International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 04 | Apr 2019 www.irjet.net p-ISSN: 2395-0072

Real-time Object Detection: A Survey

Tareek Pattewar[1], Amit Chaudhari[2], Mrunal Marathe[3], Moushumi Bhol[4]

[1]Assistant Professor, Dept. of Information Technology, R. C. Patel Institute of Technology, Maharashtra, India
[2,3,4]Student, Dept. of Information Technology, R. C. Patel Institute of Technology, Maharashtra, India

---

Abstract - *In recent trends, Image Processing domain plays a vital part in detecting a moving objects in a Video/Image. Also there are various techniques which resembles to be the tool for Image Processing which involves noise free images and compressing the original images. There are plenty of techniques available for detecting and recognizing the objects and analyzing the frame accordingly, but in this survey paper we studied various object detection methods.*

Key Words: **Image processing, background subtraction, mask, morphological operations, object, labelling**

1. INTRODUCTION

Image Processing is any form of signal processing in which the input will be given as an image, such as a photograph or video frame; the output of image processing will be either an image or a set of characteristics or parameters that are related to given image. Image processing involves processing or altering an existing image in a desired manner and also helps in obtaining the image in the readable format.

Most techniques of image-processing involve treating the image as 2 -dimensional signal and applying standard signal-processing techniques to it. The MATLAB and Mathcad are the two environments which suits for image processing. In this, MATLAB based on matrix-oriented language and well suited for manipulating images. The result produces very clarity image and economical way of expressing image processing operations.

1.1 Benefits of image processing:

1. Visualization helps in identification of the objects that are not visible.
2. Image processing is faster and cost effective.
3. Noise free.
4. Image sharpening and restoration - To create a better image.
5. Images can be retrieval easily from the database.

1.2 Problems involved in Detecting Objects:

1. Illumination variation :-
   Lighting conditions of the scene and the target might change due to motion of light source, different times of day, reflection from bright surfaces, weather in outdoor scenes, partial or complete blockage of light source by other objects etc. Thus our method selection should be in appropriate manner so that it must be capable of adjusting variable amount of Light.

Fig -1: An example of Illumination Challenge [1].

2. Moving Object Appearance Changes :-
   While working with whole video objects are projected and work in 3D spaces but when we apply our algorithm on individual frames than we came to know that objects have been adjusted in 2D planes.

Fig -2: An Example of Appearance change Challenge [1].

3. Presence of Abrupt Motion :-
   Sudden changes in the speed and direction of the object's motion or sudden camera motion are another challenges in object detection and tracking. If the object or camera moves very slowly, the temporal differencing methods may fail to detect the portions of the object coherent to background.

Fig -3: An Example for Abrupt Motion [2].

4. Occlusion :-

   The object may be occluded by other objects in the scene. In this case, some parts of the Object can be camouflaged or just hidden behind other objects (partial occlusion) or the object can be completely hidden by others (complete occlusion). As an example, consider the target to be pedestrian walking in the sidewalk, it may be occluded by trees, cars in the street, other pedestrians, etc.

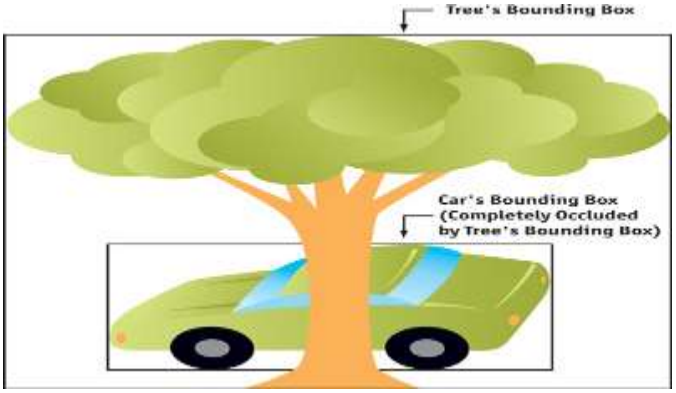

Fig -4: An Example for Occlusion in Image [3].

