Event-Based Visual Teach-and-Repeat via Fast Fourier-Domain Cross-Correlation

Gokul B. Nair Alejandro Fontan Michael Milford Tobias Fischer

Centre for Robotics

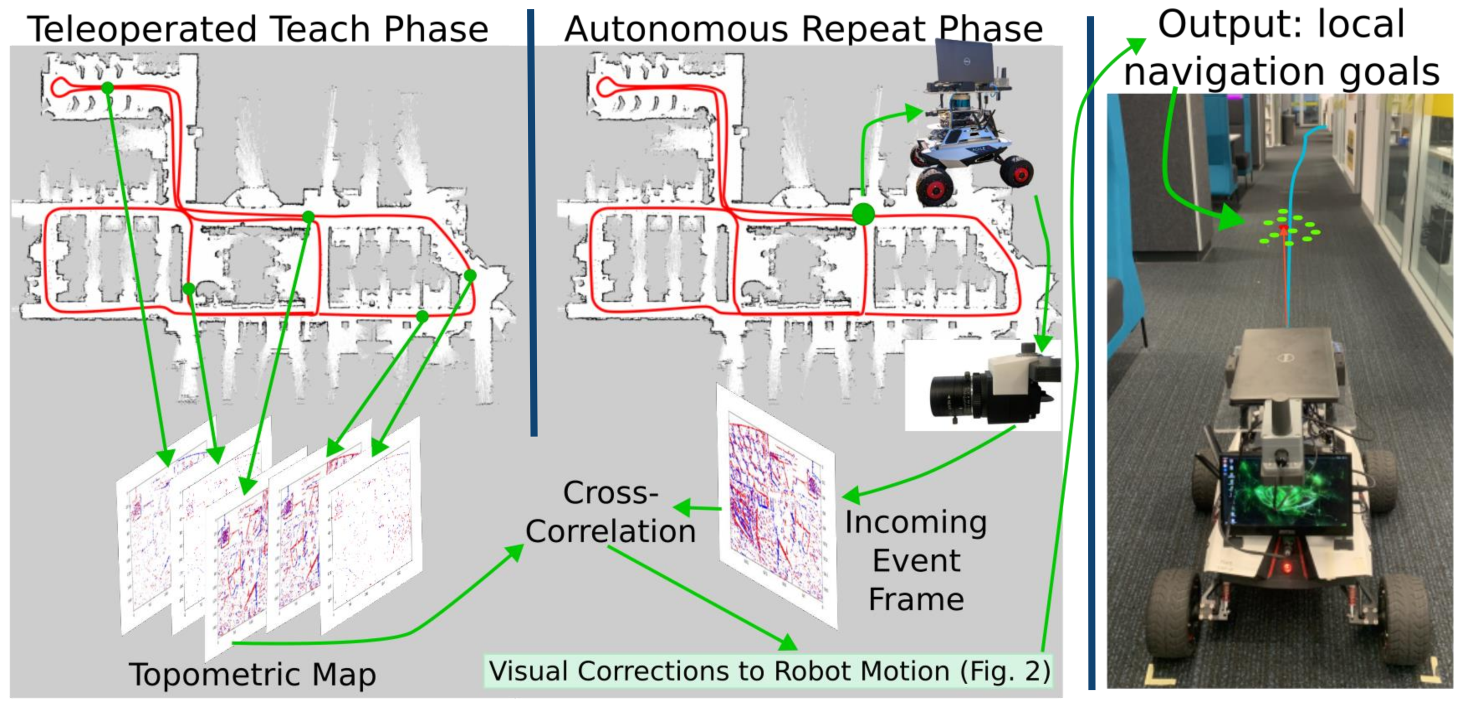
github.com /gokulbnr/event-vtr

gokulbnr @gmail.com

IROS 2026 PITTSBURGH

Motivation

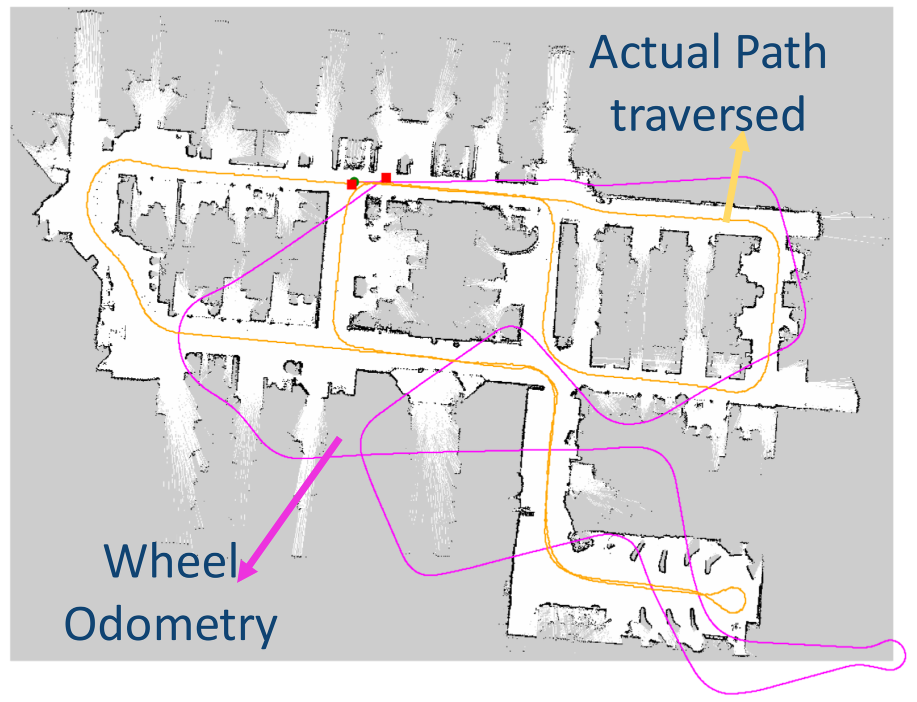
Teach-and-Repeat Navigation



Teach Phase
- Robot is teleoperated over a traverse.
