EasyChair Preprint

№ 3221

Deep Learning and OpenCV in Real time object Detection and Tracking Techniques: In Live Video Surveillance

Saurabh Patel

April 22, 2020

Deep Learning and opencv in Real time object Detection and Tracking Techniques

Abstract

Moving object detection and tracking are the lot of vital and difficult task in video police work and pc vision applications. Object detection is that the procedure of finding the non-stationary entities within the image sequences. Detection is the first step towards following the moving object within the video. Object illustration is that the next vital step to trace. Tracking is the technique of distinguishing, the position of the moving object within the video. Distinguishing the position is far more difficult task then detection the moving object in a very different video. Object tracking is applied in various applications like in automaton vision, monitoring the traffic, Video police work, Video in-painting and Simulation. Here we have a tendency to be planning to gift a quick review of diverse object detection, object classification and object following algorithms obtainable.

Keywords

Object Detection, Object Tracking, Object Representation, Performance

1. Introduction

Video scrutiny could be a more difficult task in today's environment. Pursuit the moving object has attracted several researchers within the field of pc vision and image processing. Video police investigation is that the method of perceptive the behavior, events and alternative necessary proof, usually of the people for management and protective them. Police investigation is mainly employed by governments for gathering data, for investigating and preventing the crime. Video police investigation activity may be divided into 3 varieties specifically. Manual video surveillance, Semi-autonomous police investigation and fully autonomous surveillance. In manual police investigation human is responsible for analyzing the contents of the video. Semiautonomous surveillance system involves video process by system with the interaction of the human once necessary. Fully autonomous surveillance system, the system can perform each task like detection of motion, pursuit etc., while not the necessity of human interaction. Automated observation procedure additionally called the Intelligence Visual police investigation (IVS) involves the scrutiny and interpretation of objects activities, additionally to object detection and pursuit to acknowledge the visual actions of the scene. The main task of IVS consists of wide space police investigation management and scene interpretation. Object pursuit should wear down many illumination changes and well-known challenges. Principally video analysis is categorized into 3 basic phases: moving entity detecting, finding the mechanical phenomenon of object from one frame to another frame and scrutiny of entity tracks to spot their performance. Pursuit objects in an exceedingly static atmosphere is way easier than pursuit objects in dynamic atmosphere. In common, in an exceedingly dynamic ecological system along background and entity varies [1]. In essence, to resolve this common free downside is rigid. One will place a gaggle of impulsion to form this downside answerable. The lot of the impulsions, the matter is informal to unravel. The rest of the paper is systematic as follows. Section II presents methodology during which object detection, representation and pursuit Techniques is illustrated. Performance analysis is delineate in Section III. Conclusion about the paper is given in Section IV.

2. Methodology

Detection of objects in motion is that the opening move towards non-stationary object chase. Object detection is that the method of finding the non-stationary object during a video sequence{a number of vital the key the foremost} and important ways of detecting the moving objects square measure Frame differencing, Optical flow, Background subtraction and Double distinction etc. Object illustration is that the method of demonstrating the objects. Object illustration will be categorized as form representation, color illustration, texture primarily based representation and Motion familiarized illustration. Object tracking is that the method crucial the position of the moving entity during a sequence of video.

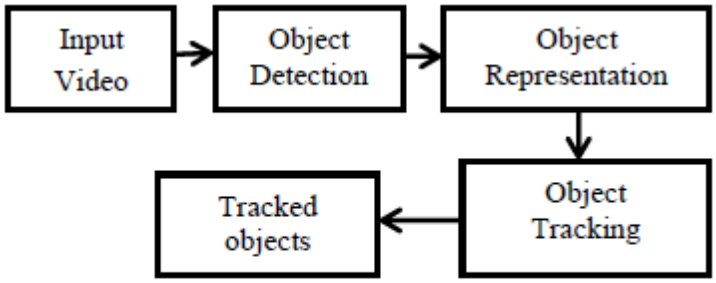

Figure 1: Phases of Moving Object Tracking

Bound sorts of chase algorithms point based tracking, Motion based tracking, feature based tracking, Kernel based tracking and shape based tracking. Step by step detail of the nonstationary object chase is as shown in figure one.

2.1 Object Detection

Object detection is that the methodology for recognizing the nonstationary or moving object in an exceedingly video sequence. This is often the primary and main step towards moving object following.

