Volume 4, Issue 1, January 2017 ISSN: 2348-8565 (Online)
International Journal of Modern Engineering and Research Technology
Website: http://www.ijmert.org Email: editor.ijmert@gmail.com

# A Review on Object Detection and Tracking in Video Image

**Ashish Choure**
*Research Scholar M.Tech. (CSE)*
*UIT BU*
*Bhopal (M.P.), [INDIA]*
*Email:ashishchoure007@rediffmail.com*

**Amit Kumar Jha**
*Assistant Professor*
*UIT BU*
*Bhopal (M.P.), [INDIA]*
*Email:akj1389@gmail.com*

## ABSTRACT

*In recent days, capturing images with high quality and good size is so easy because of rapid improvement in quality of capturing device with less costly but superior technology. Videos are a collection of sequential images with a constant time interval. So video can provide more information about our object when scenarios are changing with respect to time. Therefore manually handling videos are quite impossible. So we need an automated devise to process these videos. In this thesis one such attempt has been made to track objects in videos. Many algorithms and technology have been developed to automate monitoring the object in a video file.*

*Simple object detection compares a static background frame at the pixel level with the current frame of video. The existing method in this domain first tries to detect the interest object in video frames. One of the main difficulties in object tracking among many others is to choose suitable features and models for recognizing and tracking the interested object from a video. Some common choice to choose suitable feature to categories, visual objects are intensity, shape, color and feature points. In this thesis, we studied about mean shift tracking based on the color pdf, optical flow tracking based on the intensity and motion; SIFT tracking based on scale invariant local feature points. Preliminary results from experiments have shown that the adopted method is able to track targets with translation, rotation, partial occlusion and deformation.*

***Keywords***:—*Object detection, Frame difference, Vision and scene understanding, Background subtraction, Scale invariant feature transform (SIFT)*

## I. INTRODUCTION

Object detection and tracking is an important challenging task within the area in Computer Vision that try to detect, recognize and track objects over a sequence of images called video. It helps to understand, describe object behavior instead of monitoring computer by human operators. It aims to locating moving objects in a video file or surveillance camera. Object tracking is the process of locating an object or multiple objects using a single camera, multiple cameras or given video file. Invention of high quality of the imaging sensor, quality of the image and resolution of the image are improved, and the exponential increment in computation power is required to be created of new good algorithm and its application using object tracking.

In Object Detection and Tracking we have to detect the target object and track that object in consecutive frames of a video file.

***Challenges in object tracking in image sequence***

1. Non-uniformity in intensity variation across target.
2. Intensity variations across subsequent frames.
3. Changes due to moving camera.
4. Target size, orientation etc change.
5. Occlusion partial or full.

It is therefore essential to adopt principled over the last few years, particle filters, also known as condensation or sequential have been used for tracking.

***A) Region-based methods****:* These strategies give a productive approach to decipher and investigate movement in a frame sequence of video. A region in frame might be characterized as a set of pixels having homogeneous attributes. It could be determined by image segmentation, which might be focused around different object characteristics like color, edges and so forth. Basically, a region would be the image range secured by the projection of the object of investment onto the frame plane. After segmenting the region a bounding box is created that identifies the target being track.

***B) Contour-based methods****:* in this object is represented utilizing contour shape information is used the tracking time to time, hence recovering both its position and shape. Then again, contour based tracking are typically more robust than region based object tracking algorithm, on the grounds that it could be adjusted to adapt to halfway impediments. Additionally the outline contour information is insensitive intensity variation.

***D) Template-based methods****:* Template-matching procedures are utilized by numerous researchers to perform object tracking. Template based tracking is nearly identified with region based tracking on the idea that a template is basically a model of the target area to be tracked. These routines include two steps for tracking. In the first step template might be instated by different off-line and online strategies. Throughout matching, it includes the procedure of seeking the interested object to focus the image region that looks like the template, taking into account a likeness or separation measure.

***Object poses in the video frame:*** In a video file, since the object is moving so the appearance of an interested object may vary its projection on a video frame plane.

