Centroiding Point-Objects with Event Cameras

Dr. Scott McCloskey

kitware

International Conference on Computational Photography 2025 (to appear)
Authors: Connor Hashemi, Dennis Melamed, Albert W. Reed, Nitesh Menon, Keigo Hirakawa, Scott McCloskey

Presentation Outline: EBS-base Centroiding

- Problem definition
- Motivating applications of EBS-based centroiding
- Analytical model and theoretical bounds
- Dataset collection
- LSTM-based Centroiding approach
- Experimental results
- Summary & conclusions

kitware

Technology: **Event-Based Sensors**

Advantages

- Low latency (<100us)
- Low power consumption/less wasted energy
- High dynamic range

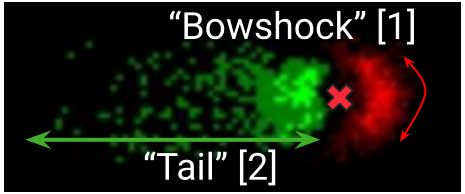
Challenges:

- Sparse/few events in small time batches
- Non-Gaussian spatial event distribution
- Event distribution depends on object brightness and velocity

Image Frame

Intensity

Event "Frame"



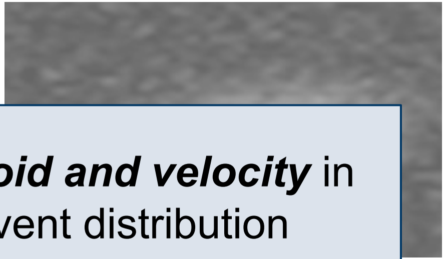
Positive Events

Negative Events

[1] Benson et al. "Simulation and analysis of event camera data for non-resolved objects." The Journal of the Astronautical Sciences 71.1 (2023): 3.
[2] Hu et al. "v2e: From video frames to realistic DVS events." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.

Technology: **Event-Based Sensors**

Image Frame

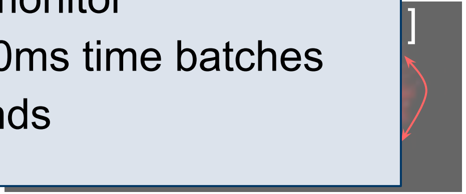


Intensity

- **Advantages**
  - Low latency (<100us)

- **Ch**

and velocity



Positive Events

Negative Events

<u>Approach</u>: Estimate an unresolved object's ***centroid and velocity*** in ***<20ms*** by leveraging the ***spatial cues*** in the event distribution

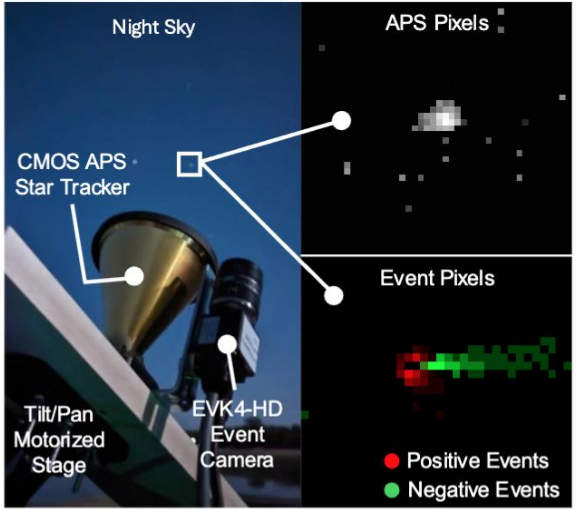
1. Develop an analytical model of event distributions
2. Create centroiding dataset from a high-speed monitor
3. Train a SOTA centroiding model for 5ms and 20ms time batches
4. Compare performance against theoretical bounds

[1] Benson et al. "Simulation and analysis of event camera data for non-resolved objects." The Journal of the Astronautical Sciences 71.1 (2023): 3.
[2] Hu et al. "v2e: From video frames to realistic DVS events." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.

