Semi-Dense 3D Reconstruction with a Stereo Event Camera

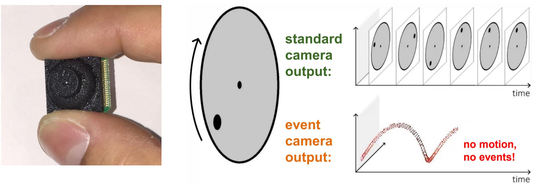
Yi Zhou, Guillermo Gallego, Henri Rebecq, Laurent Kneip, Hongdong Li, Davide Scaramuzza

Australian National University · Shanghaitech University · University of Zurich (Department of Informatics) · ETH zürich (Institute of Neuroinformatics)

Motivation: Address the problem of **stereo 3D reconstruction** for VO/SLAM using event cameras alone.

Goal: Unlock the potential of event cameras by exploiting the **temporally asynchronous** and **spatially sparse** nature of event data.

What is an event camera?



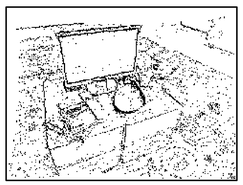
- Only transmits **brightness changes**.
- Output is a stream of **asynchronous events**.
- **Advantages**: low latency, no motion blur, HDR.

Watch video!

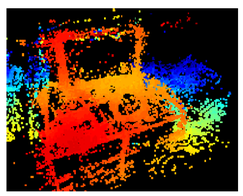
Stereo event-camera setup

Cameras	DAVIS
Width	240 pix
Height	180 pix
FOV	62.9°
Baseline	14.7 cm

Scene



Events on the left view



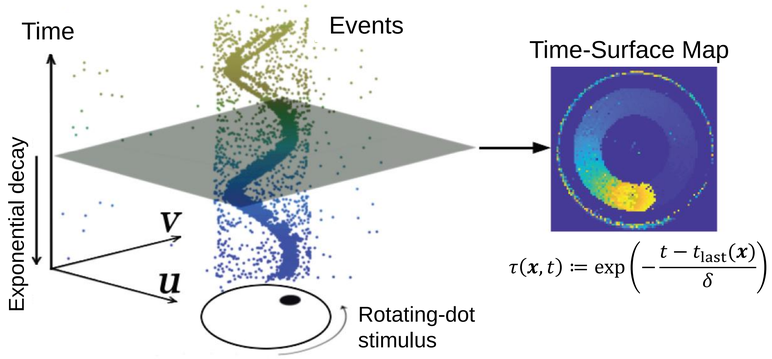
Inverse Depth Map

3D Reconstruction

Approach

Optimize the **spatio-temporal consistency of events** across stereo image planes in small-baseline setups.

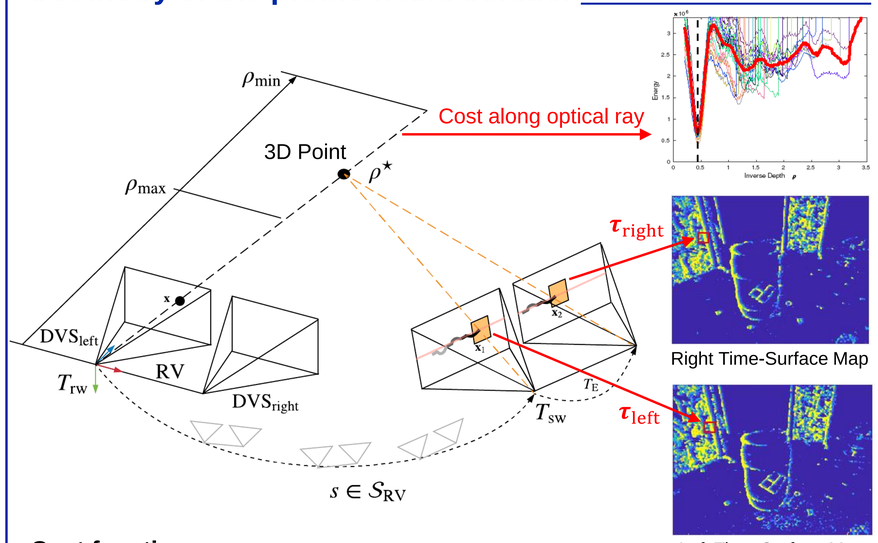
Event representation: Time-Surface Maps



$$\tau(\boldsymbol{x}, t) := \exp\left(-\frac{t - t_{\text{last}}(\boldsymbol{x})}{\delta}\right)$$

- An exponential decay kernel on the last spiking time t_{last} at $\boldsymbol{x} = (u, v)^T$.
- The decay rate parameter $\boldsymbol{\delta}$ is a small constant (~ 30 ms).