5. Shadow :-
   The presence of shadows in video image sequences complicates the task of moving object detection. Shadows occur due to the block of illumination from the light source by objects. If the object does not move during the sequence, resulted shadow is considered as static and can effectively be incorporated into the background.

Fig -5: An Example of Shadow Challenge [4].

6. Problems related to Camera :-
   Many factors related to video acquisition systems, acquisition methods, compression Techniques, stability of cameras can directly affect the quality of a video sequence. In some cases, the device used for video acquisition might cause limitation for designing object detection and tracking.

Fig -6: An Example for Problems in Camera [5].

2. LITREATURE SURVEY

1. Shashank Prasad et.al, have developed and implemented object detection and tracking system operational in an unknown background, using real-time video processing and a single camera. The system was been extensively tested to operate in complex, real world, non-plain, light variant, changing background. Many remarkable algorithms were been developed for object detection and tracking, including color segmentation, edge tracking and many more. However, all that algorithms faced the limited success in their implementation in the real world and were also bounded by the constraints such as white/plain background [6].
2. P. Khurana et.al, has analyzed the Object Recognition and Segmentation techniques in context with images and videos. According to their study object Recognition can be used in various fields such as Robot navigation, Medical diagnosis, Security, Industrial inspection and automation, Human-computer interface, Information retrieval. Object segmentation is today used in diversified fields such as: image processing, video recognition, shadow detection, Human Activity Recognition and many more. They had done systematic analysis of various existing object recognition and segmentation skills, with precise and arranged representation. They identified that recognizing and segmenting an object can be scrutinized from the perspective of static and moving objects [7].
3. Liming Wang et.al, have developed an object detection method combining top-down recognition with bottom-up image segmentation. . There were

two main steps in there method: a hypothesis generation step and a verification step. In the top-down hypothesis generation step, they design an improved Shape Context feature, which is more robust to object deformation and background clutter. The improved Shape context was used to generate a set of hypotheses of object locations and figure ground masks, which had high recall and low precision rate. In the verification step, they first compute a set of feasible segmentations that were consistent with top-down object hypotheses, then they proposed a False Positive Pruning (FPP) procedure to prune out false positives. They exploit the fact that false positive regions typically do not align with any feasible image segmentation [8].

