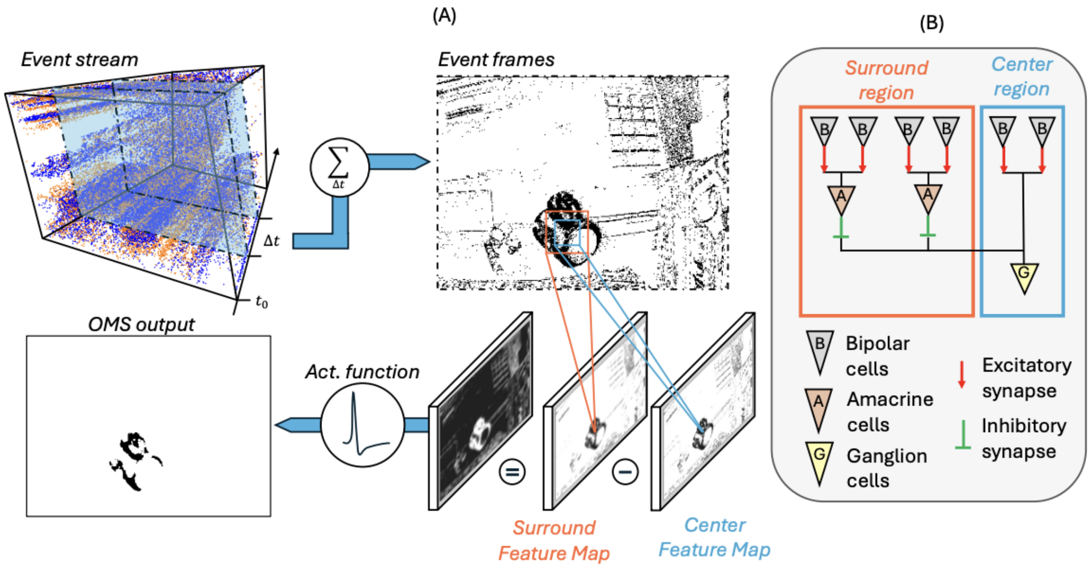
Retina-Inspired Object Motion Segmentation for Event-Cameras

Victoria Clerico, Shay Snyder, Arya Lohia, Md Abdullah-Al Kaiser, Gregory Schwartz, Akhilesh Jaiswal, Maryam Parsa

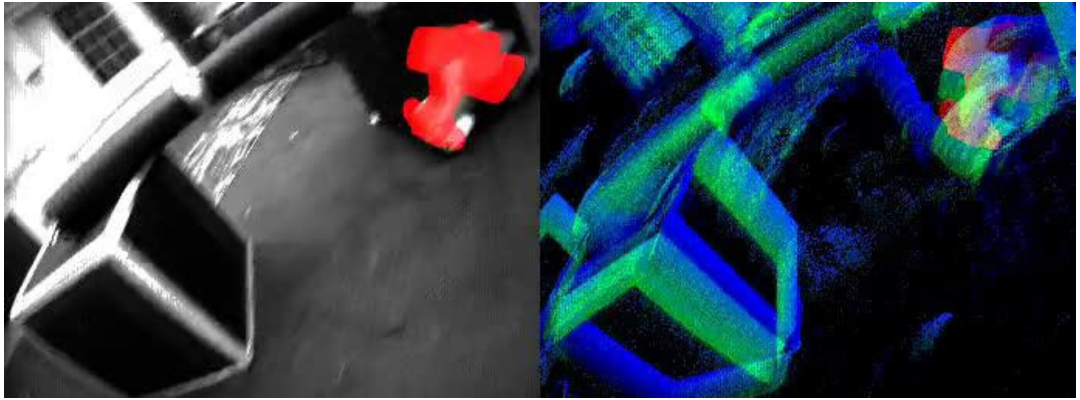


GEORGE MASON UNIVERSITY
NORTHWESTERN UNIVERSITY
WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

1. Ego-motion in Dynamic Vision Sensors (DVS)

Bio-inspired visual systems that capture per-pixel brightness changes asynchronously

