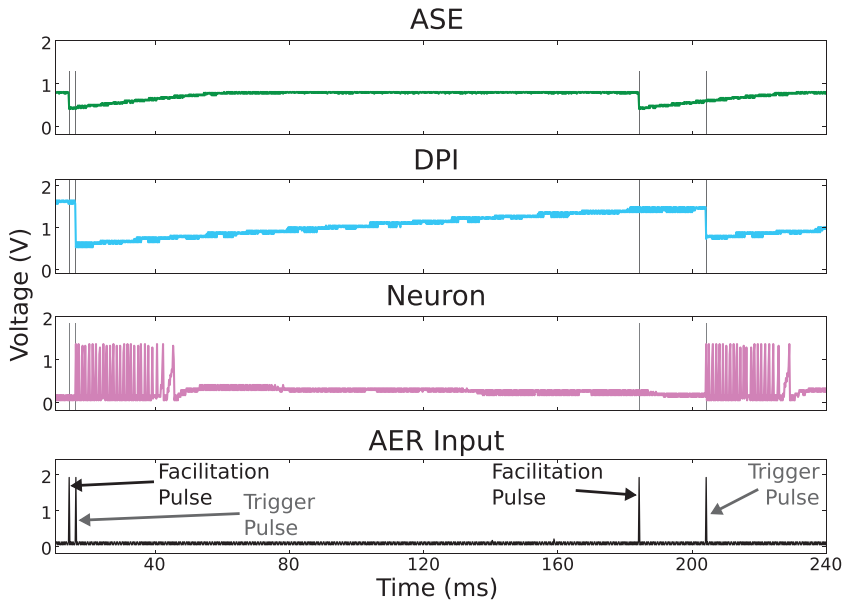


Figure 5: Membrane potential of an example neuron in response to two different Δt between facilitation and trigger pulse. Note that not only the number of spikes but also the burst length is set by the relative timing of the pulses. The biases of the sEMD circuit were set to emphasize its response. Burst duration usually does not exceed 10 ms, and the DPI time constant is in range of few tens of milliseconds.

Figure 5, DPI). Since the ASE circuit is not connected to the neuron, the facilitation pulse has no effect on the membrane potential (see Figure 5, Neuron). Only the DPI can inject current into the neuron when M_4 receives the trigger pulse at its gate. The larger EPSC amplitude for smaller Δt results in more spikes (26 within 35 ms at $\Delta t = 2$ compared to 22 within 25 ms at $\Delta t = 20$).

We conducted this experiment to demonstrate the mode of operation of the sEMD in its simplest form and provide an intuitive understanding of the impact of different voltage levels on the sEMD's response. The amplitude of the EPSC that is provided to the neuron as I_{in} was found to be scaled depending on the output voltage V_{ase} , as intended by design. The gate voltage V_{tau} of M_8 is fixed, and thus the amount of current injected into the neuron is determined only by the amplitude of the EPSC. The neuron integrates this current and varies its response in terms of the number of spikes accordingly.

To further characterize the sEMD circuit and obtain the full tuning curve (neuron response versus time-to-travel), we systematically varied the relative time between the facilitation and trigger pulse (Δt) and measured the

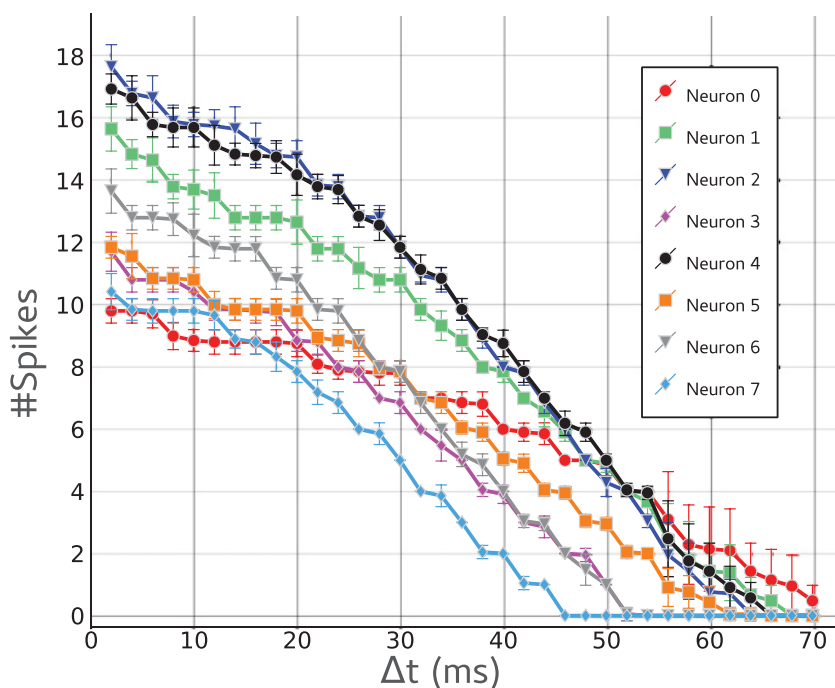


Figure 6: Tuning curves of the sEMDs. Number of spikes within a burst for each sEMD neuron is calculated for 20 stimulus presentations. The respective mean and standard deviation are calculated. The mean represents the deviations in responses across the neuron array resulting from device mismatch in our chip. The error bars indicate intraneuron response variability and could be due to thermal noise. The intraneuron response variability is less than 0.4 spikes during a burst.

number of spikes within a burst, the burst's duration, as well as the ISI distribution within a burst. The relative timing of the two pulses was varied from 2 ms to 70 ms with a 2 ms step size. The biases, which are shared among all eight sEMD circuits, were tuned for neuron 0 and kept fixed through all recordings (a detailed list of biases used can be found in the supplementary information, Table 3). Most sEMD neurons show a flattened tuning profile for short Δt , whereas all of them exhibit a linear encoding for intermediate Δt and a saturation region for large Δt (see Figure 6). All eight test circuits consistently show higher spiking activities for smaller Δt , in contrast to larger Δt which generate fewer spikes. We find that the slope for small time-to-travel (approximately smaller than 20 ms) is less steep compared to the slope for intermediate time-to-travel. This provides an optimal resolution at the operation regime of the circuit while maintaining a wide

dynamic range. It should be noted that the operation range is mainly set by the time constants of the ASE circuit and the DPI synapse, V_{rec} and V_{tau} , respectively, which makes it easy to adjust the dynamic range to meet the operation range needed for any given task. With the current bias setting the sEMD has a dynamic range of 34 dB—in other words, 2.4 to 85°/s. Overall we tested the circuit with a variety of biases and were able to distinguish Δt ranging from 10 ns up to hundreds of ms. The time constants of the ASE and DPI circuit ultimately determine the current that is injected into the post-synaptic neuron. The refractory period of the neuron defines the maximum number of spikes and the shortest interspike interval the neuron could potentially generate given the input current. Thus, the refractory period and other parameters, such as the neuron's leakage or threshold, modulate the sEMD's response. The instantaneous frequency, which is defined as the number of spikes within a burst divided by the duration of the burst, is not informative since the neuron is only sparsely active and only for a very short amount of time (less than 10 ms).

