Weekly Report (AUG 28, 2023 - Sep 3, 2023)

Zixun Xiong
Manning College of Information and Computer Sciences, UMASS Amherst
140 Governors Dr, Amherst, MA 01002
zixunxiong@umass.edu

Abstract

This week, I read some papers on

1 EVDodgeNet: Deep Dynamic Obstacle Dodging with Event Cameras

This paper [1] introduces a deep learning-based solution to dynamic object avoidance. By using a series of shallow neural networks, researchers estimate both the ego-motion and the motion of independently moving objects (IMO). This method only needs the algorithm to be trained in simulation and can transfer to the real world without any fine-tuning or retraining. For the evaluation and testing, researchers test this approach in scenes with obstacles that have different shapes and sizes. The proposed method achieves an overall 70% success rate in scenes including objects that are not in the training set and scenes with low light.

1.1 Background and Significance

Independent Motion Detection and Ego-Motion Estimation. In visual Inertial Odometry (IMO), information from Inertial Measurement Units (IMU) is utilized to accomplish tasks like Simultaneous localization and mapping (SLAM). Works have been proposed by introducing event cameras to present a low-latency VIO algorithm to estimate ego-motion [2]. Most works before this paper focused on static scenes, which are rarely met in the real world. [3], however, mentioned that by carefully modeling, one can both estimate ego-motion and IMOs.

Image Stabilization. Recently, image stabilization is the most robust algorithm to make IMO more evident. Works inspired by this idea have been done on event cameras [4].

Obstacle avoidance on aerial robots. Works mentioned above are used to aid obstacle avoidance on aerial robots. Event camera has also been used for dodging high-speed obstacles in [5].

1.2 Challenges and Advantages

Although dynamic object detection using traditional cameras has been studied extensively, they are either of high latency or computationally expensive or don't fit for generalized missions including novel objects. Following [6], this paper applied deep learning to generalize the object detections for novel objects after being trained only on simulation.

EVDeBlurNet The input for event camera-aided VIO systems are event frames, which are generated by projecting event data to 2D arrays. This projection will cause misalignment [7]. Thereby, motion-compensations are needed. In this part, researchers use a shallow network named **EVDeBlurNet** to complete this task.

First, we get the event frames in equation 1 from a set of event data (shaped like $e = (\mathbf{x}, t, p)$), where $\mathbf{x}$ is the location of the triggered point in raw latent image I of the event camera.

$$
\begin{aligned}
E(x, \delta t)_+ &= \sum_{t=t_0}^{t_0+\delta t} \mathbb{I}(x, t, p = +1) \\
E(x, \delta t)_- &= \sum_{t=t_0}^{t_0+\delta t} \mathbb{I}(x, t, p = -1) \\
E(x, \delta t)_\tau &= (\sum_{t=t_0}^{t_0+\delta t} \mathbb{I}(x, t, p = \pm 1))^{-1} \mathbb{E}(t - t_0)
\end{aligned}
\tag{1}
$$

These generated event frames of given frequency δt are inputs of the model **EVDeBlurNet**, which is designed as an encode decoder structure. Thus the object model can remove stray events (which are generally noise) while preserving events corresponding to contours. Thereby, the loss function is defined as the equation 2.

$$\underset{E}{\operatorname{argmin}} - C(\overline{E}) + \lambda D(E, \overline{E}) \tag{2}$$

$$C(E) \triangleq \mathbb{E}(||\mathbf{Var}(\nabla E)||) \tag{3}$$

$$D(E_1, E_2) \triangleq \mathbb{E}(||E_1 - E_2||_1) \tag{4}$$

Thus **EVDeBlurNet** can learn to denoise event frames, while the debulered images are not too far from the origin image by the second term in equation 2.

1.3 Reasults and Experiments

1.4 Disadvantages of the Paper

References

[1] Sanket, Nitin J., et al. ”Evdodgenet: Deep dynamic obstacle dodging with event cameras.” 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020.

[2] Vidal, Antoni Rosinol, et al. ”Ultimate SLAM? Combining events, images, and IMU for robust visual SLAM in HDR and high-speed scenarios.” IEEE Robotics and Automation Letters 3.2 (2018): 994-1001.

[3] Rebecq, Henri, Daniel Gehrig, and Davide Scaramuzza. ”ESIM: an open event camera simulator.” Conference on robot learning. PMLR, 2018.

[4] Stoffregen, Timo, et al. ”Event-based motion segmentation by motion compensation.” Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

[5] Mueggler, Elias, et al. ”Towards evasive maneuvers with quadrotors using dynamic vision sensors.” 2015 European Conference on Mobile Robots (ECMR). IEEE, 2015.

[6] Falanga, Davide, Suseong Kim, and Davide Scaramuzza. ”How fast is too fast? The role of perception latency in high-speed sense and avoid.” IEEE Robotics and Automation Letters 4.2 (2019): 1884-1891.

[7] Gallego, Guillermo, Henri Rebecq, and Davide Scaramuzza. ”A unifying contrast maximization framework for event cameras, with applications to motion, depth, and optical flow estimation.” Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.