Object motion **with** ego-motion



Microsecond temporal resolution

Low-power consumption

High dynamic range > 120 dB

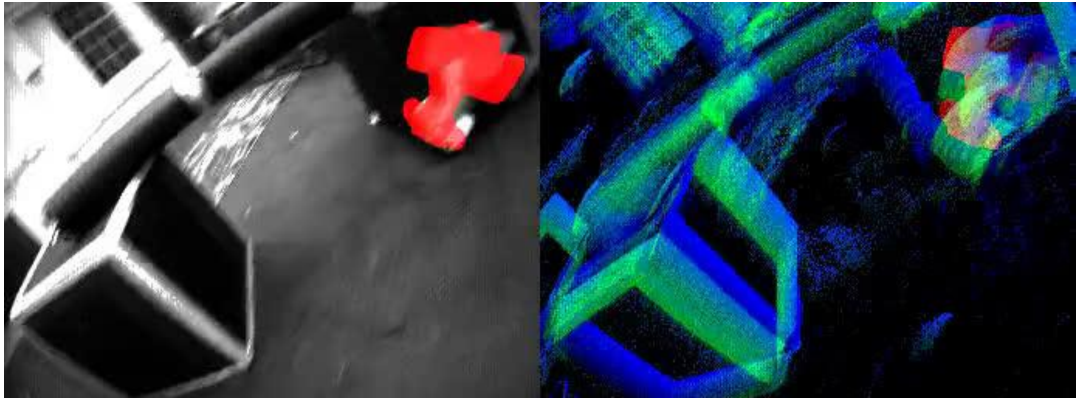
Reduced motion blur

GEORGE MASON UNIVERSITY

1. Ego-motion in Dynamic Vision Sensors (DVS)

How to suppress ego-motion to identify independent moving objects ?

→ Object Motion Sensitivity (OMS)



