Chapter:1

The field of visual tracking is one of the most prominent ones in computer vision. The purpose of visual tracking is to locate the target in the first frame and maintain the track in subsequent frames. In general, tracking a target involves three steps, namely, feature extraction, target localization and target optimization. Feature extraction first involves describing the target in the scene. Handcrafted features such as color, texture, histogram of gradient can be extracted from target. Apart from vision information, thermal profile, infrared, and audio can also be extracted for target representation. In addition, deep features can be extracted either from single layer or multiple layer of a deep learning based architecture. Next step in visual tracking is target localization. Localization of the target in a video stream is tedious due to the dynamic environmental variations that originate in real-time. Numerous algorithms have been proposed under traditional approach due to the computational benefits and simplicity in implementation. Now a day, there is a paradigm shift in tracking algorithm to deep learning based algorithms. Deep learning based algorithm provides more precise estimation of target in terms of accuracy even in tough tracking scenarios. In the end, the target optimization step involves for optimal localization of the target in presence of challenges to prevent the tracker's drift.

In the recent year, with the advancement in technology visual tracking extends to track multi-target present in the scene rather estimating location for single target only. In contract to single object tracking, multi-target introduces one extra step of detection. Tracking multi-target includes detecting and categorizing the target into multiple classes in the first frame and provides each individual target an ID to keep its track in the subsequent frames of a video stream. One category of multi-target algorithms exploits global information in order to track the target of the detected target. On the other hand, some algorithms consider present and past information of the target to provide efficient tracking solutions. Apart from these, deep leaning based algorithms provide reliable and accurate solutions. But, these algorithms are computationally slow when applied in real-time.

In recent years visual object tracking has become a very active research area. An increasing number of tracking algorithms are being proposed each year. It is because tracking has wide applications in various real world problems such as human-computer interaction, autonomous vehicles, robotics, surveillance and security just to name a few. In the current study, we review latest trends and advances in the tracking area and evaluate the robustness of different trackers based on the feature extraction methods. The first part of this work comprises a comprehensive survey of the recently proposed trackers. We broadly categorize trackers into Correlation Filter based Trackers (CFTs) and Non-CFTs. Each category is further classified into various types based on the architecture and the tracking mechanism. In the second part, we experimentally evaluated 24 recent trackers for robustness, and compared handcrafted and deep feature based trackers. We observe that trackers using deep features performed better, though in some cases a fusion of both increased performance significantly. In order to overcome the drawbacks of the existing benchmarks, a new benchmark Object Tracking and Temple Color (OTTC) has also been proposed and used in the evaluation of different algorithms. We analyze the performance of trackers over eleven different challenges in OTTC, and three other benchmarks. Our study concludes that Discriminative Correlation Filter (DCF) based trackers perform better than the others. Our study also reveals that inclusion of different types of regularizations over DCF often results in boosted tracking performance. Finally, we sum up our study by pointing out some insights and indicating future trends in visual object tracking field.

Computer vision has received a significant attention in recent year, which is one of the important parts for robots to obtain information about the external environment. Visual trackers can provide the necessary physical and environmental parameters for the mobile robot, and their performance is related to the actual application of the robot. This study provides a comprehensive survey on visual trackers. Following a brief introduction, we first analyzed the basic framework and difficulties of visual trackers. Then the structure of generative and discriminative methods is introduced, and summarized the feature descriptors, modeling methods, and learning methods which be used in tracker. Later we reviewed and evaluated the state-of-the-art progress on discriminative trackers from three directions: correlation filter, deep learning and convolutional features. Finally, we analyzed the research direction of visual tracker used in mobile robot, as well as outlined the future trends for visual tracker on mobile robot