

# Bio-Inspired Optic Flow from Event-Based Neuromorphic Sensor Input

Stephan Tschechne, Roman Sailer, and Heiko Neumann

Inst. for Neural Information Processing, Ulm University,  
Albert-Einstein-Allee 1, 89069 Ulm, Germany  
{stephan.tschechne,roman.sailer,heiko.neumann}@uni-ulm.de  
<http://www.uni-ulm.de/in/neuroinformatik.html>

**Abstract.** Computational models of visual processing often use frame-based image acquisition techniques to process a temporally changing stimulus. This approach is unlike biological mechanisms that are spike-based and independent of individual frames. The neuromorphic Dynamic Vision Sensor (DVS) [Lichtsteiner et al., 2008] provides a stream of independent visual events that indicate local illumination changes, resembling spiking neurons at a retinal level. We introduce a new approach for the modelling of cortical mechanisms of motion detection along the dorsal pathway using this type of representation. Our model combines filters with spatio-temporal tunings also found in visual cortex to yield spatio-temporal and direction specificity. We probe our model with recordings of test stimuli, articulated motion and ego-motion. We show how our approach robustly estimates optic flow and also demonstrate how this output can be used for classification purposes.<sup>1</sup>

**Keywords:** Event-Vision, Optic Flow, Neural Model, Classification.

## 1 Introduction

Event-based representation of visual information differs from a frame-based approach by providing a continuous stream of changes in the recorded image instead of full frames at fixed intervals. This approach provides numerous benefits: For one, local changes are indicated with very low latency and temporal resolution is immensely increased. The response to local relative luminance changes allows the construction of sensors with an extremely high dynamics range. Simultaneously, data output is reduced to relevant information only, ignoring image regions where no changes occur. Neural modellers profit from the biologically inspired transmission of information based on events or spikes, which allows new and more realistic models of visual processing, because the visual system also operates with a series of spikes generated by light-sensitive retinal neurons. This neural activity is integrated by neurons at subsequent processing stages

---

<sup>1</sup> This work has been supported by grants from the Transregional Collaborative Research Center SFB/TRR62 'A Companion Technology for Cognitive Technical Systems' funded by the German Research Foundation (DFG).

utilizing increasingly larger receptive fields to build feature representations of higher complexity. In the dorsal pathway of visual processing in cortex, neurons are tuned to patterns of motion in the visual field while they are tuned to static form patterns along the ventral pathway [Ungerleider and Haxby, 1994]. [Lichtsteiner et al., 2008, Liu and Delbrück, 2010] present a neuromorphic approach for a vision sensor that models the mechanisms of spike-generation of a mammalian retina. Their Dynamic Vision Sensor (DVS) generates events whenever a local change of illumination occurs without being bound to temporal frames. Events are marked with the spatial position, phase of illumination change (*on* or *off*) and time-stamp of appearance. This approach provides an energy-efficient and temporally very accurate method to represent temporal changes.

In this work, we propose a novel approach to estimate the apparent motion of visual features (optic flow) using the events generated by the DVS. We model a simplified version of the initial stages of cortical processing along the visual dorsal pathway in primates. In particular, initial responses are generated to represent movements in the spatio-temporal domain, corresponding to V1 direction sensitive cells. This introduces a new approach how biologically inspired models of motion estimation process input.

## 2 Previous Work

Several research investigations took advantage of the high-speed response properties of event-based sensor design for applications in various domains [Lichtsteiner et al., 2008]. For example, in an assisted living scenario a bank of such sensors can successfully be utilized to detect and classify fast vertical motions as unintended occasional falls, indicative of an preemptive alarm situation [Fu et al., 2008]. Tracking scenarios have been investigated to steer real-world interaction [Delbrück, 2012] as well as in microbiology set-ups [Drazen et al., 2011]. Basic research investigations have been reported as well. For example, stereo matching approaches have been suggested that consider partial depth estimation on the basis of calibrated cameras and the associated epipolar geometry constraints [Rogister et al., 2011]. Using monocular event-based sensing different approaches to optical flow computation have been investigated. In [Benosman et al., 2012] a standard optical flow integration mechanism has been applied [Lucas and Kanade, 1981]. Their approach approximates spatial and temporal derivatives from the event stream to estimate a least-squares solution of flow vectors from intersection of constraint lines. The results show precise tracking of sample inputs, like a bouncing ball or waving hands. More recently, the authors propose to estimate flow directly by evaluating the event-cloud in the spatio-temporal space [Benosman et al., 2014]. In a nutshell, the surface geometry of a small patch of a surface fitted to the cloud of sensor events is evaluated. The local gradient information directly yields an estimate of the ratio of space-time changes and, thus, a local speed measure. In sum, these approaches to motion detection rely on mechanisms that eliminate outliers in order to get reliable speed estimates, either indirectly or directly. [Benosman et al., 2014] rely

on local fitting of the data by planar patches such that, movement components orthogonal to the grey value structure can be measured (aperture problem).