To characterize the variability of the responses due to thermal noise, we calculated the mean and standard deviation of each neuron across 20 stimulus sweeps (see Figure 6). We also looked into the variations in the circuit responses across the array due to device mismatch effects and plotted the population tuning curve (see Figure 7). It is worth mentioning that circuit blocks 0 and 7 are physically located at the border of the silicon area, and the mismatch tends to be larger in these areas due to corner effects (Pavasović, Andreou, & Westgate, 1994). The effect of mismatch is clearly visible in neurons 0 and 7, which show different profiles of spiking activity to Δt compared to the rest (see Figure 6).

The instantaneous bursting response shows clearly how the precise timing of incoming spikes can be translated in order to further process the motion information. We operate the circuit and its transistors in the sub-threshold regime, which, as stated earlier, yields an exponential relationship between the gate voltages and the current flowing through each transistor. Additionally, the efficacy modulation of the synapse follows a multiplicative effect. We expected to see this nonlinear scaling reflected in the tuning profile of the sEMD; however, we could only find it slightly reflected at the population response level. We thus investigated the burst more carefully in terms of its ISI distribution and found that the nonlinear scaling is indeed preserved in the temporal evolution of the ISI within a burst (see Figure 8).

A very short time-to-travel ($\Delta t < 20$ ms) saturates the neural response to its refractory period. After a few spikes into the burst, the ISI tends to increase exponentially until all provided current is integrated and translated into spikes. The ISI increases exponentially over time, longer times-to-travel elicit larger ISIs, fewer spikes, and shorter burst durations. For longer time-to-travel ($\Delta t > 20$ ms), the first two spikes carry enough information to estimate motion on a coarse scale (see Figure 8, gray and black curve). This way of encoding information in the neuronal response has not been exploited so

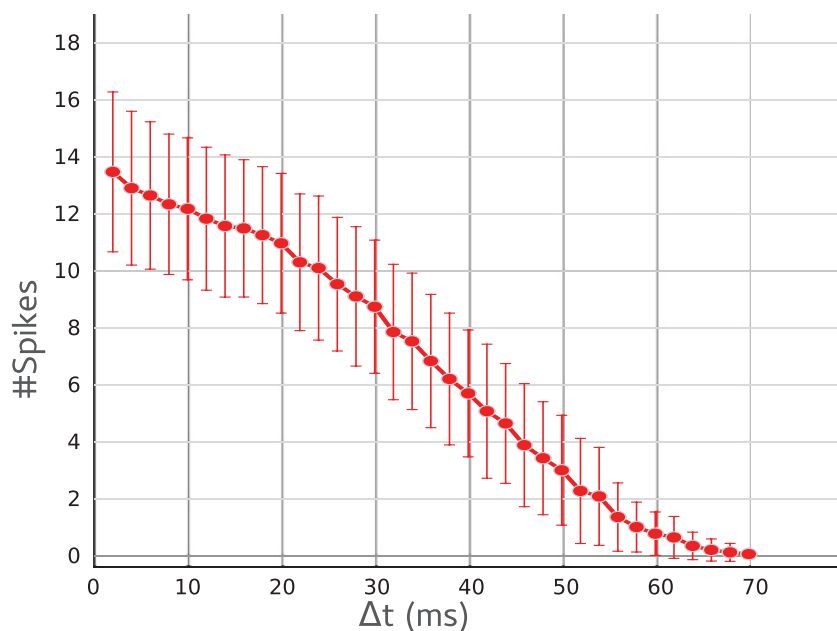


Figure 7: Tuning curve of the sEMD population. Population response of an array of 8 sEMD neurons averaged over 20 stimuli presentations and 8 neurons. The mismatch (i.e., inter-neuron variability), represented by the error bars (~ 2 spikes), is larger than the intraneuron response variability (0.4 spikes). The population tuning curve, in contrast to a single tuning profile, shows an exponential decrease for large Δt , which results from mismatch. For intermediate Δt , the tuning curve becomes linear and steeper compared to short Δt . The different slopes of the tuning curve offer good velocity resolution while preserving high dynamic range.

far by any studies to our knowledge. We hypothesize that the two described information encoding schemes, namely, the number of spikes within a burst and the ISI distribution, can be seen as complementary; potential benefits from this superimposed scheme are discussed below (see section 5).

To summarize, we have shown that the sEMD mechanism can be implemented in a mixed-signal analog/digital neuromorphic chip, and we characterized it in both software and hardware implementations in terms of single-neuron responses. Our sEMD responds to any two events occurring spatially and temporally close to each other, even if the two events are not linked to physical motion. Second-order motion stimuli give, for example, an impression of movement, although nothing is physically moving. This class of stimuli results from changes in contrast, texture, or some other quality that does not result in an increase in luminance or motion

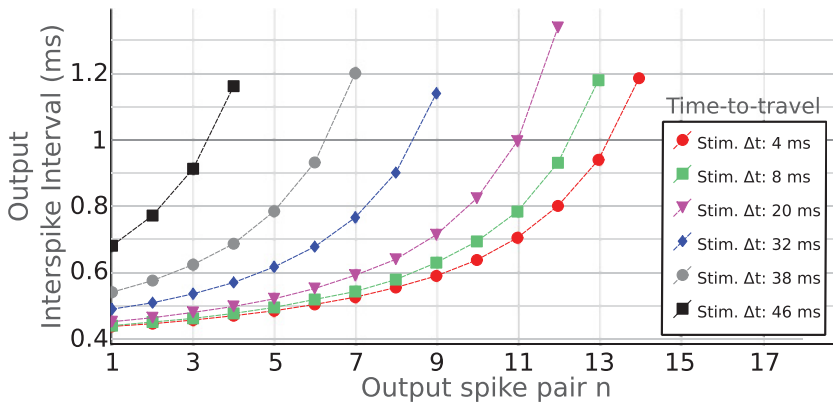


Figure 8: Interspike interval (ISI) distribution within a burst. ISI distribution for six different time-to-travel (4 ms, red dot; 8 ms, green square; 20 ms, pink triangle; 32 ms, blue diamond; 38 ms, gray dot; 46 ms, black square) within the burst of activity. The ISI is calculated as $ISI_n = t_n - t_{n-1}$. Thus, the x-axis indicates the respective spike pair. The first three stimuli (red dots, green squares, pink diamonds) make the underlying neuron fire at its maximum frequency right after the stimulus onset. However, the ISI tends to increase faster for longer Δt 's compared to smaller ones. Note that the neuron's activity triggered by longer Δt (blue diamonds, gray dots, and black squares) is slower (longer ISI of the first two spikes). This suggests that the precise timing is encoded not only in the instantaneous spiking activity but also in the relative distribution of the spikes within the burst and that the first two spikes are already informative about the presented velocity.

energy in the Fourier spectrum of the stimulus. Animals usually respond to second-order stimuli—for example, fly (Theobald, Duistermars, Ringach, & Frye, 2008), monkey (O'Keefe & Movshon, 1998), and human (Nishida et al., 2003). A large number of second-order motion stimuli exist but are not directly relevant for real-world tasks, such as collision avoidance, since for such a task, objects physically move. Thus, the systematic investigation of the response of the sEMD to second-order motion goes beyond the scope of this letter. Nevertheless, we can anticipate the response of the detector for a simple type of second-order motion, the Mu-line. For a Mu-motion (Lelkens & Koenderink, 1984), for each frame, a successive column of pixels is refreshed by random values (dark or bright). Our sEMD will interpret this as a time-to-travel if the pixel brightness of a given column and the successive column changes, and not otherwise. In the next section, we test this computational block in a real-world task, using a computer simulation of the sEMD, and testing the motion estimation mechanism on a moving wheeled robotic platform in open-loop.

