Event-based Mosaicing Bundle Adjustment

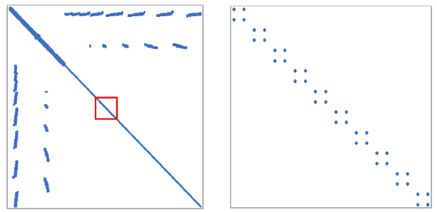
Shuang Guo and Guillermo Gallego

Project page

Summary

We tackle the problem of **simultaneous refining the camera orientations and scene map** for a **purely rotating event camera**.
We formulate the problem as **regularized non-linear least squares** (NLLS) optimization. The loss function is defined using the **linearized event generation model** in the camera orientations and the panoramic gradient map of the scene.

Problem sparsity (Normal Eqs)



$$\begin{pmatrix} \mathtt{A}_{11} & \mathtt{A}_{12} \\ \mathtt{A}_{12}^{\top} & \mathtt{A}_{22} \end{pmatrix} \begin{pmatrix} \Delta \mathbf{P}_{\alpha}^{*} \\ \Delta \mathbf{P}_{\beta}^{*} \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}$$

Quantitative Results

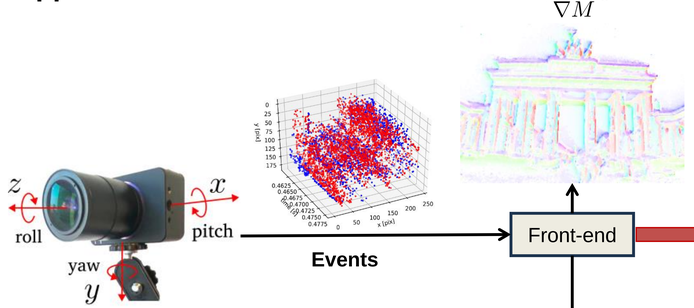
- Rotation RMSE [deg]

Front-end	playroom	bicycle	city	street	town	bay
EKF-SMT	5.86	1.47	1.69	3.44	4.32	2.50
	6.09	1.18	1.68	3.46	4.40	2.41
CMax-GAE	4.63	1.65	—	—	4.66	—
	4.42	1.50	N/A	N/A	4.53	N/A
CMax-ω	3.22	1.69	1.53	0.97	1.91	1.80
	2.86	0.92	0.97	0.74	0.86	1.41

- Squared Event-based Photometric error [$\cdot 10^6$]

Front-end	playroom	bicycle	city	street	town	bay
EKF-SMT	0.35	0.52	2.62	1.82	1.88	2.26
	0.23	0.30	2.13	1.52	1.51	1.96
CMax-GAE	0.35	0.53	—	—	1.90	—
	0.19	0.31	N/A	N/A	1.54	N/A
CMax-ω	0.33	0.55	2.71	1.90	1.92	2.30
	0.15	0.30	1.98	1.34	1.43	1.83

Approach



EMBA back-end

Methodology:

- We start from the Linearized Event Generation Model (LEGM):

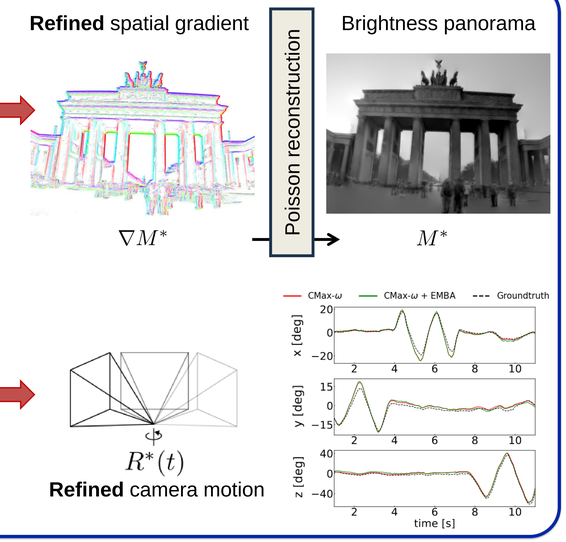
$$\Delta L \approx -\nabla L(\mathbf{x}_k, t_k) \cdot \mathbf{v} \Delta t_k = s_k C$$

- The final objective function with L^2 map regularization:

$$\min_{\{\mathtt{R}_i\}, \nabla M} \underbrace{\sum_{k=1}^{N_{\text{events}}} \Big(\underbrace{\nabla M \cdot \mathbf{v}(\{\mathtt{R}_i\}) \Delta t_k}_{\text{LEGM predicts } \Delta L} - \underbrace{s_k C_{\text{th}}}_{\text{Measured } \Delta L} \Big)^2}_{\text{Photometric error (NLLS)}} + \underbrace{\eta \, \|\nabla M\|^2}_{\text{Regularizer}}$$

Key properties:

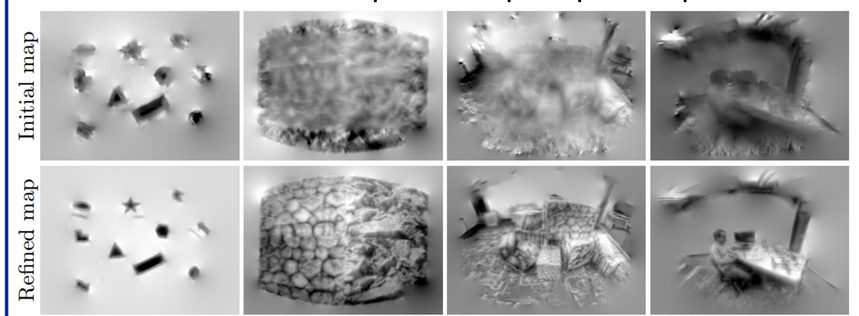
- Refines camera rotations and panoramic gradient map by minimizing the photometric error conveyed by all events.
- Leverages map regularization to improve iteration convergence.
- Exploits the sparsity of the Normal Equations for speed-up.
- Can be initialized by rotations estimated by various front-ends.
- Works well with various event camera resolutions, such as DAVIS, DVXplorer, and Prophesee's 1Megapixel camera.



Qualitative Results

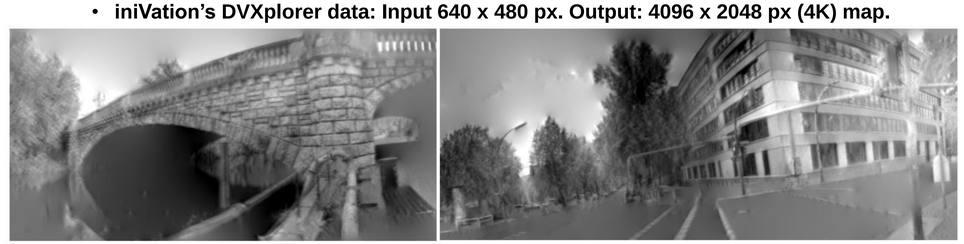
Synthetic data: Input 240 x 180 px. Output: 2048 x 1024 px (2K) map.

- **Real-world data: Input 240 x 180 px. Output: 1K map.**



- **Prophesee's EVK4 data: Input 1280 x 720 px. Output: 4096 x 2048 px (4K) map.**

- **iniVation's DVXplorer data: Input 640 x 480 px. Output: 4096 x 2048 px (4K) map.**



Guo and Gallego, "Event-based Mosaicing Bundle Adjustment", European Conference on Computer Vision (ECCV), 2024.