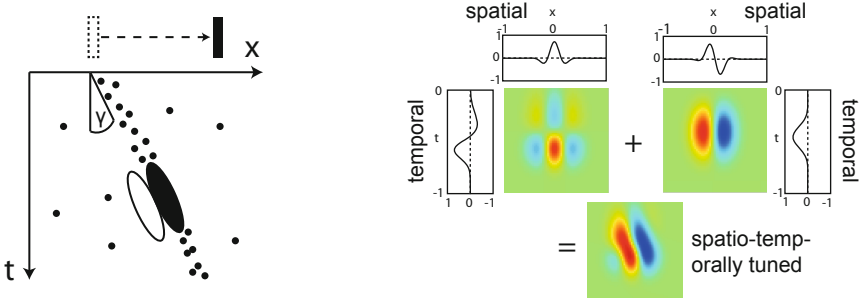
We, here, suggest an alternative approach to initial motion detection that is motivated by the first stages of spatio-temporal processing in biological vision. We make use of data about non-directional and directional cells in primary visual cortex and how these generate a first representation of local movements [De Valois et al., 2000]. Below, we outline the general mechanism that makes full use of the event-based input. We demonstrate that the approach is capable to generate spatio-temporal flow estimates for extended boundaries and for intrinsic features. Such estimates can already be used to classify different motion patterns in space and time without the need of sophisticated processing.

### 3 Methods

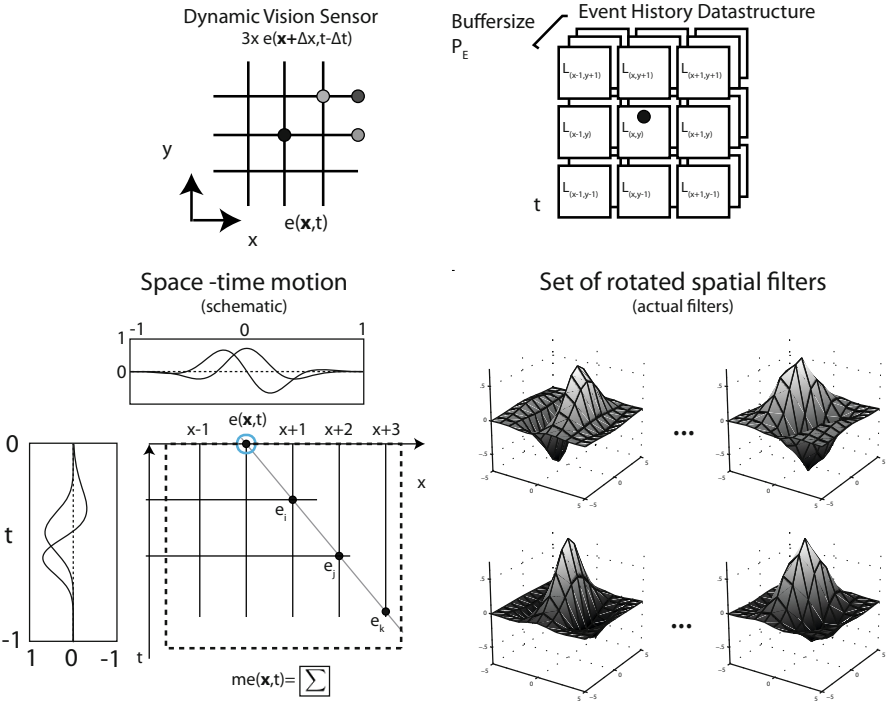
**Estimation of Optical Flow Using Spatio-Temporal Filters.** Spatio-temporal changes of different speed generate oriented structure in x-y-t-space with varying off-axis angles as measured against the temporal axis. These spatio-temporal structures need to be analysed in order to estimate optic flow. The method that we propose herein is inspired by mechanisms found in visual cortex. Our contribution is the adaption of these principles to be compatible with event-based representation of visual events.