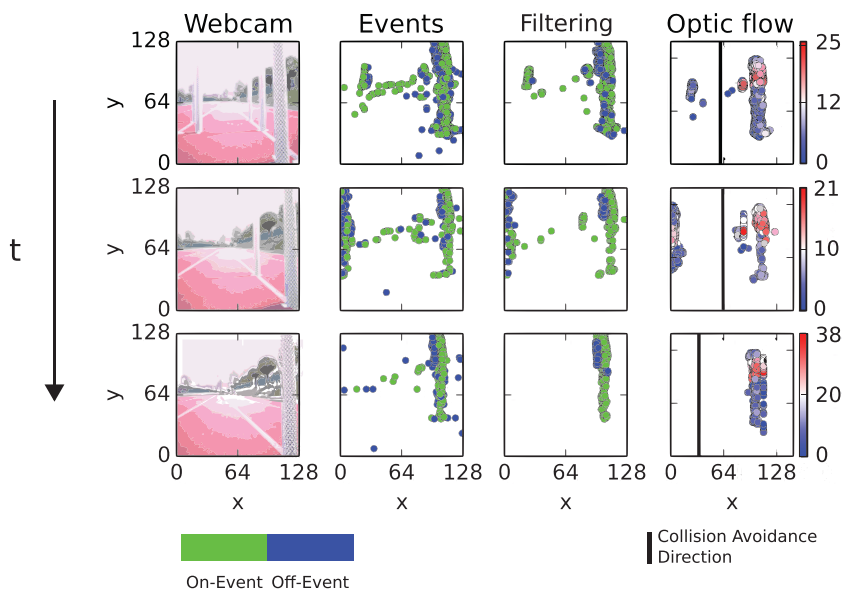


Figure 9: Neuron activity of the collision avoidance network in an open-loop for three time points. The robot moves in an outdoor environment; here, eight objects were present. The time increases from the top to the bottom row. (Column 1) Webcam frames. (Column 2) The projection of the events within a time $t_i \pm 35$ ms on the $(x - y)$ plane (i.e., the camera coordinates). ON- and OFF-events are represented by green and blue dots, respectively. Object boundaries can be seen by neighboring on and off edges. (Column 3) Spatiotemporal filtering of the events detected by the DVS. (Column 4) Spiking activity of sEMD neurons and extracted collision avoidance direction. Number of spikes of activated sEMD neuron is color-coded. Output of sEMD layer is vertically integrated (e.g., along the y -axis). The averaged activity is used to suppress the soft-inverse winner-take-all. The soft inverse winner represents the collision-avoidance direction (black bar). Note that the collision avoidance direction always points away from close-by objects along different points of the trajectory.

4 Collision Avoidance in Outdoor Cluttered Environments

Since the current implementation on the test chip of the sEMD circuit had only eight circuits, we used computer simulations (see section 2.2) to verify the working behavior of our motion detection model in a real-world, open-loop robotic scenario.

A robot was remote controlled and steered on a straight line in the center between obstacles placed in an outside environment (see supplementary information, Figure S1). A standard webcam (see Figure 9, first column) and

a DVS (see Figure 9, second column) were mounted on the robot recording data during translational movement (see Figure 9, rows 1–3). Positive and negative contrast changes (i.e., ON- and OFF-events from the DVS) were elicited by boundaries of objects. The data file sizes of the events and the webcam images were 1 MB and 9 MB, respectively, with a total recording time of 10 seconds.² The differences in file size illustrate the advantage of event versus frame-based methods in terms of the amount of data being generated and processed. To emphasize the edges of obstacles and reduce noise events as well as texture-induced events, we connected the DVS output to a layer of spatiotemporal filtering neurons (see Figure 9, third column, and section 2.2 for details). The activity of this layer was used by the sEMD neurons to extract optic flow (see Figure 9, fourth column) along the horizontal axis. The sEMD is capable of estimating rotational optic flow due to, for examples rotational ego motion. However, rotational optic flow only scales with the angular velocity and does not provide information about the distance to objects (Egelhaaf, Kern, & Lindemann, 2014). This information is crucial in the context of collision avoidance tasks; therefore, only translational optic flow can be used for ensuring collision-free movements. Each sEMD neuron encoded the time-to-travel across the retina into an instantaneous burst of spikes. The sEMD responded to time-to-travel from 1 ms to 80 ms. The lower limit was due to the simulation time step of 1 ms. The upper limit is set by the chosen parameter setting (see supplementary information Table 2). The maximal number of spikes within a burst was 38. The sEMD neurons emitted more spikes when the time-to-travel was short.

To test sEMD responses in the context of robotic navigation, we projected the neuronal activity to a winner-take-all (WTA). Since a high input rate to the WTA signals a close-by object, given constant translatory ego motion, we reversed the WTA response to select the inputs with the lowest activity. We call this a soft-inverse WTA layer, similar to Horiuchi (2009), in order to extract a collision avoidance direction. The relative position of an object in the visual field, as well as its relative nearness, was used to suppress neuronal activity in the soft inverse WTA. The activity in the soft inverse WTA layer is used to determine a steering direction, that is, a collision avoidance direction (see Figure 9, fourth column black bar). To evaluate the performance of the sEMD more systematically, we compared the output of the soft-inverse WTA using the sEMD (see Figure 10, dot-dashed orange line) as described above with the center-of-mass average nearness vector (COMANV) algorithm (Bertrand et al., 2015) along the entire robot's trajectory. The COMANV algorithm has been shown to successfully estimate the collision avoidance direction in open- as well as closed-loop scenarios using

²The webcam recorded 25 frames per second, and each frame was stored as PNG. The events are stored as a 32-bit floating-point python numpy array. No compression was used to store the events (.npy file format).

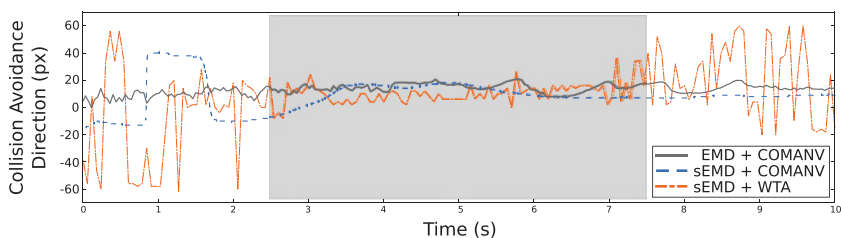


Figure 10: Comparison of conventional frame-based (solid dark gray), hybrid (dashed blue), and fully asynchronous event-based (dot-dashed orange) collision avoidance systems. Due to the retinotopic organization of the motion detector array, the collision-avoidance direction is given in pixels along the horizontal axis of the camera. A collision avoidance direction of zero indicates a straightforward motion. The obstacle corridor is indicated by the gray box. The conventional approach (solid dark gray) uses the well-established motion detection model EMD as proposed by Hassentstein and Reichardt (1956) and implemented by Meyer et al. (2016) and the center-of-mass average nearness vector (COMANV) algorithm (Bertrand et al., 2015). The hybrid approach uses optic flow as extracted by the sEMD and the COMANV algorithm to determine the collision avoidance direction. All three collision-avoidance approaches estimate a reasonable collision avoidance direction as soon as the robot enters the obstacle corridor. The huge fluctuations in collision-avoidance direction of the fully asynchronous event-based approach (sEMD + WTA, dot-dashed orange line) before and after the obstacle corridor are due to the Poisson noise, which biases the WTA and indicates missing prominent visual input, that is, objects are too far away.