4. Divya Patel et.al, had presented different techniques and methods for detecting or recognizing object with various benefits like efficiency, accuracy, robustness etc. They stated that Object detection is a computer technology that connected to image processing and computer vision that deal with detecting instance objects of certain class in digital images and videos. Object detection is a challenging problem in vision based computer applications. It is used to identifying that whether in scene or image object is been there or not [9].
5. Xinyi Zhou et.al, had discussed application of deep learning in object detection task. There were simple summary of the datasets and deep learning algorithms commonly used in computer vision. On the other hand, a new dataset are building according to those commonly used datasets, and choose one of the network called faster r-cnn to work on this new dataset. They carried out that experiment to strengthen the understanding of these networks & through the analysis of the results they learned the importance of deep learning technology, and the importance of the dataset for deep learning [10].
6. Meera M K et.al, have formulate a Object detection technique into two stages. For the first stage, the query image was categorized using a classifier. For classifier optimization they had implemented two types of classifiers- Support Vector Machine(SVM) classifier that make use of GIST features and k-nearest neighbor(kNN) classifier that make use of Scale Invariant Feature Transform(SIFT). GIST based SVM classification was done using different kernels such as linear kernel, Polynomial kernel and Gaussian kernel. SIFT features are invariant to affine transformations of the images. SIFT features of the images were extracted and a similarity matrix was formed by matching these SIFT features. Then a k-nearest neighbor(kNN) classifier was implemented on the similarity matrix. GIST feature based SVM classifier with Gaussian kernel showed better classification accuracy than SIFT feature based kNN classifier. The image datasets considered for this work were Coil-20P and Eth80 [11].
7. Shuai Zhang et.al, had proposed that Object detection and tracking were two fundamental tasks in multicamera surveillance. They also proposed a framework for achieving these tasks in a no overlapping multiple camera network. A new object detection algorithm using mean shift (MS) segmentation was introduced, and occluded objects were further separated with the help of depth information derived from stereo vision. The detected objects were then tracked by a new object tracking algorithm using a novel Bayesian Kalman filter with simplified Gaussian mixture (BKF-SGM). It employs a Gaussian mixture (GM) representation of the state and noise densities and a novel direct density simplifying algorithm avoids the exponential complexity growth of conventional Kalman filters (KFs) using GM. When coupled with an improved MS tracker, a new BKF-SGM with improved MS algorithm with more robust tracking performance was obtained. Furthermore, a non-training-based object recognition algorithm is employed to support object tracking over non overlapping network [12].
8. Sheng Ding et.al, have applied the deep learning algorithm to the detection of daily objects, and some progress has been made in that direction. Compared with traditional object detection methods, the daily objects detection method based on deep learning were faster and more accurate. The main research work of their article were : 1. collect a small data set of daily objects; 2. in the Tensor Flow framework to build different models of object detection, and use this data set training model; 3. the training process and effect of the model are improved by fine-tuning the model parameters. They also studied with the rapid development of deep learning, great breakthroughs were been made in the field of object detection [13].
9. Liming Wang et.al, had developed an object detection method combining top-down recognition with bottom-up image segmentation. There were two main steps in their method: a hypothesis generation step and a verification step. In the top-down hypothesis generation step, they designed an improved Shape Context feature, which was more robust to object deformation and background clutter. The improved Shape Context was then used to generate a set of hypotheses of object locations and figure ground masks, which had a high recall and low precision rate. In the verification step, they first compute a set of feasible segmentations that are consistent with top-down object hypotheses, then they proposed a False Positive Pruning (FPP) procedure to prune out false positives. They also exploit the fact that false positive regions typically

do not align with any feasible image segmentation. Experiment showed that simple framework was capable of achieving both high recall and high precision with only a few positive training examples and that their method can be generalized to many object classes [14].

10. Palak Khurana at.al, have analyzed the Object Recognition and Segmentation techniques in context with images and videos. They stated that Object segmentation used today in diversified fields such as: image processing, video recognition, shadow detection, Human Activity Recognition and many more. Their research consists of systematic analysis of various existing object recognition and segmentation skills, with precise and arranged representation. While doing the research they had identified that recognition and segmenting an object can be scrutinized from the perspective of static and moving objects. Most of the used techniques were based on mathematical and algorithmic models. They wind up their results with the merits and demerits of existing methods and the liabilities of future scope in this area [15].
11. Xiaofeng Ning et.al, had addressed the recognition, object detection and segmentation issues in white background photos with deep learning method. In particular, they first trained a recognition model based on Google Net to judge whether a photo is white background. Then they proposed a main object detecting algorithm to eliminate unnecessary elements such as logos, characters with Faster R-CNN. Eventually a main object segmentation method combining both CRF-RNN network and Grabcut was adopted to smoothly eliminate the shadow area and obtain the fine segmentation results. All exploring algorithms were implemented in real time with Caffe and Tesla K80 from Nvidia [16].
12. Sanjana Yadav et.al, had proposed a precise approach for image matching in real time scenario and also in the field of ROBOTICS. They had constructed an application based Image matching model that was able to detect images that are exactly the same, as well as images that have been edited in some ways. Implementation of that Image Matching and object recognition system was based on tracking an object, calculating its feature Points, and classification with the help of trained Data Sets. The system was capable to perform matching of images, both automatically and manually. On the other hand, to operate the proposed system manually, user itself takes the images of an object or something, and will store it in database. After that system will through some steps, and image matching will be performed. Black & White point calculation of an image, Chamfer matching Algorithm, 3-4 Distance Transformation with canny edge detector was applied in the application which was well suited for calculating the pixel values [17].
13. Palak Khurana et.al, have analyzed the Object Recognition and Segmentation techniques in context with images and videos. Object Recognition can be used in various fields such as Robot navigation, Medical diagnosis, Security, Industrial inspection and automation, Human-computer interface, Information retrieval. Object segmentation is today used in diversified fields such as: image processing, video recognition, shadow detection, Human Activity Recognition and many more. Their work consists of systematic analysis of various existing object recognition and segmentation skills, with precise and arranged representation. While doing the research they have identified that recognizing and segmenting an object can be scrutinized from the perspective of static and moving objects. Most of the used techniques were based on mathematical and algorithmic models. They wind up their results with the merits and demerits of existing methods and the liabilities of future scope in that area [18].

