Recent advances in neuromorphic vision sensors: A survey

Jianing Li

lijianing@pku.edu.cn

https://jianing-li.github.io/

Multimedia Learning Group

School of EE & CS, Peking University

Dec.18th , 2019

Overview

- Introduction
  - Conventional camera
  - Biological visual sensing system
  - Neuromorphic vision milestones
  - Neuromorphic vision Sensors
- Event-based signal processing
  - Spatial-temporal filter
  - Spike metric
  - Spike coding
- Feature representations (Applications)
  - Rate-based image
  - Hand-crafted feature
  - End-to-end deep network
  - Spiking neural network
- Discussion
  - Better input representations for event data
  - Event-based vision in the future

Overview

- Introduction
  - Conventional camera
  - Biological visual sensing system
  - Neuromorphic vision milestones
  - Neuromorphic vision Sensors
- Event-based signal processing
  - Spatial-temporal filter
  - Spike metric
  - Spike coding
- Feature representations (Applications)
  - Rate-based image
  - Hand-crafted feature
  - End-to-end deep network
  - Spiking neural network
- Discussion
  - Better input representations for event data
  - Event-based vision in the future

Conventional camera

- Disadvantage of conventional camera
  - **Over-sampling, data redundancy**
  - **Under-sampling, motion blur**
  - **Low dynamic range**
  - **High power**

Too expensive

16000FPS

118,000 $

Short storage

720p frame only

Storing 10s

Light requirement

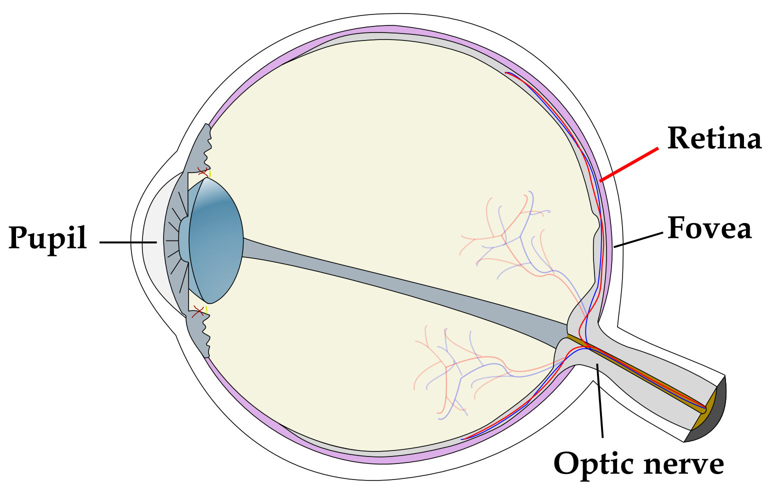
Exposure time $< 60\mu s$

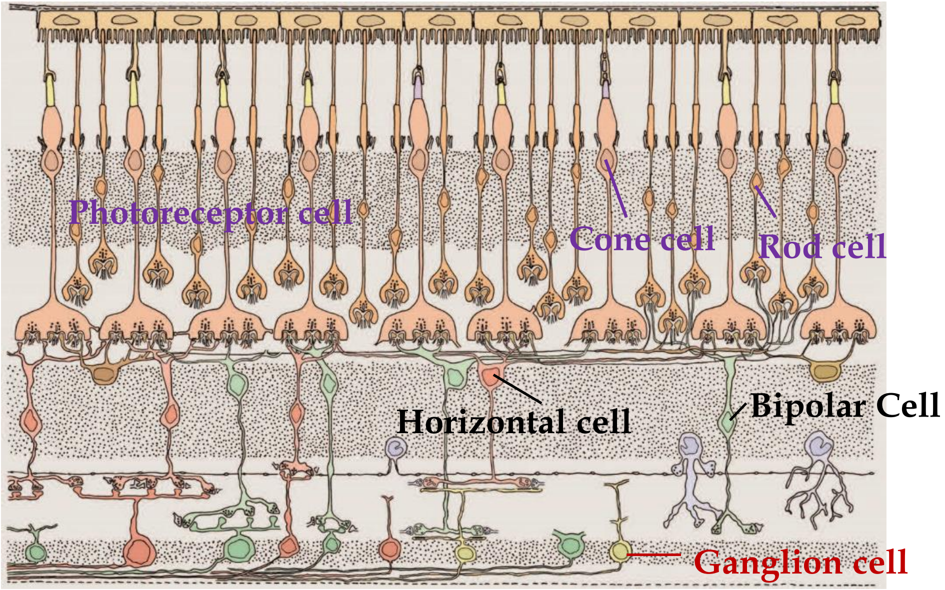
High illuminance

~~High-speed motion~~

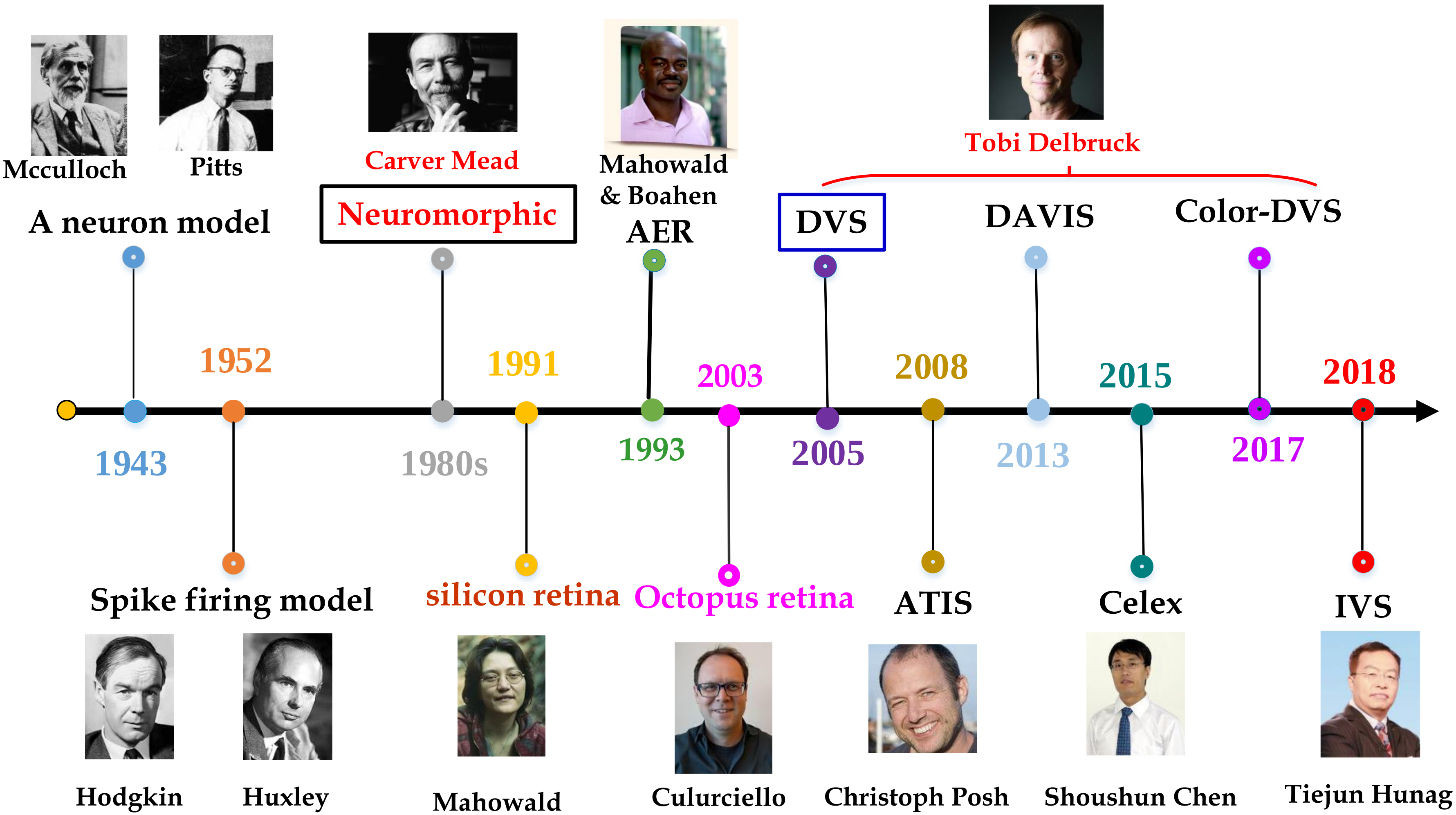
Biological visual sensing system

- Retina sampling
  - Three layers structures
  - Fovea – visual texture
  - Peripheral – high time resolution





Neuromorphic vision milestones

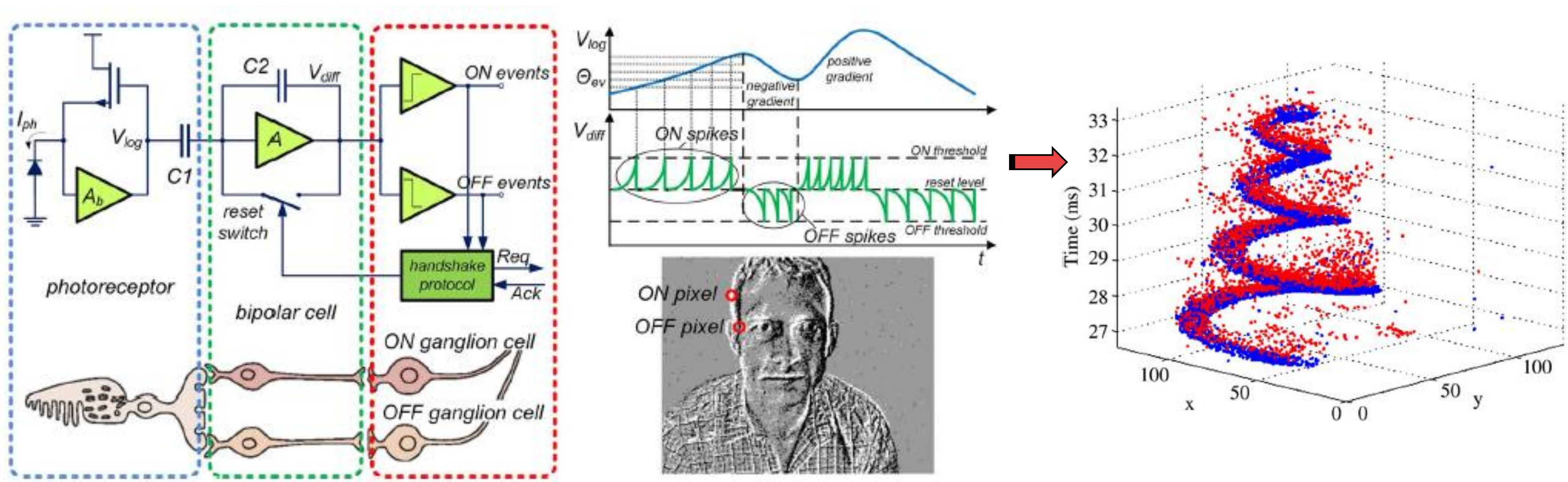


Neuromorphic sampling

- Differential sampling model
  - **Address event representation (AER)**
  - **Illumination change**