walking and wheeled robots (Meyer et al., 2016; Bertrand et al., 2015; Milde et al., 2015). We used two inputs to the COMANV algorithm: (1) the output of an array of conventional EMDs (Meyer et al., 2016), which extracted optic flow from the webcam images (see Figure 10, solid dark gray line) and (2) the output of the sEMD array (see Figure 10, dashed blue line).

To assess the performance of each of the approaches and make comparison easier, we calculated the collision avoidance direction in image coordinates. Since the robot was remotely controlled and steered in the center of the obstacle corridor with a slight left bias due to the objects' arrangement (see supplementary information, Figure S1), a reasonable collision avoidance direction would point right along the middle of the obstacle corridor—about the 64th pixel. We rescaled the collision avoidance direction (CAD) to be 0, meaning that the robot would move straight ahead: $CAD_{rescale} = 64 - CAD$. A positive collision-avoidance direction indicates that the robot would steer to the left, whereas a negative collision avoidance direction indicates that the robot would steer to the right if it were to be operated in closed-loop.

After the robot starts moving, the sEMD + WTA shows huge fluctuations in its collision-avoidance direction, which are due to the background input activity of the WTA population consisting of a Poisson spike train. As soon as the robot starts entering the obstacle corridor, indicated by the gray box, the estimated collision avoidance direction converges toward zero, with a slight bias toward positive collision avoidance directions. After leaving the obstacle corridor, the noise in the Poisson input again dominates the collision avoidance direction due to missing obstacles in the visual field. The fully conventional approach (solid dark gray line) does not suffer from noisy input or noisy computation; rather, it shows a smooth collision avoidance direction, especially when moving along the obstacle corridor. The maximum difference in collision-avoidance direction between the fully conventional and the fully asynchronous approach is less than 10 pixels. The small bias toward positive collision avoidance directions (i.e., to the left) at the end of the obstacle corridor is due to the different size in the field of view of the DVS and the webcam. While the last object on the right was already at the right border of the visual field of the webcam, the object's position in the field of view of the DVS was rather more centric (compare Figure 9, last row of the first and second columns). The combination of asynchronous motion perception using the sEMD and the COMANV algorithm shows the best collision avoidance direction, apart from a brief period before entering the obstacle corridor. This big deviation might be due to noise in the visual sensory input stream (see section 2) or to the activity onset of sEMD neurons.

Overall, all three approaches are able to predict a collision avoidance direction that ensures a collision-free navigation path within the obstacle corridor. The collision avoidance direction as extracted by the spiking neural network performs equally well as conventional frame-based approaches, but it is suitable for a neuromorphic implementation based on sub-threshold analog circuits. Such implementation offers several advantages, including significant lower power consumption and much higher temporal resolution, while producing a sparse coding of the output. Thus, the fully asynchronous event-based approach might be worth pursuing in the context of closed-loop collision-avoidance in cluttered environments. Closed-loop control systems, which rely on neuromorphic sensory systems, have been proposed, ranging from vision-based pencil balancing (Conradt, Cook et al., 2009) to auditory-based source following (Klein, Conradt, & Liu, 2015). However, post-processing of sensory information was done with conventional CPUs, and the sensor itself was stationary. While tremendous research effort has been put to post-process the sensory information on neurally inspired hardware, such as neuromorphic processors, to model saccadic eye movements and visual attention (Indiveri, 2001; Bartolozzi & Indiveri, 2009; Horiuchi, Bishofberger, & Koch, 1994), the sensor, which was mounted on a stationary pan-tilt unit, was not translating in the environment. In the domain of closed-loop neuromorphic robotic navigation,

application researchers have used either conventional cameras and post-process the sensory information by neuromorphic processors (Hwu, Isbell, Oros, & Krichmar, 2017; Hwu, Krichmar, & Zou, 2017) or neuromorphic vision sensors but postprocessed the sensory information using conventional CPUs (Clady et al., 2014; Indiveri & Verschure, 1997; Hoffmann, Weikersdorfer, & Conradt, 2013; Lubner, Biedermann, & Conradt, 2015; Moeys et al., 2016). So far, fully neuromorphic (i.e., neuromorphic sensory-processing) closed-loop robotic navigation systems have only recently been shown to work in real-world scenarios (Denk et al., 2013; Galluppi et al., 2014; Milde, Blum et al., 2017). However, these systems rely on hand-engineered features to navigate (e.g., blinking LED or laser pointer) and do not use information about the object's distance while navigating. While closed-loop implementation is beyond the scope of this letter, we already implemented the sEMD on a SpiNNaker (Furber, Galluppi, Temple, & Plana, 2014) board, and we plan to set up a closed-loop experiment in the future (see supplementary information, Figure S3).

5 Discussion

We have presented the spiking elementary motion detector (sEMD) and its application to the extraction of motion information from event-based vision sensor data. We showed how the mechanism can be simulated with spiking neurons in software, emulated in mixed-signal neuromorphic hardware, and implemented on a digital neuromorphic processor (see supplementary information, Figure S3). We characterized the silicon sEMD neurons and showed that their output reliably encodes time-to-travel. The software implementation revealed that the proposed mechanism can be used to extract a collision avoidance direction in the context of robotic navigation.

The sEMD encodes the time-to-travel of an event traveling across the retina in the absolute number of spikes within a burst and the ISI distribution within a burst. The novelty of the sEMD is twofold. First, we use the sparse bursting behavior of a neuron to provide a fast response lasting less than 10 milliseconds. Second, the sEMD encodes information about the velocity in the first two spikes of the burst (VanRullen, Guyonneau, & Thorpe, 2005; Thorpe, Fize, & Marlot, 1996; Thorpe, Delorme, & Van Rullen, 2001). These two spikes can be used by a network to distinguish fast versus slow motion on a coarse scale in a quick way. In the next 10 ms, an accurate velocity estimation can be obtained from the ISI because the absolute number of spikes within a burst depends ultimately on the time-to-travel (see Figures 6 and 8). This way of encoding the velocity estimation is possible because we modulate the synaptic efficacy based on the precise spike timing. The silicon implementation of the sEMD presented in this work consists of a first prototype comprising 8 cells, designed for testing the functionality of the circuit. Given the very good results obtained, we are now in a position to design a large-scale array of sEMD circuits for building multichip systems suitable

for autonomous navigation. The design will require a thorough analysis of device mismatch through Monte Carlo simulations to allow the implementation of consistent motion flow estimation within the array.