2. Biological inspiration

1) Photoreceptors

Contain light sensitive pigments and produce electrical currents

2) Bipolar cells

Adapt rapidly to changes in luminance

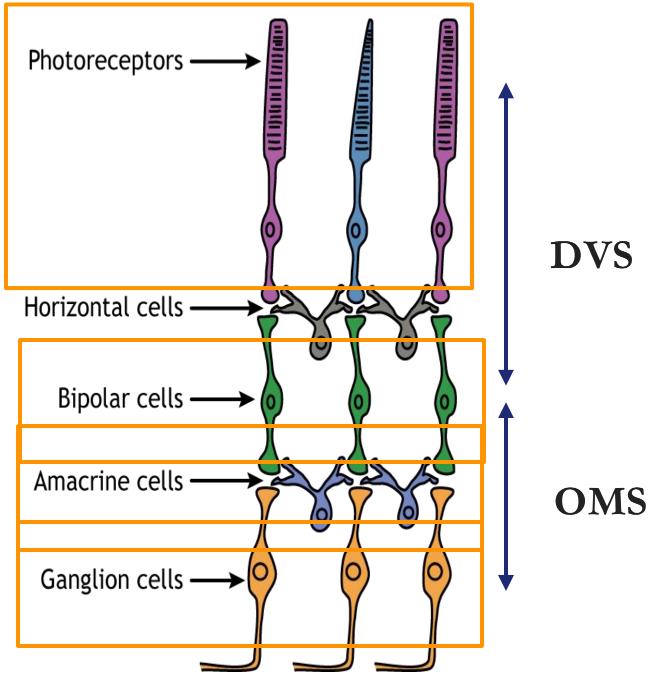
3) Amacrine cells

Perform lateral inhibition from adyactent RGCs

4) Ganglion Cells

Send information to the optical nerve

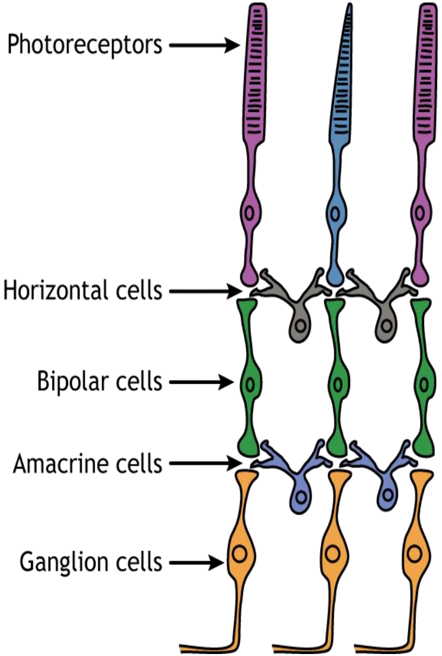
Retinal layers



GEORGE MASON UNIVERSITY

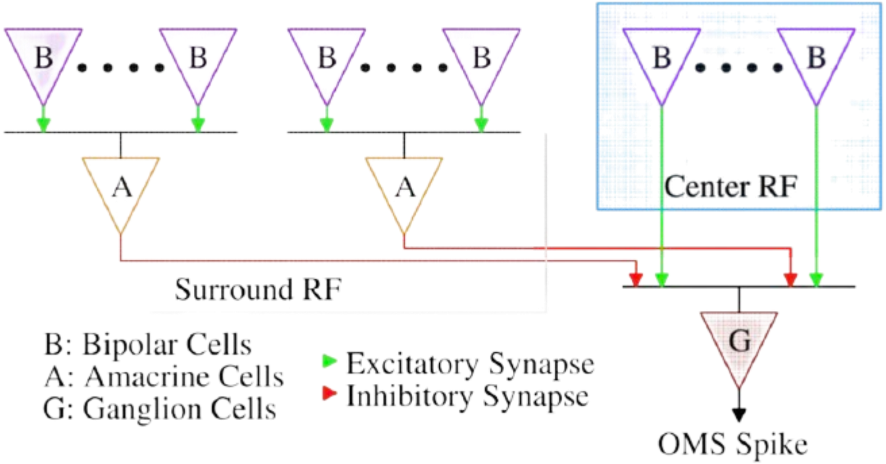
2. Biological inspiration

Neuroscience



Visual Receptive Fields (RFs)

