A New Approach For Tracking Object In Videos

Mohit Kumar Jindal[1], Er. Deepika Arora[2]
[1, 2] Dept of CSE
[1, 2] RPIIT, Bastara

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. Therefore, manually handling videos are quite impossible. So we need an automated devise to process these videos. Our work is related to the broad subject of automatic motion detection and analysis in video surveillance image sequence. Besides, proposing the new algorithm, some of the previous algorithms are analysed, where some of the approaches are noticeably complementary sometimes.*

In real time surveillance, detecting and tracking multiple objects and monitoring their activities in both outdoor and indoor environment are challenging task for the video surveillance system. Presence of a good number of real time problems limits scope of our work since the beginning. The problems are namely, illumination changes, moving background and shadow detection. An improved tracking algorithm has been proposed . The algorithm is applied on to a number of practical problems to observe whether it leads us to the expected solution. Test result shows that the proposed algorithm shows the better quality result.

Keywords- Object Detection, Background Subtraction, Gaussian mixture model, Optical flow, Foreground Detection, Median Filtering, Morphological operations, Object Detection, Object Tracking.

I. INTRODUCTION

Human quest for an automatic detection system of everyday occurrence lead to the necessity of inventing an intelligent surveillance system which will make lives easier as well as enable us to compete with tomorrows technology and on the other hand it pushes us to analyze the challenge of the automated video surveillance scenarios harder in view of the advanced artificial intelligence. Nowadays, it is seen that surveillance cameras are already prevalent in commercial establishments, with camera output being recorded to tapes that are either rewritten periodically or stored in video archives. To extract the maximum benefit from this recorded digital data, detect any moving object from the scene is needed without engaging any human eye to monitor things all the time.

1.1 Basic Steps for Tracking an Object

There are three basic steps for the object tracking :

1. Object Detection
2. Object Classification
3. Object Tracking

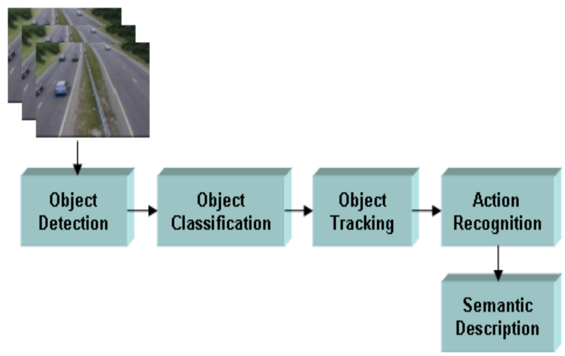


Fig 1: A generic frame work of video surveillance system

Object Detection

Object Detection is to identify objects of interest in the video sequence and to cluster pixels of these objects. Object detection can be done using various techniques like frame differencing, Background subtraction and Optical flow.

Object Classification

Object can be classified as birds, vehicles, floating clouds, swaying tree and other moving objects. The approaches to classify the objects are Motion-based classification, Shape-based classification, Color based classification and texture based classification.

Object Tracking

Tracking can be defined as the problem of approximating the path of an object in an image plane as it moves around a scene.

The methods to track the objects are point tracking, kernel tracking and silhouette. Here are some of the challenges that should be taken care in object tracking as described in:

1. Loss of evidence occurred by estimate of the 3D realm on a 2D image,
2. Noise in an image,
3. Difficult object motion,
4. Imperfect and entire object occlusions,
5. Complex objects structures.

II. OBJECT DETECTION

The detection of an object in video sequence plays a significant role in many applications specifically as video surveillance applications. The different types of object detection are shown in figure 2.

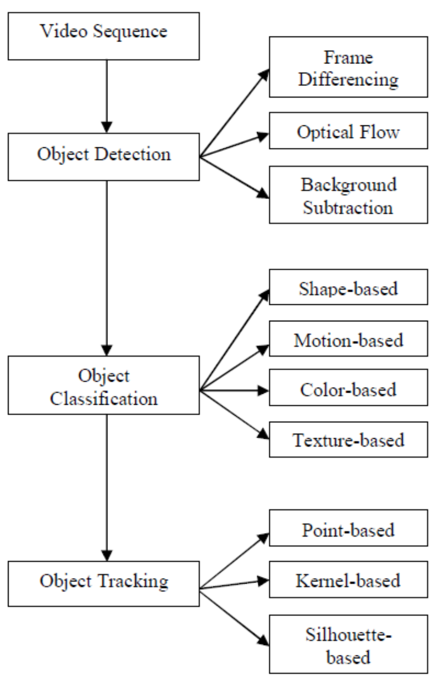


Fig 1: Basic steps for tracking an object [8]

