Supplementary Notes for Neuromorphic Optical Flow and Real-time Implementation with Event Cameras

Yannick Schnider[1,2], Stanisław Woźniak[1], Mathias Gehrig[3], Jules Lecomte[4], Axel von Arnim[4], Luca Benini[2,5], Davide Scaramuzza[3], Angeliki Pantazi[1]

[1]IBM Research – Zurich [2]ETH Zurich [3]University of Zurich [4]fortiss GmbH [5]Università di Bologna

1. WAEE metric definiton

The weighted average end point error (WAEE) combines the AEEs of the four sequences:

- outdoor_day1 (od1)
- indoor_flying1 (if1)
- indoor_flying2 (if2)
- indoor_flying3 (if3)

and is defined as:

$$\text{WAEE} = \left(\frac{\text{AEE}_{\text{od1}}}{w_{\text{od1}}} + \frac{\text{AEE}_{\text{if1}}}{w_{\text{if1}}} + \frac{\text{AEE}_{\text{if2}}}{w_{\text{if2}}} + \frac{\text{AEE}_{\text{if3}}}{w_{\text{if3}}}\right)/4.$$

The weights w_{od1}, w_{if1}, w_{if2}, w_{if3} are the sequence specific average AEEs of the spiking EV-FlowNet variants: LIF, ALIF, PLIF and XLIF for the modes $\text{dt} = 1$ and $\text{dt} = 4$ published in [1]:

$$\begin{aligned}
&\text{dt} = 1: \\
w_{\text{od1}} &= (0.53 + 0.57 + 0.60 + 0.45)/4.0 &= 0.5375 \\
w_{\text{if1}} &= (0.71 + 1.00 + 0.75 + 0.73)/4.0 &= 0.7975 \\
w_{\text{if2}} &= (1.44 + 1.78 + 1.52 + 1.45)/4.0 &= 1.5475 \\
w_{\text{if3}} &= (1.16 + 1.55 + 1.23 + 1.17)/4.0 &= 1.2775 \\
&\text{dt} = 4: \\
w_{\text{od1}} &= (2.02 + 2.13 + 2.24 + 1.67)/4.0 &= 2.0150 \\
w_{\text{if1}} &= (2.63 + 3.81 + 2.80 + 2.72)/4.0 &= 2.9900 \\
w_{\text{if2}} &= (4.93 + 6.40 + 5.21 + 4.93)/4.0 &= 5.3675 \\
w_{\text{if3}} &= (3.88 + 5.53 + 4.12 + 3.91)/4.0 &= 4.3600
\end{aligned} \tag{1}$$

2. Additional comparison

Table 1 includes an extended comparison with additional prior art non-spiking models. In particular, $\text{EV-FlowNet}_{\text{PM}}$ [3] was trained in comparable setting to ours, but used a photometric loss (PM). The results had been only reported for $\text{dt} = 1$ mode. Furthermore, several prior art architectures were trained in a different setup using directly the MVSEC dataset, as opposed to our architectures that were trained on the UZH-FPV Drone Racing Dataset and evaluated on the MVSEC dataset. Results for models trained directly on MVSEC, delimited by dashed lines, include:

- $\text{EV-FlowNet}_{\text{PM-MVSEC}}$ [4], trained in a self-supervised manner with the photometric loss (PM),
- $\text{EV-FlowNet}_{\text{CM-MVSEC}}$ [5], trained in a self-supervised manner with a contrast maximisation loss (CM),
- $\text{Hybrid-EV-FlowNet}_{\text{MVSEC}}$ [2], trained in a self-supervised manner with the photometric loss.

Considering the extended comparison with non-spiking ANN prior art models, the $\text{EV-FlowNet}_{\text{CM-MVSEC}}$ [5] yields the best performance on all MVSEC sequences for $\text{dt} = 1$ with regard to WAEE and percentage of outliers. Its WAEE of 0.67 is 8.2% lower than 0.73 of our sSNU-Timelens. In turn, the $\text{Hybrid-EV-FlowNet}_{\text{MVSEC}}$ [2] is outperformed by our sSNU-Timelens by 26.0% (0.73 vs. 0.92).

However, when evaluating in mode $\text{dt} = 4$, the $\text{Hybrid-EV-FlowNet}_{\text{MVSEC}}$ [2] yields the best overall performance with an WAEE of 0.68 compared to our sSNU-Timelens with 0.71 (+4.4%). The sSNU-Timelens shows on par performance in terms of WAEE with the $\text{EV-FlowNet}_{\text{CM-MVSEC}}$ [5] (also 0.71) in this mode.

In summary, the $\text{EV-FlowNet}_{\text{CM-MVSEC}}$ [5] and the $\text{Hybrid-EV-FlowNet}_{\text{MVSEC}}$ [2] perform best for MVSEC evaluations with $\text{dt} = 1$ and $\text{dt} = 4$, respectively. Remarkably, our sSNU-Timelens is a runner-up in both cases, despite being trained without access to the examples from the MVSEC dataset.