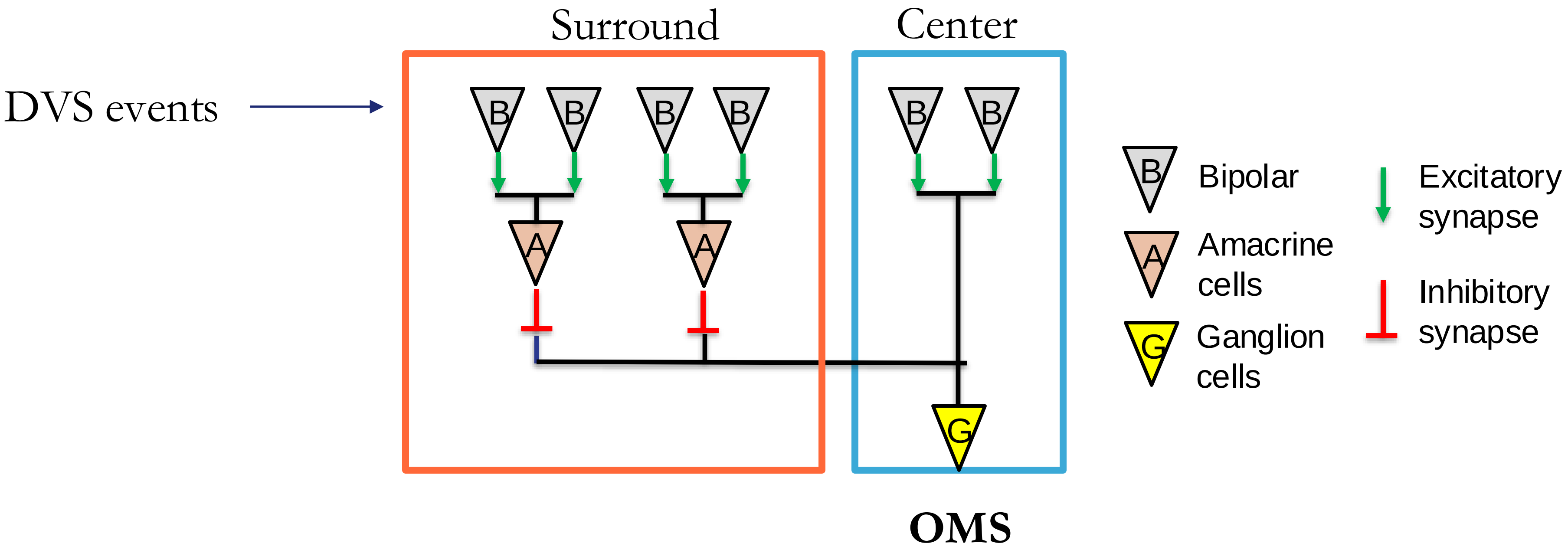
Surround and Center Regions



GEORGE MASON UNIVERSITY

3. Methodology

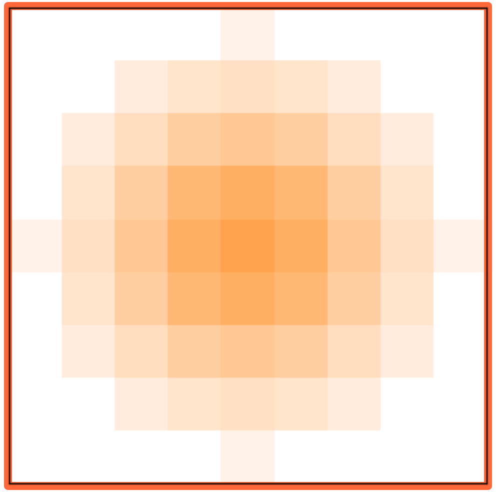
Modeled as convolutional kernels



3. Methodology

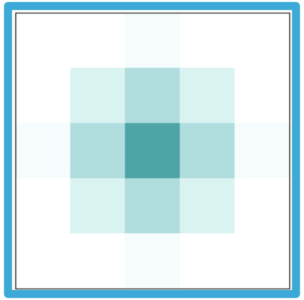
Modeled as convolutional kernels

Surround



8x8

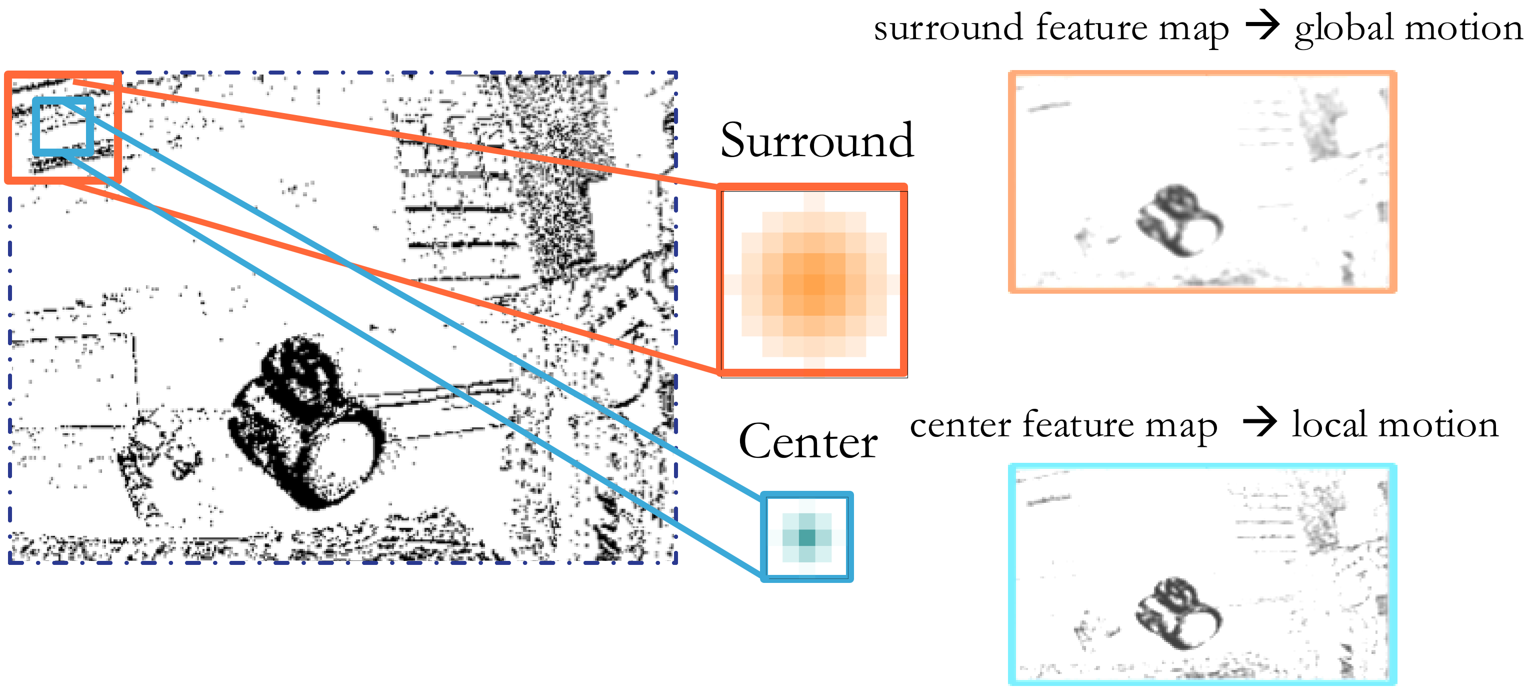
Center



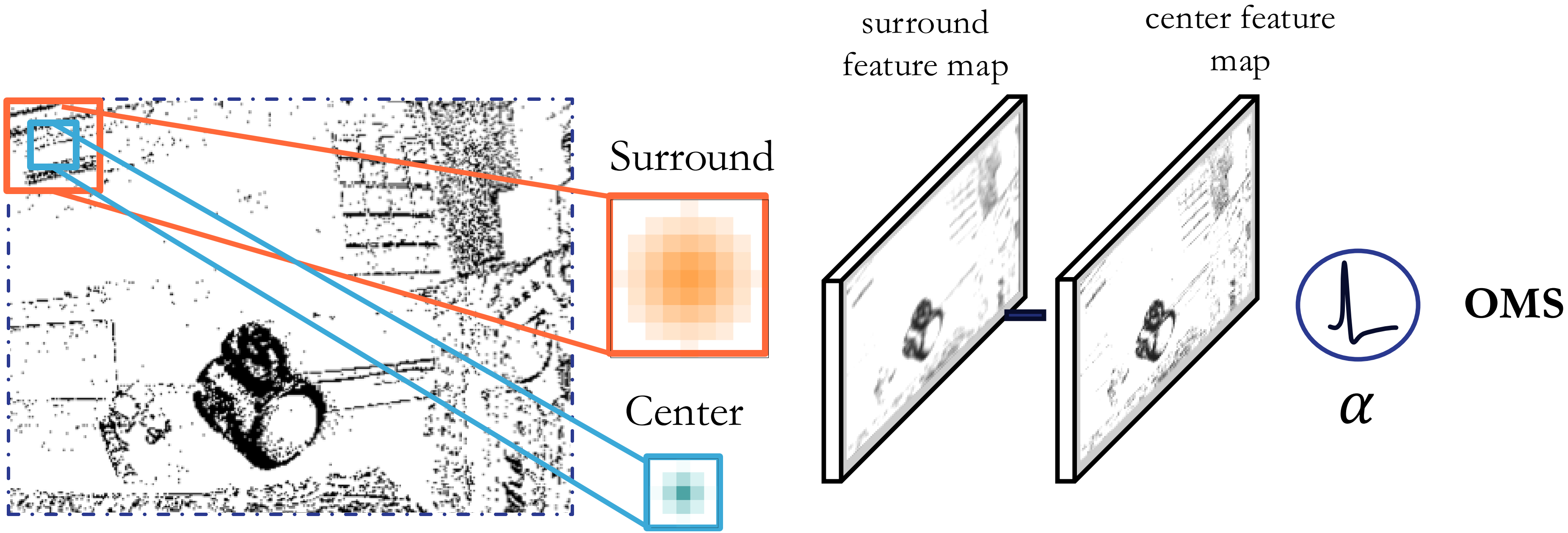
4x4

GEORGE MASON UNIVERSITY

3. Methodology