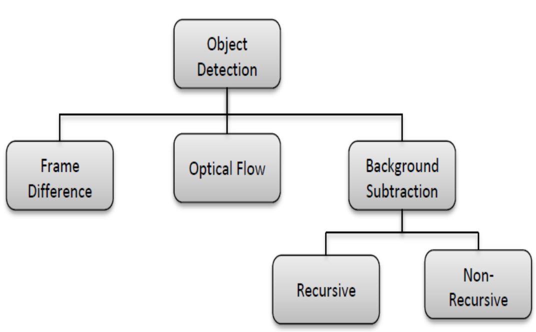


Fig 3: Types of Object Detection Method

Object detection includes detecting objects and recognizing patterns in the frame of a video sequence. An object detection mechanism is needed in any tracking method either in every frame or when the object first appears in the video. Using information in single frame is the most common method for object detection. Although some object detection methods use the temporal information computed from analyzing a sequence of frames in order to reduce the number of false detections and increase accuracy rate [26] Few methods of object detection are described as follows:

A. Frame differencing

The presence of moving objects is determined by calculating the difference between two consecutive images. Its calculation is simple and easy to implement. For a variety of dynamic environments, it has a strong adaptability, but it is generally difficult to obtain complete outline of moving object, responsible to appear the empty phenomenon, as a result the detection of moving object is not accurate [9].

B. 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.

C. Background subtraction

First step for background subtraction is background modeling. It is the core of background subtraction algorithm. Background Modeling must sensitive enough to recognize moving objects [1]. Background Modeling is to yield reference model. This reference model is used in background subtraction in which each video sequence is compared against the reference model to determine possible Variation. The variations between current video frames to that of the reference frame in terms of pixels signify existence of moving objects [12]. Currently, mean filter and median filter [2] are widely used to realize background modeling. The background subtraction method is to use the difference method of the current image and background image to detect moving objects, with simple algorithm, but very sensitive to the changes in the external environment and has poor anti- interference ability. However, it can provide the most complete object information in the case background is known. As describe in [1], background subtraction has mainly two approaches:

1. Recursive Algorithm

Recursive techniques [1] do not maintain a buffer for background estimation. Instead, they recursively update a single background model based on each input frame. As a result, input frames from distant past could have an effect on the current background model. Compared with non-recursive techniques, recursive techniques require less storage, but any error in the background model can linger for a much longer period of time. This technique includes various methods such as approximate median, adaptive background, Gaussian of mixture

2. Non-Recursive Algorithm

A non-recursive technique [16] uses a sliding-window approach for background estimation. It stores a buffer of the previous L video frames, and estimates the background image based on the temporal variation of each pixel within the buffer. Non-recursive techniques are highly adaptive as they do not depend on the history beyond those frames stored in the buffer. On the other hand, the storage requirement can be significant if a large buffer is needed to cope with slow-moving traffic.

III. OBJECT CLASSIFICATION

Different moving regions may correspond to different moving targets in natural scenes. To further track objects and analyze their behaviors, it is essential to correctly classify moving objects. The classification of objects [15] can be done under humans, vehicles, trees, birds, etc. For analyzing behavior of objects in the frame classifications among several objects is particular frame is needed. Classification approaches are based on shape, motion, color and texture.

A. Shape based classification

Different descriptions of shape information of motion regions such as points, boxes, silhouettes and blobs is available for classifying moving objects. Input features to the system of object classification 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. Classification is performed on each blob at every frame and results are kept in histogram [10]. These histograms sometimes become the basis for classification of objects for further processing.

B. Motion based classification

In general, non-rigid articulated human motion shows a periodic property, hence this has been used as a strong cue for classification of moving objects. Residual flow is used to analyze rigidity and periodicity of moving objects [10]. It is expected that rigid objects present little residual flow, whereas a non-rigid moving object such as a human being has a higher average residual flow and even display a periodic component. Based on this useful cue, human motion is distinguished from motion of other objects, such as vehicles, pedestrians or herds.

C. Color based classification

Unlike many other image features, color is relatively constant under viewpoint changes and it is easy to be acquired. Although color is not always appropriate as the solo means of detecting and tracking objects, but the low computational cost of the algorithms proposed makes color a desirable feature to exploit when appropriate. To detect and track vehicles or pedestrians in real-time, color histogram based technique is used. Color histograms have become extremely popular to describe a large image region [10]. Since it does not vary much to target translation, rotation, or the target scale variation, it has been used in many tracking applications. Color histogram describes the color distribution in a given region, which is robust against partial occlusions.