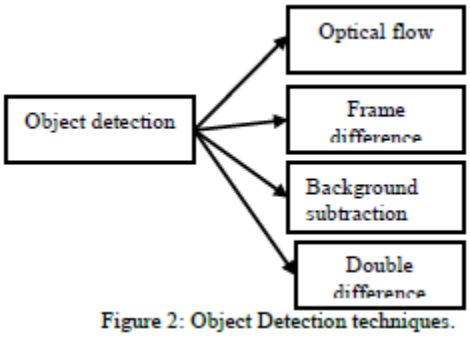

Figure 2: Object Detection techniques.

2.1.1 Background Subtraction

Background subtraction is that the most generally used method for moving object detection. It are often of 2 varieties firstly by considering 1st frame because the arrangement or background image. Second by considering average of „n‟ frames because the background image. During this background subtraction method each component of on-going frame is subtracted with the pixels of the background image. The equation (1) and (2) shows the background subtraction technique for 1st frame because the background image.

$$B(a, b) = A(a, b) \qquad (1)$$

Where B (a, b) represents background image component by pixel. The background subtraction technique splits the video frames into foreground and background object, where the foreground object is discovered by matching this frames A (a, b) through the background image B (a, b). The equation used is

$$C(a,b) = \begin{cases} 1 & if\ B(a,b) - A(a,b) > threshold \\ 0 & otherwise \end{cases} \qquad (2)$$

Where C (a, b) is that the foreground component, threshold price can be set manually or will selected mechanically as per video input. This technique consumes less memory. Accuracy of detection is moderate. However it'll not suit for multimodal backgrounds. Results of the background subtraction strategies area unit as pictured in figure three.

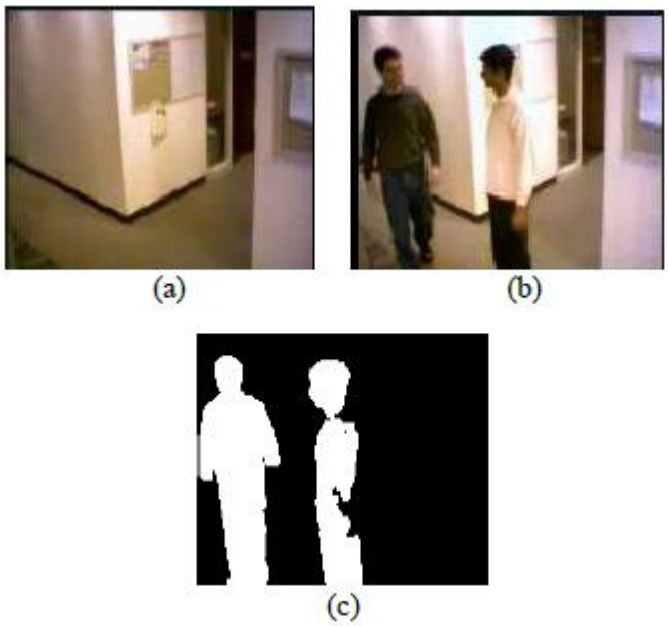

Figure 3: Background subtraction result for hall monitoring video (a) Background model (b) Current frame #89 (c) Resulting frame.

2.1.2 Frame Difference

The frame distinction theme is additionally called the temporal distinction, during which every current frame constituent is subtracted with its previous frame constituent. If the transformation is superior to the manually set threshold price than that constituent is reflected because the foreground constituent else the constituent is mirrored as the background constituent. Equation (3) presents the means for frame difference

$$F(a,b) = \begin{cases} 1 & if\ I_n(a,b) - I_{n+1}(a,b) > T \\ 0 & otherwise \end{cases} \qquad (3)$$

Wherever in that is the previous frame constituent and In+ 1 is that the constituent value of the present frame. T are the edge price that is manually outlined by the user. Calculation of this method is modest and straightforward. For non-static environments, it is very challenging to realize whole define of the moving entity. So it is terribly cumbersome to get accuracy. Results of the frame difference ways are as pictured in figure four.

