Comparing Representations in Tracking for Event Camera-based SLAM

Jianhao Jiao†, Huaiyang Huang†, Liang Li‡, Zhijian He†, Yilong Zhu†, Ming Liu

†The Hong Kong University of Science and Technology
‡The University of Hong Kong

CVPR 2021 Workshop on Event-based Vision
Motivation: Event Camera-only SLAM
Related Work: Event Camera-only SLAM (needs GPU)

Kim, Hanme, Stefan Leutenegger, and Andrew J. Davison. "Real-time 3D reconstruction and 6-DoF tracking with an event camera." European Conference on Computer Vision. Springer, Cham, 2016.
Related Work: Event Camera-only SLAM (CPU-only)

EVO: monocular event camera-based VO

Rebecq, Henri, et al. "Evo: A geometric approach to event-based 6-dof parallel tracking and mapping in real time." IEEE Robotics and Automation Letters 2.2 (2016): 593-600.
**Related Work: Event Camera-only SLAM (CPU-only)**

Rebecq, Henri, et al. "Evo: A geometric approach to event-based 6-dof parallel tracking and mapping in real time." IEEE Robotics and Automation Letters 2.2 (2016): 593-600.

Zhou, Yi, Guillermo Gallego, and Shaojie Shen. "Event-based stereo visual odometry." IEEE Transactions on Robotics (2021).
Different Event Representations in Tracking

• Tracking Problem Formulation:

\[ \sum_x \left[ I(W(x; p)) - T(x) \right]^2 \]

• Two event representations:

  1. Event Map (EM) in EVO: asynchronous output; fixed number of events; fast generation
Different Event Representations in Tracking

- **Tracking Problem Formulation:**
  \[ \sum_x \left[ I(W(x; p)) - T(x) \right]^2 \]

- **Two event representations:**
  1. Event Map (EM) in EVO: asynchronous output; fixed number of events; fast generation
  2. Time Surface Map (TS) in ESVO: synchronous output; implicit distance field for tracking

\[ I(x, t) = \exp \left( - \frac{t - t_{last}(x)}{\delta} \right) \]
Different Event Representations in Tracking

• Tracking Problem Formulation:

\[
\sum_x \left[ I(W(x; p)) - T(x) \right]^2
\]

• Two event representations:
  1. Event Map (EM) in EVO: asynchronous output; fixed number of events; fast generation
  2. Time Surface Map (TS) in ESVO: synchronous output; implicit distance field for tracking

Our interest: how do the different event representations influence the tracking performance?
Combine EM with TS with the degeneracy check

• Degeneracy factor $\lambda$: the minimum eigenvalue of the Hessian matrix

Zhang, Ji, Michael Kaess, and Sanjiv Singh. "On degeneracy of optimization-based state estimation problems." 2016 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2016.
Experimental Results

1. Simulated planar sequences + simulated 6DoF sequences [1]
2. RPG handheld 6DoF sequences [2]
3. UPenn UAV 6DoF sequences [3]

[1] Rebecq, Henri, Daniel Gehrig, and Davide Scaramuzza. "ESIM: an open event camera simulator." Conference on Robot Learning. PMLR, 2018.
[2] Zhou, Yi, Guillermo Gallego, and Shaojie Shen. "Event-based stereo visual odometry." IEEE Transactions on Robotics (2021).
[3] Zhu, A. Z., Thakur, D., Ozaslan, T., Pfommer, B., Kumar, V., & Daniilidis, K. (2018). The Multi Vehicle Stereo Event Camera Dataset: An Event Camera Dataset for 3D Perception. IEEE Robotics and Automation Letters, 3(3), 2032-2039.
Experimental Results: Figure Explanation

Real-time 3D map

Intensity image

Inverse depth of the 3D map

3D map aligned on Representation (TS or EM)

Estimated trajectories w.r.t. GT
Comparison on rpg_bin

Trajectories: Estimated w.r.t. GT
Comparison on upenn_indoor_flying3

Trajectories: Estimated w.r.t GT

TS (unreliable)

$EM_{4000}$

TS + $EM_{4000}$
## Quantitative Results

| Sequence                        | TS  | EM<sub>2000</sub> | EM<sub>3000</sub> | EM<sub>4000</sub> | EM<sub>5000</sub> | TSEM<sub>4000</sub> (λ<sub>th</sub> = 31) |
|---------------------------------|-----|-------------------|-------------------|-------------------|-------------------|------------------------------------------|
| simu_office_planar              | 4.7 | 4.0               | **3.9**           | **3.7**           | 4.1               | **4.9**                                  |
| simu_poster_planar              | 4.7 | **3.7**           | 4.3               | 4.6               | 5.0               | **4.6**                                  |
| simu_checkerboard_planar        | 4.2 | 2.9               | 2.2               | 2.3               | 2.4               | **4.7**                                  |
| simu_office_6DoF                | **9.1** | 25.3             | 21.0              | 16.6              | 15.8              | **18.7**                                 |
| simu_poster_6DoF                | 18.2 | **15.4**          | **16.3**          | 16.8              | 17.4              | **17.3**                                 |
| simu_checkerboard_6DoF          | 23.0 | 17.0              | **14.0**          | 15.1              | **13.4**          | 28.1                                     |
| rpg_bin_6DoF                    | **3.4** | 22.4             | 16.6              | 8.0               | 14.1              | **3.8**                                  |
| rpg_box_6DoF                    | **6.5** | **5.3**          | 17.1              | 13.7              | 9.8               | **7.1**                                  |
| rpg_desk_6DoF                   | 3.4  | **2.9**           | 3.3               | 3.2               | **2.9**           | 3.8                                      |
| rpg_monitor_6DoF                | 7.2  | **5.3**           | **5.2**           | 7.4               | 7.3               | **7.0**                                  |
| upenn_indoor_flying1_6DoF       | 18.5 | 22.0              | 16.7              | **16.0**          | 22.1              | **14.8**                                 |
| upenn_indoor_flying3_6DoF       | 20.9 | **10.8**          | 11.9              | 14.0              | 15.0              | **10.9**                                 |
Conclusion

- Extensive comparisons of two representations: event map and time surface map
- Enhanced tracker to make use of their complementary strengths
- Six tracker variations
- Indicate possible ways to improve the state-of-the-art (SOTA) methods.
Thank you!

Email: jijiao@connect.ust.hk
Page: https://gogojjh.github.io
Code: https://github.com/gogojjh/ESVO_extension