$$\Delta L \doteq \ln L(u,t) - \ln L(u,t-\Delta t) = p\theta$$

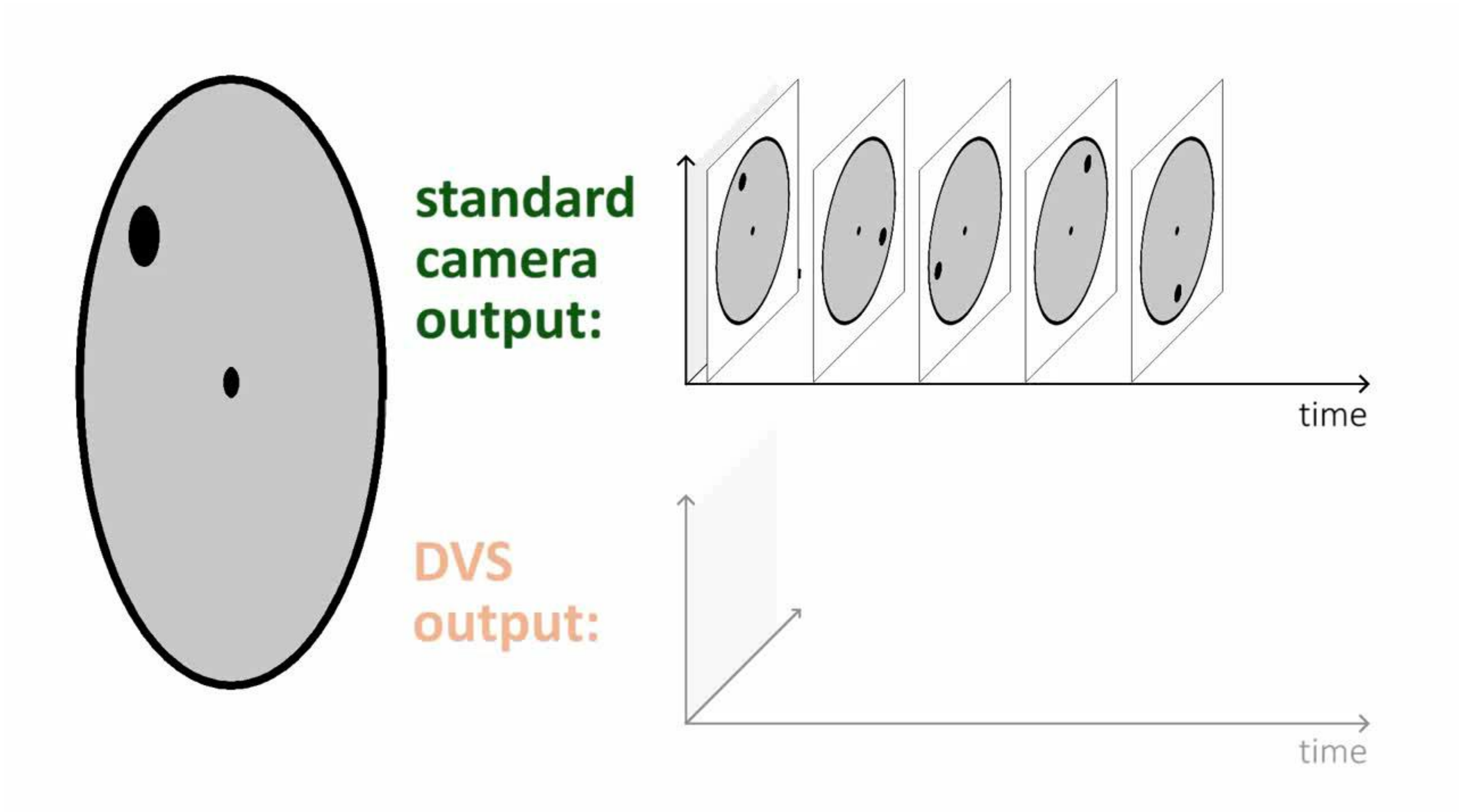
  - **DVS, DAVIS, ATIS, CeleX**



C Posch et al., Retinomorphic Event-Based Vision Sensors: Bioinspired Cameras With Spiking Output. *Proceedings of IEEE*, 2014.

Neuromorphic vision sensors

- Dynamic vision sensor (DVS) – Differential sampling



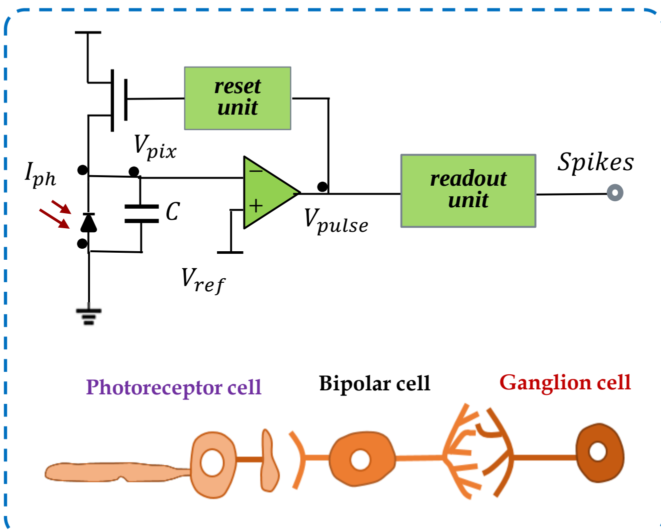
Neuromorphic vision sensors

Integral sampling model

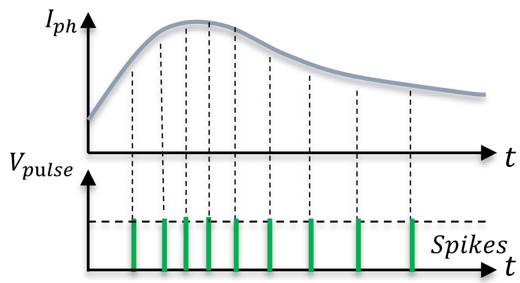
- **Leaky integrate and fire (LIF) model**