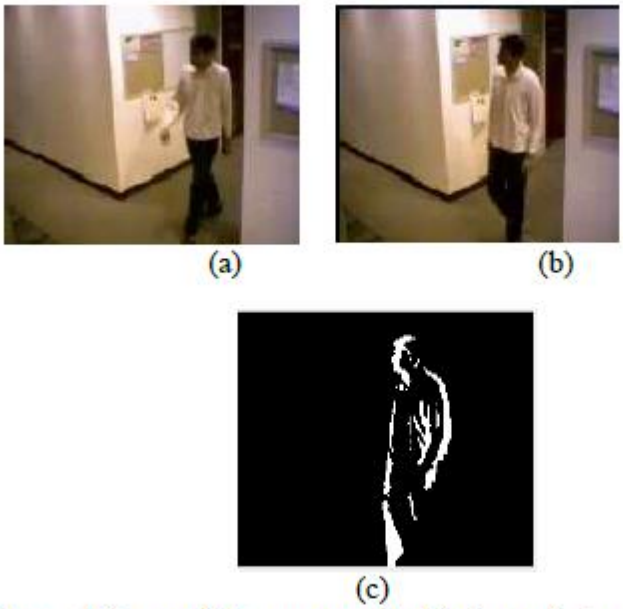

Figure 4: Frame difference result of hall monitoring video. (a) frame #47 (b) frame #48 (c) Resulting frame.

2.1.3 Optical Flow

Optical flow is substitute customary from of object detection during which the optical flow arena of the image is calculated and grouping of these arenas is completed rendering to appearances of the image. The motion among twin video frames occupied at time t and t+ δt at each single location is calculable in optical flow method. This system provides the broad information concerning the movement of the item. And also detects the item accurately compared to it of background technique. This technique isn't wide used attributable to its Brobdingnagian calculation and it's terribly sensitive to noise. It’s not sensible for real-time occlusion condition.

2.1.4 Double Difference

The frame distinction theme is additionally called the temporal distinction, within which every current frame picture element is subtracted with its previous frame picture element and immediate next frame pixel. If the renovation is quite outlined threshold price then that picture element is reproduced because the foreground picture element else, the pixel is replicated because the background picture element.

$$C_n(a,b) = D_n(a,b) - D_{n+1}(a,b) \quad (4)$$

$$C_{n+1}(a,b) = D_{n+1}(a,b) - D_{n+2}(a,b) \quad (5)$$

$$DD(a,b) = C_n(a,b) - C_{n+1}(a,b) \quad (6)$$

Where Cn (a, b) is that the ensuing foreground picture element. Dn denotes the present frame of the video sequence. Dn+1 indicates the next frame. Equally in equation (5) Dn+1 is that the current frame, Dn+2 is that the next frame. Finally DD (a, b) specifies the resulting double distinction frame picture element price.

$$R(a,b) = \begin{cases} 1 \; if \; DD(a,b) > Th \\ 0 \qquad otherwise \end{cases} \quad (7)$$

Where Th is that the threshold price. If the picture element of the absolute distinction is larger than the brink price than the pixel is mirrored as black otherwise it's mirrored as white picture element. This technique produces correct movement of the objects. However it consumes giant memory and its takes longer to calculate.

2.2) Object representation

The extracted moving object could also be of any varieties like human, vehicles, trees, floating clouds, birds and different nonstationary objects. Ways to represent moving objects are shown in figure five.

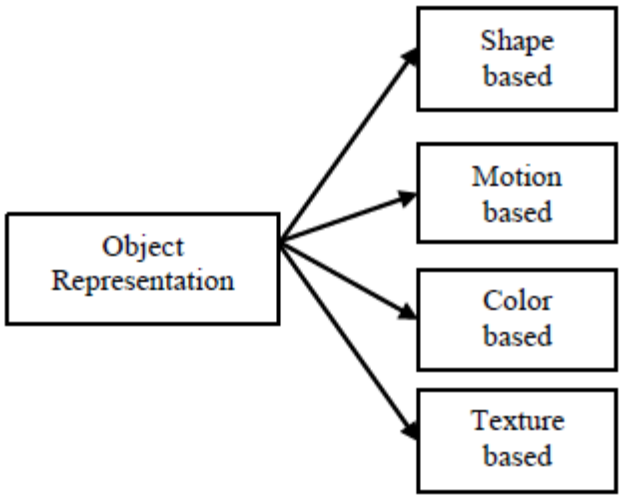

Figure 5: Object Representation Methods.

2.2.1) Shape-based representation

Dissimilar imageries form information of motion regions such as depictions of points, Blob and boxes area unit accessible for categorizing objects in motion [2]. Mixture of image and scene object constraint like blob region of image, deceptive feature ratio of blob parallelogram box space can type the input feature for network of pictures. Ordering is accomplished on each blob at respective frame and outcomes area unit maintained as bar graph.