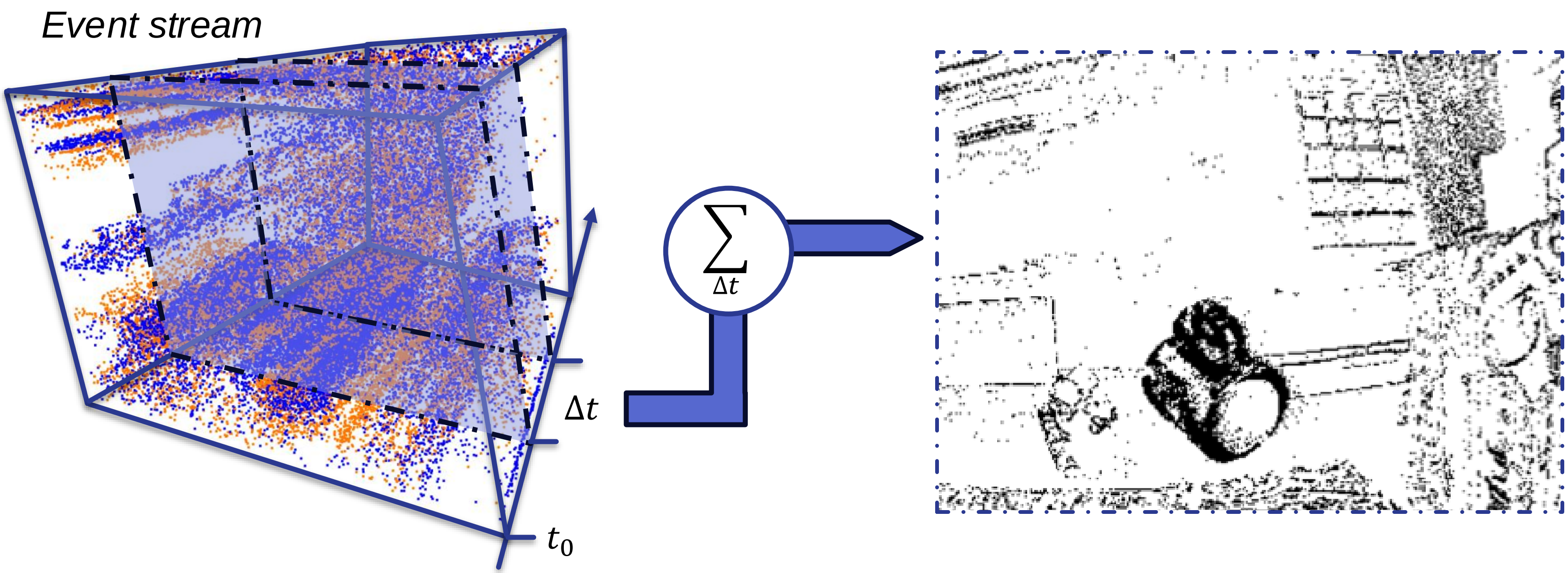


3. Methodology



3. Methodology

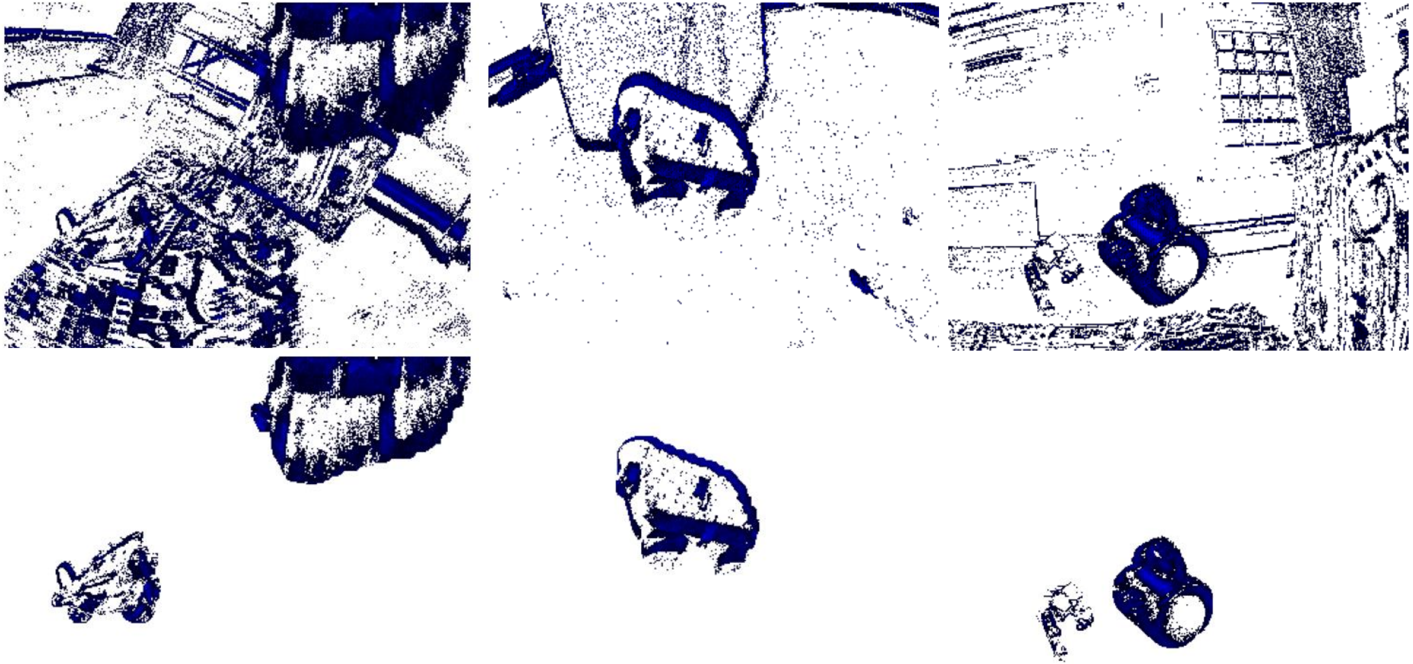
Time window accumulation



3. Methodology

Input

Groundtruth label

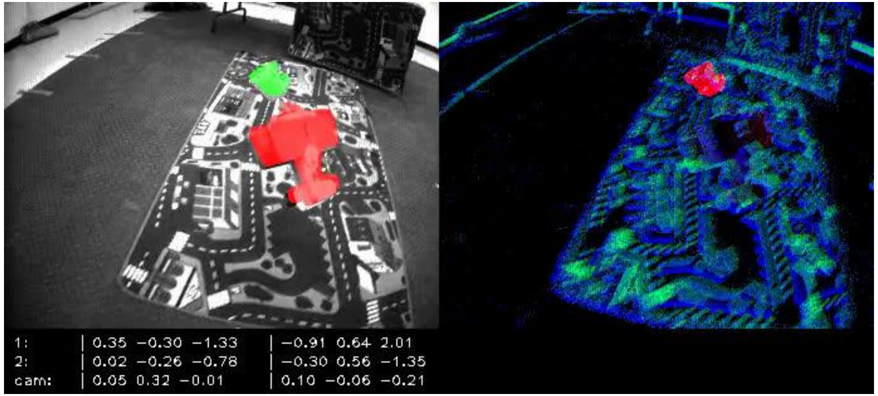


3. Methodology

Datasets: Event streams and motion masks (346x260 resolution)

EV-IMO [2]: real event dataset
- DAVIS 346C at 200 Hz
- Masks at 40 Hz

MOD [3]: synthetic dataset
- DVS simulator



Metrics