$$A(t) = \int_0^t I(t)dt \geq \varphi$$

- **Octopus retina, FSM**

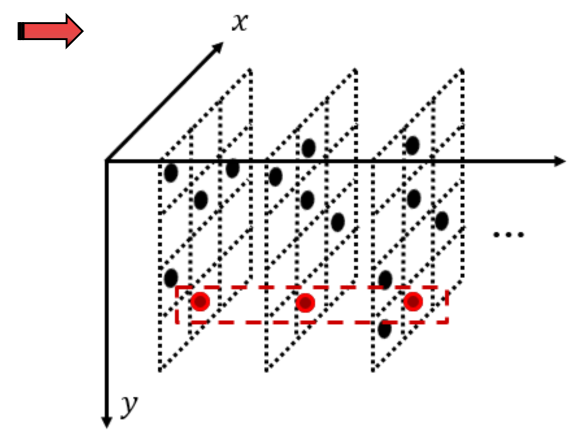


FSM



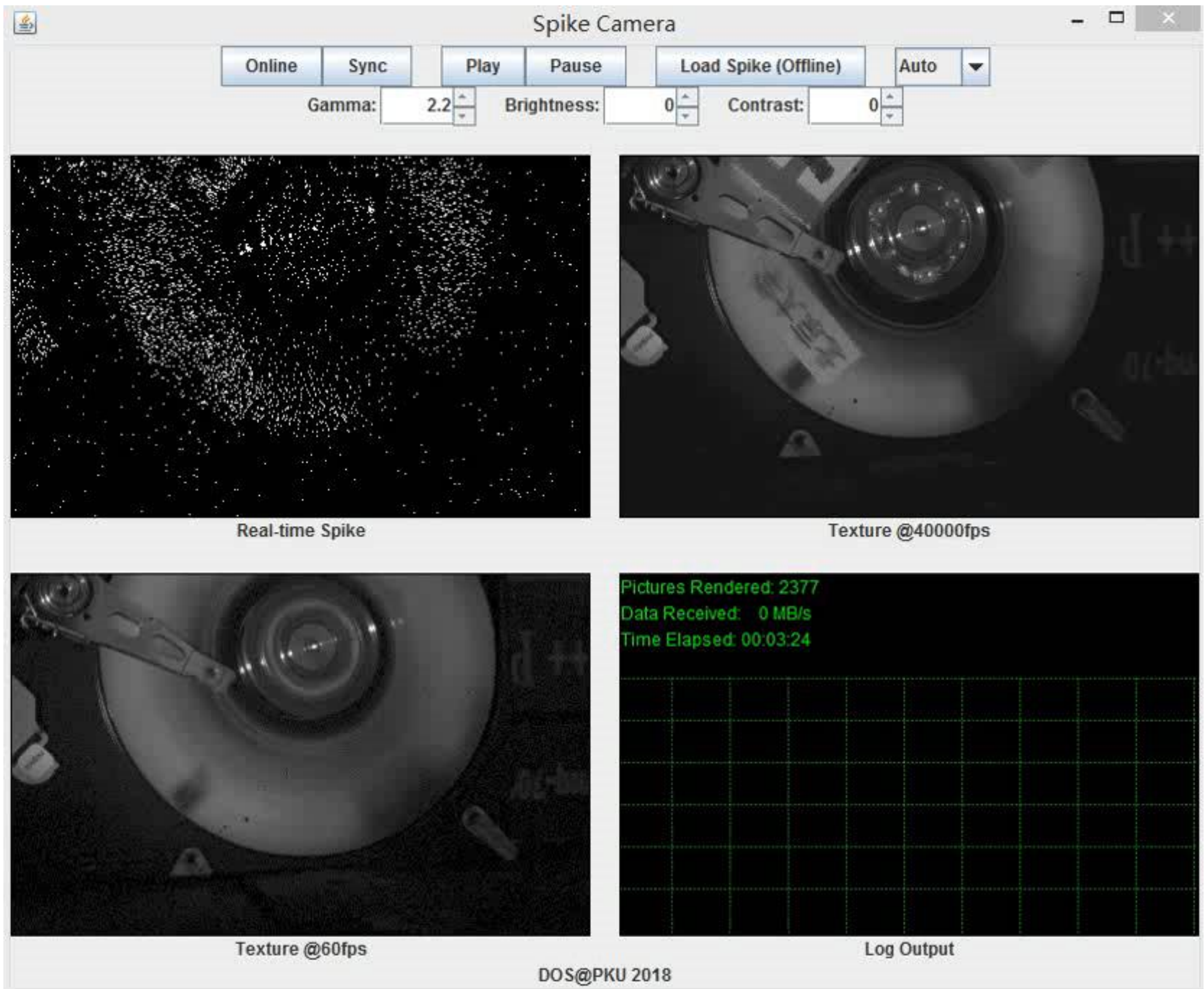
Integral firing

Texture image



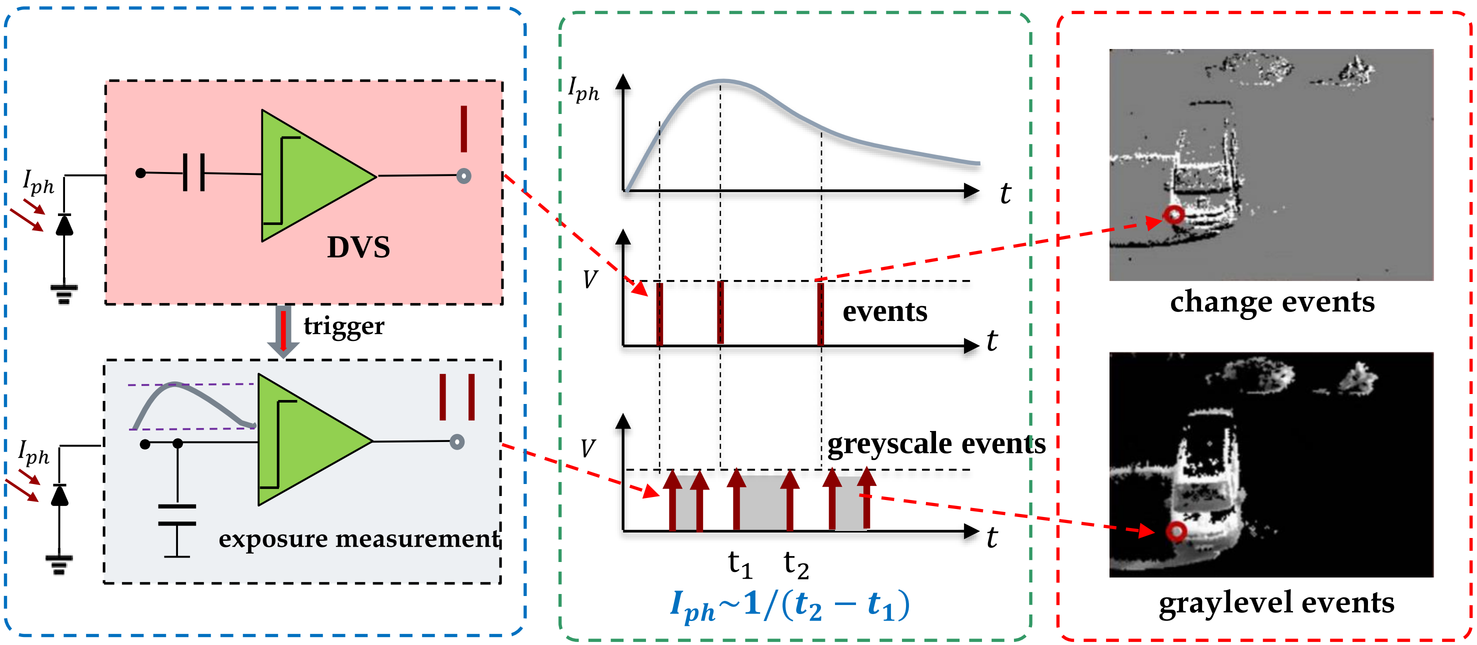
Neuromorphic vision sensors

- ☐ **Fovea-like sampling model (FSM) – Integral sampling**



Neuromorphic vision sensors

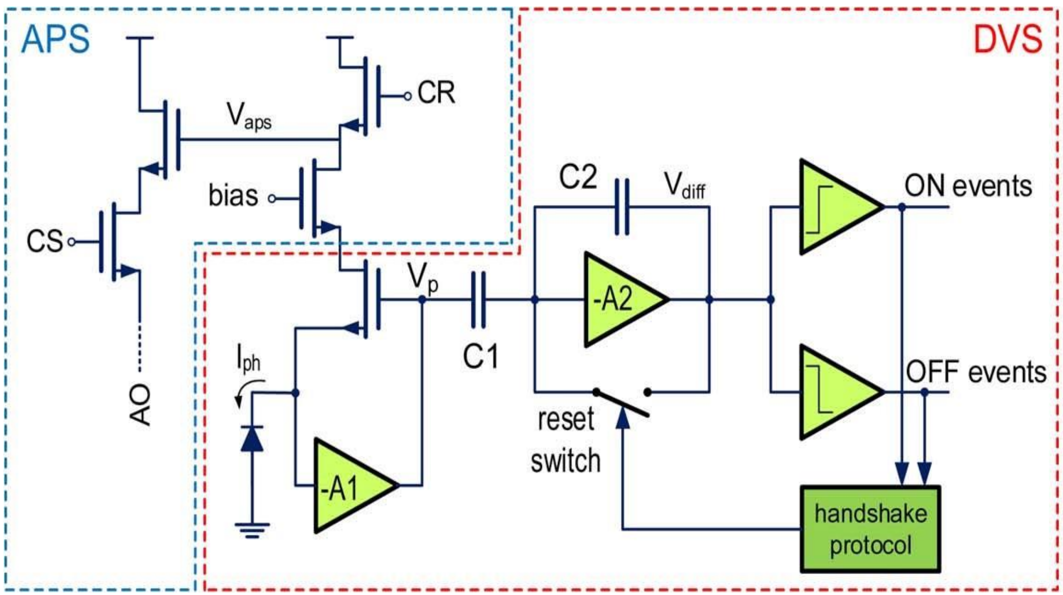
- Asynchronous time-based image sensor (ATIS)
  - DVS: change detector
  - Time-based: greyscale events



C Posch et al., Retinomorphic Event-Based Vision Sensors: Bioinspired Cameras With Spiking Output. *Proceedings of IEEE*, 2014.

Neuromorphic vision sensors

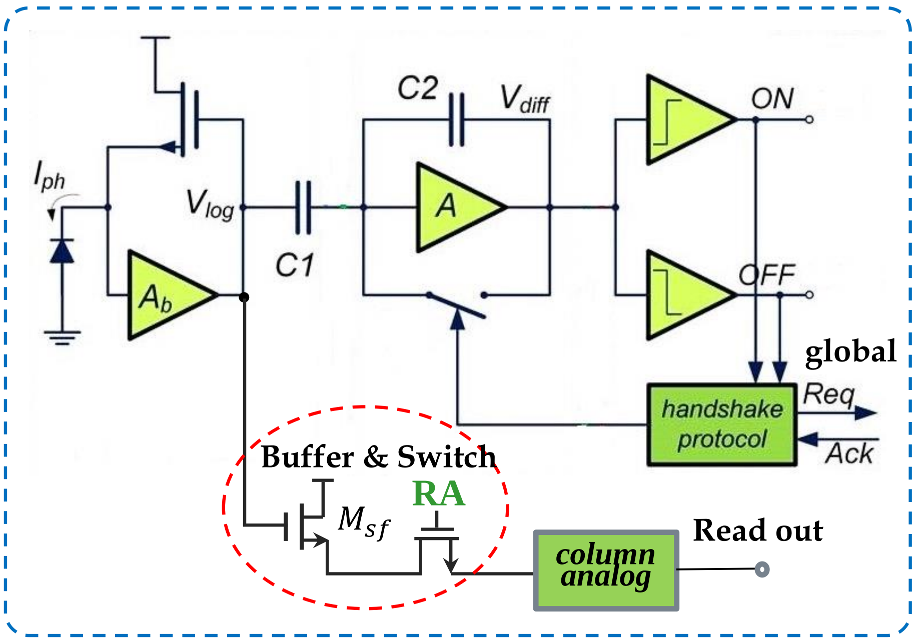
- Dynamic and active pixel vision sensor (DAVIS)
  - **DVS**
  - **APS : 50 Hz texture image**
  - **Two stream: asynchronous**

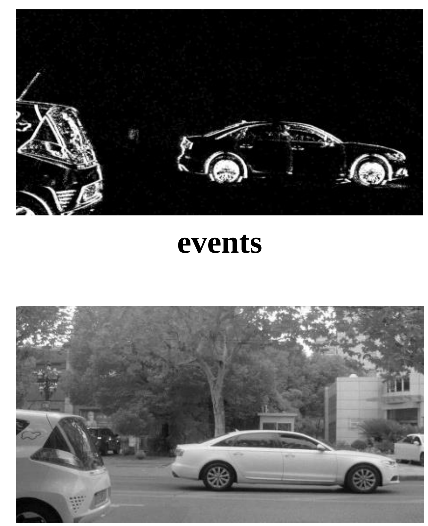


C Posch et al., Retinomorphic Event-Based Vision Sensors: Bioinspired Cameras With Spiking Output. *Proceedings of IEEE*, 2014.