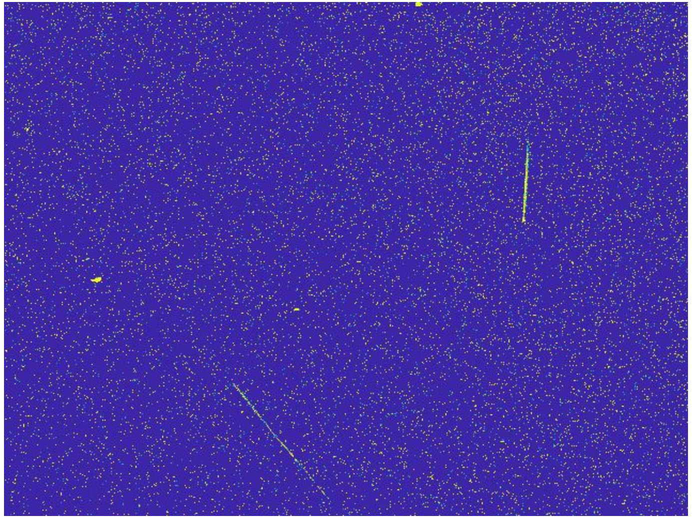
Kitware's EBS Work & Motivating Applications

Event-based star tracking

Poster #108, Friday afternoon



Meteor detection and classification [1]



[1] L. Tinch, N. Menon, K. Hirakawa, S. McCloskey, **Event-based Detection, Tracking, and Recognition of Unresolved Moving Objects.** Proc. of AMOS 2022

Closed-Form Event Distribution Model

- **In low-light**, each EBS pixel acts as a low-pass filter (LPF) [3]:

“Event Likelihood”

LPF Cutoff Frequency

Pixel Voltage

$$E_{\mathrm{LL}}(\mathbf{x}, t) = 2\pi \cdot f_c(\tilde{I}(\mathbf{x}, t)) \cdot \left[\tilde{I}(\mathbf{x}, t) - V(\mathbf{x}, t)\right]$$

Pixel

Time

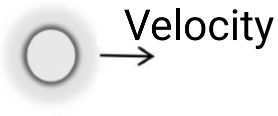
Intensity on Image Plane

- With some simple assumptions, we can achieve a **closed-form solution** that **facilitates advanced analysis**

[3] Reed, Albert W., et al. "EBS-EKF: Accurate and High Frequency Event-based Star Tracking." Proceedings of the Computer Vision and Pattern Recognition Conference. 2025.

Event Distribution Comparisons

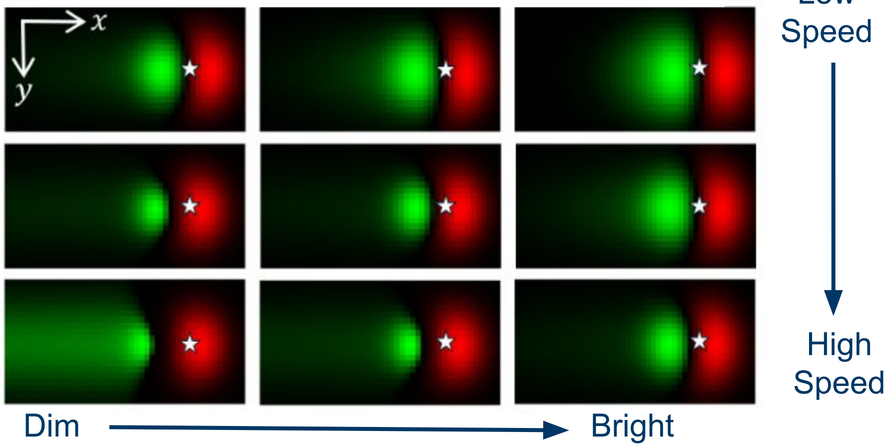
Unresolved Object



Conventional Models

- Speed, position, and intensity are all ***spatially coupled***
- Therefore, ***spatial cues*** lend important information
- Analytical results align with captured dataset

Proposed Closed-Form Solution

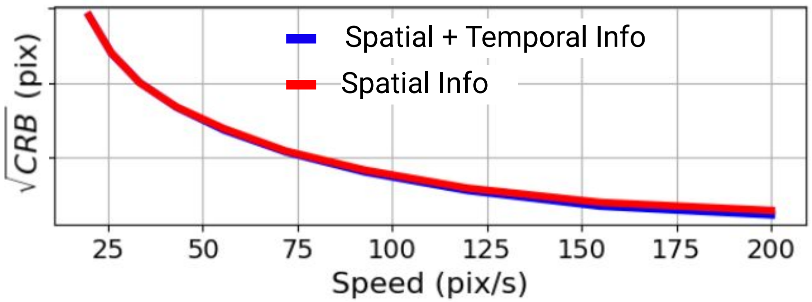


[4] Shah, Sachin, et al. "Codedevents: optimal point-spread-function engineering for 3d-tracking with event cameras." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2024.