***Ambient illumination:*** In a video, it is possible to change in intensity, direction and color of ambient light in appearance of interested objects in a video frame plane.

***Noise:*** In the acquisitions process of video, it may possible to introduce a certain amount of noise in the image or video signal. The amount of noise depends upon sensor qualities which are used in acquitting the video.

***Occlusions:*** In a video file, moving object may fall behind some other objects which are present in the current scene. In that case tracker may not observe the interested object.

## II. LITERATURE SURVEY

The research conducted so far for object detection and tracking objects in video surveillance system. Tracking is the process to locating the interested object within a sequence of frames, from its first appearance to its last. The type of object and its description within the system depends on the application. During the time that it is present in the scene it may be occluded by other objects of interest or fixed obstacles within the scene. A tracking system should be able to predict the position of any occluded objects.

In [1], the author suggests an algorithm to isolate the moving objects in video sequences and then presented a rule-based tracking algorithm. The preliminary

experimental results demonstrate the effectiveness of the algorithm even in some complicated situations, such as new track, ceased track, track collision, etc. A tracking method without background extraction is discussed in [2]. Because while extracting background from video frame if there are small moving things in that frame they form a blob in thresholding which create confusion in case of tracking that blob as they aren't of any use that can be reduced here. The author introduces a video tracking in computer vision, including design requirements and a review of techniques from simple window tracking to tracking complex, deformable objects by learning models of shape and dynamics in [3].

In [22], there are various studies identified with the MS-based CMS (or Camshift) trackers. In the investigation of Stern and Efros, they created a strategy that adaptively switches shade space demonstrates all around the transforming of a feature. Additionally, they proposed another execution measure for assessing following calculation. Their proposed technique is utilized to discover the ideal color space and shade appropriation models fusion in the configuration of versatile shade following frameworks. Their color exchanging strategy was performed Inside the skeleton of the CAMShift tracking algorithm, they consolidated various methodology to develop an improved face tracking methodology. At every cycle of the CAMShift algorithm, given image is changed over into a likelihood image utilizing the model of shade dispersion of the skin color being tracked.

In the study of Li et al. [23], they proposed a novel methodology for global target following focused around MS strategy. The proposed technique speaks to the model and the applicant regarding background and shade weighted histogram, separately, which can get exact object estimate adaptively with low computational unpredictability. Likewise, they actualized the MS technique by means of a coarse-to-_ne path for global greatest looking for. This system was termed as versatile pyramid MS, in light of the fact that it utilizes the pyramid examination procedure and can focus the pyramid level adaptively to diminishing the amount of iteration needed to attain merging. The trial consequences of the study of Li et al. [23] indicate that the proposed system can effectively adapt to distinctive circumstances, for example, camera movement, camera vibration, camera zoom and center, high velocity moving item following, halfway impediments, target scale varieties, and so forth.

In the study of Yuan et al. [24] Proposed another moving object tracking algorithm, which joins together enhanced nearby binary pattern texture surface and tone information to portray moving object and embraces the thought of CAMShift algorithm. With a specific end goal to diminish matching unpredictability on the reason of fulfilling the correctness, numerous sorts of neighborhood twofold example and tint are chopped down. As per Yuan et al. [24], the experiment demonstrate that the proposed algorithm can track adequately moving interested object, can fulfill continuous and has preferred execution over others.

In the study of Mazinan and Amir-Latifi [25], an enhanced curved part capacity was proposed to defeat the fractional impediment. Hence, to enhance the MS calculation against the low immersion and additionally sudden light, changes are created out of movement data of the fancied succession. By utilizing both the color feature and the motion information all the while, the competence of the MS calculation was correspondingly expanded. In their study [25], by accepting a steady speed for the article, a hearty estimator, i.e., the Kalman channel was acknowledged to tackle the full impediment issue. As indicated by Mazinan and Amir-Latifi [25], the trial

results checked that the proposed technique has an ideal execution progressively protest following, while the aftereffect of the first MS calculation may be unsatisfied.