Neuromorphic vision sensors

- CeleX - NTU
  - DVS
  - Voltage -> grayscale
  - Optical flow




events
texture image

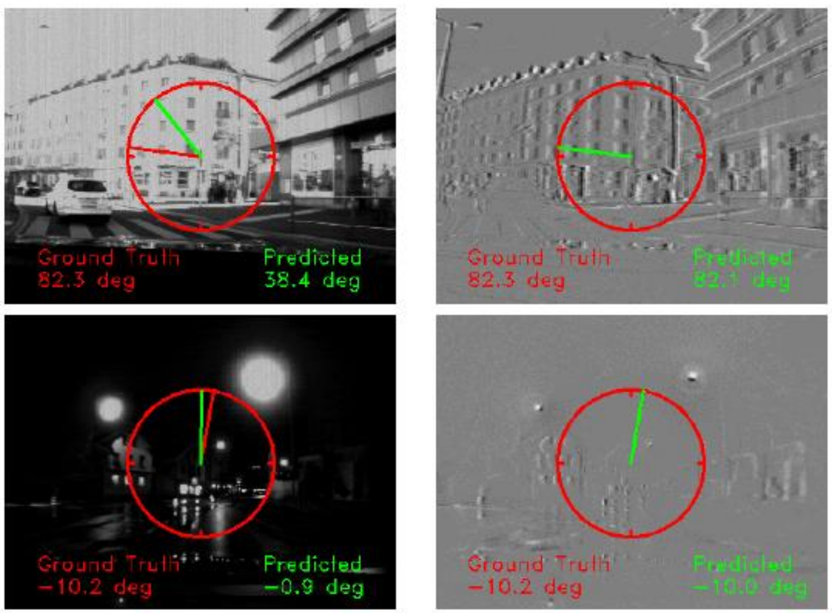
Neuromorphic vision sensors

Comparison between different commercialized neuromorphic vision sensors

Sensors	DAVIS128	ATIS	DAVIS346	DVS-G2	CeleX-V	FSM
Supplier	iniVation	Prophesee	iniVation	Samsung	CelePixel	PKU
Year	2008	2011	2017	2017	2018	2018
Resolution	128×128	304 ×240	346×260	640×480	**1280×800**	400×250
Sampling (Hz)	1×10^6	1×10^6	1.2×10^7	$\mathbf{3\times10^9}$	1.6×10^8	4×10^4
DR (dB)	120	143	120	90	120	70
Power (mW)	23	50-175	10-170	27-50	390-470	370
Chip Size (mm^2)	6.3×6	9.9×8.2	8×6	8×5.8	14.3×11.6	10×6
Pixel Size (μm^2)	40×40	30×30	18.5×18.5	9×9	9.8×9.8	20×20
Fill factor	8.1%	20%	22%	100%	9%	13.75%
Latency (μs)	12	3	20	65-410	1	25
Voltage (V)	3.3	1.8&3.3	1.8&3.3	1.2&3.3	1.2&3.3	1.5&3.3
Texture image	No	gray	**color**	No	gray	gray