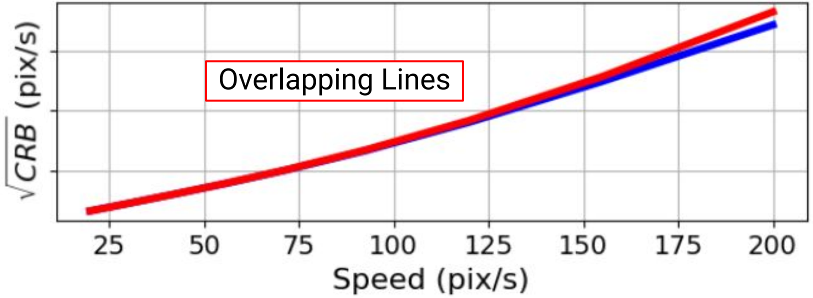
Cramér-Rao Bounds (CRB)

- Sample-normalized CRB can be analytically calculated with the closed-form model

Object Centroid Sample-Normalized CRB



Object Speed Sample-Normalized CRB



(lower CRB = more accurate estimator)

Two main takeaways:

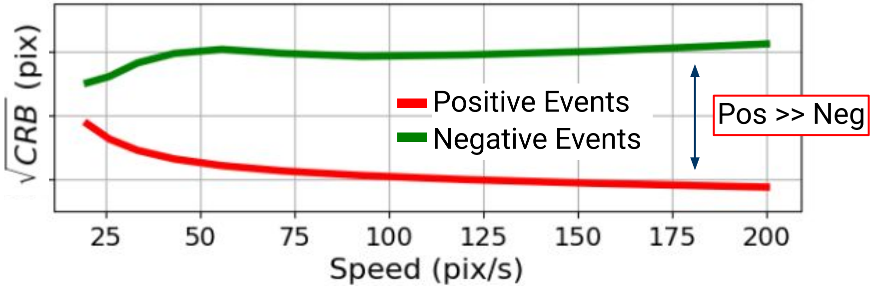
- Temporal information is **not as important** for small time batches

Cramér-Rao Bounds (CRB)

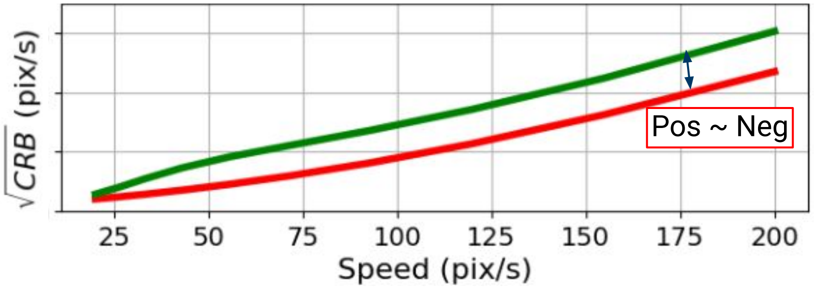
[4] Shah, Sachin, et al. "Codedevents: optimal point-spread-function engineering for 3d-tracking with event cameras." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2024.