3. METHODOLOGY

Based on our literature survey we observe following methods can be proved useful in detecting the objects.

1. Modelling based Background Subtraction :-

One of the most common techniques used to detect moving objects in a video sequence Captured by a moving camera is background subtraction based on background modelling. The General concept of these techniques is shown in Fig. 12. As can be seen, first a background model is initialized using a set of first frames of the sequence. Typically the model is created based on statistical features extracted from the background. Next, some feature points are extracted from the current frame and then their correspondences in the background are found.

a) Background model creation: Wren et.al [19] used a simple Gaussian model, which was not robust to a complex background. Hayman and Eklundh [20] used a mixture of Gaussian (MoG) Which was more appropriate for representing a complex background.
b) Feature point selection: Zivkovic and Van der Heijden [21] used Laplacian gradient to find the interesting points in the current frame, however it is very sensitive to noise, which is present most of the time in the frame. Setyawan et al. [22] proposed to use the Harris corner detector to extract the feature points. This approach is less sensitive to the noise and robust to illumination changes and served commonly by other methods.

2. Trajectory Classification :-

Trajectory classification has been used in many works to obtain moving objects in the video sequences captured by a moving camera. The general concept of this technique is shown in Fig.07. Preliminary, the interesting points are chosen from the first image of the sequence.

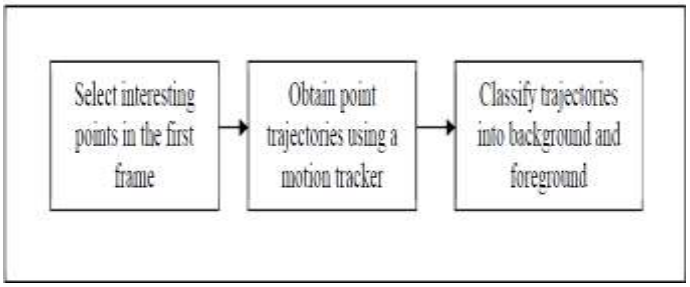

Fig -7: General Diagram for Trajectory Classification [23].

3. Low Rank and Sparse Matrix decomposition :-
The Low rank and sparse decomposition is currently considered to be one of the leading techniques for video background modelling, which consists of segmenting the moving objects from the static background. A common solution is to embed a global motion compensation model into matrix decomposition optimization. Figure. 08 shows the general concept of this strategy.

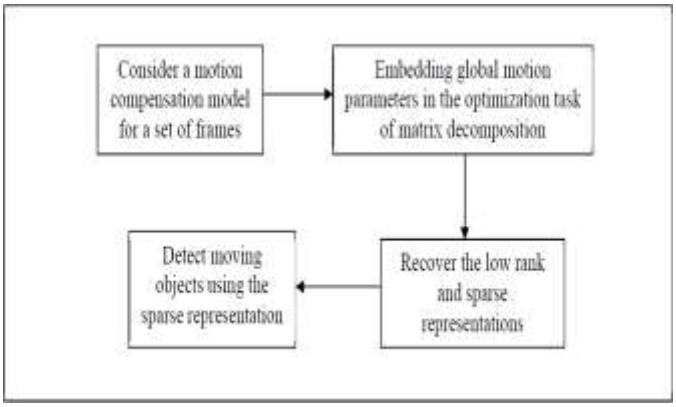

Fig -8: Concept of Low Rank & Sparse decomposition technique for moving camera [24].