2.2.2) Motion-based classification

Non-rigid object motion illustrates an interrupted assets, since it's been used as a strong indication for dynamic object organization. Optical flow method is convenient method for object grouping [3]. Residual flow could also be cast-off to scrutinize inflexibility and cyclicity of moving objects. It is predictable that inflexible entities could describe minute lasting flow whereas nonrigid objects like creature has composite average residual flow and even displays a intermittent element.

2.2.3) Color-based Representation

Unlike many options of a picture the relatively constant feature is color info that doesn't change and simple to be developed [4]. even if color isn't permanently applicable because the individual resources of identifying and trailing objects, even the processes that contain tiny procedure value makes color as Associate in Nursing significant feature to observe once appropriate. The most significant technique is color bar graph for detection and tracking non-stationary objects in real time[5]. A Gaussian mixture model is recycled for example the colour scattering surrounded by the structure of pictures.

2.2.4) Texture-based Representation

Texture primarily based system counts the existences of gradient alignment in confined components of a picture, then calculates the knowledge on a condensed grid of systematically spread out cells and uses overlapping slender inequality standardization for increased accuracy. Texture feature is important to live the intensity inequality of surfaces and are apprehensive with object pattern demonstration [1].

2.3) Object tracking

Object illustration is that the method of finding the route of the moving objects in an exceedingly sequence of the photographs [4]. Object chase is accomplished to seek out or manufacture, the path for entity by discovering its location in every frame. Major categories of the chase area unit purpose based mostly chase, Kernel based chase, form based mostly chase, Motion based mostly chase etc. as shown in figure half dozen. Some of the foremost wide used chase techniques are Mean shift chase, CAM-Shift chase, KLT chase etc.

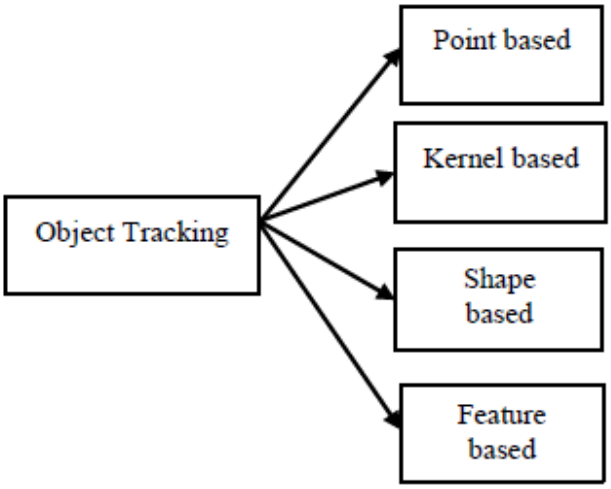

Figure 6: Object tracking methods.

2.3.1) Point based tracking

Moving objects area unit sometimes denoted by their feature point within the image sequence throughout chase. purpose chase will not suit for occlusion condition and false object detection. Some of the points chase algorithms area unit as represented below.

a) Kalman Filter

The Kalman filter is effective for chase distinct sorts of moving objects [4]. It absolutely was at first created by Rudolf Kalman at National Aeronautics and Space Administration to trace the trail of ballistic capsule. Kalman filters will be utilized by numerous differing kinds of linear driving systems. They are created on most Periodic Documents Dispensation algorithmic program. The Kalman Filter accomplishes the obstructive likelihood compactness propagation. It uses a gaggle of scientific equations that has a cost-effective computational thanks to appraise the objects state of incidence. The Kalman filter measures a procedure through a type of opinion management. The Kalman filter may be a algorithmic two-stage filter. The two phases area unit predict and update [5]. Within the predict part, current location of the moving object is calculable or foretold based on the previous observation created. For illustration, if an object is moving with consistent dispatch, current location of object at interference will be supported its previous location, Xt-1. Within the update phase the capability of the capability current location is pooled with the predicted location and acquires the posteriori projected current position of the item.

Frame #65 Frame #98

Figure 7: Kalman filter tracking results for video with ball moving in frame #65 and #98.

b) Particle Filtering

Particle Filter is most generally used filter to trace single and multiple moving objects [2]. It's a hypothesis huntsman, filtered future distribution area unit calculable by a collection of weighted particles. One restriction obligatory by Kalman filter is the predefined state variables area unit commonly spread (Gaussian). This algorithmic program normally uses, or texture feature mapping, contours, color options etc. The particle filter may be a Bayesian written record Trial procedure, that algorithmic approach. It conjointly contains 2 phases like that of Kalman filter, predict and update part.