- Sample-normalized CRB can be analytically calculated with the closed-form model

Object Centroid Sample-Normalized CRB



Object Speed Sample-Normalized CRB



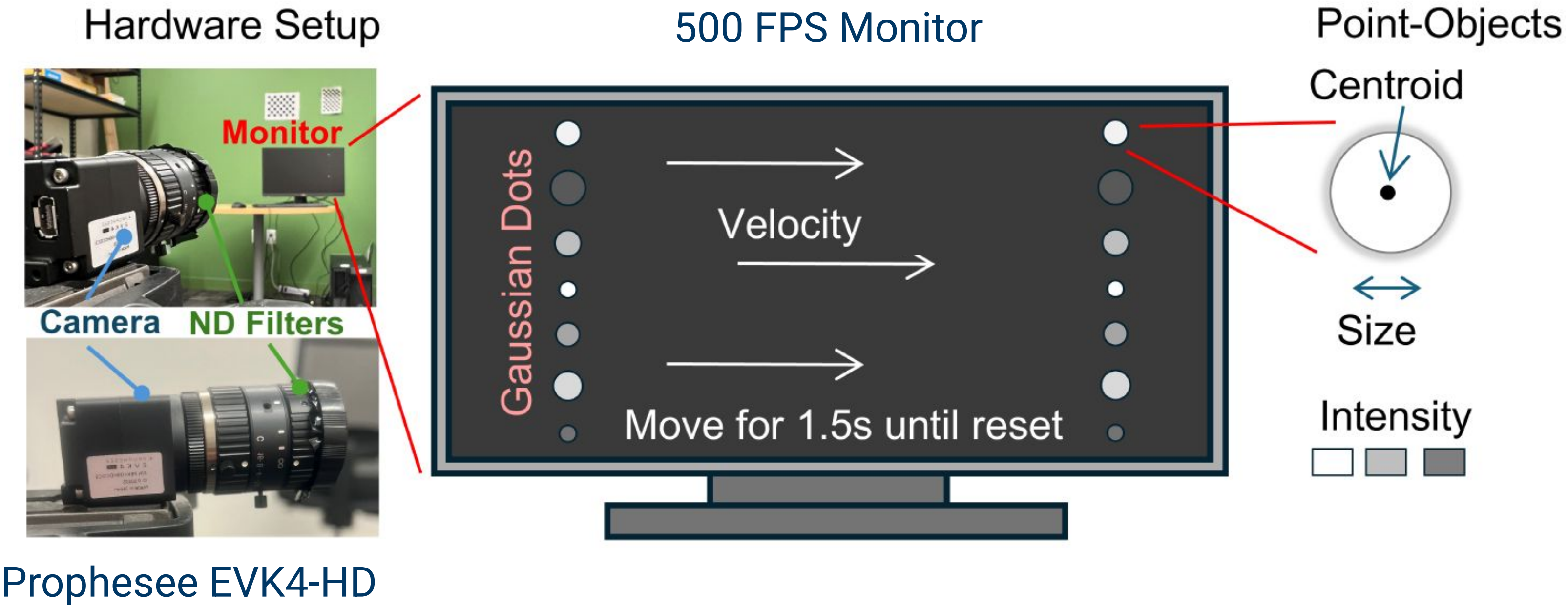
(lower CRB = more accurate estimator)

Two large takeaways:

- Temporal information is **not as important** for small time batches
- **Negative events are useful** in estimating velocity

Centroiding Dataset

- To train a model to leverage these spatial cues, we collect a point-object centroiding dataset



Centroiding Dataset

- To train a model to leverage these spatial cues, we collect a point-object centroiding dataset

Camera Resolution	Length (time)	# Unique Objects	# Unique 5ms Event Streams
1280x720	3 hours	10,588	2,117,600

Calibration Error	Obj Event Rates	Obj Inten Range	Obj Size Range	Obj Vel Range
$\sim$0.2 pixels	2 to 20kHz	8 to 300 μlux	1 to 8 pixels	15 to 350 pix/s