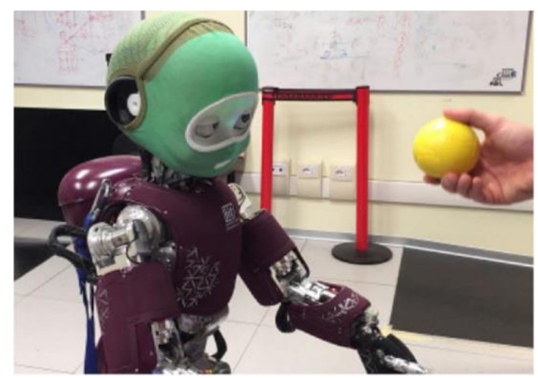
Neuromorphic vision sensors

- **Machine vision**
  - **High speed**
  - **Challenging illumination**



Autonomaous driving

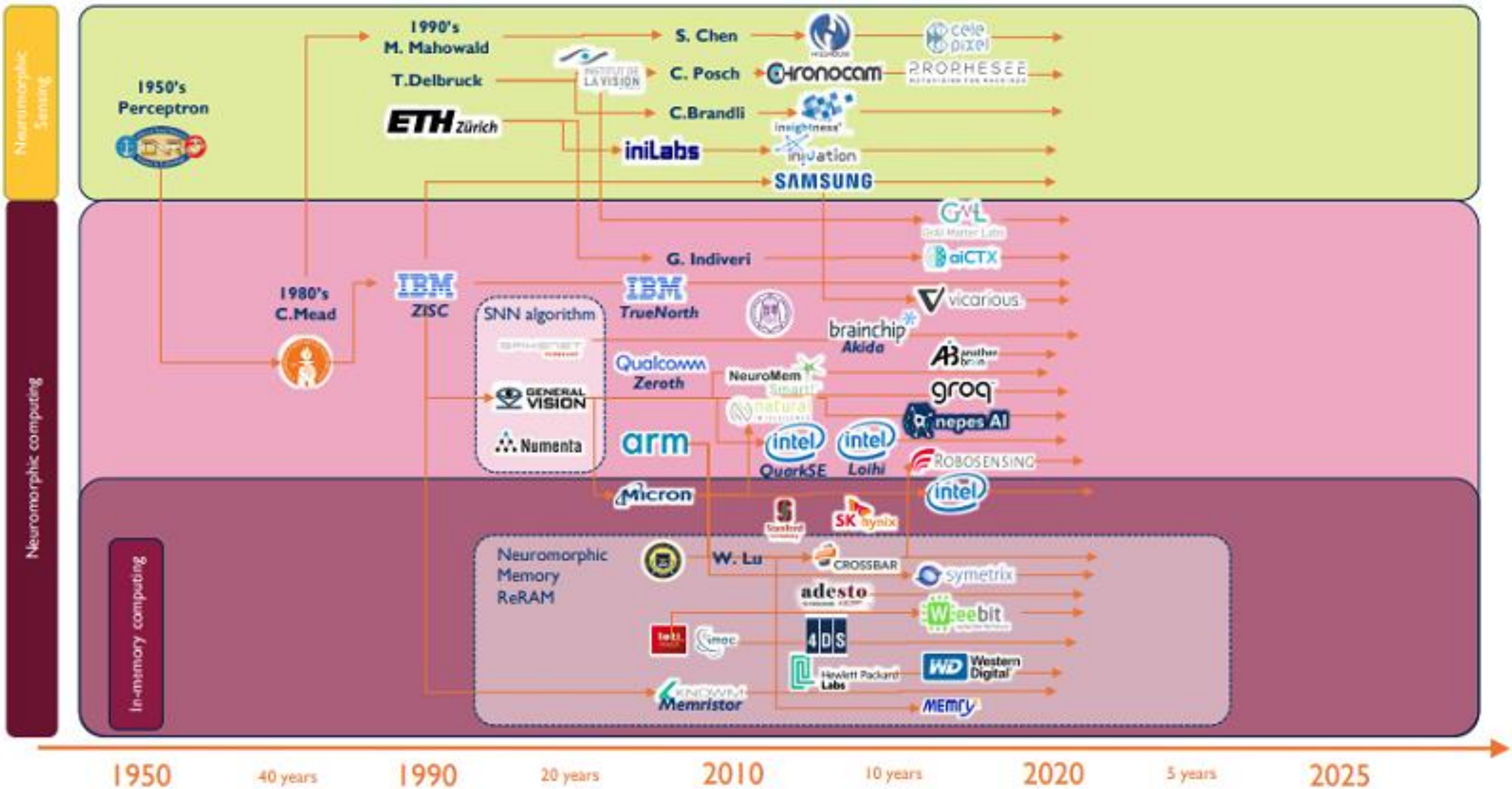
Drone



iCub robot

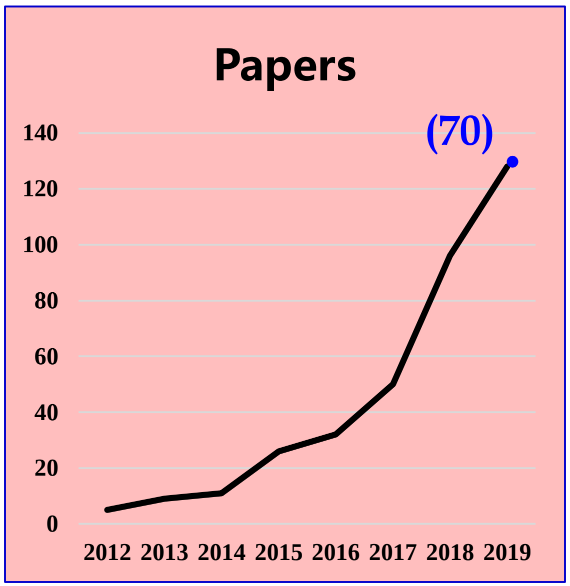
Neuromorphic vision sensors

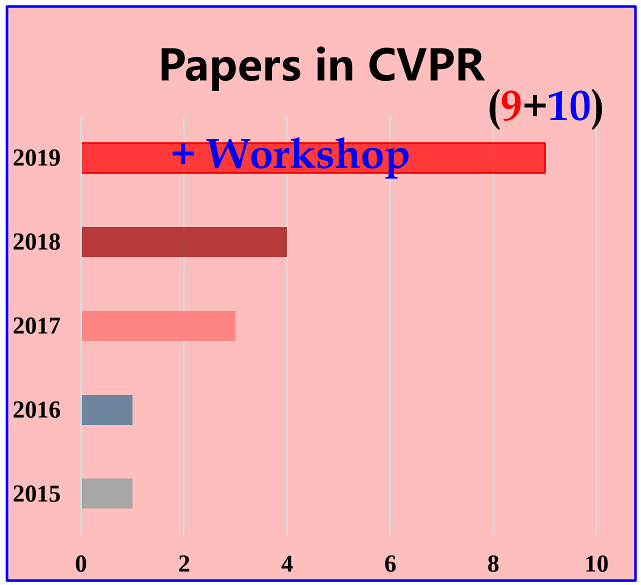
- Research institutes on neuromorphic

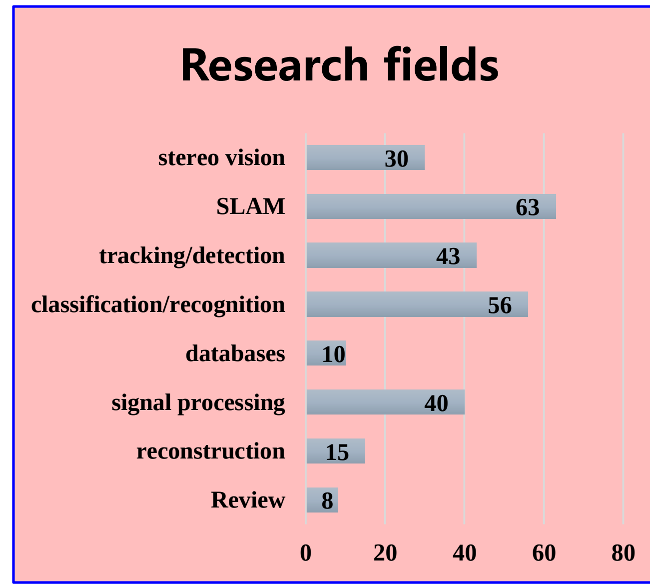


Neuromorphic vision sensors

- Related works
  - Paper numbers
  - Papers in CVPR
  - Research fields





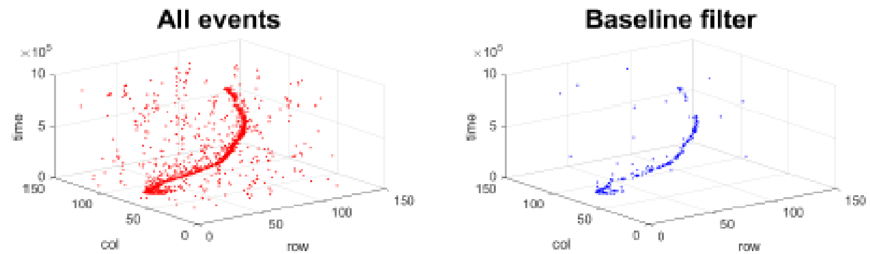


Overview

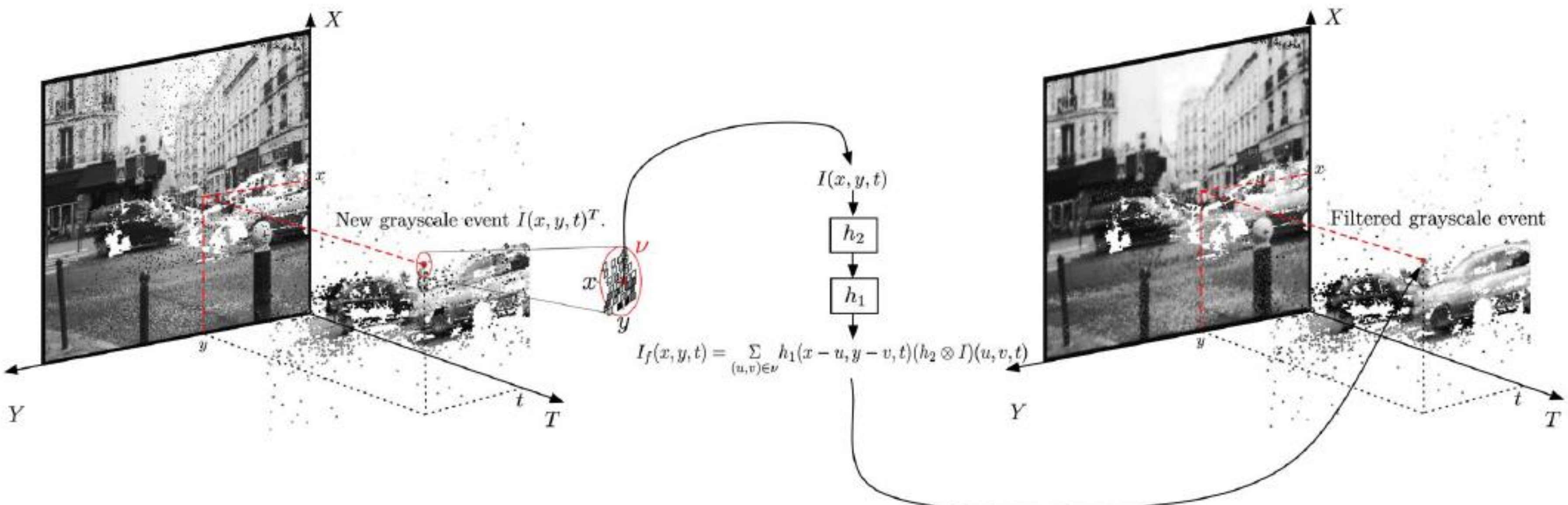
- Introduction
  - Conventional camera
  - Biological visual sensing system
  - Neuromorphic vision milestones
  - Neuromorphic vision Sensors
- **Event-based signal processing**
  - **Spatial-temporal filter**
  - **Spike metric**
  - **Spike coding**
- Feature representations (Applications)
  - Rate-based image
  - Hand-crafted feature
  - End-to-end deep network
  - Spiking neural network
- Discussion
  - Better input representations for event data
  - Event-based vision in the future

Event-based signal processing

- Spatial-temporal filter
  - **Denoising**
  - **Intensity estimation**



denoising



Intensity estimation

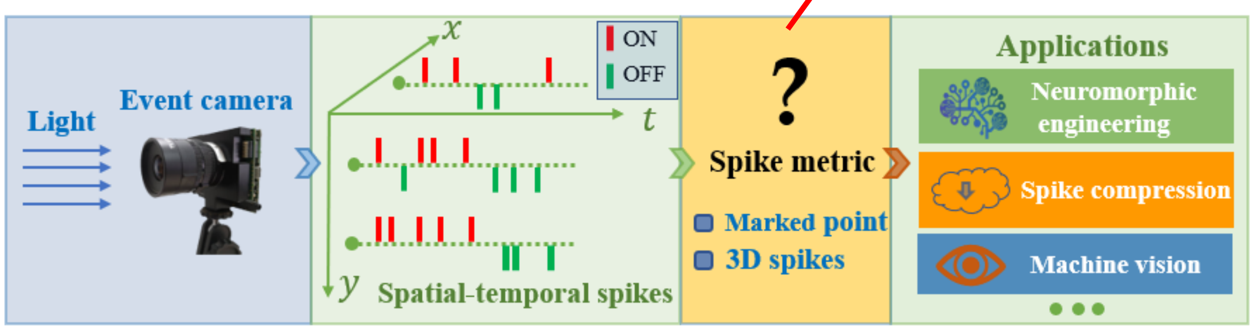
[1] O(N)-Sapce spatiotemporal filter for reducing noise in neuromorphic vision sensors. IEEE Transactions on Emerging Topics in Computing, 2018
[2] Asynchronous neuromorphic event-driven image filtering, Proceedings of the IEEE, 2014.

北京大学

Event-based signal processing

- Spike metric --- the basis of event-based signal processing
  - Spike coding
    - Distortion measurement
    - Motion estimation
  - Neuroscience
    - Retinal prostheses
    - Multi-neuron synchrony
  - Event-based vision
    - Retrieval and tracking
    - Matching in 3D stereo
    - *Loss function for ANN or SNN*

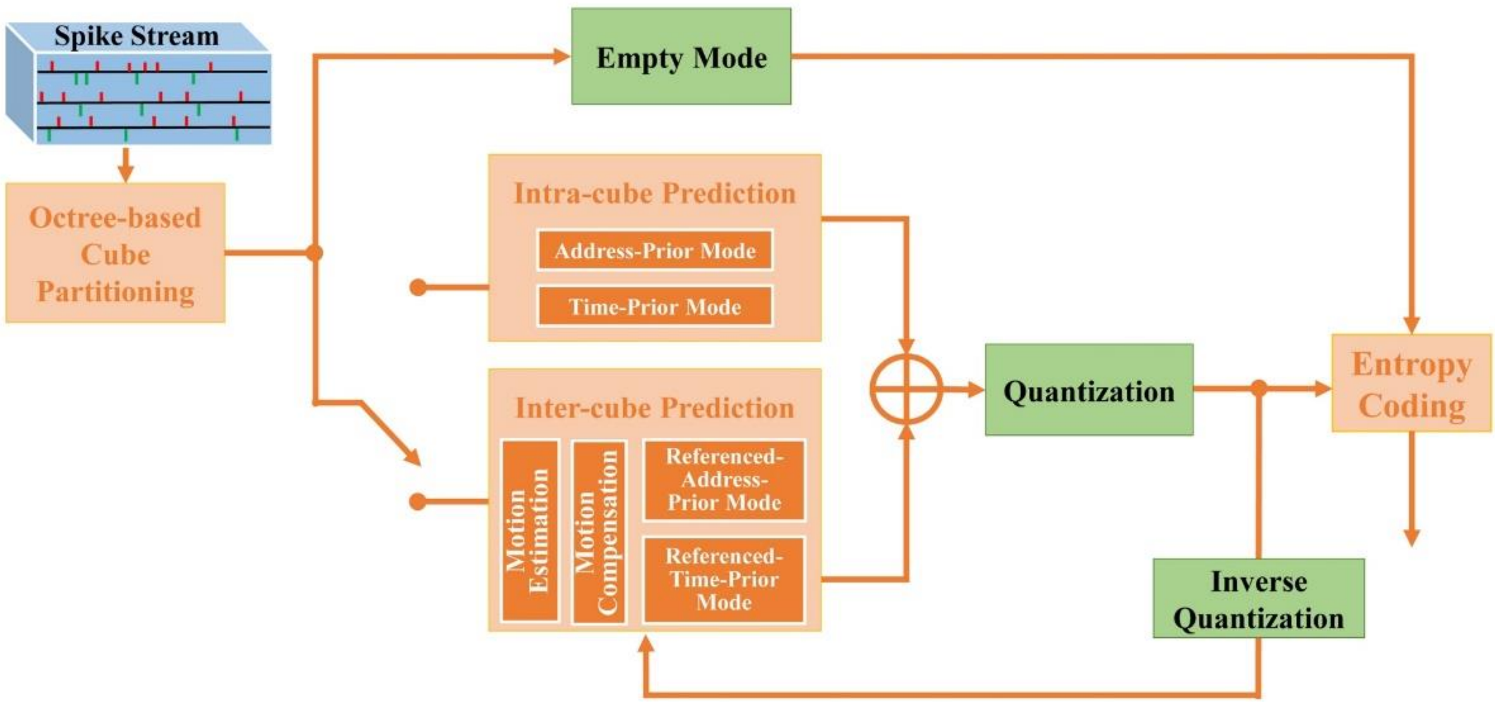
1. No taking polarity into account
2. No spatial structure information



[1] Learning a neural response metric for retinal prosthesis, Nishal P. Shah et al., *ICLR*, 2018.
[2] Focus is all you need: loss functions for event-based vision, Guillermo Gallego et al., *CVPR* 2019.