Movement detection in visual cortex is based on cells that are either directionally selective or non-selective [De Valois et al., 2000]. Directional selectivity in cortical cells is generated by linear combination of spatial even-symmetric bi-phasic and odd-symmetric mono-phasic cells as measured orthogonal to the locally oriented contrast. Based on theoretical considerations, [Adelson and Bergen, 1985] have argued in favor of combining cells with even and odd spatial symmetry each having different temporal band-pass characteristics. The final calculation of the response energies lead to demonstrate that such filtering is formally equivalent to spatio-temporal correlation detectors as proposed by [Hassenstein and Reichard, 1956] as a model of motion detection in the fly visual system. We focus here on the findings of [De Valois et al., 2000] and apply a set of spatio-temporal filters on the event stream to yield a selectivity to different motion speeds and directions while maintaining the sparse representation provided by the address-event coding. Figure 2 shows the principle of this estimation process. Two spatial filter functions of different class (even and odd-symmetric) and two temporal filter functions (monophasic and biphasic) are combined to build two different spatio-temporal filters. Those can be added to generate a spatio-temporally tuned filter that is able to respond to spatio-temporal structures. Parameters of the filters contribute to the speed- and direction selectivity of the resulting filter. In the following we describe how such filters are modelled and applied to the event-based representation.

**Event Representation and Data Structure.** A conventional camera projects light onto an image sensor (CCD or CMOS) and reads out the measurement of all light- and colour-sensitive pixels at a fixed frame rate. Several limitations concerning the representation can be identified: The temporal sampling rate is limited, thus



**Fig. 1.** *Left:* In an event based representation, an object leaves a trail of events when moving across the visual array, along with some spurious events. Depending on the speed, the resulting structure is slanted with angle  $\gamma$  towards the  $t$  axis. *Right:* Filters tuned to structures can be generated using two spatially and two temporally tuned functions, motivated by recordings of individual cells in visual cortex.



**Fig. 2.** *Top:* The DVS produces set of events which are stored in an Event History Datastructure for efficient processing. *Bottom:* Spatial and temporal filters are applied to yield sensitivity for motion caused by events.

when all pixel values are sampled in a fixed time interval, fast local changes are integrated and thus lost for further processing. The photo sensitivity of the sensor is limited, thus different illumination conditions force the camera system to globally adapt the operating range using different exposure times or apertures. Finally, from the perspective of sampling the plenoptic function  $P(x, y, \lambda, t, V_{-x}, V_{-y}, V_{-z})$  [Adelson and Bergen, 1991] the corresponding continuous volume is regularly sampled along the  $t$ -axis, irrespective of the structure along this dimension. In case that no changes happen to occur redundant information is generated that is carried to the subsequent processing steps.

In an *event-based* system individual pixels generate events in the moment they measure a change in contrast. A single DVS event encodes the characteristic phase of luminance change, namely *on* or *off*, the pixel coordinates of its origin and a timestamp (in microseconds). In this way only luminance changes in the observed scene are recorded and bandwidth can be saved while the dynamic range can be increased [Lichtsteiner et al., 2008]. This happens with temporal sampling rate that is orders of magnitude higher than those of a conventional camera. The way it represents the captured information is referred to as address-event representation [Lichtsteiner et al., 2008].

For the discrete implementation, we suggest the following simple buffering structure to achieve an event-based sampling: On a basis of the  $x, y$ -resolution of the sensor ( $128 \times 128$  in case of the DVS) lists with space for  $P_E$  timestamps are managed in a first in, first out principle. Whenever a DVS event  $e(\mathbf{x}, t)$  (with  $\mathbf{x} = (x, y)$ ) occurs, its timestamp  $t$  is pushed to the list  $L_{(x,y)}$  at the reference grid-position. In case that  $P_E$  elements are already held in  $L_{(x,y)}$ , the last (and hence, oldest) event is pruned. The buffering structure subsequently holds the last  $P_E$  timestamps per DVS pixel in a chronological order at each point of time.

