Mosaicing and 3-DoF Tracking with an Event Camera *based on the work of Kim et al.* [1]

Céline Nauer[1], Nik Dennler[1], Joël Bachmann[2]
[1]Institute of Neuroinformatics, University of Zurich; [2]ETH Zürich

1 Event Cameras

Event-based cameras output a stream of asynchronous spikes, recorded and transmitted by each pixel individually. The events are triggered through a certain log intensity change and consist of a timestamp, the pixel coordinates and the polarity of the change.

Key advantages are the low latency and thus virtually no motion blur, a high dynamic range and a reduced bandwidth.

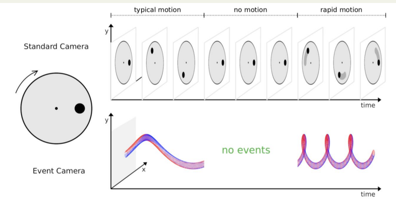

Fig. 1. Functionality of an event camera [1]

2 Method overview

Tracking of the rotational pose of an event camera and simultaneously reconstructing the intensity map of an indoor scene.

This project's aim was to reconstruct and apply the algorithm proposed by [1] to a simulated DVS dataset and to a self-recorded DAVIS data set.

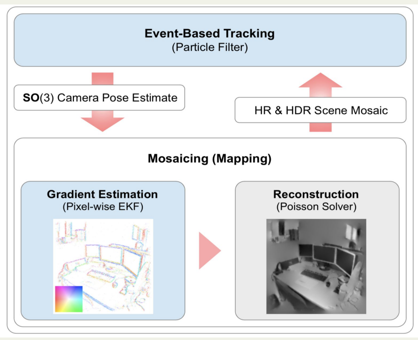

Fig. 2. Mapping and Tracking algorithm [1]

Intensity Reconstruction

1. Extended Kalman Filter (EKF) for the gradient map reconstruction.
2. Poisson solver for the intensity map reconstruction.

Tracking

Particle Filter to track the 3 DoF of an event camera, assuming constant position motion model, rotation matrices as particles and one initial pose. Iterating over the following steps:

1. Motion Update: Perturb particles randomly with a given standard deviation.
2. Measurement Update: For every firing pixel, obtain previous and current pose and get corresponding intensity change from intensity map. Weight particles according to sensor specific event likelihood.
3. Resampling: Collect a new set of particles by sampling the previous according to the weights.

3 Data

Simulated Data Set
Simulated data set of a cluttered indoor scene. Prepared by the Robotics & Perception Group at UZH [3]. Includes event stream as well as ground truth poses.

Recorded Data Set
Recording of the event stream, RGB images and the poses with a Dynamic and Active Pixel Vision Sensor (DAVIS). Calibration of the with checkerboard technique.

4 Results

In a first step, the reconstruction was built based on the event stream and the ground truth poses of the simulated dataset. The output is shown in Fig. 3.

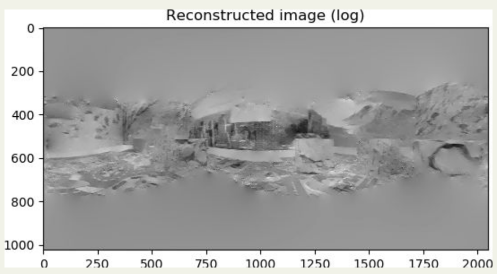

Fig. 3. Intensity reconstruction of simulated data set

In a second step, the tracker was run based on the event stream and the reconstructed map. The output (yellow) and the ground truth (red) are shown in Fig. 4.

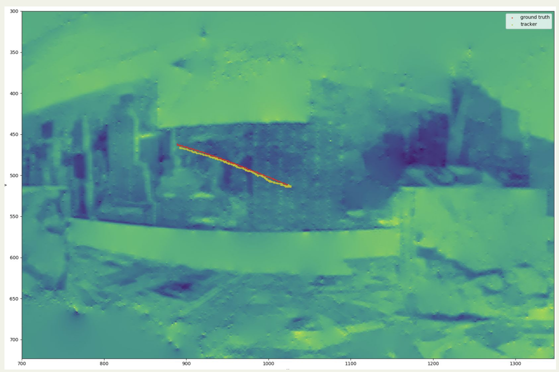

Fig. 4. Ground truth and estimated positions on reconstructed map

Currently we are working on obtaining the reconstruction and pose estimation simultaneously. Also, we try to run the algorithm on the recorded dataset. Due to the noisy nature of the camera the feasibility of this is not guaranteed.

5 Evaluation

To compare the output of the tracker to the ground truth, their resulting rotation matrices were applied to an arbitrary vector and plotted on the unit sphere, which is shown in Fig. 5.

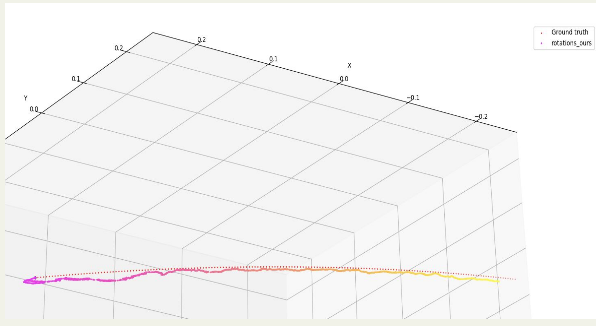

Fig. 5. Ground truth and estimated positions on unit sphere

Fig. 6. shows the calculated score based on the absolute difference between the individual quaternions as well as the RMSE.

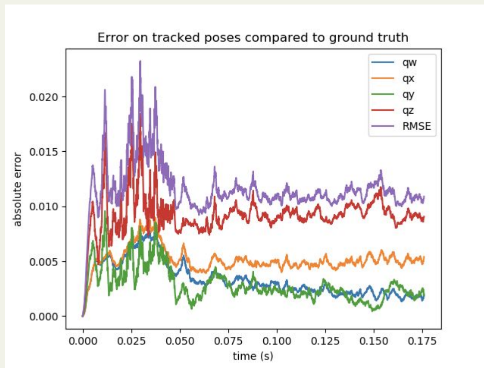

Fig. 6. Score on estimated quaternions if compared with ground truth.

6 Conclusion

The Extended Kalman Filtering and Particle Filtering approach for scene reconstruction and tracking seems to perform reasonably well for simulated data without noise. For real data, a pure Kalman Filtering approach, which is not as susceptible to noise, might be the better choice.

The Particle Filter is a statistical method using hundreds of particles being processed for each event, thus the computational performance is extremely expensive and unsuitable for real-time implementation.

To optimize the model, one could further tune the parameters or filter the event stream in a preprocessing step.

7 References

1. H. Kim, A. Handa, R. Benosman, S.-H. Ieng, A.J. Davison, *Simultaneous Mosaicing and Tracking with an Event Camera*. British Machine Vision Conference, 2014.
2. Hanme Kim, *Real-time visual SLAM with an event camera*, Imperial College London, UK, 2017.
3. G. Gallego, C. Forster, E. Mueggler, D. Scaramuzza, *Event-based Camera Pose Tracking using a Generative Event Model*. arXiv:1510.01972, 2015.