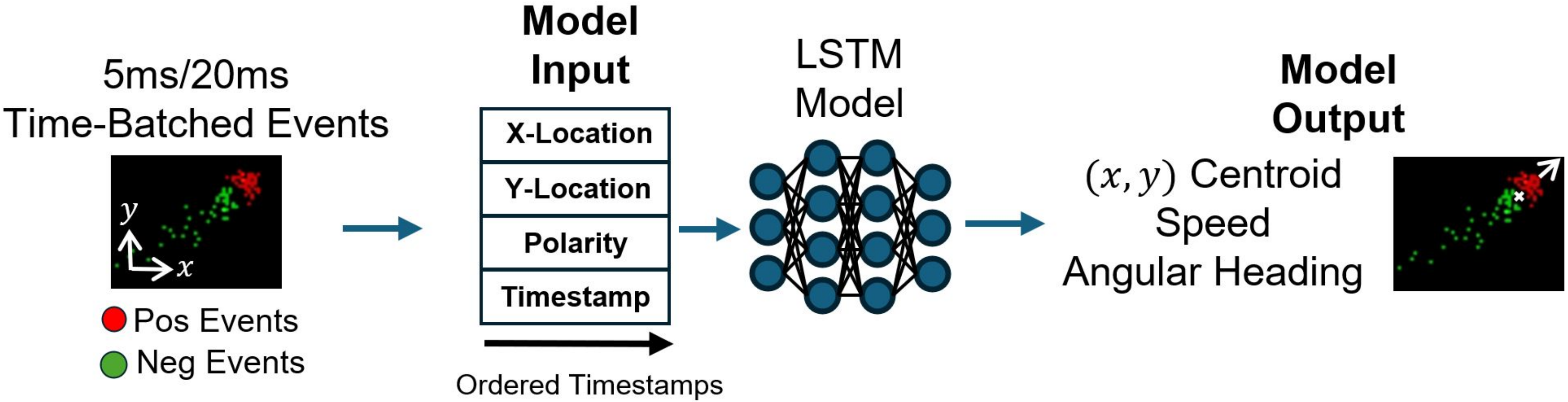
Examples of Dataset Capture

Monitor Video

Resulting Events

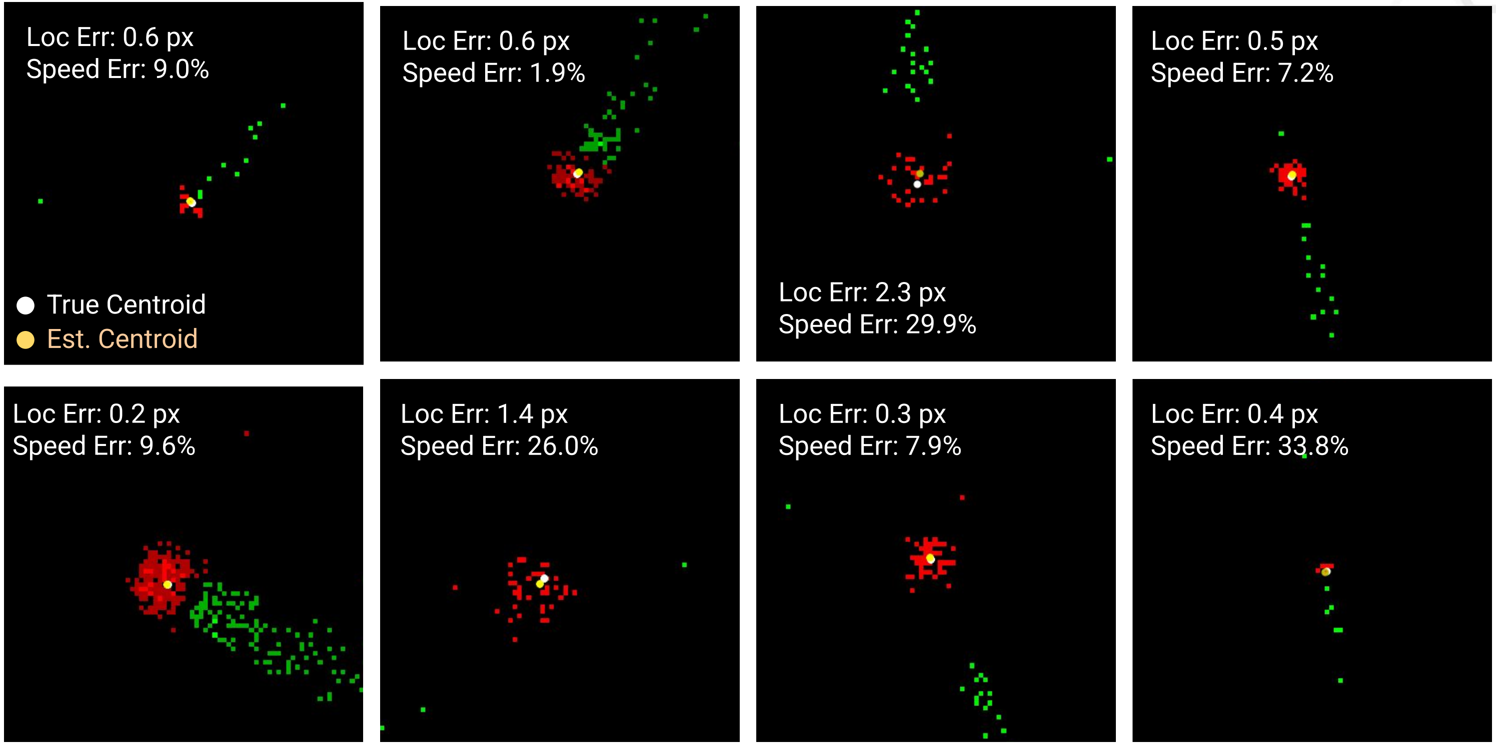
LSTM-based EBS Centroiding Approach

- **Model**: 3-layer LSTM, 500k total parameters
- Feed each event in one at a time, observe the output after the last event



kitware

Centroiding Results



Centroiding Results

Tested on 85,000 5ms and 20ms event streams
Best results **bolded**, second-best <u>underlined</u>

- **Centroiding Results:** Compared to previous methods, the proposed model has…
  - ~2x location accuracy
  - ~5x speed accuracy

(a) 5ms time batch

Method	Events Used	Location Error (pix)	SMAPE Speed Error (%)	Angular Error (deg)
Mean Cent.	Pos	<u>1.74 ± 1.45</u>	112.30 ± 56.60	72.42 ± 48.89
Mean Cent.	Pos, Neg	6.60 ± 3.00	136.82 ± 51.08	<u>15.91 ± 36.80</u>
Inten. Recon.	Pos, Neg	1.76 ± 1.47	111.27 ± 56.31	72.23 ± 48.76
Gauss MLE	Pos	2.34 ± 2.00	<u>102.34 ± 57.13</u>	67.93 ± 48.74