**Generation of Spatio-Temporal Filters.** To model spatial filters, we use a population of rotated two-dimensional Gabor functions, that are calculated using the common definition:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\pi \left[ \frac{(\hat{x} - x_0)^2}{\sigma^2} + \frac{(\hat{y} - y_0)^2}{\sigma^2} \right]\right) \cdot \exp(i[\xi_0 x + \xi_0 y]) \quad (1)$$

$$\text{with } \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2)$$

Even and odd components of this function are used and parameterized to generate  $11 \times 11$  filter kernels ( $x, y \in [-2\pi..2\pi]$ ) using standard deviation  $\sigma = 2.5$ , frequency tuning  $\xi_0 = 2\pi$  and rotation  $\theta = [0, \frac{1}{4}\pi, \dots, \pi]$ . Figure 2 depicts some of the filters used. Temporal filter functions are generated using the following equation:

$$f_t = w_{m1} \cdot \exp\left(-\frac{x - \mu_{m1}^2}{2\sigma_{m1}^2}\right) - w_{m2} \cdot \exp\left(-\frac{x - \mu_{m1}^2}{2\sigma_{m2}^2}\right) \quad \text{for } x \in [0..1] \quad (3)$$

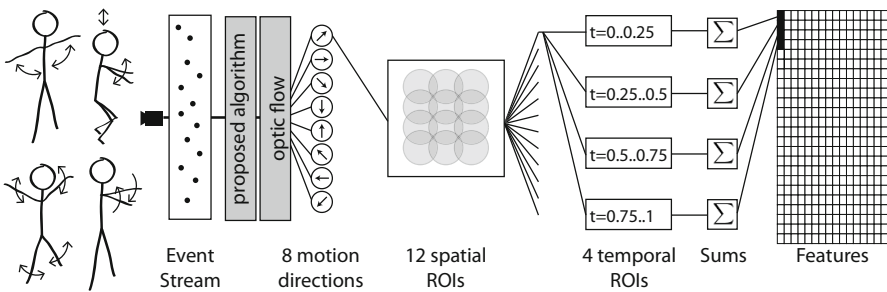
For monophasic temporal kernels  $T_1$  these values are set to [ $w_{m1} = 1.95$ ,  $\mu_{m1} = 0.55$ ,  $\sigma_{m1} = 0.10$ ,  $w_{m2} = 0.23$ ,  $\mu_{m1} = 0.55$ ,  $\sigma_{m2} = 0.16$ ]. For biphasic

temporal kernels  $T_2$  the values are set to  $w_{m1} = 0.83$ ,  $\mu_{m1} = 0.44$ ,  $\sigma_{m1} = 0.12$ ,  $w_{m2} = -0.34$ ,  $\mu_{m2} = 0.63$ ,  $\sigma_{m2} = 0.21$ . The resulting spatio-temporal filter is then generated by  $F = G \cdot T_1 + G \cdot T_2$  and applied to incoming events. Due to the sparse representation, a full convolution with the filter is not necessary. Weights are only calculated for the events in the spatial and temporal neighborhood of the newest event, thus saving computational power.

For each application of a spatio-temporal filter the model generates a confidence for the occurrence of this motion direction and speed  $C_{s,\theta}$ . To maintain a single motion hypothesis, we generate a weighted sum of the fundamental  $N$  directions using the estimated confidence value:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \sum_{s=1}^N \sum_{\theta=0}^{2\pi} C_{s,\theta} \cdot \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix} \quad (4)$$

**Flow Features for Action Recognition.** To demonstrate the usability of event-based optic flow features for classification purposes, we perform a small classification scenario. We train classifiers on features extracted from optic flow for the classification of sequences of articulated motions. Our feature extraction roughly follows concepts proposed by [Escobar and Kornprobst, 2012]. Optic flow is estimated using the methods developed in this paper, directly resulting in hypotheses for 16 motion directions, which are binned into 8 directions for classification. From this input, 12 spatial and 4 temporal regions of interest (ROIs) are selected and flow is summed in these regions to calculate motion available motion energy. We perform a principal component analysis (PCA) on the feature vectors to identify the dominant dimensions of the distribution. Energy responses are projected onto the first 10 eigenvectors for a more compact representation. We trained four support vector machines (SVM), one for each class using 32 training samples.



**Fig. 3.** Left: Four actions used for classification demo: *Butterfly*, *Jumpfrog*, *JumpJack* and *Harvester*. Right: Schematic of feature extraction. For the classification we extract 384 feature values from the sequences. Eight directional parts are extracted from the motion estimation stage. These are subsequently split into 12 spatial and 4 temporal ROIs, before a summation is performed.