2.3.2) Kernel based tracking

Kernel based mostly chase is usually performed to see objects in motion, that is symbolized by elementary object area, since frame to frame[3]. The item movement is usually in the methodology of constant movement like transformation, affine, etc. this kind of chase will be applied for each rigid and non-rigid moving object. Kernel based mostly chase sometimes works supported the item illustration, appearance, silhouette of the item.

a) Mean shift tracking

In mean shift chase the position of the nonstationary object is caterpillar-tracked centred on the bar graph [5]. In this methodology region of the item will be copied created on the similarity computation of the item. A gradient rise scheme is used to transfer the huntsman to position that exploits a similarity price among the model and therefore the current frame region. Target region area unit sometimes designated in parallelogram form or elliptical kind. This chase methodology contains target model and candidate model [6]. Color bar graph is sometimes chosen to characterize the target. Likelihood Density operate is used to represent the target model. If the similarity scores matches than the item is caterpillar-tracked. Results of mean shift tracking, initial figure shows the target choice and second shows its chase outcome.

Figure 8: Mean-shift Tracking results.

b) Simple Template Matching

Template matching may be a heuristic search method of scrutinizing the vital areas within the video [6]. A background image is matched with the sequence of pictures that area unit detached from a video. This algorithmic program suite well for single moving object chase. Example matching is that the method for handling sequence of pictures to find minor portions of a picture that matches model with a example image in each frame. The matching technique encompasses of image pattern for all probable locations within the basic image and computes a geometrical key that identifies however accurately the model suits the image position. This method is wide used once single object or particle occluded object is to be tracked.

c) Layering based tracking

This is the foremost common mechanism for chase multiple non-stationary objects in an exceedingly video [7]. Each layer comprises of form illustration, movement like rotation and translation, and layer presence, relying upon power. Layering is earned chiefly by recompensing the background movement, such movement will be calculated by the content image via second constant gesture. All pixel's prospect is calculable supported such previous gesture and form structures. It may track multiple pictures and complete occlusion condition of the item.

2.3.3) Shape based tracking

Certain objects would have composite shapes like hand, finger, shoulder, nose etc. that cannot be demonstrated by precise geometric shapes [6]. The goal of this technique is to spot the moving object state in every and every one frame through object model created within the previous frame. This technique is practiced of chase advanced form objects. Occlusion condition, split and merge condition may be handled during this methodology. Data of the item region is utilized by this method to trace the objects.

a) Contour Tracking

Contour chase ways, repeatedly method a principal contour of the previous structure to its new space within the present image sequence [6][7]. This silhouette procedure desires specific live relating to things gift within the current frame intersection with the entity district within the previous frame. Contour Tracking could also be enforced utilizing 2 distinct methodologies. The principal methodology utilizes formal house models to demonstrate the shape, form and movement. The subsequent methodology foursquare develops the form by reducing the contour vitality utilizing direct reduction

ways, likewise gradient descent. This can be versatile to handle immense variation of item form.

b) Shape Based Matching

These methodologies examine for the item model in the extant frame. Form identical algorithmic program works analogous to that of example based mostly chase methodology. This methodology is employed to identify the matching shapes detected in a pair of consecutive frames. Form matching, will be measured appreciate purpose matching [3][4]. Background subtraction mechanism is used for police investigation the form of the item. Object reproductions area unit in the structure of density gatherings, form frontier, edges of the objects. Form based mostly matching area unit practiced of coping with solitary object and Hough renovate technique is used to handle occlusion condition.

2.3.4) Feature Based tracking

Feature based mostly chase is that the most generally used chase technique currently a days. It's chiefly divided into a pair of steps. First phase is to extract the options of the entity like centre of mass, shape, color etc. Second step is to match those options in each frame. One or additional feature will be combined to get higher results or outcome. Cluster.

a) Color feature based tracking

The color data of each object in an exceedingly video frame is assembled group-by-group [8]. Specific weights area unit appointed for every cluster for more analysis method. The color evidence is obtained from the movement chunks within the current frame area unit divided into regions of nearly alike color as cluster. Color data will be either RBG or HSV [1]. Consequent step is to acknowledge matching color data. This is accomplished by matching the cluster of color knowledge of the motion chunk within the existing frame with the cluster color proof motion chunks within the previous frames exploitation weighted matching. The maximum analysis score of each cluster within the existing frame is achieved and huntsman is initialized to it position [8]. This method repeats till finish of the video sequence.

