Object Detection and Tracking in Real Time Video

Rajkumar Kaushik*, Subhash Chander Swami, Dev Sharma, Peeyush Ahluwalia

Abstract:

Object tracking is the process of locating moving objects over time using the camera in video sequences. The objective of object tracking is to associate target objects in consecutive video frames. Object tracking requires the location and shape or features of objects in the video frames. So, object detection and object classification are the preceding steps of object tracking in computer vision applications. To detect or locate the moving object in the frame, Object detection is the first stage in tracking. It is a challenging or difficult task in image processing to track the objects into consecutive frames. Various challenges can arise due to complex object motion, irregular shape of object, occlusion of object to object and object to the scene, and real-time processing requirements. This paper presents the various techniques of object tracking in video sequences.

Keywords: Object Detection, Object tracking, Real-Time Video

Introduction:

Object detection and tracking are fundamental and often challenging tasks within the realm of computer vision, finding wide-ranging applications across various industries and domains. Object detection specifically involves the identification and localization of objects within images or video streams. Typically, this process entails the continuous capture of images through a webcam, with a computer performing intricate image processing to display the object's trajectory on a monitor. Conversely, object tracking is the process of monitoring the movement of objects over time using a camera in video sequences. The central objective of object tracking is to establish associations between target objects across consecutive video frames, necessitating the accurate determination of their positions, shapes, and distinctive features. In the context of computer vision, object detection and object classification serve as vital initial steps in the workflow of object tracking. Practically every tracking method depends on an object detection mechanism, which may operate in every frame or be triggered when an object initially appears in real-time video. The rapid advancements in technology, characterized by high-performance computers and the widespread availability of affordable, high-quality video cameras, have generated considerable interest in the development of object tracking algorithms. Video analysis, as a broader field, encompasses three pivotal steps: the detection of interesting moving objects, the tracking of these objects across successive frames, and the analysis of object trajectories to gain insights into their behavior. Object tracking finds particular relevance in the context of motion-based recognition, facilitating various critical tasks such as automated object detection, tracking, and counting. These tasks are indispensable across numerous applications, offering solutions to a variety of real-world challenges. The underlying concept in this process revolves around the identification of key regions, referred to as key points, that exhibit robustness to various transformations. For each of these detected key points, an invariant feature vector, known as a descriptor, is generated to represent the image data surrounding that specific point. These image features can be categorized into two primary types: global features and local features. Global features, encompassing aspects like color and texture, are geared towards describing the image as a whole, capturing overarching properties that involve all pixels. In contrast, local features focus on the identification of key points or regions of interest within the image and provide more specific descriptions. Of particular significance is the role of local features, which, when employed in an image, generate numerous vectors describing each key point's characteristics, including its shape, color, orientation, texture, and more. These local features are especially useful in object classification, the recognition of similar objects, or identifying instances of the same object or scene. In summary, object detection and tracking are pivotal components of computer vision, with object detection forming the foundation for subsequent tracking endeavors. These processes have gained increasing significance in recent times due to technological advancements, making automated video analysis and the solutions they offer for motion-based recognition, object detection, tracking, and counting vital in various practical scenarios.

Corresponding Author: Rajkumar Kaushik
Assistant Professor, Electrical Engineering, Arya Institute of Engineering And Technology, Jaipur, Raj.
Assistant Professor,Electrical Engineering, Arya Institute of Engineering, Technology and Management,
Science Student, M.M.J.S.V Sr. Sec. School, Mahendragarh, Haryana
Science Student, Stepping Stone School, Civil Lines, Jaipur, Raj.

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

The approach includes the identification of key regions, the construction of invariant descriptors, and the differentiation between global and local image features to achieve a comprehensive understanding of object behavior in video sequences.