4. Object Tracking :-

The aim of object tracking is to associate different regions belonging to the same object in consecutive frames of a video sequence. In other words, tracking is the localization of a moving object in frames of the sequence, so we can consider it as a process of moving object detection. The general concept of this technique is shown in Fig. 09. First, a number of connected pixels in the first image of the sequence are marked as the desired object (target).

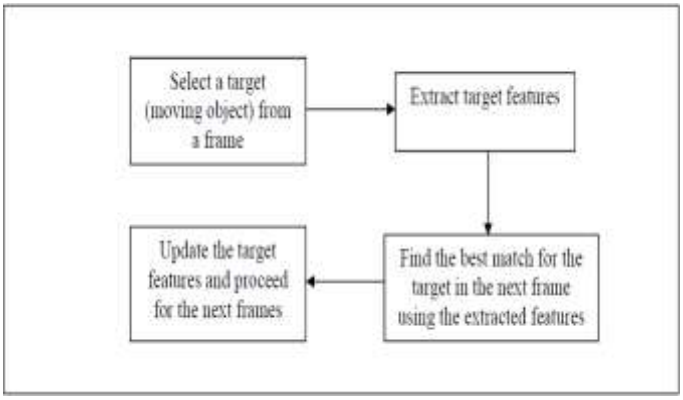

Fig -9: General Concept of Object Tracking for Moving Camera [25].

5. Optical Flow :-

Optical flow method [26] is to calculate the image optical flow field, and do clustering processing according to the optical flow distribution characteristics of image. This method can get the complete movement information and detect the moving object from the background better, however, a large quantity of calculation, sensitivity to noise, poor anti-noise performance, make it not suitable for real-time demanding occasions.

Given Below is a comparison chart for Optical Flow & Background Subtraction Method.

Methods		Accuracy	Computational Time	Comments
Background Subtraction	Gaussian Of Mixture	Moderate	Moderate	+ Low memory requirement
				- It does not cope with multimodal background
	Approximate Median	Low to Moderate	Moderate	+ It does not require sub sampling of frames for creating an adequate background model
				- It computation requires a buffer with the recent pixel values
Optical Flow		Moderate	High	+ It can produce the complete movement information
				- Require Large amount of calculation

Fig -10: Comparison chart between Optical Flow & Background Subtraction [27].

Based on our survey we observe following methods can be proved useful in classifying the objects.

Object can be classified as vehicles, birds, floating clouds, swaying tree and other moving objects. The approaches to classify the objects are

1. Shape-based classification:
Different descriptions of shape information of motion regions such as representations of points, box and blob are available for classifying moving

objects. Input features to the network is mixture of image-based and scene-based object parameters such as image blob area, apparent aspect ratio of blob bounding box and camera zoom.

2. Motion-based classification:
Non-rigid articulated object motion shows a periodic property, so this has been used as a strong cue for moving object classification. Optical flow is also very useful for object classification. Residual flow can be used to analyse rigidity and periodicity of moving entities. It is expected that rigid objects would present little residual flow where as a non-rigid moving object such as human being had higher average residual flow and even displayed a periodic component [28].

3. Color-based classification:
Unlike many other image features (e.g. shape) color is relatively constant under viewpoint changes and it is easy to be acquired. Although color is not always appropriate as the sole means of detecting and tracking objects, but the low computational cost of the algorithms proposed makes color a desirable feature to exploit when appropriate.

4. Texture-based classification:
Texture based technique [29] counts the occurrences of gradient orientation in localized portions of an image, is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

Based on our survey we observe following features can be considered while recognizing objects.