- Intersection over Union (%)
- Detection rate (%)

GEORGE MASON UNIVERSITY

4. Experiments: SOTA Comparisons

Comparable performance

Performance: mean IoU (%) over EV-IMO validation set

Method	box	floor	wall	table	fast
EV-IMO [1]	70 ± 5	59 ± 9	78 ± 5	79 ± 6	67 ± 3
PointNet++ [3]	80 ± 15	76 ± 10	74 ± 10	68 ± 23	20 ± 6
Gconv [4]	60 ± 18	55 ± 19	80 ± 7	51 ± 16	39 ± 19
SpikeMS [5]	65 ± 8	53 ± 16	63 ± 6	50 ± 8	38 ± 10
MSRNN [6]	80	85	82	85	75
Ours	$\mathbf{72 \pm 16}$	**94**	$\mathbf{82 \pm 6}$	$\mathbf{88 \pm 10}$	$\mathbf{69 \pm 3}$

GEORGE MASON UNIVERSITY

4. Experiments: SOTA Comparisons

mean IoU (%)

Method	EV-IMO	MOD
EV-IMO [1]	77	–
EVDodgeNet [2]	65.76	75
0-MMS [7]	80.37	–
SpikeMS [5]	–	68
Ours	**80.17**	**70**

Method	# of Parameters
EVDodgeNet [2]	30k – 3M
EV-IMO [1]	40k
SpikeMS [5]	46k
Gconv [4]	117k
PointNet++ [3]	1.4M
MSRNN [6]	4.1M
Ours	**80**