In [26], the of strategy is oftentimes utilized as a part of picture movement investigation. The reckoning of from a picture succession gives exceptionally imperative data to movement investigation. This issue includes moving object detection and tracking, moving object segmentation, and movement distinguishment. In the study of Lai, another movement estimation calculation was displayed that it gives faultless of processing under non uniform brilliance varieties.

Lipton et al. [27] proposed edge contrast that utilization of the pixel-wise contrasts between two casing pictures to concentrate the moving locales. In an alternate work, Stauffer & Grimson et al. [28] proposed a Gaussian mixture model focused around foundation model to distinguish the item. Liu et al. [29], proposed foundation subtraction to recognize moving districts in a picture by taking the contrast between present and reference foundation picture in a pixel-by-pixel.

Collins et al. [30], created a half breed system that joins three-edge differencing with a versatile foundation subtraction model for their VSAM (Video Surveillance\ and Monitoring undertaking. Desa & Salih et al [31], proposed a mixture of foundation subtraction and casing contrast that enhanced the past consequences of foundation subtraction and edge distinction. Cheng & Chen, 2006 proposed a color and a spatial characteristic of the item to recognize the track object. The spatial characteristic is concentrated from the jumping box of the article. In the interim, the color characteristics concentrated is mean and standard worth of each one article. Czyz et al., 2007 proposed the color conveyance of the object as perception model. The comparability of the object estimated by Bhattacharya distance. The low Bhattacharya distance relates to the high similitude.

## III. Conclusion

Object detection and tracking is an important task in computer vision field. In object detection and tracking it consist of two major processes, object detection and object tracking. Object detection in video image obtained from single camera with static background that means fixing camera is achieved by background subtraction approach. In this thesis, we tried different videos with fixed camera with a single object and multiple objects to see it is able to detect objects. Motion based systems for detecting and tracking given moving object of interest can be created. Using SIFT feature extraction first feature of the object and the frame has detected to match the interested object. Since for feature extraction, SIFT algorithm has been used so tracker is invariant to representation of interested object.

## IV. Future Work

In the future, we can extend the work to detect the moving object with non-static background, having multiple cameras which can be used in real time surveillance applications.

## REFERENCES:

[1] Yiwei Wang, John F. Doherty and Robert E. Van Dyck, "Moving Object Tracking in Video", in proceedings of 29th applied imagery pattern recognition workshop, ISBN 0-7695-0978-9, page 95,2000.

[2] Bhavana C. Bendale, Prof. Anil R. Karwankar, "Moving Object Tracking in Video Using MATLAB", International Journal of Electronics, Communication and Soft Computing Science and Engineering ISSN: 2277-

9477, Volume 2, Issue 1.

[3] Marcus A. Brubaker, Leonid Sigal and David J. Fleet, "Video-Based People Tracking", hand book of ambient intelligence under smart environments 2010, pp 57-87.

[4] Emilio Maggio and Andrea Cavallaro, "Video Tracking: Theory and Practice", first edition 2011, John Wiley and Sons, Ltd.

[5] Y.Alper, J.Omar, and S.Mubarak. "Object Tracking: A Survey" ACM Computing Surveys, vol. 38, no. 4, Article 13, December 2006.

[6] B. Triggs, P.F. Mc Lauchlan, R.I. Hartley and A.W. "Fitzgibbon. Bundle adjustment – a modern synthesis". In Proceedings of the International Conference on Computer Vision, London, UK, 1999, 298?372.

[7] G.C. Holst and T.S. Lomheim. "CMOS/CCD Sensors and Camera Systems". Bellingham, WA, SPIE Society of Photo-Optical Instrumentation Engineering, 2007.

[8] E. Maggio, M. Taj and A. Cavallaro. "E_cient multi-target visual tracking using random finite sets". IEEE Transactions on Circuits Systems and Video Technology, 18(8), 1016? 1027, 2008.

[9] G. David Lowe. Object recognition from local scale-invariant features. Proceedings of the International Conference on Computer Vision. 2. pp. 1150?1157,1997.

[10] G. David Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), pp, 91-110, 2004.