b) Centroid feature tracking

The key feature centre of mass of every object is extracted. The centre of mass data is accessed from the gesture chunks in the recent frame to classify identical centre of mass data among gesture blocks within the existing frame and previous frames [8]. Consequently, a bounding box with centre of mass is allotted for the gesture blocks within the existing frame. New centre of mass is detected by averaging the add of previous centre of mass as in equation (8) Edge feature chase is analogous to the colour based mostly feature chase [8].

$$Cen_n = \frac{\sum_{i=1}^{n-1} cen_i}{n} \qquad (8)$$

Where *Cenn* is the new centroid point and *Ceni* prior centroid and *n* is the overall number of frames processed. Results of the centroid feature based tracking is as shown below

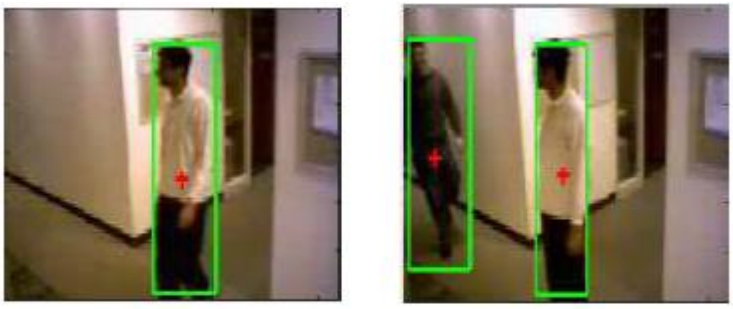

Figure 9: Tracking results for centroid based object tracking frame # 85 and frame # 131

c) Edge feature tracking

Fringe of the moving object will be determined by numerous edge detection technique like smart edge detection, sobel edge operator etc. in each frame. The next step is to spot matching edge data. This is obtained by matching the burden data of the gesture block of the current frame with the burden knowledge of the gesture chunks within the previous frames. The very best score comparison of every cluster within the current frame is achieved and huntsman is initialized to it position. This methodology is perennial till last frame of the video sequence.

III. Performance Evaluation

Performance of the video analysis is achieved by extracting manually non-stationary object as ground-truth image and compares it with the results obtained. Some parameter like True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FP) are calculable. TP refers to total quantity of pixels wherever each ground-truth and proposed outcome overlap. Volunteer State refers to total variety of pixels that doesn't comprise object, each in ground truth and proposed. FP is add of pixels within which planned system contains object however ground truth doesn't. FN refers to total amount of constituent wherever ground-truth contains entity and projected system doesn't have objects. Precision and Recall are a number of the necessary factors for performance analysis. Recall provides the percentage of truth positive rate that's the same as groundtruth image. It's given by the equation

Recall= TP/ (TP+FN) (9)

Precision offers the proportion of truth positive that doesn't match with the ground-truth image. The precision is found by victimization the equation

Precision = TP / (TP+FP) (10)

A) MOTP (Multiple object tracking precision)

Total faults in known space for compared object hypothesis among all frames, averaged via add of comparisons made. It illustrates the capability of hunter to work out specific object space, liberated of its proficiency at distinctive objects formation.

$$MOTP = \frac{\sum_{i,t} d_i^t}{\sum_i c_t} \qquad (11)$$

MOTA (Multiple object tracking accuracy)

Where m_i, fp_i indicates number of misses related to false positive. Time mismatches at an interval of „i" is demonstrated by mme_i.

$$MOTA = 1 - \frac{\sum_i (m_i + fp_i + mme_i)}{\sum_i g_i} \qquad (12)$$

MOTA provides various error percentages that are

$$m' = \frac{\sum_i m_i}{\sum_i g_i} \qquad (13)$$

Is the proportion of errors in the group, and is calculated around the sum of quantity of object existing in every frame.
Percentage of false positive and percentage of disparities is given by the equation (14) and (15) respectively