Event-based signal processing

- Spike coding
  - **Prediction framework**
    - Intra-cube
    - Inter-cube



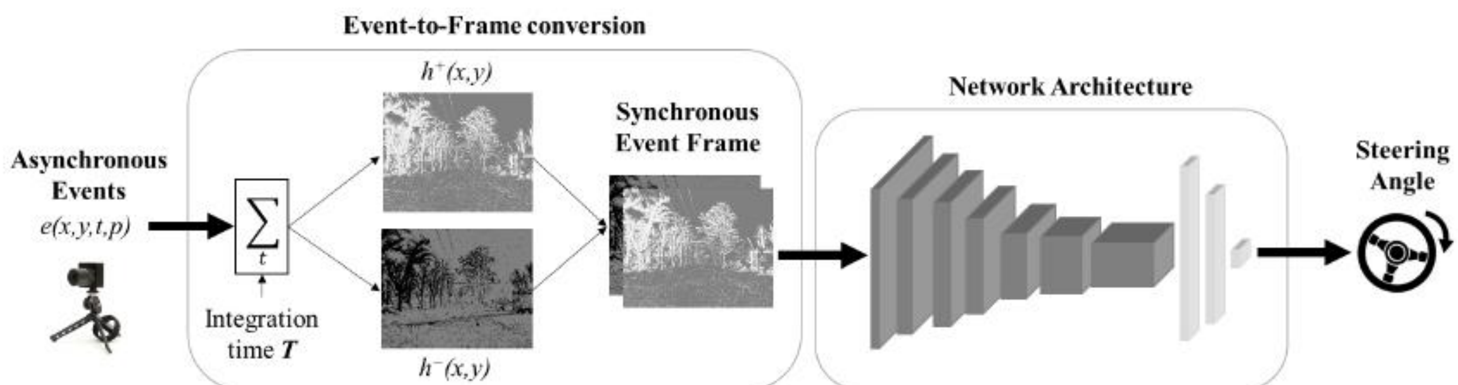
[1] Siwei Dong et al. Spike Coding for Dynamic Vision Sensors in Intelligent Driving, *IEEE Internet of Things Journal (IOT)*, 2018.
[2] Yihua Fu et al. Spike Coding: Towards Lossy Compression for Dynamic Vision Sensor, *Data Compression Conference (DCC)* , 2019.

Overview

- Introduction
  - Conventional camera
  - Biological visual sensing system
  - Neuromorphic vision milestones
  - Neuromorphic vision Sensors
- Event-based signal processing
  - Spatial-temporal filter
  - Spike metric
  - Spike coding
- **Feature representations (Applications)**
  - **Rate-based image**
  - **Hand-crafted feature**
  - **End-to-end deep network**
  - **Spiking neural network**
- Discussion
  - Better input representations for event data
  - Event-based vision in the future

Feature representation for event streams

☐ Rate-based image

<table>
<tr><td>Framework</td><td>

Rate-based image + Deep Network</td></tr>
<tr><td>Representation works</td><td>1. Event-based vision meets deep learning on <u>steering prediction</u> for self-driving cars, CVPR 2018
2. EKLT: Asynchronous photometric <u>feature tracking</u> using events and frames, IJCV 2019
3. EV-Gait: Event-based robust <u>gait recognition</u> using dynamic vision sensors, CVPR 2019</td></tr>
<tr><td>Reviews</td><td>1.Directly applied in image-based method;
2.<u>Failed to exploit spatial-temporal attribute.</u></td></tr>
</table>

Jianing Li, Siwei Dong, Zhaofei Yu,* **Yonghong Tian**, Tiejun Huang. Event-based vision enhanced: A joint detection framework in autonomous driving. ***ICME 2019***.

Feature representation for event streams

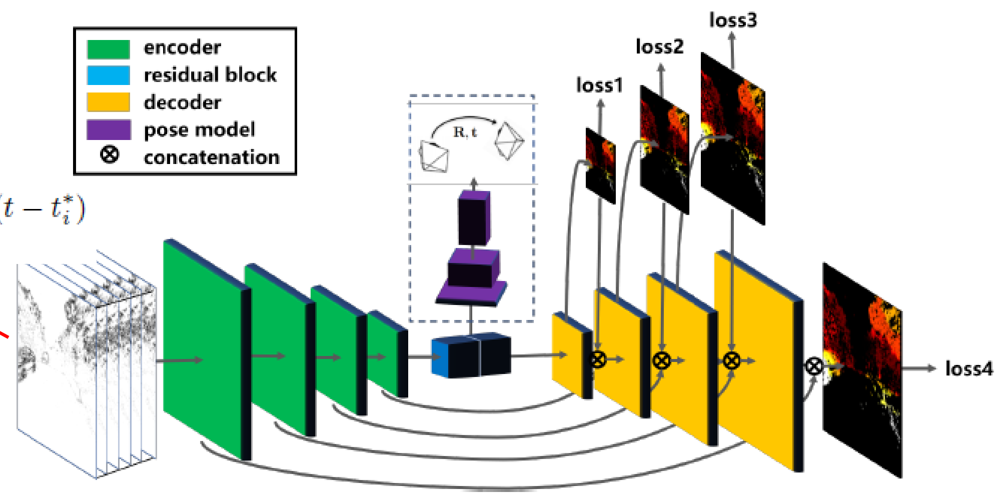
Event volume

<table>
<tr><td>Framework</td><td>

$t_i^* = (B-1)(t_i - t_0)/(t_N - t_1)$

$V(x,y,t) = \sum_i p_i k_b(x-x_i) k_b(y-y_i) k_b(t-t_i^*)$

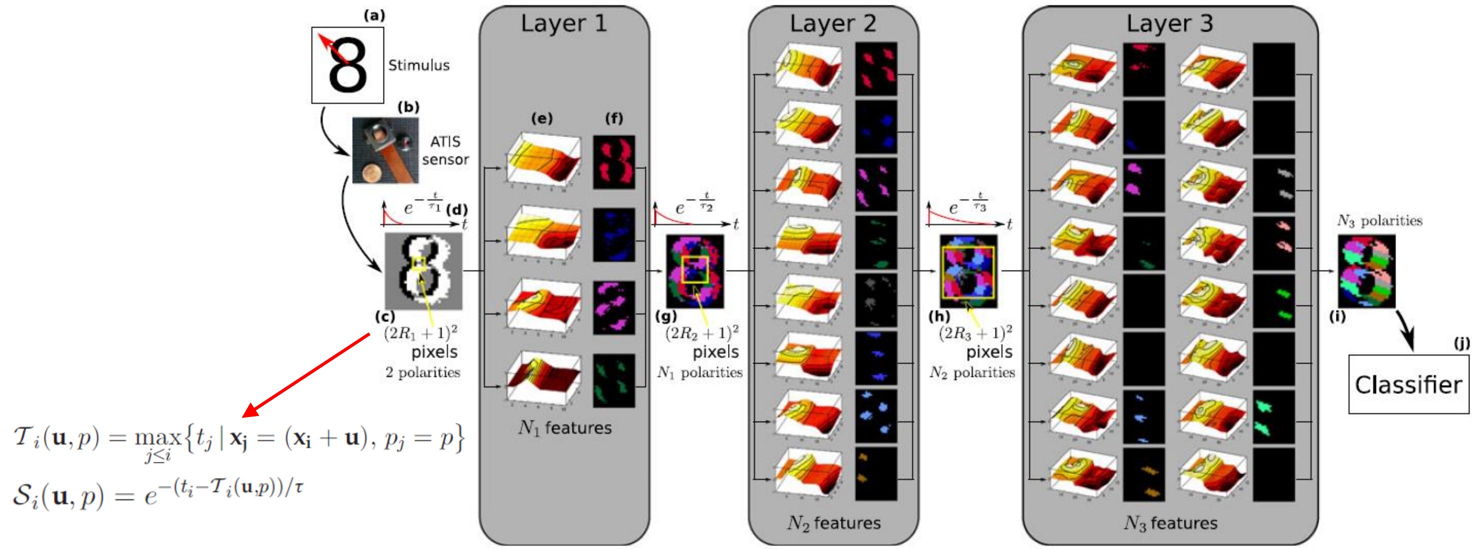
$k_b(a) = \max(0, 1-|a|)$



Event volume+ Deep Network
</td></tr>
<tr><td>Representation works</td><td>1. Unsupervised event-based learning of <u>optical flow, depth, and ego-motion</u>, CVPR 2019
2. <u>Events-to-video</u>: Bring modern computer vision to event cameras, CVPR 2019</td></tr>
<tr><td>Reviews</td><td>1.Transform image by sampling function;
2. <u>Failed to exploit spatial-temporal attribute.</u></td></tr>
</table>

Feature representation for event streams

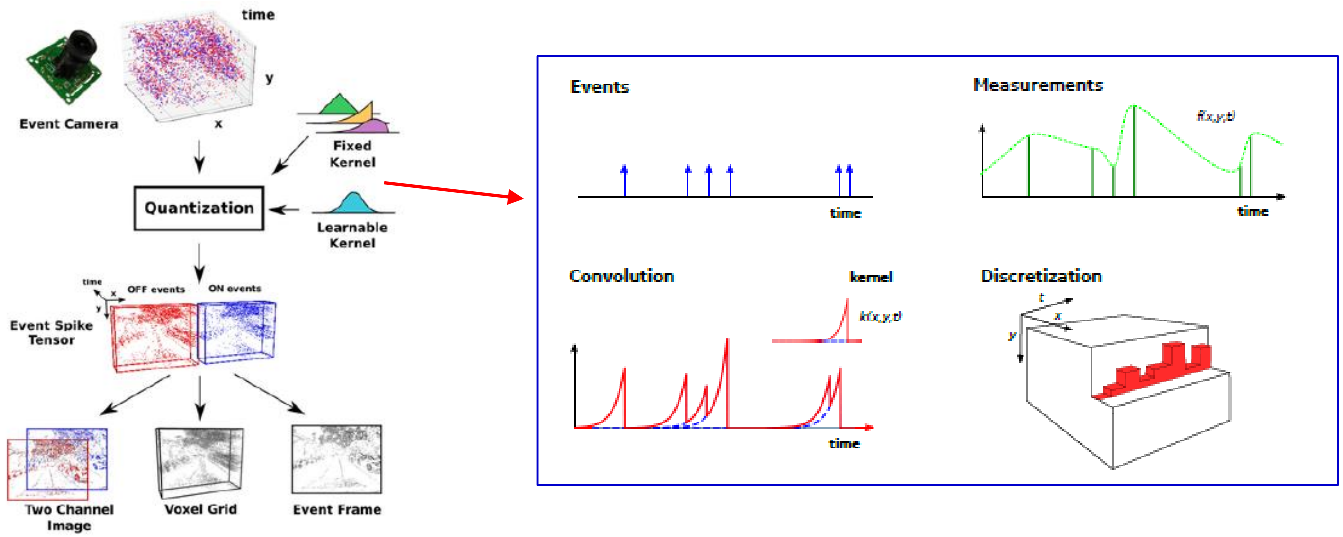
□ Hand-crafted feature

<table>
<tr><td>Framework</td><td>

Time surface</td></tr>
<tr><td>Representation works</td><td>1. HOTS: A hierarchy of event-based time-surfaces for <u>pattern recognition</u>, PAMI 2017
2. HATS: Histograms of averaged time surfaces for robust event-based <u>object classification</u>, CVPR 2018</td></tr>
<tr><td>Reviews</td><td>1. Strong prior performance；
2. <u>Time-consuming by hand-crafted debug</u>.</td></tr>
</table>