[11] Bruce D. Lucas and Takeo Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. International Joint Conference on Artificial Intelligence, pages 674-679, 1981.

[12] Carlo Tomasi and Takeo Kanade. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.

[13] Jianbo Shi and Carlo Tomasi. Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition, pages 593-600, 1994.

[14] Stan Birchfield. Derivation of Kanade -Lucas-Tomasi Tracking Equation. Unpublished, January 1997.

[15] Y. Cui, S. Samarasekera, Q. Huang. Indoor Monitoring Via the Collaboration Betweena Peripheral Senson and a Foveal Sensor, IEEE Work-shop on Visual Surveillance, Bomba y, India, 2-9, 1998.

[16] G. R. Bradski, Computer Vision Face T racking as a Component of a Perceptual User Interface, IEEE Work. on Applic. Comp. Vis., Princeton, 214-219, 1998.

[17] S.S. Intille, J.W. Davis, A.F. Bobick, Real-Time Closed-World T racking. IEEE Conf. on Comp. Vis. and Pat. Rec., Puerto Rico, 697-703, 1997.

[18] C. Wren, A. Azarbayejani, T. Darrell, A. Pentland, P_nder: Real-Time Tracking of the Human Body, IEEE Trans. Pattern Analysis Machine Intell, 19:780-785, 1997.

[19] A. Eleftheriadis, A. Jacquin. Automatic Face Location Detection and Tracking for Model-Assisted Coding of Video Teleconference Sequences at Low Bit Rates, Signal Processing- Image Communication, 7 (3): 231-248, 1995.

[20] D. Fuiorea, V. Gui, D. Pescaru, and C. Toma. Comparative study on RANSAC and Mean shift algorithm, International Symposium on Electronics and Telecommunications Edition 8. vol. 53(67) Sept. 2008, pp. 80-85.

[21] Y.Cheng. Mean Shift, Mode Seeking, and Clustering, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 17, No 8, 790-799,1995.

[22] Stern H, Efros B (2005) Adaptive color space switching for tracking under varying illumination. Image Vis Comput 23(3):353364. doi : 10.1016 /j. imavis 2004.09.005

[23] Li S-X, Chang H-X, Zhu C-F (2010) Adaptive pyramid mean shift for global real-time visual tracking. Image Vis Comput 28(3):424437. doi : 10.1016 / j. imavis 2009.06.012

[24] Yuan G-W, Gao Y, Xu D (2011) A moving objects tracking method based on a combination of local binary pattern texture and Hue. Procedia Eng 15:39643968. doi:10.1016/j. proeng 2011.08.742

[25] Mazinan AH, Amir-Lati_ A (2012) Applying mean shift, motion information and Kalman filtering approaches to object tracking. ISA Trans 51(3):485497. doi: 10.1016/j. isatra 2012.02.002

[26] Lai S-H (2004) Computation of optical ow under non-uniform brightness variations. Pattern Recognit Lett 25(8):885892. doi : 10.1016 / j. patrec 2004.02.001

[27] Alan J Lipton, Hironobu Fujiyoshi, and Raju S Patil. Moving target classification and tracking from real-time video. In Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop on, pages 814. IEEE, 1998.

[28] Chris Stauffer and W Eric L Grimson. Adaptive background mixture models for real-time tracking. In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., volume 2. IEEE, 1999.

[29] Ya Liu, Haizhou Ai, and Guang-you Xu. Moving object detection and tracking based on background subtraction. In Multispectral Image Processing and Pattern Recognition, pages 6266. International Society for Optics and Photonics, 2001.

[30] Changick Kim and Jenq-Neng Hwang. Fast and automatic video object segmentation and tracking for content-based applications. Circuits and Systems for Video Technology, IEEE Transactions on, 12(2):122129, 2002.

[31] Shahbe Mat Desa and Qussay A Salih. Image subtraction for real time moving object extraction. In Computer Graphics, Imaging and Visualization, 2004. CGIV 2004. Proceedings. International Conference on, pages 4145. IEEE, 2004.

* * * * *