LSTM Model	Pos, Neg	**0.95 ± 0.70**	**17.75 ± 17.10**	**15.54 ± 30.60**

(b) 20ms time batch

Method	Events Used	Location Error (pix)	SMAPE Speed Error (%)	Angular Error (deg)
Mean Cent.	Pos	2.19 ± 1.36	42.04 ± 44.62	28.59 ± 35.87
Mean Cent.	Pos, Neg	7.51 ± 2.68	64.47 ± 48.53	<u>12.94 ± 35.25</u>
Inten. Recon.	Pos, Neg	<u>1.68 ± 1.33</u>	41.77 ± 44.34	28.56 ± 35.91
Gauss MLE	Pos	1.76 ± 1.38	<u>36.79 ± 41.38</u>	26.01 ± 34.80
LSTM Model	Pos, Neg	**0.76 ± 0.54**	**10.66 ± 10.35**	**6.27 ± 10.48**

5ms Time-Batch Centroiding Ablation Study

Event Information			Metrics		
Pos Events	Neg Events	Event Order	Loc Err (pix)	Speed Err (%)	Angular Err (deg)
✓	✓	✓	0.95	17.7	15.5
✓	✓	X	0.95	17.4	15.5
✓	X	✓	1.08	37.3	45.4
X	✓	✓	2.81	25.9	17.8

- Removing timestamps/event order *does not* affect model accuracy
  - ***Majority of information is in the spatial distribution*** for small time batches
- **Positive** events are important for *location estimation*
- **Negative** events are important for *velocity estimation*