x $10^3 – 10^6$ parameter reduction

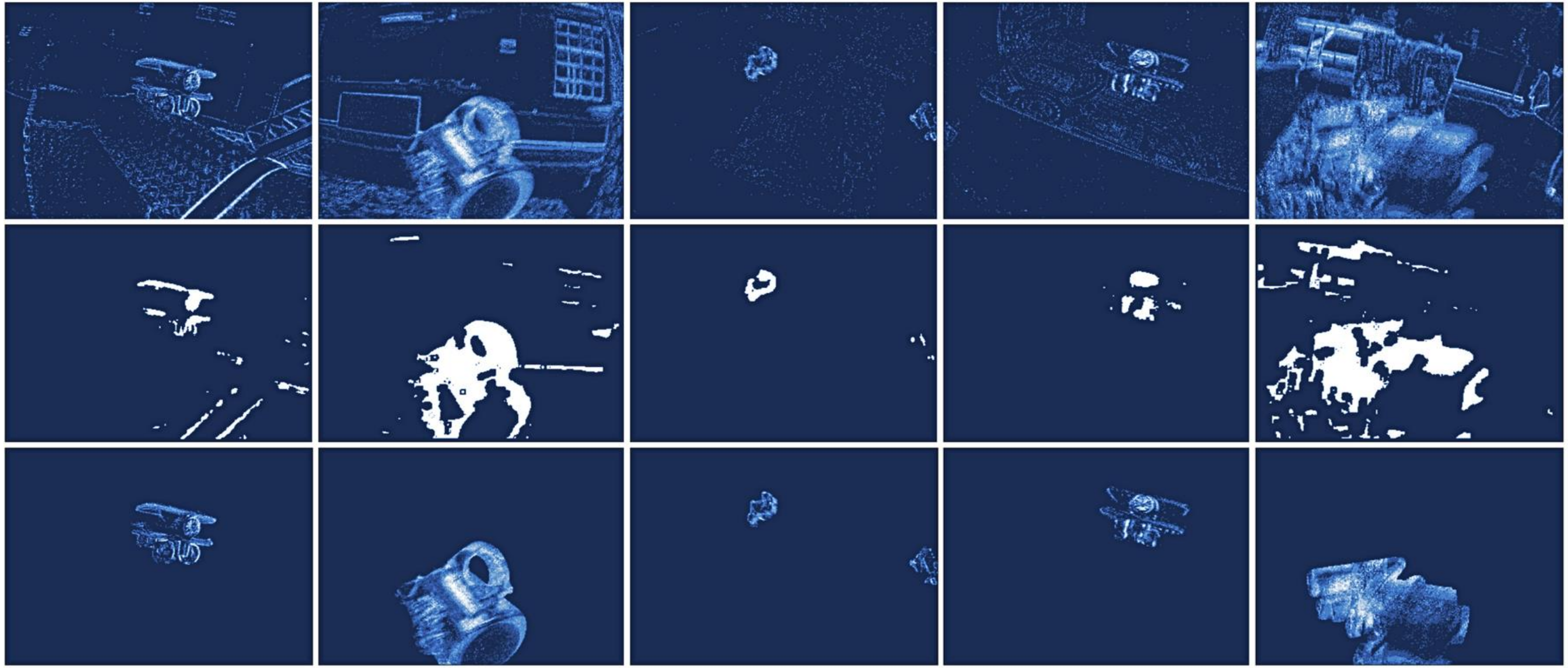
GEORGE MASON UNIVERSITY

4. Experiments

DVS

OMS

Mask



Conclusions

- Our method achieves top performance with **10^3–10^6× fewer parameters** than existing methods (SOTA)
- **Learning-free bio-inspired** method, enabling robust motion segmentation across domains.
- Expanding to **additional visual features** can enhance intelligent visual systems.

Acknowledgements

The research was funded in part by National Science Foundation through awards CCF2319617 and CCF2319619.

GEORGE MASON UNIVERSITY

PARSA RESEARCH LABORATORY

NORTHWESTERN UNIVERSITY

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Bibliography

[1] Mitrokhin, A., Ye, C., Fermuller, C., Aloimonos, Y., Delbruck, T.: Ev-imo: Motion segmentation dataset and learning pipeline for event cameras (2020)
[2] Sanket, N.J., et al.: Evdodgenet: Deep dynamic obstacle dodging with event cam-eras. In: 2020 IEEE International Conference on Robotics and Automation (ICRA). pp. 10651–10657. IEEE (2020)
[3] Qi, C., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In: Neural Information Processing Systems (2017)
[4] Mitrokhin, A., Hua, Z., Fermüller, C., Aloimonos, Y.: Learning visual motion segmentation using event surfaces. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
[5] Parameshwara, C.M., Li, S., Fermüller, C., Sanket, N.J., Evanusa, M.S., Aloimonos, Y.: Spikems: Deep spiking neural network for motion segmentation. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems
[6] Zhang, S., Sun, L., Wang, K.: A multi-scale recurrent framework for motion segmentation with event camera. IEEE Access 11, 80105–80114 (2023)
[7] Parameshwara, C.M., Sanket, N.J., Gupta, A., Fermüller, C., Aloimonos, Y.:MOMS with events: Multi-object motion segmentation with monocular event cameras. CoRR abs/2006.06158 (2020)
[8] Yin, Z., Kaiser, M.A.A., Camara, L.O., Camarena, M., Parsa, M., Jacob, A.,Schwartz, G., Jaiswal, A.: Iris: Integrated retinal functionality in image sensors. Frontiers in Neuroscience 17 (2023)
[8] Sinaga, J. Clerico, V., Kaiser, M.A.A., Snyder, S., Lohia, A., Schwartz, G., Parsa, M., Jaiswal, A: Hardware-Algorithm Re-engineering of Retinal Circuit for Intelligent Object Motion Segmentation – ICONS'24