- Wheel odometry and frames are stored as a topometric map.

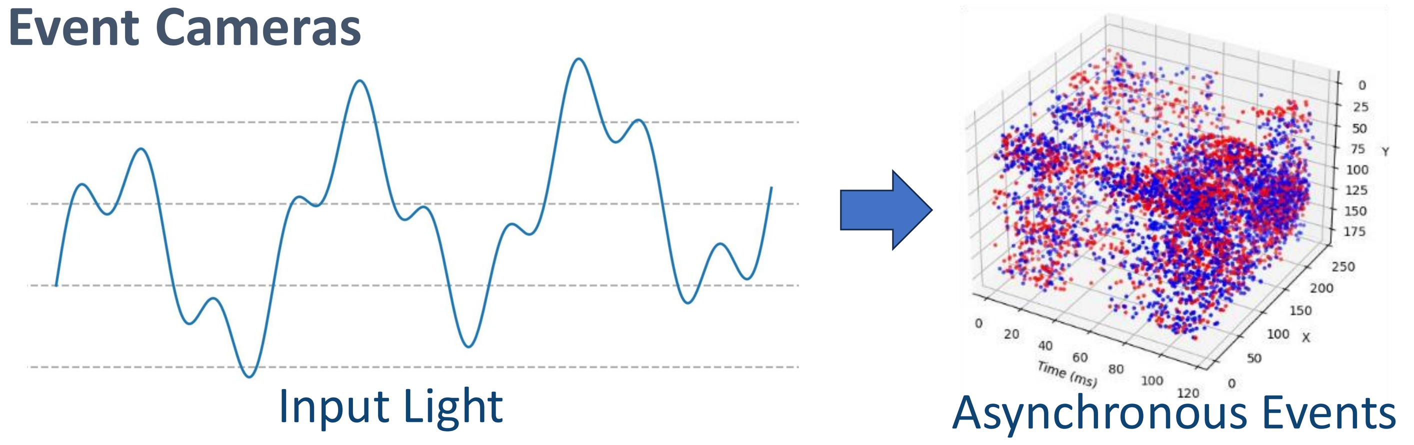
Repeat Phase
- Robot repeats traverse autonomously by tracking wheel odometry sequence.
- Frame-matching with the map generates corrections keeping the robot from deviating from path.

Why incorporate visual corrections?



- Wheel-odometry has an incremental drift.
- Visual corrections keep robot on track.