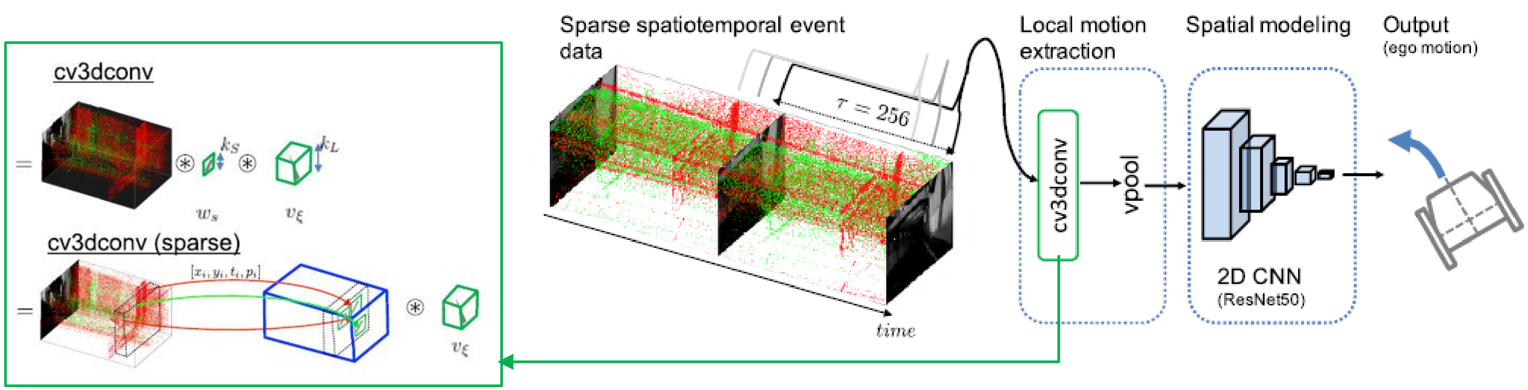
Feature representation for event streams

☐ Learning feature

Framework	 **Learning feature via deep network**
Representation works	1. End-to-end learning of representations for asynchronous event-based data, *ICCV 2019*
Reviews	1. **Obtaining better performance than rate-based image**; 2. **Multi-stage design.**

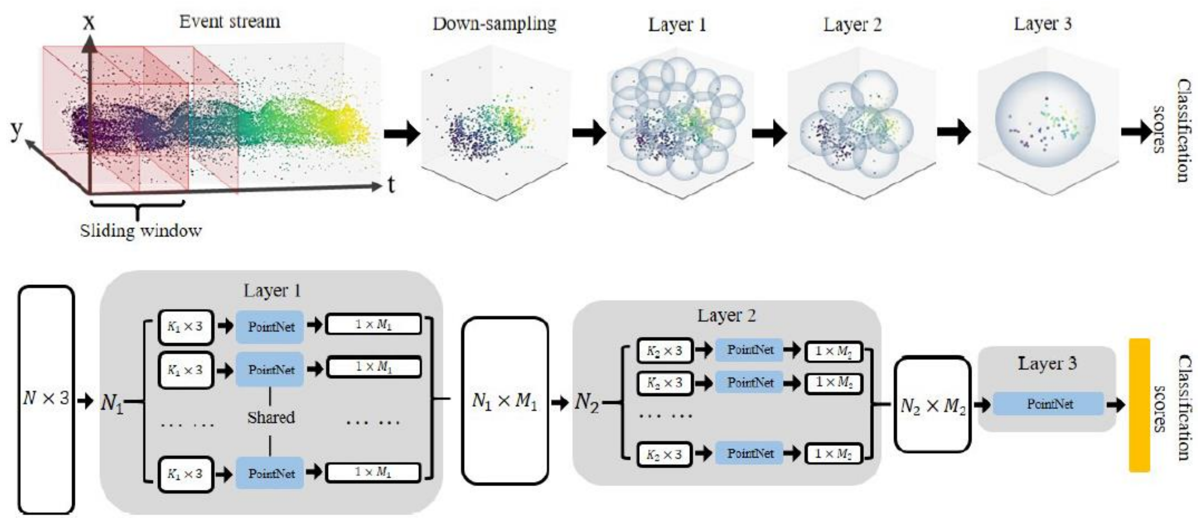
Feature representation for event streams

□ 3D convolution

Framework	 **3D convolution (deep network)**
Representation works	1. Constant velocity 3D convolution, *3D Vision 2018* 2. Constant velocity 3D convolution, *IEEE Access 2018*
Reviews	1. **Obtaining better performance than rate-based image**; 2. **High computing complexity.**

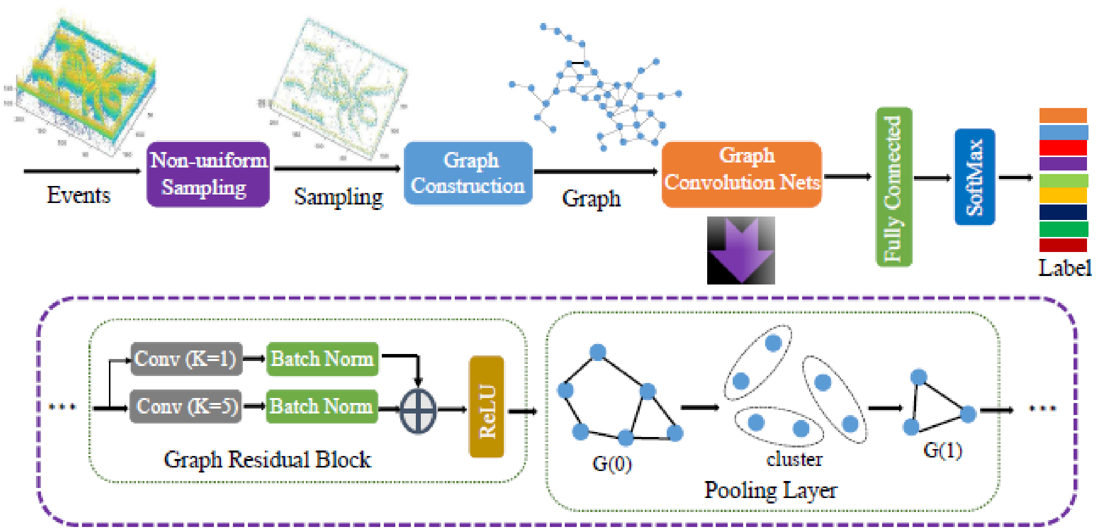
Feature representation for event streams

☐ Event clouds

Framework	 Event clouds + **PointNet**
Representation works	1. Space-time event clouds for **gesture recognition**: from RGB cameras to event cameras, *WACV 2019* 2. **EventNet**: Asynchronous recursive event processing, *CVPR 2019* 3. Modeling **point clouds** with self-attention and Gumbel subset sampling, *CVPR 2019*
Reviews	1. **Research hotspot**; 2. **The state-of-the-art method**.

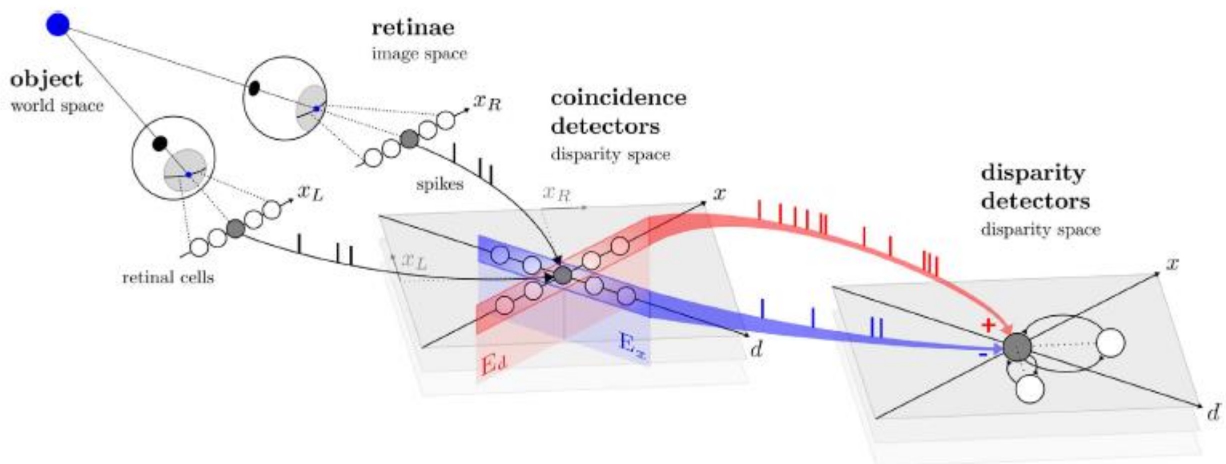
Feature representation for event streams

□ Graph-based

Framework	 Event as graph+ **GNN**
Representation works	1. Graph-based <u>**object classification**</u> for neuromorphic vision sensing, *ICCV 2019* 2. Graph-based <u>**spatial-temporal feature learning**</u> for neuromorphic vision sensing, *arXiv 2019* 3. <u>**Graph based event processing**</u>, *Imaging and Applied Optics, 2019*
Reviews	1. **First work using graph-based neural network**; 2. **The state-of-the-art methods**。

Feature representation for event streams

☐ Spiking neural network

Framework	 **Spiking Neural Networks**
Representation works	1. SLAYER: Spike layer **error reassignment** in time, *NIPS 2019* 2. **Direct training** for spiking neural networks: Faster, larger and better, ***AAAI 2019*** 3. A spiking neural network model of **3D perception** for event-based neuromorphic stereo vision, *Scientific Reports, 2017* 4. A spiking neural network model of **depth from defocus** for event-based neuromorphic vision, ***Scientific Reports, 2019***
Reviews	1. **Focus on learning theory (such as supervised learning) and computing ability (GPU)**; 2. **Low performance, especially in complex tasks**。