## 4 Results

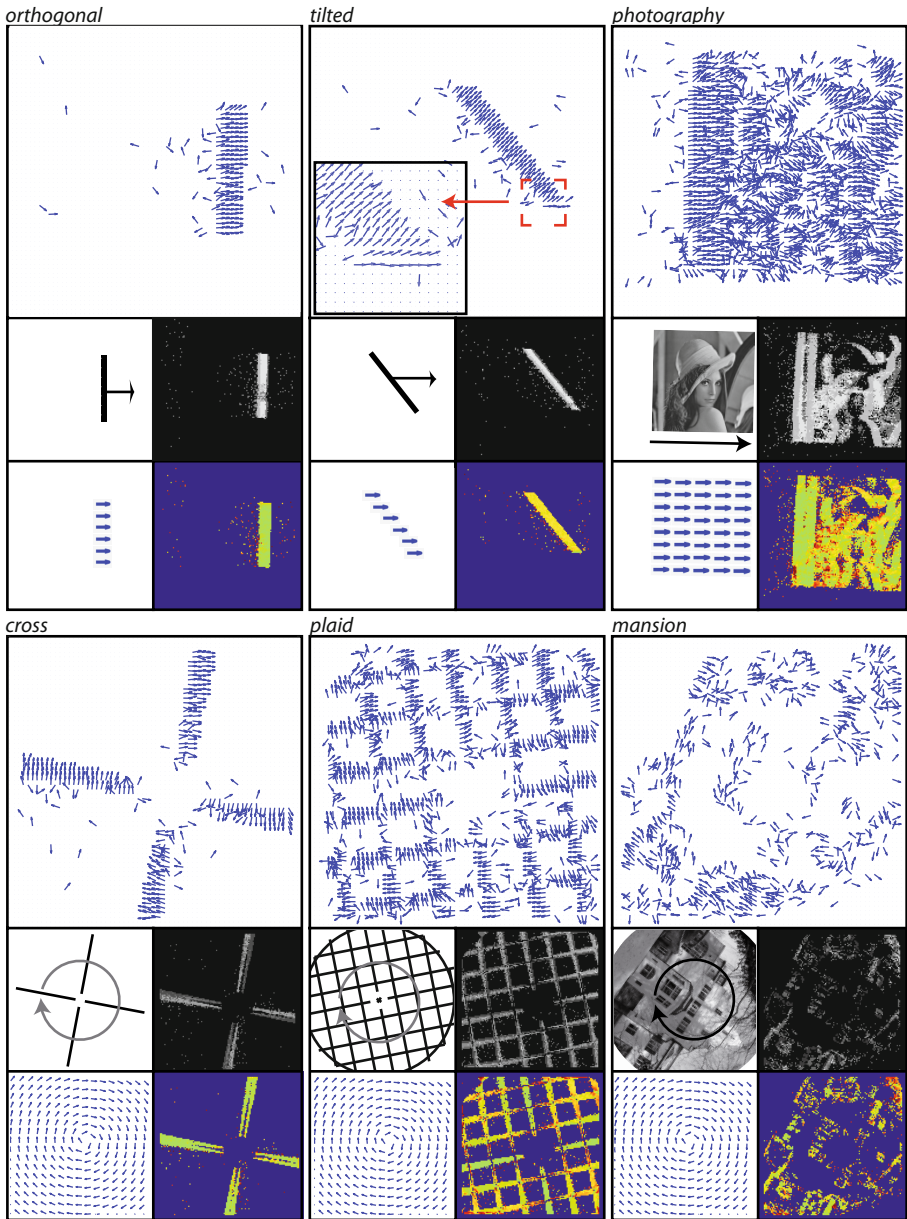
We recorded a set of different stimuli to test the performance of our model. Probes were mounted on simple mechanical devices that produced linear or rotational motion. The DVS sensor was mounted on a tripod and the sensor surface aligned parallel and centered to the motion plane. The recorded event streams were caused by translational and rotational motion as well as articulated motion. The sensor as well as our algorithm does not process complete frames but individual sensor events. Individual events cannot sensibly be visualized in printed form. For this reason we integrate estimated motion events, however, over a short period of time. All results presented here used an integration window of  $T = 50 \cdot 10^{-3}s$ .  $T$  therefore lies within a range of commonly used integration times (per frame) used in video technologies which is long enough to ensure that also tiny structures produce enough DVS events for a rich flow visualization. Where adequate we calculated an error measure for our results as follows: With the type of motion known, we synthesize a ground truth vector field of a linear or rotational motion for each sequence. The estimated error is the angular error between  $0^\circ$  and  $180^\circ$  between the synthetic vector field and the estimated motion direction at this position.