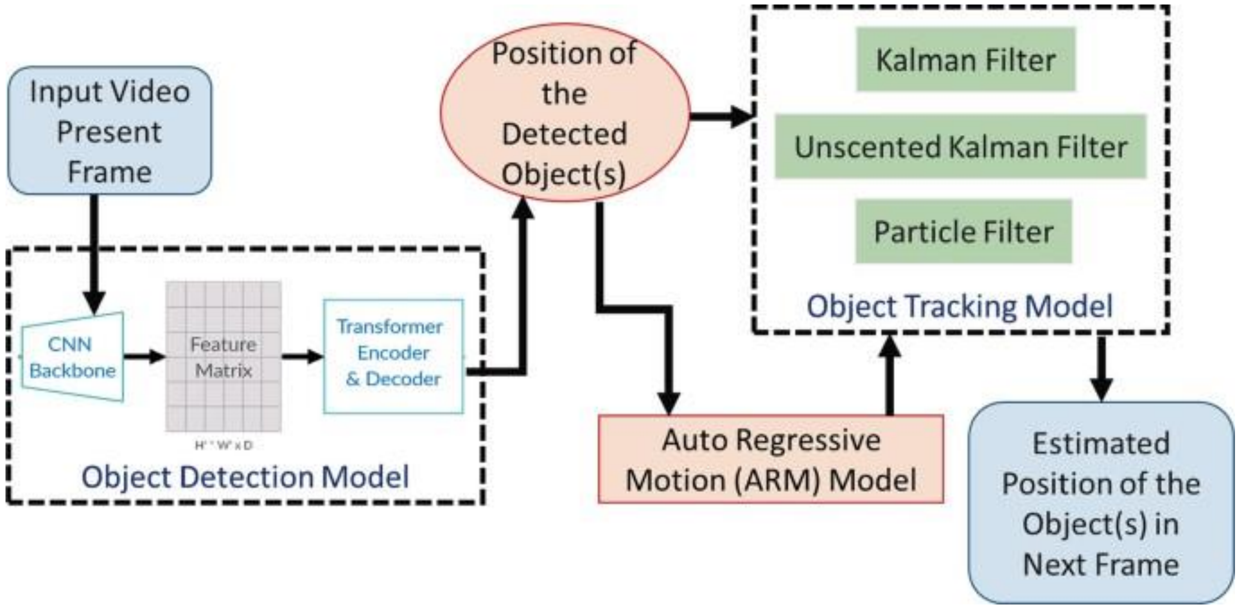

Figure 1

Previous work:

Jean-Philippe Jodoin proposed an urban tracker to track various a priori unknown road users. The method is based on tracking the resulting foreground blobs of pixels. Each blob is modelled by a collection of key points. Data association is performed from frame to frame, and a finite state machine (FSM) corrects the associations by handling blob merging, splitting, and fragmenting.

Harpreet Kaur proposed a Fractional order gain Kalman filter (FOGKF) to avoid divergence of the extended Kalman filter. It uses the fractional derivative-based method to improve performance by modifying the Kalman gain so that the sensitivity for large perturbations can be increased. The fractional order gain Kalman filter performs better even in the presence of abrupt variations in the inputs. The gain value will never be too large due to the feedback loop. **Junlin Hu** proposed a deep metric learning (DML) approach for robust visual tracking under the particle filter framework. The DML tracker can explicitly learn several hierarchical nonlinear transformations to map data points into another subspace via feed-forward neural network architecture so that these nonlinear transformations are explicitly solved by maximizing the interclass variations of negative pairs and minimizing the intra-class variations of positive pairs simultaneously. It overcomes both the nonlinearity and scalability problems of conventional metric learning methods and kernel-based method.

Nan Jiang proposes a nonparametric data-driven local metric adjustment method. It finds a spatially adaptive metric that exhibits different properties at different locations in the feature space, due to the differences of the data distribution in a local neighborhood. It minimizes the deviation of the empirical misclassification probability to obtain the optimal metric such that the asymptotic error as if using an infinite set of training samples can be approximated. **Bohan Zhuang** proposed a reversed multitask sparse tracking framework that projects the templates matrix (both positive and negative templates) into the candidate's space. By selecting and weighing the discriminative sparse coefficients, the DSS map and pooling method lead to the best candidate. With this DSS map, candidates are evaluated in both directions: not only how similar it is to the target object but also how different it is from the background.

AI-Powered Video Object Detection and Tracking

The title "AI-Powered Video Object Detection and Tracking" encapsulates the core features and capabilities of the system it represents. In this context, "AI-powered" signifies the utilization of artificial intelligence techniques, particularly machine learning algorithms, to enable automated and intelligent object detection and tracking within video streams. The term "video object detection" refers to the system's ability to identify and locate specific objects or entities within the video content, such as people, vehicles, or animals. This process involves analyzing frames in the video to recognize objects of interest.

"Object Tracking" signifies the system's capacity to follow and monitor these objects as they move through the video frames, maintaining their identities over time. This is particularly valuable in applications like video surveillance, where it enables continuous monitoring and analysis of object behaviour. In summary, the title suggests a system that leverages artificial intelligence for the automated identification, localization, and tracking of objects within video data, offering valuable insights and applications in fields ranging from security to autonomous systems.

Result and conclusion:

How Good It Is: We'd summarize how effective this method is. Think of it as giving it a grade, like saying it's an "A" because it's good at finding and following objects.

Where It Can Be Used: We'd talk about where this method can be handy. For instance, it could be useful in security, self-driving cars, or robots that need to watch things in real time.