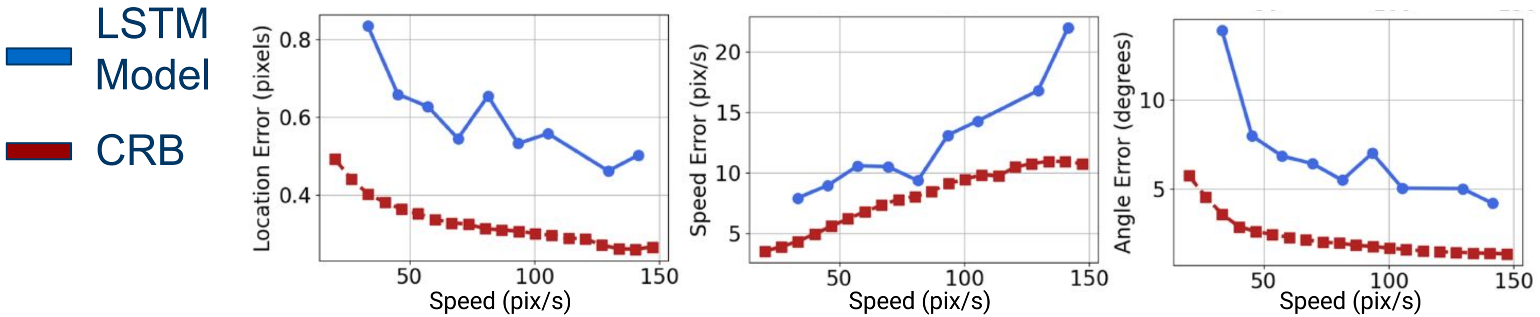
5ms Time-Batch Centroiding Ablation Study

Object Priors		Metrics		
Size + Inten Info	Vel Info	Loc Err (pix)	Speed Err (%)	Angular Err (deg)
✓	✓	0.81	NA	NA
✓	X	0.88	16.0	16.3
X	X	0.95	17.7	15.5

- Priors (size, intensity, and velocity) only offer minor improvement
- This suggests that the network can **successfully decorrelate these effects**

Cramér-Rao Bounds (CRB)

Constant object size, intensity



- We can compare the model's accuracy to CRB limitations
- Unlike previous analytical CRB, this is empirically calculated via the dataset
- Our model's accuracy approaches that limit or at least adopts a similar shape

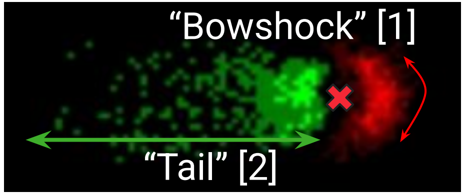
Summary

- **Objective**: leverage the spatiotemporal distribution of events to better localize unresolved objects in <20ms

Image Frame

Intensity

Event “Frame”



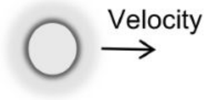
Positive Events

Negative Events

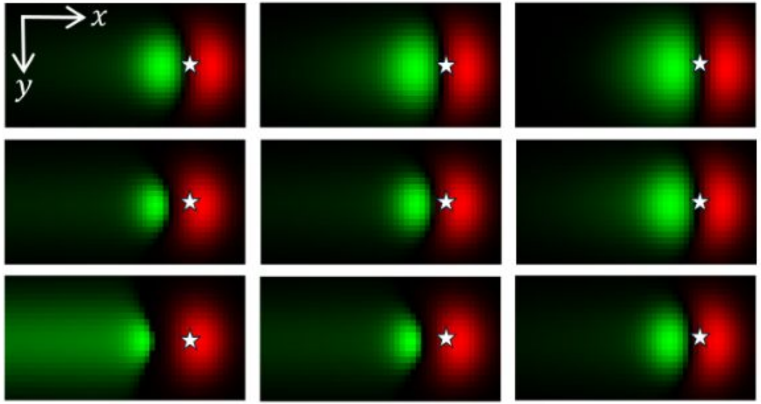
Summary

- **<u>Objective</u>**: leverage the spatiotemporal distribution of events to better localize unresolved objects in <20ms
- Analytical model demonstrates that the spatial distribution is dependent on the object speed, intensity, and location