1. Edge – based features :

Most object detection and recognition methods can be classified into two categories based on the feature type they use in their methods. The methods that use edge-based feature type extract the edge map of the image and identify the features of the object in terms of edges. Some examples include [29]. Using edges as features is advantageous over other features due to various reasons. They also represent the object boundaries well and represent the data efficiently in the large spatial extent of the images. In this category, there are two main variations: use of the complete contour (shape) of the object as the feature and use of collection of contour fragments as the feature of the object FIGURE .11 shows an example of complete contour and collection of contours for an image.

Fig -11: Edge based Feature Extraction [30].

2. Patch – based Features :-

In this feature there are two variations:
a. Patches of rectangular shapes that contain characteristic boundaries describing the features of object [31]. Usually this features are known as local features.
b. Irregular patches, in which each patch is homogeneous in terms of intensity or texture.

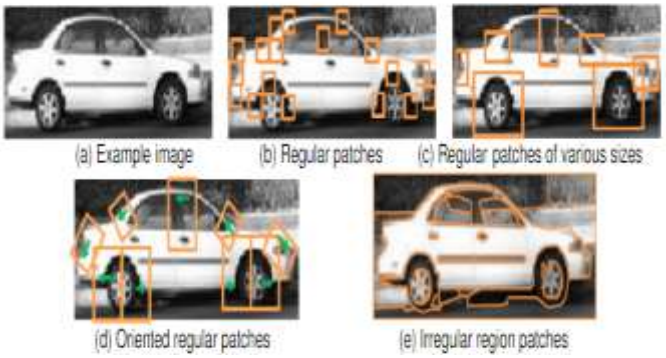

Fig -12: Patch based Feature Extraction [30].

4. CONCLUSION

We studied Optical, Sparse Matrix Distribution, Trajectory Paths methods can be found useful not only for better recognition of objects in the Video Streams but also with noisy inputted video & missing pixel intensities, which are further removed through morphological operations. We also come to know about features that can be considered for object and detected motion among them which can help in better classification. We also studied how problem of occlusion can be avoided so that better recognition of object over any other object can be done. Use of better classifier for object detection can be done and maximum number of features extraction can help in future to detect objects quickly.

REFERENCES

[1] Incremental learning for robust visual tracking project website. http://www.cs.utoronto.ca/~dross/ivt/ (2007)
[2]https://www.sciencedirect.com/science/article/pii/S092
23121730694X