D. Texture based classification

Texture based technique counts the occurrences of gradient orientation in localized portions of an image, which is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy. The texture-based approaches also work similar to motion-based approaches but with the help of texture pattern recognition. It provides better accuracy, but may require more time, which can be improved using some fast techniques.

An average or normal trend of the techniques has been discussed to give an overall understanding of each category. However Table 2 below describes about comparative study made among different techniques used for classification of moving objects in visual surveillance.

Object Classification Method	Computational Time	Accuracy	Comments
Shape Based	Low	Moderate to high	Simple pattern-matching approach can be applied with appropriate templates. It does not work well in dynamic situations and is unable to determine internal movements well [12]
Motion Based	High	Moderate	Does not require predefined pattern templates but struggles to identify a non-moving human [12]
Color based	High	High	Provides improved quality with the expense of additional computation time [14]
Texture Based	High	High	It creates a Gaussian Mixture Model to describe the color distribution within the sequence of images & to segment the image into background and objects [14]

IV. OBJECT TRACKING

The object tracking is the term which used to identify the moving object position as well as tracking them from video sequences (Balasubramanian et al., 2014). The tracking method is classified into three types such as kernel, point and silhouette based tracking (Yilmaz et al., 2006). Compared to silhouette method, existing most of them have focused on kernel-based method due to high accuracy with less computational cost. However, the point tracking method has less computational cost with reduce in accuracy (Weng et al., 2013). The various types of object tracking techniques are shown in figure.

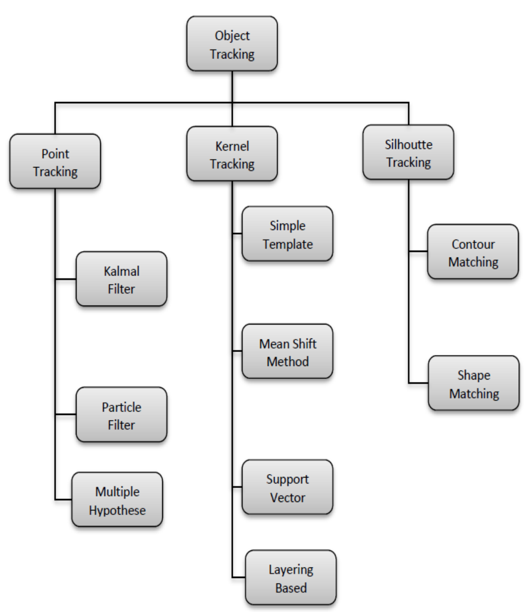


Figure 3 Types of object tracking approach

V. PROPOSED METHODOLOGY

GMM (Gaussian Mixture Model) and the Optical Flow Model are state of art algorithms for object tracking in video surveillance. In our dissertation work we have to combine the advantages of GMM and Optical Flow. Better understanding of object tracking system is essential for this purpose.

First of all have a brief review of the concept of GMM and Optical Flow

GMM (GAUSSIAN MIXTURE MODEL)

The Gaussian mixture model is a single extension of the Gaussian probability density function. As the GMM can approximate any smooth shape of the density distribution, so often used in image processing in recent years for good results. Assuming the Gaussian mixture model consists of and the combination of Gaussian probability density function, the Gaussian probability density function of each has its own mean, standard deviation, and weight, the weights can be interpreted by the corresponding Gaussian model of the frequency, they more often appear in the Gaussian model the higher the weight. The higher frequency of occurrence, then find the maximum weight on the Gaussian probability density function, Finally, the Gaussian probability density function of the means 3pixel value is background image.[20]

Optical Flow

Optical flow or optic flow is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer and the scene. The concept of optical flow was first studied in the 1940s and ultimately published by American psychologist James J. Gibsoas part of his theory of affordance. Optical flow techniques such as motion detection, object segmentation, time-to-collision and focus of expansion calculations, motion compensated encoding, and stereo disparity measurement utilize this motion of the objects' surfaces and edges. Sequence of ordered images allows the estimation of motion as either instantaneous image velocities or discrete image displacements. Fleet and Weiss provide a tutorial introduction to gradient based optical flow. John L. Barron, David J. Fleet, and Steven Beauchemin provide a performance analysis of a number of optical flow techniques. It emphasizes the accuracy and density of measurements. [29], [11] The optical flow method try to calculate the motion between two image frames which are taken at times t and t+ δt at every position. These methods are called differential since they are based on local Taylor Series approximation of the image signal; that is, they use partial derivatives with respect to the spatial and temporal coordinates. [30]

Proposed Methodology

Brief study and analysis of test results obtaind with GMM and Optical flow model reveals that in some cases GMM is better and in some cases Optical Flow is better. Taking the advantages and disadvantages of the both in consideration we can conclude that they are complementary to each other. This is the basic motivation behind our proposed methodology.