5.1 State-of-the-Art in Event-Based Optic Flow Estimation. The proposed sEMD mechanism is a correlation-based motion detection scheme that relies on the precise timing of neighboring pixels. It measures the time an event needs to travel across the retina, that is, the time-to-travel. Our goal was to (1) efficiently estimate the time-to-travel from event-based vision sensors and provide a scalable and highly flexible solution by decoupling the estimation from the sensor. In order to optimally exploit the sparsity and asynchronicity of event-based data, we wanted (2) an asynchronous solution that features additionally a distributed and parallel processing scheme with low power consumption and low latencies. To enable a downstream spiking neural network to further process the motion information, we needed (3) an implementation on a mixed-signal neuromorphic processor by estimating the motion using artificial spiking neurons and synapses. Finally, our objective was to (4) build a circuit that can operate on a very fast timescale in order to be used in the context of robotic navigation.

The time-to-travel algorithm was originally proposed by (Kramer et al., 1995). Kramer's facilitate-and-sample circuit encoded the time-to-travel using a constant voltage level (Kramer et al., 1995, 1997). Furthermore, the temporal edge detector was colocated on the same chip. The constant voltage level does not directly enable a downstream network of spiking neurons to further process this information, and the colocation of pixel and motion detector prevents the system from being scalable without a new chip. Conradt's (2015) implementation of the time-to-travel algorithm on a microcontroller overcomes the scalability issue by using a separate vision sensor but avoids a neural implementation by calculating the differences between neighboring event time stamps. This implementation also does not address the problem of how to further process the motion estimate in a network of spiking neurons.

Giulioni and colleagues (2016) proposed for the first time an event-driven motion estimation approach, with decoupled sensor and processor, which allows scalability and encoding of time-to-travel in neuronal activity suitable for a downstream spiking neural network. One so-called elementary motion unit (EMU) is implemented with two neurons and two synapses, plus one facilitation neuron that is shared among four EMUs. The time-to-travel is only encoded by the absolute number of spikes and not in the first ISI; thus, their system needs to wait until the respective stop neuron inhibits the direction-selective neuron. In contrast to Giulioni's work, a single sEMD needs only one neuron, one synapse, and one ASE circuit (three transistors and one capacitor). The sEMD encodes the time-to-travel partially by the number of spikes within a burst of activity; in addition, the ISIs

within a burst of the sEMD increase exponentially, and the first ISI already provides a coarse motion estimate.

In conclusion, this work proposes a novel event-based motion detection scheme, the spiking elementary motion detector. The sEMD circuit requires less silicon area to be implemented compared to other work and encodes the time-to-travel by the absolute number of spikes (providing precision), as well as by the ISI distribution (providing low latency). When a decision has to be made quickly, relative timing of the first two spikes can be used to generate a fast, coarse motion estimate. The motion can then be more accurately estimated over the next 10 ms from the absolute number of spikes within the burst.

5.2 State-of-the-Art Collision Avoidance. Our proposed spiking neural network with the sEMD model implemented to estimate optic flow from an event-based camera and a soft-inverse winner-take-all (WTA) population to estimate a collision avoidance direction works as well as conventional frame-based approaches (Meyer et al., 2016) in an open-loop setting. These results suggest that the proposed approach might also be useful in a closed-loop scenario, as it has been shown already for the frame-based approach (Meyer et al., 2016). However, closed-loop analog neuromorphic sensory-processing systems, in which the spiking activity directly affects the robot's steering behavior, have only recently been shown to be able to navigate in real-world scenarios (Milde, Blum et al., 2017; Milde, Dietmüller, Blum, Indiveri, & Sandamirskaya, 2017). A digital neuromorphic processor, SpiNNaker, has already been shown to successfully steer a robot and perform target acquisition and following (Denk et al., 2013). These studies used only simple hand-engineered features such as tracking a certain LED's flickering frequency or counted the number of spikes in certain sub-regions of the visual field to generate histograms, which determined the steering direction of the robot. A logical next step is to incorporate information about the objects' distance and use relative motion cues to steer the robot. We already implemented the sEMD model on the digital neuromorphic processor SpiNNaker (Furber et al., 2014; see supplementary information, S3) and obtained the first promising real-time sEMD array responses (data not shown). Additionally, it would be ideal to control the robot's steering with higher resolution than previous approaches. To this end, controlling the motors using pulse-frequency modulation (Perez-Peña, Leñero-Bardallo, Linares-Barranco, & Chicca, 2017), rather than a pulse-width modulation, would allow the network to directly control the robot's behavior without the need for time-averaging its output. Both real-time sEMD array implementation on SpiNNaker and motor control are subject to ongoing research.

Since mixed-signal neuromorphic sensory-processing systems can by some means interact with their environment (Denk et al., 2013; Milde, Blum

et al., 2017) and were shown to scale up nicely, while maintaining a low power consumption and a small size, these systems represent an interesting alternative to conventional frame-based solutions (Hwu, Isbell et al., 2017; Hwu, Krichmar et al., 2017). Furthermore, in the context of search and rescue operations, small and autonomous robotic agents are desired, since operation and processing time, as well as size, are three crucial constraints (Calisi et al., 2007; Ko & Lau, 2009).

5.3 Sensory-Domain Generalization. The sEMD responds to information inherent in the temporal difference of incoming stimuli. Such an encoding of Δt also appears in information processing in the brain, a prime example being the interaural time difference to localize the position of an auditory stimulus (Konishi, 1971; Finger & Liu, 2011).

To test generalization properties of the presented circuit, we connected the two inputs of the sEMD circuit, ASE and DPI, to a dynamic audio sensor (Chan, Liu, & van Schaik, 2007; Liu, Mesgarani, Harris, & Hermansky, 2010; Liu, Van Schaik, Mincti, & Delbruck, 2010). The time difference in incoming spikes originates from the two microphones on the audio sensor. After tuning the biases to fit the dynamic range of incoming stimuli ($10 \text{ ns} < \Delta t < 700 \mu\text{s}$) the sEMD could encode the input Δt into a burst of spikes and, furthermore, the sEMD activity was used to extract the position of a sound source using a network implemented on a SpiNNaker (Furber et al., 2014) board (data not shown).

5.4 Biological Plausibility. Our sEMD model exhibits performances (see Figure 10) similar to those of the well-established frame-based elementary motion detector (EMD) model proposed by Hassenstein and Reichardt (1956) more than 60 years ago. It has been recently shown by Borst and colleagues that many of the original predictions by Hassenstein and Reichardt were found to be valid, and large parts of the neural correlates of the EMD model have been identified in vivo (Mauss, Meier, Serbe, & Borst, 2014; Maisak et al., 2013; Borst & Helmstaedter, 2015). The original EMD model was used to explain motion detection mechanisms in the insect's visual pathway, which at the level of small local neurons in the peripheral visual system, as in the vertebrates retina, relies to a large extent on graded potentials to convey information. Whilst the classical EMD model explains how to estimate local motion based on graded potentials, it fails to explain how to extract motion from spiking inputs. In contrast, the sEMD model might provide insight into how to extract motion information if the underlying neural code relies on spikes, as are assumed to play a dominant role in the mammalian visual cortex. Here, we presented a biologically plausible solution of how such a computation could be performed using precise spike timing.