Event Cameras

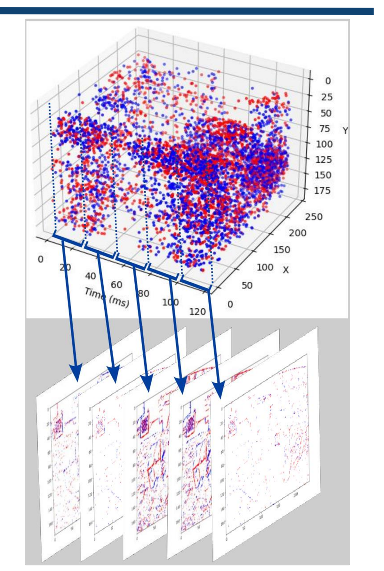


- Event cameras capture changes in input light intensity.
- Data is transmitted asynchronously from individual pixel circuits.
- They offer a higher temporal resolution than conventional frame-based cameras.

Can event-based sensing help produce high throughput navigation corrections in visual teach-and-repeat navigation?!

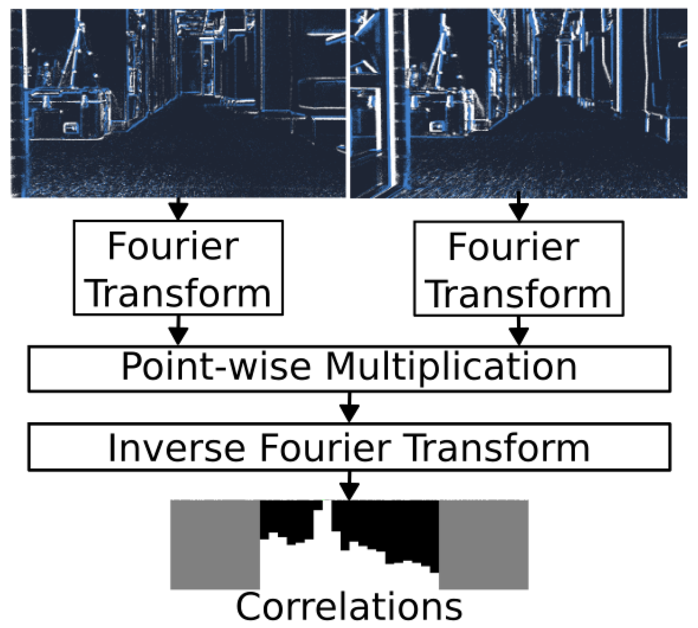
Approach

Event Frames from Asynchronous Events



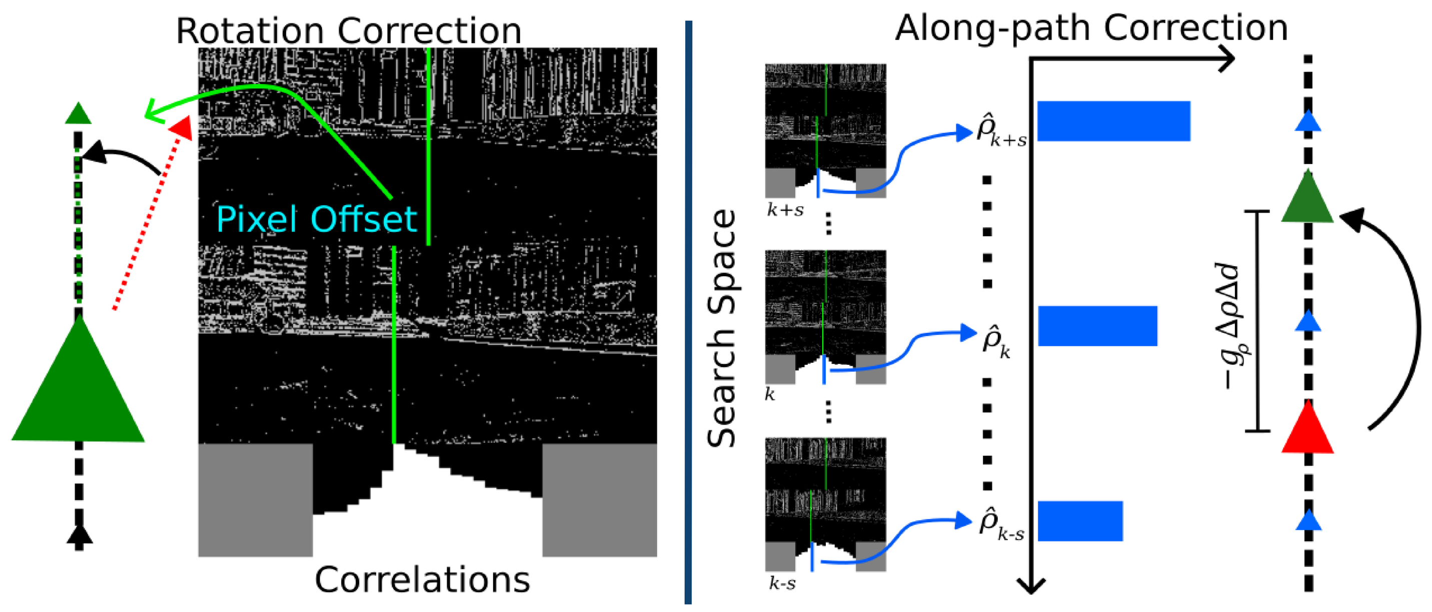
- Binary frames by accumulating events.
- A pixel value indicates occupancy of an event.
- We compare two event accumulation strategies: fixed-time and fixed-count.