$$fp' = \frac{\sum_i fp_i}{\sum_i g_i} \qquad (14)$$

$$mme' = \frac{\sum_i mme_i}{\sum_i g_i} \qquad (15)$$

IV. Conclusion

In this paper totally different phases of object pursuit technique like object detection, object illustration and object tracking square measure summarized. Among numerous detection strategies Background subtraction is that the simplest technique that provides complete info. A recent man of science chooses texture and color for representing objects. Object pursuit is achieved via various algorithms that uses purpose, shape, feature techniques. These square measure some economical algorithms that decrease calculation time. And conjointly reduces the value, used for pursuit the entities for numerous styles of the sequence of frames containing differentiated appearances. These square measure a number of basic techniques that square measure usually used. These techniques works well for static camera. Despite the actual fact that amid the foremost recent centuries, has seen a major advancement against object detection and tracking. However, pursuit objects in a very droning, informality, and vigorous things creates this issue because the analysis topic. Another key issue influencing is analysis of system powerful and good to ecological variations remains a inspiring work. Manipulation of previous and specific information in pursuit is quiet in its primary section. A serious attention ought to be paid towards pursuit objects in droning and dense video information.

References

[1] Anand Singh Jalal, Vrijendra Singh " The State of-the-Art in Visual Object Tracking", International journal of computing and scientific discipline, Vol. 36, Issue-3, pp. 227-248, May 2012.

[2] J.Joshan Athanesious, P.Suresh, "Systematic Survey on Object chase ways in Video", International Journal of Advanced analysis in laptop Engineering & Technology (IJARCET) pp. 242-247, October 2012.

[3] A Yilmaz, O. Javed, M. Shah, "Object Tracking: A Survey", ACM Computing Survey, Vol. 38, Issue 4, pp. 1-45, 2006.

[4] Gandham Sindhuja, Dr Renuka Devi S.M, "A Survey on Detection and chase of Objects in Video Sequence", International Journal of Engineering analysis and General Science Volume three, Issue 2, March-April, 2015. [5] Hitesh A Patel, Darshak G Thakore,"Moving Object Tracking exploitation Kalman Filter", International Journal of Computer Science and Mobile Computing , pg.326 – 332, April 2013.

[5] S.Bhuvaneswari, T.S.Subashini, "Tracking Manually Selected Object In Videos exploitation Color bar graph Matching", International Journal of Theoretical and Applied data Technology, Vol. 67, Issues 3, pp. 562 -568, September 2014.

[6] Mr. Joshan Athanesious J; man. Suresh P, "Implementation and Comparison of Kernel and Silhouette primarily based Object Tracking", International Journal of Advanced analysis in Computer Engineering & Technology, pp 1298-1303, March 2013.

[7] Kinjal A Joshi, Darshak G. Thakore, "A Survey on Moving Object Detection and chase in Video police investigation System", International Journal of soppy Computing and Engineering, Vol. 2, Issue-3, July 2012.

[8] Sunitha M.R, H.N Jayanna, Ramegowda, "Tracking Multiple Moving Object supported combined Color and Centroid feature in Video Sequence", in IEEE International Conference on process Intelligence and Computing Research, pp. 846-850, Dec 2014.

[9] M.Besita Augustin, Mrs. Sujitha Juliet, Mr. S. Palanikumar, "Motion and have primarily based Person chase In Surveillance Videos", in Proceedings Of ICETECT, Vol. 19, no. 7, pp. 780-785, Dec 2011.

[10] Rupali S.Rakibe, Bharati D.Patil, "Background Subtraction algorithmic rule primarily based Human Motion Detection", International Journal of Scientific and analysis Publications, May 2013.

[11] Saravanakumar, S, Vadivel A, Saneem Ahmed, C.G., " Multiple human object chase exploitation background subtraction and shadow removal techniques", International Conference on Signal and Image process, pp.79-84, Dec. 2010.

[12] W. M. Hu, T. N. Tan, L. Wang and S. Maybank, "A survey on visual police investigation of object motion and behaviour", IEEE Transactions on Systems, Man and Cybernetics, vol.34, no.3, pp. 334-352, Dec 2004.

[13] Adam, Amit, Ehud Rivlin, and Ilan Shimshoni. "Robust fragments-based chase exploitation the integral bar graph." Computer Vision and Pattern Recognition, Vol. 1. IEEE, 2006.

[14] Rahul Mishra, Mahesh K. Chouhan, Dr. Dhiiraj Nitnawwre, "Multiple Object chase by Kernel primarily based Centroid methodology for Improve Localization", International Journal of Advanced analysis in computing and Software Engineering, July-2012, pp 137-140.

[15] Hitesh A Patel, Darshak G Thakore,"Moving Object Tracking exploitation Kalman Filter", International Journal of Computer Science and Mobile Computing, April 2013, pp.326 – 332.