5.5 Conclusion. We have shown that our spiking elementary motion detector (sEMD), an event-based temporal correlation detector, can be used to asynchronously compute motion, estimate a collision-avoidance direction, or even localize the source of a sound.

Acknowledgments

We thank Matthew Cook for helpful discussions and comments on the manuscript, Stephen Nease for the FPGA firmware, Richard George for help with the PCB, and Thorben Schoepe and Philip Klein for preliminary simulations on SpiNNaker and support in the lab. We also thank Giacomo Indiveri for helpful comments on the manuscript, support with the bias tuning, and critical discussions. This work was supported by the Cluster of Excellence Cognitive Interaction Technology at Bielefeld University, which is funded by the DFG.

References

- Address Event Representation-Caltech-Memo (1993). *The address-event representation communication protocol AER 0.02* (Caltech internal memo).
- Barlow, H., & Levick, W. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, 178(3), 477–504.
- Bartolozzi, C., & Indiveri, G. (2007). Synaptic dynamics in analog VLSI. *Neural Computation*, 19(10), 2581–2603.
- Bartolozzi, C., & Indiveri, G. (2009). Global scaling of synaptic efficacy: Homeostasis in silicon synapses. *Neurocomputing*, 72(4–6), 726–731.
- Benosman, R., Clercq, C., Lagorce, X., Ieng, S.-H., & Bartolozzi, C. (2014). Event-based visual flow. *IEEE Transactions on Neural Networks and Learning Systems*, 25(2), 407–417.
- Benosman, R., Ieng, S.-H., Clercq, C., Bartolozzi, C., & Srinivasan, M. (2011). Asynchronous frameless event-based optical flow. *IEEE Transactions on Neural Networks*, 27, 32–37.
- Bertrand, O. J. N., Lindemann, J. P., & Egelhaaf, M. (2015). A bio-inspired collision avoidance model based on spatial information derived from motion detectors leads to common routes. *PLoS Computational Biology*, 11(11), e1004339.
- Borst, A., & Helmstaedter, M. (2015). Common circuit design in fly and mammalian motion vision. *Nature Neuroscience*, 18(8), 1067–1076.
- Brandli, C., Berner, R., Yang, M., Liu, S.-C., & Delbruck, T. (2014). A 240×180 130 dB 3 μ s latency global shutter spatiotemporal vision sensor. *IEEE Journal of Solid-State Circuits*, 49(10), 2333–2341.
- Brosch, T., Tschene, S., & Neumann, H. (2015). On event-based optical flow detection. *Frontiers in Neuroscience*, 9.
- Calisi, D., Farinelli, A., Iocchi, L., & Nardi, D. (2007). Multi-objective exploration and search for autonomous rescue robots. *Journal of Field Robotics*, 24(8–9), 763–777.

- Camunas-Mesa, L., Zamarreno-Ramos, C., Linares-Barranco, A., Acosta-Jimenez, A. J., Serrano-Gotarredona, T., and Linares-Barranco, B. (2012). An event-driven multi-kernel convolution processor module for event-driven vision sensors. *IEEE Journal of Solid-State Circuits*, 47(2), 504–517.
- Chan, V., Liu, S.-C., & van Schaik, A. (2007). AER EAR: A matched silicon cochlea pair with address event representation interface. *IEEE Transactions on Circuits and Systems I*, 54(1), 48–59.
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2016). *Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS*. arXiv:1606.00915.
- Clady, X., Clercq, C., Ieng, S.-H., Houseini, F., Randazzo, M., Natale, L., . . . Benosman, R. (2014). Asynchronous visual event-based time-to-contact. *Frontiers in Neuroscience*, 8(9).
- Conradt, J. (2015). On-board real-time optic-flow for miniature event-based vision sensors. In *Proceedings of the 2015 IEEE International Conference on Robotics and Biomimetics* (pp. 1858–1863). Piscataway, NJ: IEEE.
- Conradt, J., Cook, M., Berner, R., Lichtsteiner, P., Douglas, R. J., & Delbruck, T. (2009). A pencil balancing robot using a pair of aerodynamic vision sensors. In *Proceedings of the International Symposium on Circuits and Systems* (pp. 781–784). Piscataway, NJ: IEEE.
- Denk, C., Llobet-Blandino, F., Galluppi, F., Plana, L. A., Furber, S., & Conradt, J. (2013). Real-time interface board for closed-loop robotic tasks on the spinnaker neural computing system. In *International Conference on Artificial Neural Networks*. Vol. 8131 (pp. 467–474). Berlin: Springer.
- DVS. (2009). *Product of Inilabs: Dynamic vision sensor*. <http://www.inilabs.com/products/dynamic-vision-sensors>.
- Egelhaaf, M., Kern, R., & Lindemann, J. P. (2014). Motion as a source of environmental information: A fresh view on biological motion computation by insect brains. *Frontiers in Neural Circuits*, 8.
- Faessler, M., Fontana, F., Forster, C., Mueggler, E., Pizzoli, M., & Scaramuzza, D. (2016). Autonomous, vision-based flight and live dense 3D mapping with a quadrotor micro aerial vehicle. *Journal of Field Robotics*, 33(4), 431–450.
- Fasnacht, D., Whatley, A., & Indiveri, G. (2008). A serial communication infrastructure for multi-chip address event system. In *Proceedings of the International Symposium on Circuits and Systems* (pp. 648–651). Piscataway, NJ: IEEE.
- Finger, H., & Liu, S.-C. (2011). Estimating the location of a sound source with a spike-timing localization algorithm. In *Proceedings of the 2011 IEEE International Symposium on Circuits and Systems* (pp. 2461–2464). Piscataway, NJ: IEEE.
- Foster, K., Gaska, J. P., Nagler, M., & Pollen, D. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, 365(1), 331–363.
- Furber, S. B., Galluppi, F., Temple, S., & Plana, L. A. (2014). The SpiNNaker project. *Proceedings of the IEEE*, 102(5), 652–665.
- Galluppi, F., Denk, C., Meiner, M. C., Stewart, T. C., Plana, L. A., Eliasmith, C., . . . Conradt, J. (2014). Event-based neural computing on an autonomous mobile platform. In *Proceedings of the 2014 IEEE International Conference on Robotics and Automation* (pp. 2862–2867). Piscataway, NJ: IEEE.