Frame-Matching: Fourier-Domain Cross-Correlation

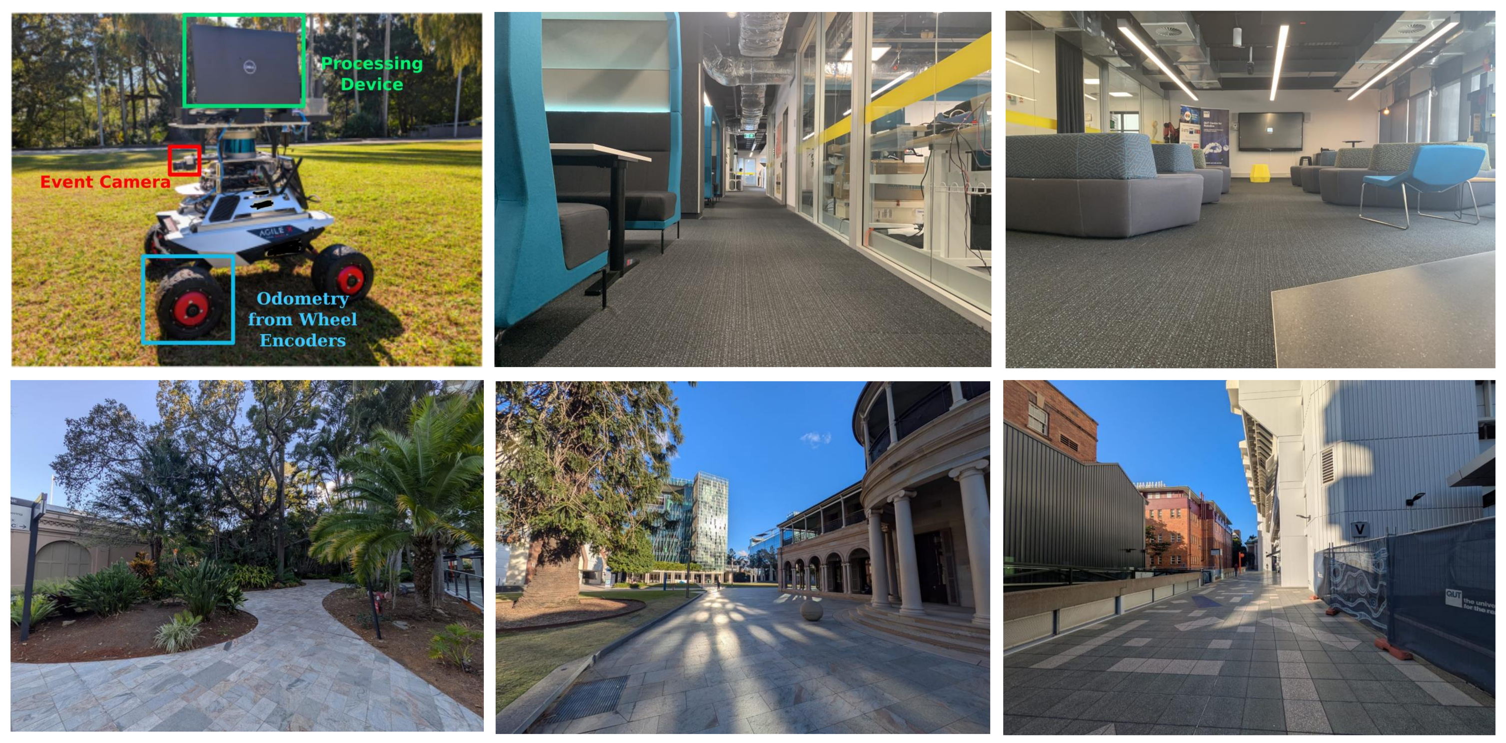


- Correlation peak indicates offset between the frames.
- Horizontal-only padding ensures no vertical offset is computed.
- Candidates are concatenated side by side and column-summed, so many-to-one matching runs in a single, cheap pass

Motion Corrections from Correlation Scores



Setup and Experiment Locations