[3]https://developer.nvindia.com/sites/all/modules/custom/gpuges/books/GPUgems/gpugems_ch29.html
[4] Zdenek Kalal's website. http://personal.ee.surrey.ac.uk/Personal/Z.Kalal/ (2011)
[5] https://digital-photography-school.com/how-to-avoid-camera-problems/
[6] S. Prasad and S. Sinha, "Real-time object detection and tracking in an unknown environment," 2011 World Congress on Information and Communication Technologies, Mumbai, 2011, pp. 1056-1061.doi: 10.1109/WICT.2011.6141394
[7]P. Khurana, A. Sharma, S. N. Singh and P. K. Singh, "A survey on object recognition and segmentation techniques," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 3822-3826.
[8] Liming Wang, Jianbo Shi, Gang Song, and I-Fan Shen. 2007. Object detection combining recognition and segmentation. In Proc of - Volume Part I . Springer-Verlag, Berlin, Heidelberg, 189-199.
[9] https://pdfs.semanticscholar.org/86da/acfca95204c1bff000da7fa5fcfb430205e5.pdf
[10] X. Zhou, W. Gong, W. Fu and F. Du, "Application of deep learning in object detection," 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), Wuhan, 2017, pp. 631-634. doi: 10.1109/ICIS.2017.7960069
[11] Meera M K and Shajee Mohan B S, "Object recognition in images," 2016 International Conference on Information Science (ICIS), Kochi, 2016, pp. 126-130.
[12] S. Zhang, C. Wang, S. Chan, X. Wei and C. Ho, "New Object Detection, Tracking, and Recognition Approaches for Video Surveillance Over Camera Network," in IEEE Sensors Journal, vol. 15, no. 5, pp. 2679-2691, May 2015.
[13] Sheng Ding, P. Meenaltha,Research on Daily Objects Detection Based on Deep Neural Network, In- ternational Journal of Advanced Computer Science and Applications, 2016, 57- 63
[14] Liming Wang, Jianbo Shi, Gang Song, "Object Detection Combining Recognition and Segmentation ," in IEEE Sensors Journal, vol. 15, no. 5, pp. 2679-2691, May 2015.
[15] Palak Khurana and Anshika Sharma, "A Survey on Object Recognition and Segmentation Techniques," 2016 International Conference on Information Science (ICIS), Banglore, 2016, pp. 130-140.
[16] Xiaofeng Ning, Wen Zhu and Shifeng Chen, "Recognition, object detection and segmentation of white background photos based on deep learning" , 2017 World Congress on Information and Communication Technologies, Mumbai, 2011, pp. 1056-1061. doi: 10.1109/WICT.2011.6141394
[17] S. Yadav and A. Singh, "An image matching and object recognition system using webcam robot," 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, 2016, pp. 282-286
[18] P. Khurana, A. Sharma, S. N. Singh and P. K. Singh, "A survey on object recognition and segmentation techniques," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 3822-3826.
[19] C.R. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland, Pfinder:Real-time tracking of the human body, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), (1997), 780–785.
[20] E. Hayman and J. O. Eklundh, Staristical background subtraction for a mobile observer, IEEE International Conference on Computer Vision (ICCV), (2003), 67-74.
[21] Z. Zivkovic and F. Van der Heijden, Efficient adaptive density estimation per image pixel for the task of background subtraction, Pattern Recognition Letters, 27(7), (2006), 773–780.
[22] F. A. Setyawan, J. K. Tan, H. Kim, and S. Ishikawa, Detection of moving objects in a video captured by a moving camera using error reduction, SICE Annual Conference, Sapporo, Japan, (Sept. 2014), 347-352.
[23] Y. Sheikh, O. Javed, and T. Kanade, Background subtraction for freely moving cameras, IEEE International Conference on Computer Vision (ICCV), (Sept. 2009), 1219-1225.
[24] X. Zhou, C. Yang, and W. Yu, Moving object detection by detecting contiguous outliers in the low-rank representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(3), (2013), 597-610.
[25] C. Stauffer and W. E. L. Grimson, Adaptive background mixture models for real-time tracking, IEEE Conference on Computer Vision and Pattern Recognition, 2, (1999), 246-252.
[26] Abhishek Kumar Chauhan, Prashant Krishan, “Moving Object Tracking Using Gaussian Mixture Model And Optical Flow”, International Journal of Advanced Research in Computer Science and Software Engineering, April 2013
[27] https://pdfs.semanticscholar.org/25a6/c5dff9a7019475daa81cd5a7f1f2dcdb5cf1.pdf
[28] Hitesh A Patel, Darshak G Thakore,“Moving Object Tracking Using Kalman Filter”, International Journal of Computer Science and Mobile Computing,April 2013, pg.326 – 332.
[29] Jae-Yeong Lee; Wonpil Yu, "Visual tracking by partition-based histogram backprojection and maximum support criteria," Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on , vol., no., pp.2860,2865, 7-11 Dec. 2011
[30] Prasad, Dilip. (2012). Survey of The Problem of Object Detection In Real Images. International Journal of Image Processing (IJIP). 6. 441.
[31] A. Opelt, A. Pinz, and A. Zisserman, "Learning an alphabet of shape and appearance for multi-class object detection," International Journal of Computer Vision, vol. 80, pp. 16-44, 2008.