	outdoor_day1		indoor_flying1		indoor_flying2		indoor_flying3		overall	
dt = 1	AEE	$\%_{\text{Out.}}$	AEE	$\%_{\text{Out.}}$	AEE	$\%_{\text{Out.}}$	AEE	$\%_{\text{Out.}}$	WAEE	$\overline{\%}_{\text{Out.}}$
LIF-EV-FlowNet [1]	0.53	0.33	0.71	1.41	1.44	12.75	1.16	9.11	0.93	5.90
XLIF-EV-FlowNet [1]	0.45	**0.16**	0.73	$\underline{0.92}$	1.45	12.18	1.17	8.35	0.90	5.40
LIF-FireNet [1]	0.57	0.40	0.98	2.48	1.77	16.40	1.50	12.81	1.15	8.02
PLIF-FireNet [1]	0.56	0.38	0.90	1.93	1.67	14.47	1.41	11.17	1.10	7.00
our SNN-Timelens	$\underline{0.44}$	0.18	$\underline{0.70}$	**0.79**	$\underline{1.30}$	$\underline{9.41}$	$\underline{1.05}$	$\underline{6.00}$	$\underline{0.84}$	$\underline{4.10}$
our SNUo-Timelens	**0.39**	$\underline{0.17}$	**0.64**	0.96	**1.17**	**7.71**	**0.96**	**4.92**	**0.76**	**3.44**
$\text{EV-FlowNet}_{\text{PM}}$ [3]	0.92	5.4	0.79	1.2	1.40	10.9	1.18	7.4	1.13	6.23
EV-FlowNet [1]	0.47	0.25	$\underline{0.60}$	$\underline{0.51}$	$\underline{1.17}$	$\underline{8.06}$	$\underline{0.93}$	5.64	0.78	$\underline{3.61}$
RNN-EV-FlowNet [1]	0.56	1.09	0.62	0.97	1.20	8.82	$\underline{0.93}$	$\underline{5.51}$	0.83	4.10
our sSNU-Timelens	$\underline{0.36}$	$\underline{0.10}$	**0.58**	0.56	1.19	8.78	0.96	6.11	$\underline{0.73}$	3.89
$\text{EV-FlowNet}_{\text{PM-MVSEC}}$ [4]	0.49	0.20	1.03	2.20	1.72	15.10	1.53	11.90	1.13	7.35
$\text{EV-FlowNet}_{\text{CM-MVSEC}}$ [5]	**0.32**	**0.00**	**0.58**	**0.00**	**1.02**	**4.00**	**0.87**	**3.00**	**0.67**	**1.75**
$\text{Hybrid-EV-FlowNet}_{\text{MVSEC}}$ [2]	0.49	-	0.84	-	1.28	-	1.11	-	0.92	-
dt = 4										
LIF-EV-FlowNet [1]	2.02	18.91	2.63	29.55	4.93	51.10	3.88	41.49	0.92	35.26
XLIF-EV-FlowNet [1]	1.67	12.69	2.72	31.69	4.93	51.36	3.91	42.52	0.89	34.57
LIF-FireNet [1]	2.12	21.00	3.72	48.27	6.27	64.16	5.23	58.43	1.17	47.97
PLIF-FireNet [1]	2.11	20.64	3.44	44.02	5.94	64.02	4.98	57.53	1.11	46.55
our SNN-Timelens	$\underline{1.65}$	$\underline{11.03}$	$\underline{2.61}$	$\underline{29.40}$	$\underline{4.50}$	$\underline{50.87}$	$\underline{3.58}$	$\underline{40.22}$	$\underline{0.84}$	$\underline{32.88}$
our SNUo-Timelens	**1.44**	**8.98**	**2.36**	**24.18**	**3.98**	**44.71**	**3.25**	**36.01**	**0.75**	**28.47**
EV-FlowNet [1]	1.69	12.50	$\underline{2.16}$	$\underline{21.51}$	3.90	**40.72**	**3.00**	**29.60**	0.74	$\underline{26.08}$
RNN-EV-FlowNet [1]	1.91	16.39	2.23	22.10	4.01	41.74	$\underline{3.07}$	$\underline{30.87}$	0.78	27.78
our sSNU-Timelens	1.34	$\underline{7.99}$	**2.15**	**20.92**	3.97	$\underline{41.31}$	3.17	32.44	$\underline{0.71}$	**25.67**
$\text{EV-FlowNet}_{\text{PM-MVSEC}}$ [4]	$\underline{1.23}$	**7.30**	2.25	24.70	4.05	45.30	3.45	39.70	0.73	29.25
$\text{EV-FlowNet}_{\text{CM-MVSEC}}$ [5]	1.30	9.70	2.18	24.20	$\underline{3.85}$	46.80	3.18	47.80	$\underline{0.71}$	32.13
$\text{Hybrid-EV-FlowNet}_{\text{MVSEC}}$ [2]	**1.09**	-	2.24	-	**3.83**	-	3.18	-	**0.68**	-

Table 1. Extended evaluation on MVSEC: AEE (the lower, the better ↓), the percentage of outliers $\%_{\text{Out.}}$(↓) per sequence, and the overall WAEE(↓) as defined in Eq. 1 as well as the average percentage of outliers $\overline{\%}_{\text{Out.}}$(↓). Best scores are in bold, while runner-ups are underlined. Horizontal lines delimit the spiking and the non-spiking models. Dashed line delimits not directly comparable prior art setups.

References

[1] Jesse Hagenaars, Federico Paredes-Vallés, and Guido de Croon. Self-supervised learning of event-based optical flow with spiking neural networks. *Advances in Neural Information Processing Systems*, 34, 2021. 1, 2

[2] Chankyu Lee, Adarsh Kosta, Alex Zhu, Kenneth Chaney, Kostas Daniilidis, and Kaushik Roy. *Spike-FlowNet: Event-Based Optical Flow Estimation with Energy-Efficient Hybrid Neural Networks*, pages 366–382. Oct 2020. 1, 2

[3] F. Paredes-Vallés and Guido C.H.E. de Croon. Back to event basics: Self-supervised learning of image reconstruction for event cameras via photometric constancy. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3445–3454, 2020. 1, 2

[4] Alex Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. EV-FlowNet: Self-supervised optical flow estimation for event-based cameras. In *Robotics: Science and Systems XIV*. Robotics: Science and Systems Foundation, Jun 2018. 1, 2

[5] Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. Unsupervised event-based learning of optical flow, depth and egomotion. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1694–1694, 2019. 1, 2