- Motion controller: Sliding Mode Controller (SMC).

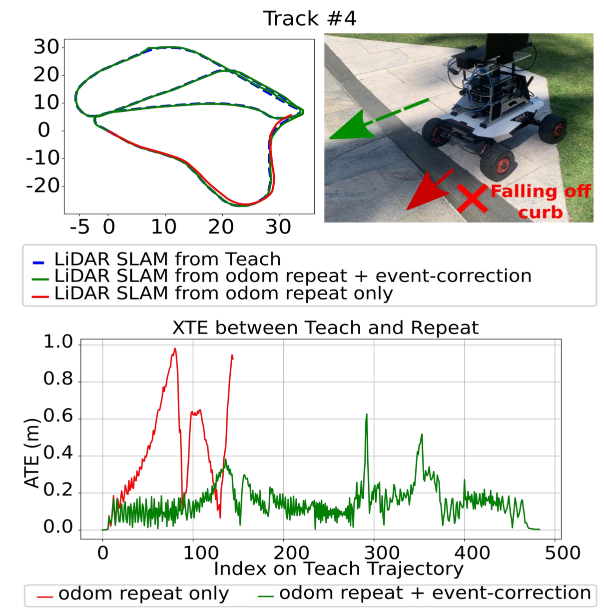
Results

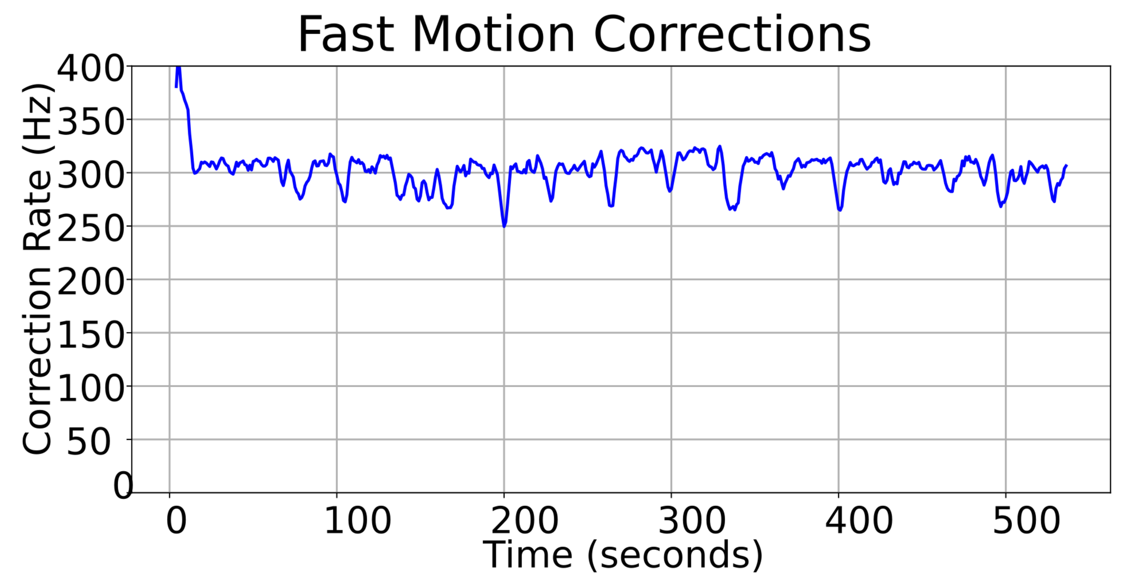
Baselines

- Dall'Osto et. al. [1]: Conventional frames matched by normalised cross correlation (NCC), run using a proportional controller.
- Nourizadeh et. al. [2]: Conventional frames matched by NCC, run using sliding mode controller (SMC).