Motion events are estimated whenever a new event is processed in our model, in contrast to frame-based approaches where all pixels are processed for one synchronised point in time. For illustration purposes, our model offers the freedom to decide how often a visual output should be generated, ranging from outputs visualising every motion event to an output integrating the events of a complete sequence.

**Translatory Motion.** We selected three stimuli for tests with a translatory motion, which we label *in italic letters* for future identification: (i) a black bar that moved *orthogonal* to its longer side, (ii) a *tilted* black bar that was rotated  $45^\circ$  relative to its movement direction and (iii) a natural *photography*. The movement speed was approximately  $170mm \cdot sec^{-1}$ . Ground truth of this sequence is a constant motion to the right. Results of our tests for translatory motion are displayed in Fig. 4, *Top*. In the figure, we displayed a sketch of the stimulus, the events used for the current motion estimation, ground truth and errors, with the large image depicting the estimated motion vectors.

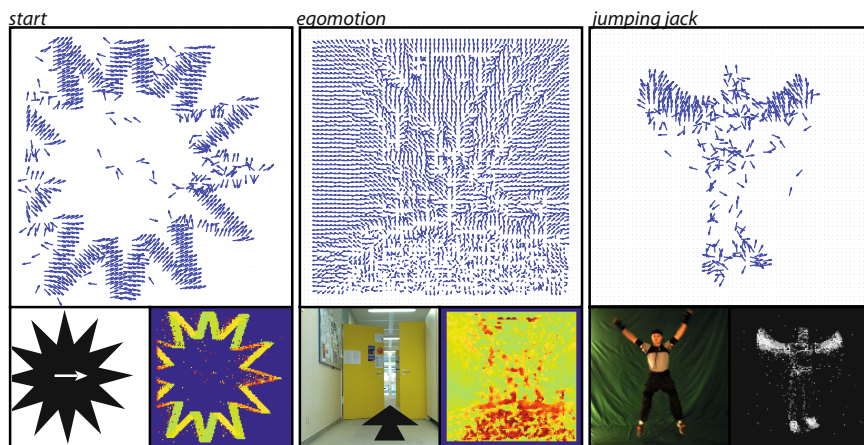
In *orthogonal*, the results correctly show motion vectors to the right. In the *tilted* case the results shows that our mechanism suffers from the aperture problem because it processes motion events only in a local surround. Motion is only estimated in a direction orthogonal to the contrast orientation in the absence of intrinsically two-dimensional features. However, our algorithm produces correct motion vectors at the ends of the bar. The movement of *photography* shows correct motion estimations but also spurious ones, which are caused by locally tilted contrasts and aliasing problems due to the low resolution of the sensor. The error measure reveals that many estimations are still correct.

**Rotational Motion.** The second test set was used to acquire events caused by rotational motion. The stimuli were rotated with approx.  $85deg \cdot sec^{-1}$  using a



**Fig. 4.** Processing results for six example stimuli. Each image contains the results of flow detection (*large image*), a stimulus sketch, events used for the estimation, ground truth and illustration of error. (*small images, from top left to bottom right*). *Top*: Three test cases for translational motion: A bar moving *orthogonal* to its orientation, a *tilted* bar and a *photographic* image. *Bottom*: Rotational motion for three test cases: *cross*, *plaid* and *mansion*. Angular errors were estimated using an artificially generated ground truth in all cases.





**Fig. 5.** Test cases for complex motion. *Left:* Correct motion direction is estimated at line endings and normal flow at lines in a complex translational pattern. *Middle:* Results for *ego-motion*. With increased distance to the center, flow quality increases. *Right:* Snippets from a sequence containing articulated motion.

DC motor with adjustable speed. The image plane of the sensor was parallel to the stimulus in a distance of  $150\text{mm}$ . We selected two artificial stimuli, a rotating *plaid* and a rotating *cross* as well as a natural photograph of a *mansion*. We also generated a rotational vector field as ground truth. The results for these patterns are more complex and desire more explanations. The arms of the rotating *cross* move orthogonally to their orientation, which is depicted in the figure. The structures of the moving *plaid* are excentric and thus motion estimation locally again suffers from the aperture problem at some positions. The rotating photograph of the *mansion* is challenging for our algorithm due to the high rotational speed and the large number of produced visual events. Our model nevertheless achieves to estimate a number of correct estimates.