Geometry of the problem and solution



Cost function:

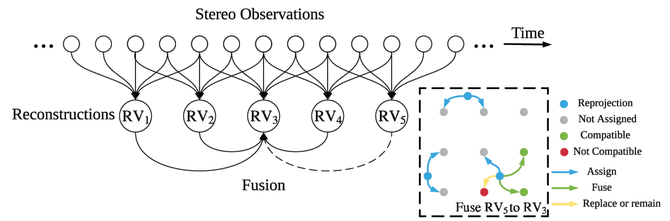
$$C(\boldsymbol{x}, \rho) := \frac{1}{|\mathcal{S}_{RV}|} \sum_{s \in \mathcal{S}_{RV}} \left\| \tau^s_{\text{left}}(\boldsymbol{x}_1(\rho)) - \tau^s_{\text{right}}(\boldsymbol{x}_2(\rho)) \right\|_2^2$$

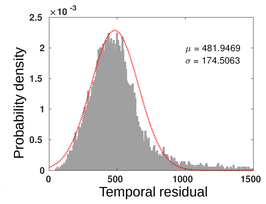
Best inverse depth: $\rho^* = \operatorname{argmin}_\rho C(\boldsymbol{x}, \rho)$

- ρ^* maximizes the spatio-temporal consistency of events.
- Optimizer: Gauss-Newton method.

Depth Map Fusion

- Improve the density or reconstruction.
- Reduce the depth uncertainty.



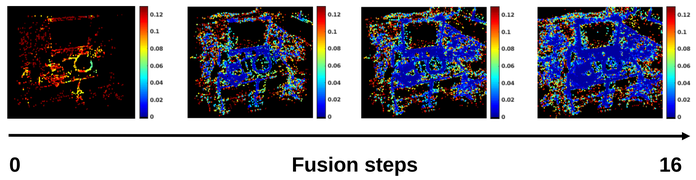


Inverse-depth uncertainty:

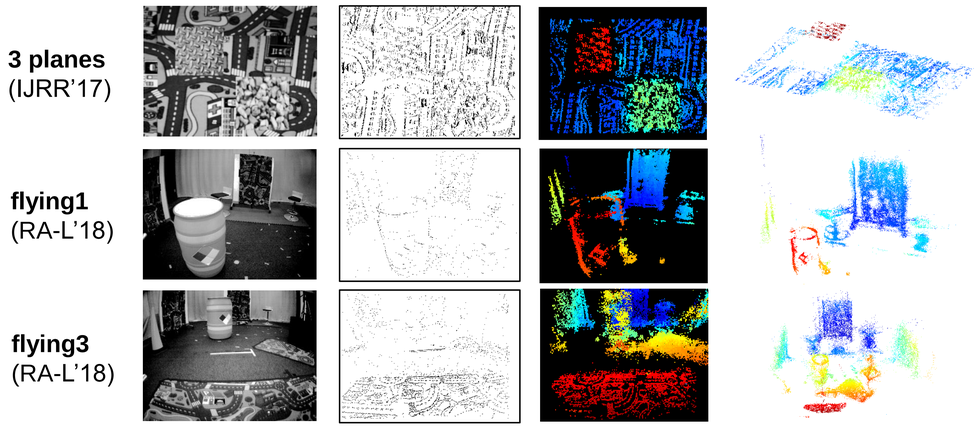
$$\sigma^2_{\rho^\star} \approx \left(\frac{\partial \rho^\star}{\partial \mathbf{r}}\right)^T (\sigma_r^2 \mathrm{Id}) \frac{\partial \rho^\star}{\partial \mathbf{r}}$$

Fused depth distribution:

$$\mathcal{N}\left(\frac{\sigma_a^2 \rho_b + \sigma_b^2 \rho_a}{\sigma_a^2 + \sigma_b^2}, \frac{\sigma_a^2 \sigma_b^2}{\sigma_a^2 + \sigma_a^2}\right)$$



Experiments



(3D Errors)	Dataset	3 planes	flying1	flying3
	Depth range	2.76 m	4.96 m	5.74 m
Our Method	Mean error	**0.03** m	**0.13** m	**0.33** m
	Median error	**0.01** m	**0.05** m	**0.11** m
	Relative error	**1.17** %	**2.65** %	**5.79** %
FCVF (PAMI'13)	Mean error	0.05 m	0.99 m	1.03 m
	Median error	0.03 m	0.25 m	**0.11** m
	Relative error	1.84 %	20.8 %	17.3 %
SGM (PAMI'08)	Mean error	0.08 m	0.93 m	1.19 m
	Median error	0.03 m	0.31 m	0.20 m
	Relative error	3.22 %	18.7 %	20.8 %

Sponsors: Australian Centre for Robotic Vision · robotics Swiss National Centre of Competence in Research · FNSNF Swiss National Science Foundation