What It Struggles With: We'd also be honest about its limitations. Like, there are times when it might not work well or when it needs a powerful computer.

What's Next: We'd suggest what researchers could do in the future. Maybe they can try different approaches, make it work even faster, or tackle the limitations we mentioned.

Why It Matters: Finally, we'd say why this research is important. It's like explaining why it's a big deal for the world of computers, technology, and the specific areas where this can be used.

Future scope:

Improved Accuracy: Future developments will likely focus on improving the accuracy of object detection and tracking algorithms. This may involve the use of more advanced neural network architectures, larger and more diverse training datasets, and better optimization techniques. Real-time Performance: As hardware continues to advance, real-time object detection and tracking will become more achievable for a wider range of applications. This includes not only traditional computers but also edge devices like drones, autonomous vehicles, and IoT devices. Multi-Object Tracking: The ability to track multiple objects in a video stream simultaneously is a challenging problem. Future research will focus on developing more robust and efficient algorithms for multi-object tracking, particularly in crowded or complex environments. Drones and Robotics: Object detection and tracking will continue to be vital for drones and robotics. These technologies will enable drones to navigate complex environments and robots to interact with objects and humans more effectively.

References

1. Lijing Zhang, Yingli Liang," Motion human detection based on background subtraction," Second International Workshop on Education Technology and Computer Science. 2010 IEEE.
2. Tao Jianguo, Yu Changhong, "Real-Time Detection and Tracking of Moving Object," Intelligent Information Technology Application 2008 UTA '08. Second International Symposium on Volume 2, 20-22 Dec2008 Page(s):860 - 863
3. Carlos R. del-Blanco, Fernando Jaureguizar, and Narciso García, " An Efficient Multiple Object Detection and Tracking Framework for Automatic Counting and Video Surveillance Applications," IEEE Transactions on Consumer Electronics, Vol. 58, No. 3, August 2012.
4. K.Kinoshita, M.Enokidani, M. Izumida and K.Murakami, "Tracking of a Moving Object Using One-Dimensional Optical Flow with a Rotating Observer," Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on 5-8 Dec. 2006 Page(s): 1 - 6
5. Niu Lianqiang and Nan Jiang, "A moving objects detection algorithm based on improved background subtraction," Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on Volume 3, 26-28 Nov. 2008 Page(s):604 – 607
6. M. Mignotte and IKonrad, "Statistical Background Subtraction Using Spatial Cues," Circuits and Systems for Video Technology, IEEE Transactions on Volume 17 Issue 12, Dec. 2007Page(s):1758-1763.
7. Zhen Tang and Zhenjiang Miao, "Fast Background Subtraction and Shadow Elimination Using improved Gaussian Mixture Model," Haptic, Audio and Visual Environments and Garnes, 2007. IEEE International Workshop on 12-14 Oct. 2007 Page(s):38 – 41
8. Wang Weiqiang, Yang Jie and Gao Wen, "Modeling Background and Segmenting Moving Objects from Compressed Video, " Circuits and Systems for Video Technology, IEEE Transactions on Volume 18, Issue 5, May 2008 Page(s):670 – 681
9. M.Dimitrijevic, "Human body pose detection using Bayesian spatio temporal templates," 2007 International Conference on Intelligent and Advanced Systems, 2008, pp.764-9.
10. R. K. Kaushik Anjali and D. Sharma, "Analyzing the Effect of Partial Shading on Performance of Grid Connected Solar PV System", *2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*, pp. 1-

11. R. Kaushik, O. P. Mahela, P. K. Bhatt, B. Khan, S. Padmanaban and F. Blaabjerg, "A Hybrid Algorithm for Recognition of Power Quality Disturbances," in *IEEE Access*, vol. 8, pp. 229184-229200, 2020.
12. Kaushik, R. K. "Pragati. Analysis and Case Study of Power Transmission and Distribution." *J Adv Res Power Electro Power Sys* 7.2 (2020): 1-3.
13. Kaushik, M. et al. (2015) "Availability analysis for embedded system with N-version programming using fuzzy approach," International Journal of Software Engineering Technology and Applications, 1(1), p. 90. doi: 10.1504/ijseta.2015.067533.
14. Sharma, R., Kaushik, M. and Kumar, G. (2015) "Reliability analysis of an embedded system with multiple vacations and standby", International Journal of Reliability and Applications, Vol. 16, No. 1, pp. 35-53.
15. Kaushik, M. and Kumar, G. (2015) "Markovian Reliability Analysis for Software using Error Generation and Imperfect Debugging" , International Multi Conference of Engineers and Computer Scientists 2015, vol. 1, pp. 507-510.