# X-maps: Direct Depth Lookup for Event-based Structured Light Systems


Fraunhofer HHI

[1] Fraunhofer Heinrich Hertz Institute

[2] Humboldt University

Wieland Morgenstern[1], Niklas Gard[1,2], Simon Baumann[1], Anna Hilsmann[1,2], Peter Eisert[1,2]



JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

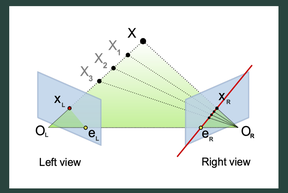

https://fraunhoferhhi.github.io/X-maps/

This project enables you to utilize **event cameras** to carry out **live depth estimation** from images projected with a laser projector.

We've streamlined the depth estimation process by creating a **lookup image with one spatial and one temporal axis** (y and t), forming an X-map.

This idea enables speedy calculation-free depth lookup (taking **less than 3 ms per frame**), but also maintains the accuracy of depth estimation through disparity search in time maps.

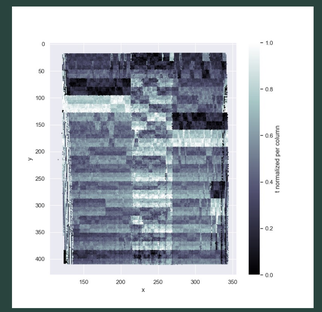
The end result is an efficient, reactive tool for designing real-time **Spatial Augmented Reality** experiences.



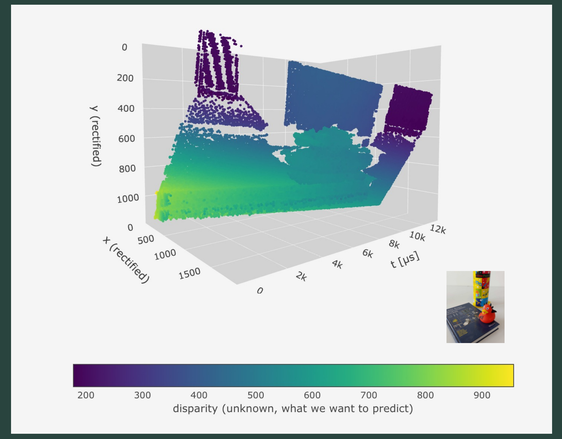

To estimate depth, we want to match the signal from the laser projector scanning over the scene (left view) with incoming events in the event camera (right view).

The system is calibrated, so events need to be matched on their epipolar line.

For matching, we get the projector x from the event time stamp.

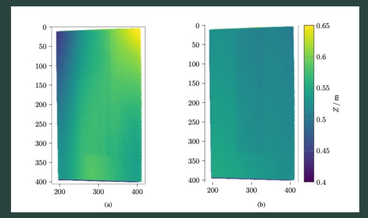



The time stamps in the faster scanning axis of the projector are too noisy to be useful. Event timestamps can only be used to determine the current line, not the position of the laser within that line. We rotate the projector upright to align the slow axis with the epipolar lines.




Events produced by the projector scanning over a scene visualized in time and space. Previous methods built time maps (x, y) → t from the data, and search along epipolar lines to match the time.

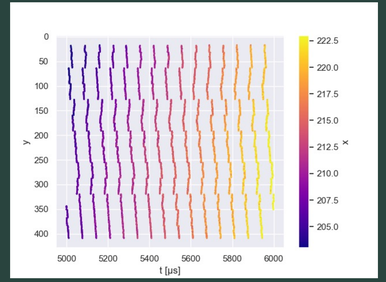
Our idea is to mix space and time, producing an X-map that maps (y, t) → x for the projector.



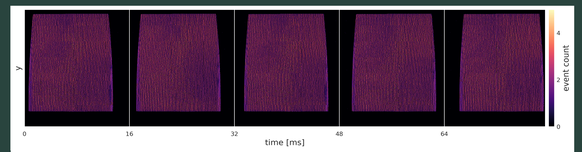

Projector scanning speed may not be constant. We create a time-calibrated X-map by measuring the laser scanning behavior projecting on a flat plane.

## Highlights (TLDR)

- Depth from event timestamp
- Real-time estimation on CPU
- Cheap laser projector
- No hardware sync
- Project arbitrary content: no pattern required




A zoomed-in view of an X-map of incoming events. Different scan lines from the projector become visible in this spatio-temporal visualization.



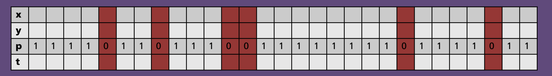

A sequence of camera X-maps

Duck dataset: M. Muglikar, G. Gallego, and D. Scaramuzza, **ESL: Event-based Structured Light**, In 2021 International Conference on 3D Vision (3DV)
Epipolar geometry visualization: By Arne Nordmann (norro) - Own work (Own drawing), CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1702052

Our real-time depth estimation and reprojection demonstrator reads events from the camera, then processes them in these steps:

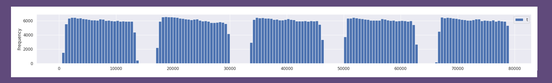
**Processing sync**
Keep sync if processing lags behind at very high frame rates

**Polarity filter**




**Activity noise filter**
Remove isolated events outside projected area

**Trigger finder**




**Duplicate coordinate filter**


**X-Maps disparity lookup**

$x_{p_r} = x(y_{c_r}, t)$

$d = x_{p_r} - x_{c_r}$

Lookup rectified camera y coordinate and normalized event time in the X-map to retrieve rectified projector x coordinate.

**Disparity map**

**Depth coloring**

| Method | Runtime (abs.) | Runtime (rel.) |
|---|---|---|
| ESL (CPU) | 174.68 s (±26.97) | > 10000× |
| ESL-init (CPU) | 11.87 s (±2.13) | > 1000× |
| ESL-init (CUDA) | 18.99 ms (±0.88) | 7.12× |
| X-maps (ours, CPU) | **2.67 ms** (±0.31) | **1.0**× |

Our processing is pretty quick, while keeping the quality comparable to ESL-init (temporal epipolar search).