- Gibson, J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Gibson, J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Giulioni, M., Lagorce, X., Galluppi, F., & Benosman, R. (2016). Event-based computation of motion flow on a neuromorphic analog neural platform. *Frontiers in Neuroscience*, 10(9).
- Goodman, D., & Brette, R. (2008). Brian: A simulator for spiking neural networks in Python. *Frontiers in Neuroinformatics*, 2(5), 1–10.
- Hassentstein, B., & Reichardt, W. (1956). Systemtheoretische analyse der Zeit-Reihenfolgen- und Vorzeichenauswertung bei der Bewegungspersonen des Rüsselkäfers chlorophanus. *Z. Naturforsch.*, 11b, 513–524.
- Hoffmann, R., Weikersdorfer, D., & Conradt, J. (2013). Autonomous indoor exploration with an event-based visual slam system. In *Proceedings of the 2013 European Conference on Mobile Robots* (pp. 38–43). Piscataway, NJ: IEEE.
- Horiuchi, T. K. (2009). A spike-latency model for sonar-based navigation in obstacle fields. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 56(11), 2393–2401.
- Horiuchi, T., Bair, W., Bishofberger, B., Lazzaro, J., & Koch, C. (1992). Computing motion using analog VLSI chips: An experimental comparison among different approaches. *International Journal of Computer Vision*, 8, 203–216.
- Horiuchi, T. K., Bishofberger, B., & Koch, C. (1994). An analog VLSI saccadic eye movement system. In J. D. Cowen, G. Tesaure, & J. Alspeter (Eds.), *Advances in neural information processing systems*, 6 (pp. 582–589). San Mateo, CA: Morgan Kaufmann.
- Horiuchi, T., Lazzaro, J., Moore, A., & Koch, C. (1991). A delay-line based motion detection chip. In D. Touretzky & R. Lippmann (Eds.), *Advances in neural information processing systems*, 3 (pp. 406–412). San Mateo, CA: Morgan Kaufmann.
- Hwu, T., Isbell, J., Oros, N., & Krichmar, J. (2017). A self-driving robot using deep convolutional neural networks on neuromorphic hardware. In *Proceedings of the International Joint Conference on Neural Networks* (pp. 635–641). Piscataway, NJ: IEEE.
- Hwu, T., Krichmar, J., & Zou, X. (2017). A complete neuromorphic solution to outdoor navigation and path planning. In *Proceedings of the 2017 IEEE International Symposium on Circuits and Systems* (pp. 1–4). Piscataway, NJ: IEEE.
- Indiveri, G. (2001). A current-mode hysteretic winner-take-all network, with excitatory and inhibitory coupling. *Analog Integrated Circuits and Signal Processing*, 28(3), 279–291.
- Indiveri, G., Chicca, E., & Douglas, R. J. (2006). A VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity. *IEEE Transactions on Neural Networks*, 17(1), 211–221.
- Indiveri, G., & Verschure, P. M. (1997). Autonomous vehicle guidance using analog VLSI neuromorphic sensors. In W. Gerstner, A. Germond, M. Hasler, & J.-D. Nicoud (Eds.), *Lecture Notes in Computer Science: Vol. 1327. Proceedings of the International Conference on Neural Networks* (pp. 811–816). Berlin: Springer Verlag.
- Klein, P., Conradt, J., & Liu, S.-C. (2015). Scene stitching with event-driven sensors on a robot head platform. In *Proceedings of the 2015 IEEE International Symposium on Circuits and Systems* (pp. 2421–2424). Piscataway, NJ: IEEE.
- Ko, A. W., & Lau, H. Y. (2009). Intelligent robot-assisted humanitarian search and rescue system. *International Journal of Advanced Robotic Systems*, 6(2), 12.

- Konishi, M. (1971). Sound localization in the barn owl. *Journal of the Acoustical Society of America*, 50(1A), 148–148.
- Kramer, J., Sarpeshkar, R., & Koch, C. (1995). An analog VLSI velocity sensor. In *Proc. IEEE Int. Symp. Circuit and Systems* (pp. 413–416). Piscataway, NJ: IEEE.
- Kramer, J., Sarpeshkar, R., & Koch, C. (1996). Analog VLSI motion discontinuity detectors for image segmentation. In *Proc. IEEE Int. Symp. Circuit and Systems* (pp. 620–623), Piscataway, NJ: IEEE.
- Kramer, J., Sarpeshkar, R., & Koch, C. (1997). Pulse-based analog VLSI velocity sensors. *IEEE Transactions on Circuits and Systems II*, 44(2), 86–101.
- Lelkens, A., & Koenderink, J. J. (1984). Illusory motion in visual displays. *Vision Research*, 24(9), 1083–1090.
- Lichtsteiner, P., Posch, C., & Delbruck, T. (2008). An 128×128 120 dB 15 μ s-latency temporal contrast vision sensor. *IEEE J. Solid State Circuits*, 43(2), 566–576.
- Liu, M., & Delbruck, T. (2017). *Block-matching optical flow for dynamic vision sensor-algorithm and FPGA implementation*. arXiv:1706.05415.
- Liu, S.-C., Mesgarani, N., Harris, J., & Hermansky, H. (2010). The use of spike-based representations for hardware audition systems. In *Proceedings of the 2010 IEEE International Symposium on Circuits and Systems* (pp. 505–508). Piscataway, NJ: IEEE.
- Liu, S.-C., Van Schaik, A., Mincti, B. A., & Delbruck, T. (2010). Event-based 64-channel binaural silicon cochlea with q enhancement mechanisms. In *Proceedings of the 2010 IEEE International Symposium on Circuits and Systems* (pp. 2027–2030). Piscataway, NJ: IEEE.
- Luber, D., Biedermann, J., & Conradt, J. (2015). Chain of small robots. Munich: Project Laboratory Computational Neuro Engineering, University of Munich.
- Lucas, B., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the Seventh Joint Conference on Artificial Intelligence* (pp. 674–679). San Mateo, CA: Morgan Kaufmann.
- Mahowald, M. (1992). *VLSI analogs of neuronal visual processing: A synthesis of form and function*. PhD diss., California Institute of Technology, Pasadena, CA.
- Mahowald, M. (1994). *An analog VLSI system for stereoscopic vision*. Boston: Kluwer.
- Maisak, M. S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., . . . Borst, A. (2013). A directional tuning map of drosophila elementary motion detectors. *Nature*, 500(7461), 212–216.
- Manen, S., Kwon, J., Guillaumin, M., & Van Gool, L. (2014). Appearances can be deceiving: Learning visual tracking from few trajectory annotations. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Proceedings of the European Conference on Computer Vision*: Vol. 8693 (pp. 157–172). Springer Berlin: Springer.
- Markram, H., Pikus, D., Gupta, A., & Tsodyks, M. (1998). Potential for multiple mechanisms, phenomena and algorithms for synaptic plasticity at single synapses. *Neuropharmacology*, 37(4), 489–500.
- Martel, J. N., Chau, M., Dudek, P., & Cook, M. (2015). Toward joint approximate inference of visual quantities on cellular processor arrays. In *Proceedings of the 2015 IEEE International Symposium on Circuits and Systems* (pp. 2061–2064). Piscataway, NJ: IEEE.
- Mauss, A. S., Meier, M., Serbe, E., & Borst, A. (2014). Optogenetic and pharmacologic dissection of feedforward inhibition in drosophila motion vision. *Journal of Neuroscience*, 34(6), 2254–2263.