Unresolved Object



Low Speed

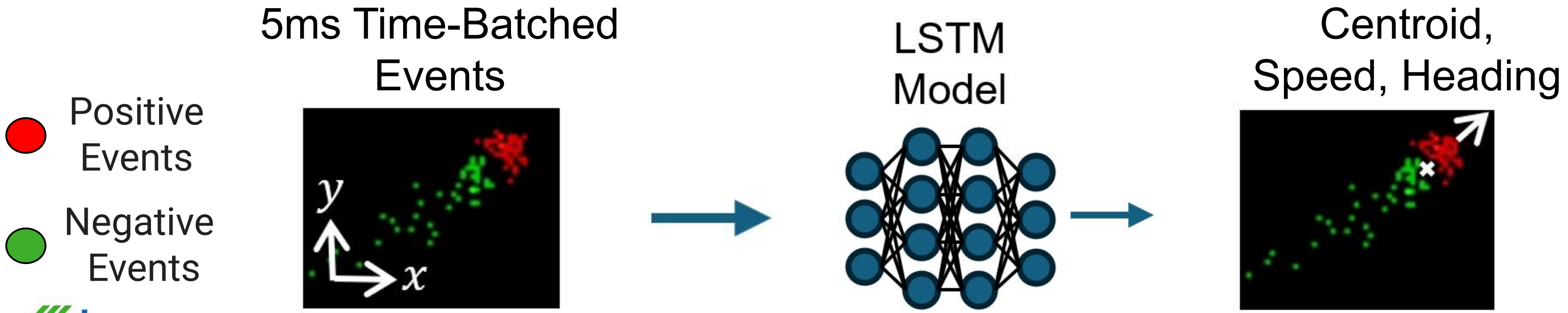


High Speed

Small Intensity → Large Intensity

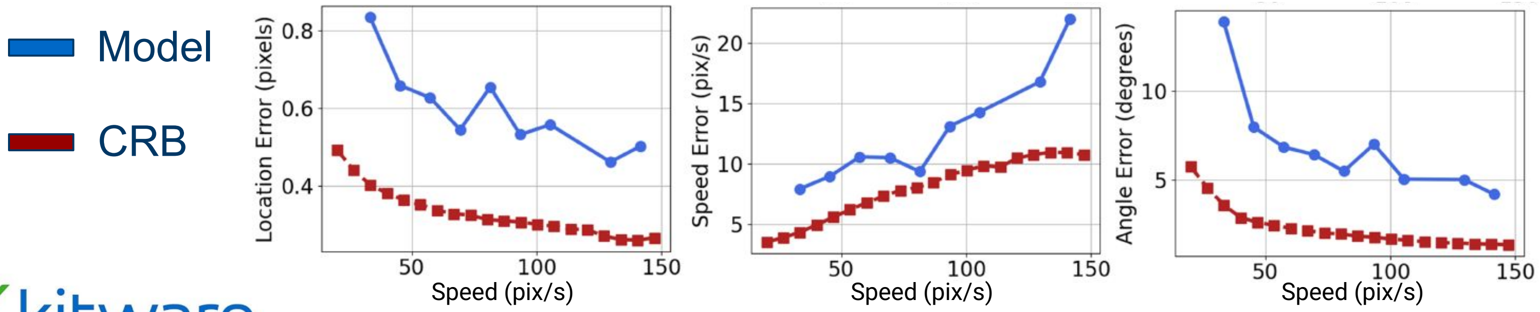
Summary

- **Objective**: leverage the spatiotemporal distribution of events to better localize unresolved objects in <20ms
- Analytical model demonstrates that the spatial distribution is dependent on the object speed, intensity, and location
- Proposed lightweight LSTM model can localize an object within **a pixel** and estimate velocity within **17%** with just 5ms of events



Summary

- **Objective**: leverage the spatiotemporal distribution of events to better localize unresolved objects in <20ms
- Analytical model demonstrates that the spatial distribution is dependent on the object speed, intensity, and location
- Proposed lightweight LSTM model can localize an object within **a pixel** and estimate velocity within **17%** with just 5ms of events
- Our model's performance can be compared to the Cramér-Rao Bounds (CRB)



kitware

Summary

- **Objective**: leverage the spatiotemporal distribution of events to better localize unresolved objects in <20ms
- Analytical model demonstrates that the spatial distribution is dependent on the object speed, intensity, and location
- Proposed lightweight LSTM model can localize an object within **a pixel** and estimate velocity within **17%** with just 5ms of events
- Our model's performance can be compared to the Cramér-Rao Bounds (CRB)
- Additional results include more ablation studies, additional analysis, and performance on star tracking
- Dataset/code: https://gitlab.kitware.com/nest-public/ebs-datadriven-centroiding