Feature representation for event streams

Comparison between different recognition methods on publication datasets

Representations	Methods	N-MNIST	MNIST-DVS	N-Caltech	CIFAR10-DVS	N-CARS	DVS-Gesture	ASL-DVS
Rate-based image	VGG_19	0.972	0.983	0.549	0.334	0.728	-	0.806
	ResNet_50	0.984	0.982	0.637	0.558	0.903	-	0.886
	LSTM	0.837	0.824	0.196	0.245	0.789	0.769	0.621
Hand-crafted feature	HOTS	0.808	0.803	0.210	0.271	0.624	0.785	0.656
	HATS	0.991	0.984	0.642	0.524	0.902	0.933	0.871
End-to-end deep network	PATs	-	-	-	-	-	**0.974**	-
	RG-CNNs	0.990	**0.986**	**0.657**	0.540	**0.914**	0.938	**0.901**
Spiking neural network	H-Frist	0.712	0.595	0.054	0.077	0.561	0.529	0.479
	Direct-SNN	**0.995**	-	-	**0.605**	-	-	-
	SLAYER	0.992	0.956	0.598	0.532	0.907	0.936	0.869

Overview

- Introduction
  - Conventional camera
  - Biological visual sensing system
  - Neuromorphic vision milestones
  - Neuromorphic vision Sensors
- Event-based signal processing
  - Spatial-temporal filter
  - Spike metric
  - Spike coding
- Feature representations (Applications)
  - Rate-based image
  - Hand-crafted feature
  - End-to-end deep network
  - Spiking neural network
- **Discussion**
  - **Better input representations for event data**
  - **Event-based vision in the future**

Discussion

- How to find better representation for event data?
  - Learning feature
  - Event clouds
  - Graph-based

Representations	Works	Vision tasks
Learning feature	End-to-end learning of representations for asynchronous event-based data, *ICCV 2019*	Recognition & optical flow
	DART: Distortion aware retinal transform event-based cameras, *TPAMI 2019*	Recognition & tracking
Event clouds	Space-time event clouds for gesture recognition: from RGB cameras to event cameras, *WACV 2019*	recognition
	EventNet: Asynchronous recursive event processing, *CVPR 2019*	segmentation
	Modeling point clouds with self-attention and Gumbel subset sampling, *CVPR 2019*	recognition
Graph-based	Graph-based object classification for neuromorphic vision sensing, *ICCV 2019*	recognition
	Graph based event processing, *Imaging and Applied Optics, 2019*	Signal processing

Discussion

- How many challenges are there in neuromorphic vision?
  - Large datasets for supervised learning
  - Asynchronous spatial-temporal spike metric
  - Better representations for spatial-temporal streams
  - Spiking neural networks for complex vision
  - High computing using neuromorphic chip

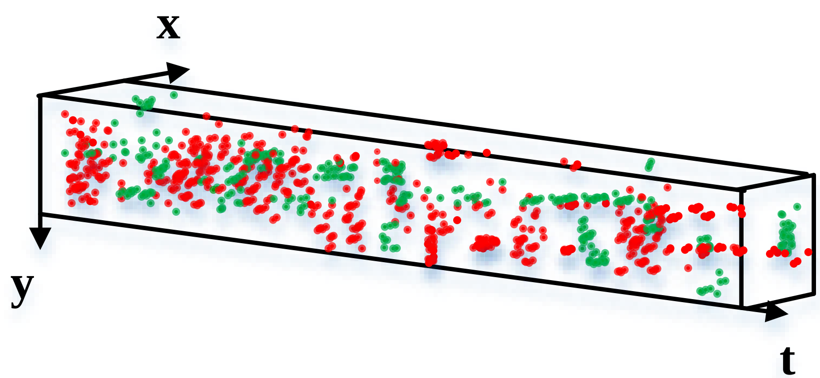
Loihi, Intel

TureNorth, IBM

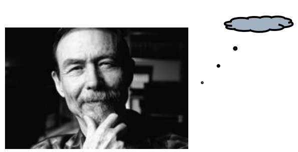
SpiNNaker, UM

Discussion

- **What is the future of neuromorphic vision sensors?**
  - **Memory mechanism using neuromorphic vision sampling**
  - **Multi-spectral sampling using neuromorphic concept**
  - **Multi-sensors fusion, such as DVS, DAS, neuromorphic torch**
  - **Neuroscience & Neuromorphic Engineering**



Asynchronous spatial-temporal point processes



Carver Mead

"Listen to the technology, find out what it's telling you"

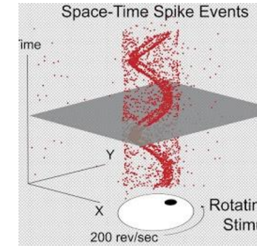
北京大学

THANKS

Q&A?

Telluride 2018 Neuromorphic Cognition Engineering Workshop

July 1-20, 2018. Telluride, Colorado, USA



ESP18: Fundamentals of Event Senor Signal Processing

- 1 . Can we lay a practical mathematical foundation that allows deriving efficient event-driven signal processing algorithms, analogous to the Z-transform of DSP?
- 2. Can we find better noise reduction (NR) algorithms than existing ones?
- 3. Can we find general methods for adaptively controlling sensor parameters like threshold, bandwidth, and refractory period?
- 4. Can we find better input representations for event cameras data for CNN?
- 5. What can we do to combine DVS events with color vision?

Tobi Delbruck
ETH

Ryad Benosman
UMPC

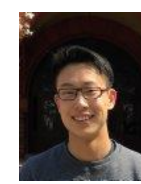
Garrick Orchard
NUS

Cornelia Fermuller
Univ.Maryland

David Mascarenas
LANS

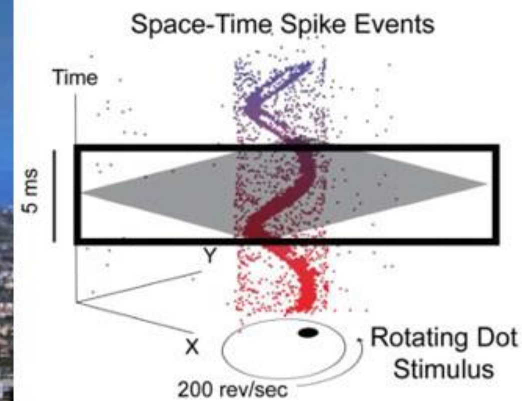
Yiannis Andreopoulus
UCL

Francisco
Univ, Grenada



Alex Zhu
Univ, Penn.

CVPR 2019 Workshops



- **Organizers:**

Davide Scaramuzza
UZH

Guillermo Gallego
UZH

Kostas Daniilidis
UPenn

CVPR 2019 Workshops

- Call for papers and demos
  - Event-based / neuromorphic vision.
  - Algorithm: Visual odometry, SLAM, 3D reconstruction, Optical flow estimation, image intensity reconstruction, recognition, stereo depth reconstruction, feature/ object detection and tracking, calibration, sensor fusion.
  - Model based, embedded or learning approaches.
  - Event-based signal processing, control, bandwidth control.
  - Event-based active vision.
  - Event-based camera datasets and/or simulators.
  - Applications in: robotics(navigation, manipulation, drones…), automotive, IoT, AR/VR, space, inspection, surveillance, crowd counting, physics.
  - Biologically-inspired vision and smart cameras
  - Novel hardware(cameras, neuromorphic processors, etc.) and/or software platforms.
  - New trends and challenges in event-based and/or biologically-inspired vision.

CVPR 2019 Workshops

Invited speakers

Tobi Delbruck
ETH

Garrick Orchard
NUS

Jorg Conradt
KTH

Giacomo Indiveri
ETH

Piotr Dudek
Univ. Manchester

Andrew Davision
ICL

Cornelia Fermuller
Univ.Maryland

Yulia Sandamirskaya
ETH

Chiara Bartolozzi
Italiano di Tecnlogia

Margarita Chli
ETH

Robert Mahony
ANU

Invited companies

PROPHESEE

ATIS, France

SAMSUNG

DVS(640*480), SK

intel

Loihi, USA

Insightness,
DVS, Switzerland

inivation

DVS, Switzerland

cele pixel

DVS, China

Appendix

- Paper list (1+8) in *CVPR 2019*
  - **Bring a blurry frame alive at high frame-rate with an event camera**, Liyuan Pan et. al, *ANU*. (oral)

  - **Unsupervised event-based learning of optical flow, depth and ego-motion**, Alex Z. Zhu et al, *University of Penn*.
  - **Events-to-video: bringing modern computer vision to event cameras**, Henri Rebecq et al, *UZH & ETH*.
  - **EventNet: Asynchronous recursive event processing**, Yusuke Sekikawa et al, *Denso IT Laboratory*.
  - **EV-Gait: Event-based robust gait recognition using dynamic vision sensors**, Yanxiang Wang et al, *HEU, China*.
  - **Event-based high dynamic range image and very high frame rate video generation using conditional generation adversarial networks**, S. M. Mostafavi et al, *GIST*.
  - **Speed invariant time surface for learning to detect corner points with event-based cameras**, J. Manderscheid et al, *PROPHESEE*.
  - **Focus loss functions for event-based vision**, Guilleromo Gallego et at, *UZH & ETH*.
  - **Event cameras, contrast maximization and reward functions: an analysis**, T. N. Stoffregen et al, *Monash University*.

Appendix

- Paper list (1+3) in *ICCV 2019*
  - Learning an event sequence embedding for dense event-based deep stereo, Stepan Tulyakov et. al, *EPFL*. (oral)

  - End-to-end learning of representations for asynchronous event-based data, Daniel Gehrig et al, *UZH & ETH*.
  - Events-based motion segmentation by motion compensation, Timo stoffregen et al, *UZH & ETH*.
  - Graph-based object classification for neuromorphic vision sensing, Yin Bin et al, *University of College London*.

Appendix

- Paper list (1+3) in *CCF A journal 2019*
  - **Towards spike-based machine learning intelligence with neuromorphic computing**, *Nature*, Ksushlk Roy et al, *Purdue University*.
  - **Optoelectronic resistive random access memory for neuromorphic vision sensors**, *Nature Nanotechnology*, Feichi Zhou et al, *The Hong Kong Polytechnic University*.
  - **Learning sensorimotor control with neuromorphic sensors: towards hyperdimensional active perception**, *Science Robotics*, A. Mitrokhin et al, *University of Maryland*.

  - **Event-driven sensing for efficient perception: vision and audition algorithms**, *IEEE Signal Processing Magazine*, Shi-Chii Liu et al, *UZH & ETH*.
  - **DART: distribution aware retinal transform for event-based cameras**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Bharath Ramesh et al, *NUS*.
  - **Unsupervised learning of a hierarchical spiking neural network for optical flow estimation: from events to global motion perception**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Federico Paredes-Valles et al, *Delft University of Technology*.
  - **EKLT: Asynchronous photometric feature tracking using events and frames**, *International Journal of Computer Vision*, Daniel Gehrig et al, *UZH & ETH*.