- Mead, C. (1989). *Analog VLSI and neural systems*. Reading, MA: Addison-Wesley.
- Meyer, H. G., Bertrand, O. J., Paskarbit, J., Lindemann, J. P., Schneider, A., & Egelhaaf, M. (2016). A bio-inspired model for visual collision avoidance on a hexapod walking robot. In *Proceedings of the Conference on Biomimetic and Biohybrid Systems* (pp. 167–178). Berlin: Springer.
- Milde, M. B., Bertrand, O. J. N., Benosman, R., Egelhaaf, M., & Chicca, E. (2015). Bioinspired event-driven collision avoidance algorithm based on optic flow. In *Proceedings of the 2015 International Conference on Event-Based Control, Communication, and Signal Processing* (pp. 1–7). Piscataway, NJ: IEEE.
- Milde, M. B., Blum, H., Dietmüller, A., Sumislawska, D., Conradt, J., Indiveri, G., & Sandamirskaya, Y. (2017). Obstacle avoidance and target acquisition for robot navigation using a mixed signal analog/digital neuromorphic processing system. *Frontiers in Neurobotics*, 11, 28.
- Milde, M. B., Dietmüller, A., Blum, H., Indiveri, G., & Sandamirskaya, Y. (2017). Obstacle avoidance and target acquisition in mobile robots equipped with neuromorphic sensory-processing systems. In *Proceedings of the 2017 IEEE International Symposium on Circuits and Systems* (pp. 1–4). Piscataway, NJ: IEEE.
- Moeys, D. P., Corradi, F., Kerr, E., Vance, P., Das, G., Neil, D., . . . Delbrück, T. (2016). Steering a predator robot using a mixed frame/event-driven convolutional neural network. In *Proceedings of the 2016 Second International Conference on Event-Based Control, Communication, and Signal Processing* (pp. 1–8). Piscataway, NJ: IEEE.
- Mueller, M., Bertrand, O. J. N., Lindemann, J., & Egelhaaf, M. (2018). The problem of home choice in skyline-based homing. *PLoS One*, 13(3), e0194070.
- Nishida, S., Sasaki, Y., Murakami, I., Watanabe, T., & Tootell, R. B. (2003). Neuroimaging of direction-selective mechanisms for second-order motion. *Journal of Neurophysiology*, 90(5), 3242–3254.
- Nozaki, Y., & Delbruck, T. (2017). Temperature and parasitic photocurrent effects in dynamic vision sensors. *IEEE Transactions on Electron Devices*, 64(8), 3239–3245.
- O’Keefe, L. P., & Movshon, J. A. (1998). Processing of first-and second-order motion signals by neurons in area MT of the macaque monkey. *Visual Neuroscience*, 15(2), 305–317.
- Partzsch, J., Höppner, S., Eberlein, M., Schüffny, R., Mayr, C., Lester, D. R., & Furber, S. (2017). A fixed point exponential function accelerator for a neuromorphic many-core system. In *Proceedings of the 2017 IEEE International Symposium on Circuits and Systems* (pp. 1091–1094). Piscataway, NJ: IEEE.
- Pavasović, A., Andreou, A., & Westgate, C. (1994). Characterization of subthreshold MOS mismatch in transistors for VLSI systems. *Journal of VLSI Signal Processing*, 8(1), 75–85.
- Pelgrom, M., Duinmaijer, A., & Welbers, A. (1989). Matching properties of MOS transistors. *IEEE Journal of Solid-State Circuits*, 24(5), 1433–1440.
- Perez-Peña, F., Leñero-Bardallo, J. A., Linares-Barranco, A., & Chicca, E. (2017). Towards bioinspired close-loop local motor control: A simulated approach supporting neuromorphic implementations. In *Proceedings of the 2017 IEEE International Symposium on Circuits and Systems*. Piscataway, NJ: IEEE.

- Perrone, J. A., & Thiele, A. (2002). A model of speed tuning in MT neurons. *Vision Research*, 42(8), 1035–1051.
- Posch, C., Matolin, D., & Wohlgenannt, R. (2010). A QVGA 143 dB dynamic range asynchronous address-event pwm dynamic image sensor with lossless pixel-level video compression. In *Proceedings of the International Solid-State Circuits Conference Digest of Technical Papers* (pp. 400–401). Piscataway, NJ: IEEE.
- Posch, C., Serrano-Gotarredona, T., Linares-Barranco, B., & Delbruck, T. (2014). Retinomorph event-based vision sensors: Bioinspired cameras with spiking output. *Proceedings of the IEEE*, 102(10), 1470–1484.
- Priebe, N. J., Lisberger, S. G., & Movshon, J. A. (2006). Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *Journal of Neuroscience*, 26(11), 2941–2950.
- Raggedstone2. (2017). *Raggedstone2 with spartan 6 FPGA*. <https://www.enterpoint.co.uk/products/spartan-6-development-boards/raggedstone-2/>
- Rahtu, E., Kannala, J., Salo, M., & Heikkilä, J. (2010). Segmenting salient objects from images and videos. In K. Danilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision—ECCV 2010* (pp. 366–379). Berlin: Springer.
- Ramachandran, H., Weber, S., Aamir, S. A., & Chicca, E. (2014). Neuromorphic circuits for short-term plasticity with recovery control. In *Proceedings of the 2014 IEEE International Symposium on Circuits and Systems* (pp. 858–861). Piscataway, NJ: IEEE.
- Rokszin, A., Márkus, Z., Braunitzer, G., Berényi, A., Benedek, G., & Nagy, A. (2010). Visual pathways serving motion detection in the mammalian brain. *Sensors*, 10(4), 3218–3242.
- Rueckauer, B., & Delbruck, T. (2016). Evaluation of event-based algorithms for optical flow with ground-truth from inertial measurement sensor. *Frontiers in Neuroscience*, 10(176).
- Serres, J. R., & Ruffier, R. (2017). Optic flow-based collision-free strategies: From insects to robots. *Arthropod Structure and Development*, 46(5), 703–717.
- Son, B., Suh, Y., Kim, S., Jung, H., Kim, J.-S., Shin, C., . . . Ryu, H. (2017). A 640×480 dynamic vision sensor with a $9 \mu\text{m}$ pixel and 300 meps address-event representation. In *Proceedings of the 2017 IEEE International Conference on Solid-State Circuits* (pp. 66–67). Piscataway, NJ: IEEE.
- Stefanini, F., Neftci, E. O., Sheik, S., & Indiveri, G. (2014). Pyncs: A microkernel for high-level definition and configuration of neuromorphic electronic systems. *Frontiers in Neuroinformatics*, 8(73).
- Theobald, J. C., Duistermars, B. J., Ringach, D. L., & Frye, M. A. (2008). Flies see second-order motion. *Current Biology*, 18(11), R464–R465.
- Thorpe, S., Delorme, A., & Van Rullen, R. (2001). Spike-based strategies for rapid processing. *Neural Networks*, 14(6), 715–725.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- VanRullen, R., Guyonneau, R., & Thorpe, S. J. (2005). Spike times make sense. *Trends in Neurosciences*, 28(1), 1–4.
- Weinzaepfel, P., Revaud, J., Harchaoui, Z., & Schmid, C. (2013). Deepflow: Large displacement optical flow with deep matching. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1385–1392). Piscataway, NJ: IEEE.

- Yang, M., Liu, S.-C., & Delbruck, T. (2015). A dynamic vision sensor with 1% temporal contrast sensitivity and in-pixel asynchronous delta modulator for event encoding. *IEEE Journal of Solid-State Circuits*, 50(9), 2149–2160.
- Zingg, S., Scaramuzza, D., Weiss, S., & Siegwart, R. (2010). MAV navigation through indoor corridors using optical flow. In *Proceedings of the 2010 IEEE International Conference on Robotics and Automation* (pp. 3361–3368). Piscataway, NJ: IEEE.

Received November 9, 2017; accepted April 19, 2018.