**Complex Motion Sequences.** The first sequence highlights the ability of our algorithm to estimate to real flow for local image features. We generate sequences of a moving *star*. At the line endings, the correct direction of optic flow is estimated while normal flow is estimated along the boundaries.

For the second sequence, the event sensor was mounted on a mobile platform and pushed across a hallway to imitate *egomotion*. Forward motion with parallel optical axis causes an expansion field of optic flow, with the focus of expansion centered in the image, see Fig. 5. Our results show that our model can estimate such an expansion field reliably. However, the low spatial resolution of the sensor seems inadequate to differentiate small expanding features close to the focus of expansion. Here, local transitions of width cause spurious local expansion patterns. To overcome these challenges we integrated motion events over the complete sequence length. We applied a  $3 \times 3$  median filter on the results of this sequence.

The third sequence depicts a complex motion pattern of multiple speeds and directions. Here, a person is performing a *jumping jack* action, consisting of huge and fast movements of arms and legs. This causes a complex motion pattern with multiple directions and speeds.

**Classification Results.** To round up our line of arguments, we show how we use the optic flow output for classification of simple actions. We picked 64 recordings of people performing four different actions using the event based sensor and used 32 of them for training. Tests were performed on the remaining sequences. When none or more than one classifier voted for its class, we assigned one class randomly. The evaluation was repeated 50 times. Table 1 shows the classification results we achieved with our approach. We emphasize that even with this simple approach, the classes are nicely separated. We want to highlight that our main contribution is the biologically motivated estimation of optic flow, and the classification is just a case study to demonstrate the richness of the extracted features.

**Table 1.** Confusion matrix for a simple action recognition experiment using the estimated optic flow algorithm. We trained four classifiers on four different actions to demonstrate the feasibility of the approach.

	Detection			
↓Signal	<i>butterfly</i>	<i>jumpfrog</i>	<i>jumpjack</i>	<i>harvester</i>
<i>butterfly</i> Action	100.0%	0.0%	0.0%	0.0%
<i>jumpfrog</i> Action	5.5%	73.0%	6.75%	14.75%
<i>jumpjack</i> Action	7.0%	0.0%	93.0%	12.5%
<i>harvester</i> Action	0.0%	0.0%	0.0%	100.0%

## 5 Discussion

In this work we presented a biologically inspired mechanism to estimate optic flow using an event-based neuromorphic vision sensor. Our model implements mechanisms of motion detection reminiscent of early stages of processing in visual cortex. Initial motion hypotheses are estimated by detectors tuned to a set of directions. Their responses are integrated at a subsequent level to yield a tuning in the spatio-temporal domain that allows estimation of motion in different directions with increased robustness. Many approaches exist for detecting and integrating motion, among those are biologically motivated [Hassenstein and Reichard, 1956, Adelson and Bergen, 1985] and other algorithms of error minimisation [Lucas and Kanade, 1981, Brox et al., 2004]. Our model differs from other approaches in a number of ways. First, we make use of the event-based representation of visual input instead of processing input in a frame-based manner. This allows our model to generate responses with short latency to the input, while other frame-based models process input at fixed intervals given by the frame rate. In time-critical applications like robotics or

mechatronics, immediate availability of flow signals clearly is beneficial. Second, our model profits from the sparse representation that comes with an event-based approach and only processes flow where visual changes occur in the image, leaving those regions unconsidered in the first place where no optic flow is expected. Third, our model incorporates detectors that allow a temporal tuning function with an increased temporal resolution compared to a frame-based approach and thus allows to model the temporal characteristics with increased precision.

Our model was evaluated with a series of stimuli and exhibited the desired behaviour. The model robustly estimated visual motion in a series of test cases with an identical parameter set. Along elongated visual features, the model estimates normal flow due to the aperture problem, but at line endings it generates the correct motion direction under certain conditions. These correct local motion estimates have earlier been used to solve the aperture problem in a model of motion integration [Bayerl and Neumann, 2004] and the mechanisms used therein can be used for our model as well. However, the limited spatial resolution of the sensor handicaps the estimation of accurate optic flow in conditions where shapes are sampled having small and many visual features and complex motions. In many cases, sampling artefacts generate spurious estimations. This problem is however not a conceptual one of our approach and will most likely be solved with future versions of the vision sensor. Optic flow can be used to acquire an enriched feature set because it provides the trace of temporal correspondences of image features. It thus serves well for the extraction of more meaningful features from a scene containing motion and has been used in many domains like action recognition or navigational problems. With the event based processing approach, our models builds the basis for more sophisticated processing models where even the early processing stages have a high level of biological plausibility. We also briefly demonstrated the power of motion feature extraction for classifying actions from articulated motions. The classification test case should be seen as demonstration and feasibility study, a serious evaluation would undoubtedly contain more than 64 samples in four classes.