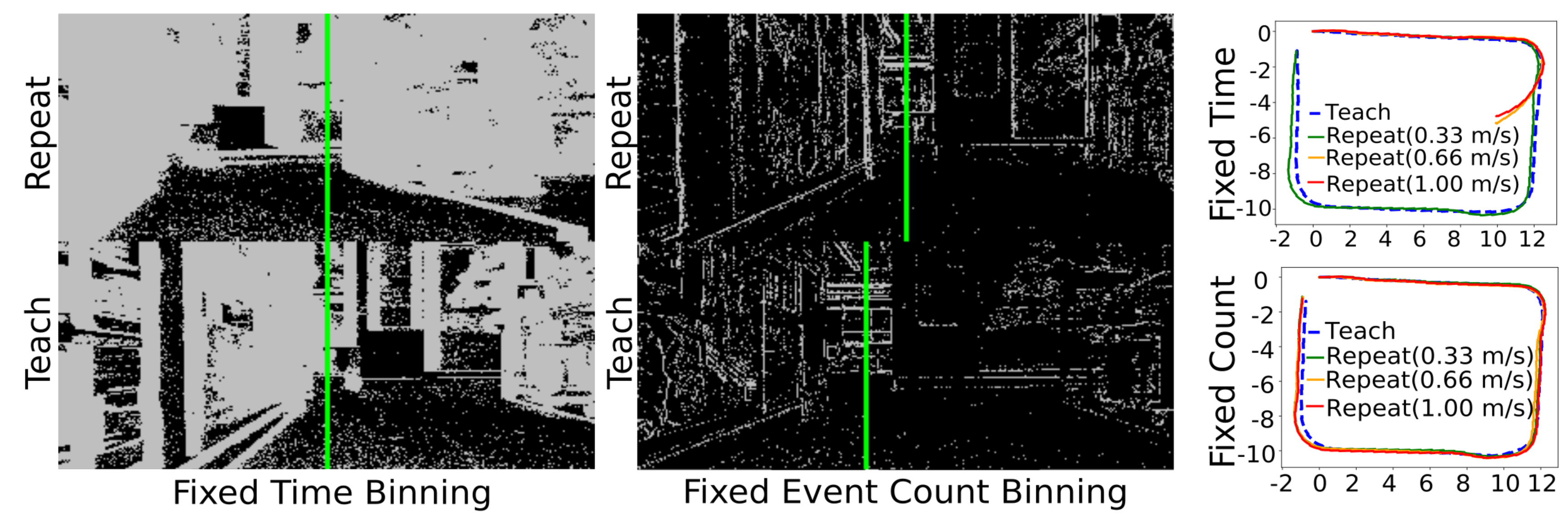
Env.	Tr.#	Len. (m)	Dall'Osto et al. [1] mean±SD	SR	Nourizadeh et al. [2] mean±SD	SR	Event VT&R (Ours) mean±SD	SR
Ind.	1	65	6.5±6.4	3/3	4.5±3.7	3/3	6.3±5.8	3/3
	2	100	8.8±9.1	3/3	9.0±8.4	3/3	8.4±8.3	3/3
	3	200	8.7±9.5	3/3	11.3±10.3	3/3	9.4±7.7	3/3
	4	120	17.1±21.3	3/3	8.6±8.0	3/3	7.8±7.2	3/3
Out.	5	223	17.4±18.7	3/3	17.0±18.2	3/3	10.7±9.6	3/3
	6*	381	15.4±18.1	3/3	8.2±9.5	3/3	11.1±14.1	3/3

*Night time traverse





Event-Accumulation Strategies and Velocity Invariance



- Accumulating events by fixed-count generates velocity invariant frames.

References

[1] Dall'Osto, Dominic, Tobias Fischer, and Michael Milford. "Fast and robust bio-inspired teach and repeat navigation." IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2021.

[2] Nourizadeh, Payam, Michael Milford, and Tobias Fischer. "Teach and repeat navigation: A robust control Approach." IEEE International Conference on Robotics and Automation (ICRA) 2024.