Using estimation of camera movement to achieve multi-target tracking

August 20 2020

Credit: CC0 Public Domain

Estimating the motion of a moving camera is a ubiquitous problem in the field of computer vision. With technology such as self-driving cars and autonomous drones becoming more popular, fast and efficient

algorithms enabling on-board video processing are needed to return timely and accurate information at a low computational cost. This estimation of camera movement, or 'pose estimation,' is also a crucial component of target tracking aboard moving vehicles or platforms.

Researchers at Brigham Young University have published their results in *IEEE/CAA Journal of Automatica Sinica*, a joint publication by the Institute of Electrical and Electronics Engineers and the Chinese Academy of Sciences. They have found a way to greatly reduce the computation time and complexity of pose estimation by cleverly 'seeding' an algorithm already in use in the computer vision industry.

Pose estimation algorithms utilize frames of a video feed from a moving camera to generate hypotheses for how the camera has moved over the course of each consecutive frame. Until now, algorithms used for pose estimation required the generation of up to five to ten hypotheses for how the camera was moving given the data in the video feed. These hypotheses were then scored by how well they fit the data, with the highest scoring hypothesis being chosen as the best pose estimate. Unfortunately, the generation of multiple hypotheses is computationally expensive and results in slower return time for robust pose estimation.

Researchers have found a way to seed, or give hints to, an already used algorithm in computer vision by feeding it information in between each frame, thus greatly reducing the need for generating many hypotheses. "At each iteration, we use the current best hypothesis to seed the algorithm." The reduction of required hypotheses directly results in reduced computation time and complexity; "we show that this approach significantly reduces the number of hypotheses that must be generated and scored to estimate the pose, thus allowing real-time execution of the algorithm."

The team then compared their seeding method to other state of the art

pose estimation algorithms to classify how reducing the number of hypotheses affected the accuracy of the computation. "After 100 iterations, error for the seeding methods that use prior information is comparable to the OpenCV five-point polynomial solver, despite the fact that only one hypothesis is generated per iteration instead of an average of about four hypotheses." Furthermore, when the two algorithms were examined in time, the teams' algorithm significantly outperformed other state of the art methods. In most cases, the new algorithm was ten times faster.

The group then modified their algorithm to enable target tracking and tested it on a multi-rotor UAV. The algorithm successfully tracked multiple targets at 640 x 480 resolution. The results were consistent with their earlier analysis. "The complete algorithm takes 29 milliseconds to run per frame, which means it is capable of running in real-time at 34 frames per second (FPS)." As for what's next, the team plans on applications to 3-D scene reconstruction and more complex tracking methods.

More information: Jacob H. White et al, An iterative pose estimation algorithm based on epipolar geometry with application to multi-target tracking, *IEEE/CAA Journal of Automatica Sinica* (2020). DOI: 10.1109/JAS.2020.1003222

Provided by Chinese Association of Automation