## References

- Adelson and Bergen, 1991. Adelson, E.H., Bergen, J.H.: The plenoptic function and the elements of early vision. In: Landy, M., Movshon, J.A. (eds.) *Computational Models of Visual Processing*, pp. 3–20. MIT Press (1991)
- Adelson and Bergen, 1985. Adelson, E.H., Bergen, J.R.: Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am.* 2(2), 284–299 (1985)
- Bayerl and Neumann, 2004. Bayerl, P., Neumann, H.: Disambiguating visual motion through contextual feedback modulation. *Neural Computation* 16(10), 2041–2066 (2004)
- Benosman et al., 2014. Benosman, R., Clercq, C., Lagorce, X., Ieng, S.H., Bartolozzi, C.: Event-Based visual flow. *IEEE Trans. on Neural Networks and Learning Systems* 25(2), 407–417 (2014)
- Benosman et al., 2012. Benosman, R., Ieng, S.-H., Clercq, C., Bartolozzi, C., Srinivasan, M.: Asynchronous frameless event-based optical flow. *Neural Networks* 27, 32–37 (2012)

- Brox et al., 2004. Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: High Accuracy Optical Flow Estimation Based on a Theory for Warping. In: Pajdla, T., Matas, J.(G.) (eds.) ECCV 2004. LNCS, vol. 3024, pp. 25–36. Springer, Heidelberg (2004)
- De Valois et al., 2000. De Valois, R., Cottaris, N.P., Mahon, L.E., Elfar, S.D., Wilson, J.A.: Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. *Vision Research* 40, 3685–3702 (2000)
- Delbrück, 2012. Delbrück, T.: Fun with asynchronous vision sensors and processing. In: Fusiello, A., Murino, V., Cucchiara, R. (eds.) ECCV 2012 Ws/Demos, Part I. LNCS, vol. 7583, pp. 506–515. Springer, Heidelberg (2012)
- Drazen et al., 2011. Drazen, D., Lichtsteiner, P., Häflinger, P., Delbrück, T., Jensen, A.: Toward real-time particle tracking using an event-based dynamic vision sensor. *Exp. Fluids* 51, 1465–1469 (2011)
- Escobar and Kornprobst, 2012. Escobar, M.J., Kornprobst, P.: Action recognition via bio-inspired features: The richness of center-surround interaction. *Computer Vision and Image Understanding* 116(5), 593–605 (2012)
- Fu et al., 2008. Fu, Z., Delbrück, T., Lichtsteiner, P., Culurciello, E.: An address-event fall detector for assisted living applications. *IEEE Trans. on Biomedical Circuits and Systems* 2(2), 88–96 (2008)
- Hassenstein and Reichard, 1956. Hassenstein, B., Reichard, W.: Functional structure of a mechanism of perception of optical movement. In: *Proc. 1st Intl. Congress on Cybernetics*, pp. 797–801 (1956)
- Lichtsteiner et al., 2008. Lichtsteiner, P., Posch, C., Delbrück, T.: A 128×128 120dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor. *IEEE Journal of Solid-State Circuits* 43(2), 566–576 (2008)
- Liu and Delbrück, 2010. Liu, S.C., Delbrück, T.: Neuromorphic Sensory Systems. *Current Opinion in Neurobiology* 20, 288–295 (2010)
- Lucas and Kanade, 1981. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: *Proceedings of the 7th International Joint Conference on Artificial Intelligence, IJCAI*, pp. 674–679 (1981)
- Rogister et al., 2011. Rogister, P., Benosman, R., Ieng, S.H., Posch, C.: Asynchronous event-based binocular stereo matching. *IEEE Transactions on Neural Networks* 22(11), 1723–1734 (2011)
- Ungerleider and Haxby, 1994. Ungerleider, L.G., Haxby, J.V.: ‘what’ and ‘where’ in the human brain. *Current Opinion in Neurobiology* 4(2), 157–165 (1994)