In our proposed methodology we have ensemble the results of GMM and Optical Flow Model which we abbreviate as EGOM (Ensemble of GMM and Optical Flow).

Working Methodology

Following are the basic steps in our dissertation work:

- Input: Captured with a fixed camera containing one or more moving objects of interest
- Processing goals: Determine the image regions where significant motion has occurred
- Obtain the results with GMM and Optical Flow Model
- Ensemble the results of GMM and Optical Flow Model.
- Output: An outline of the motion within the image sequence. Comparitive performance display of GMM , Optical Flow and proposed methodology (EGOM).

Flowchart of Proposed Methodology

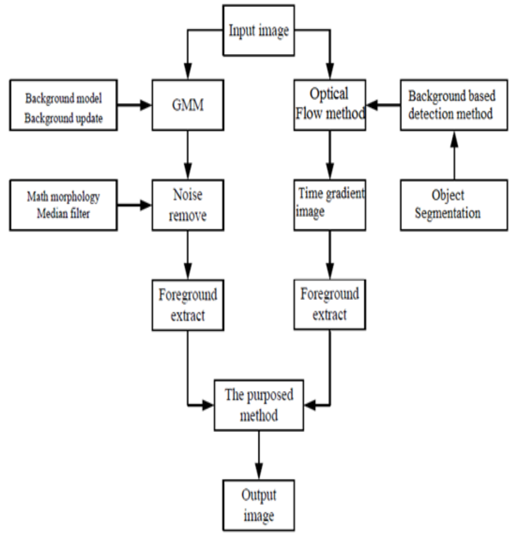


Figure 2: Flowchart of Proposed Methodology

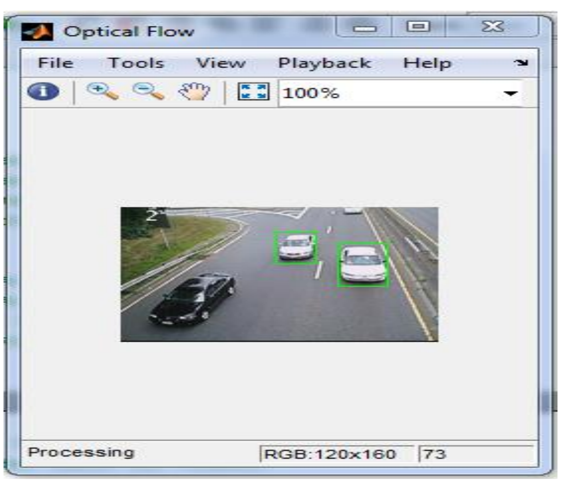
VI. SIMULATION AND RESULT ANALYSIS

In our work we have proposed EGOM (Ensemble of Gaussian Mixture Model and Optical Flow Model). We have analysed the performance of GMM (Gaussian Mixture Model) and Optical Flow Model for car tracking in videos. For the same video we have analysed the performance of our proposed methodology (EGOM).

As it is not possible to show the simulation results for the whole video in our dissertation, so we have selected certain video frames for this purpose. We have shown the simulation results for GMM (Gaussian Mixture Model), Optical Flow Model and proposed methodology (EGOM) for the same video frame. From the simulation results one can compare the performance of the three methodologies discussed in our work (GMM, Optical Flow and proposed (EGOM)).

Simulation results are shown for a given frame.
Results with proposed (EGOM) (Comparison):

1) Original Video Frame
2) Tracking Results GMM
3) Tracking Results Optical Flow
4) Tracking Results EGOM



Simulation Results Analysis

In the above pages simulation results are shown for GMM, Optical Flow Model and proposed (EGOM) for different frames. Detailed analysis of our results leads us to the following conclusions:

❖ In some cases performance of GMM is better than Optical flow while in rest of the cases performance of Optical flow is better than GMM. So we can conclude that GMM and Optical Flow Model are complementary to each other.

The performance of proposed EGOM is better than both the algorithms as clear from the simulation results. A Object detection and tracking is an important task in computer vision field. In this dissertation various phases of object tracking system viz. object detection, object classification and object tracking has been described. Methods available for these phases have been explained in details and a number of limitations were highlighted in every techniques.

VII .CONCLUSION

In our dissertation, we proposed a new algorithm EGOM (Ensemble Gaussian Mixture Model and Optical Flow model) for object tracking . In our dissertation we successfully combined GMM and Optical Flow object tracking. The advantages of Optical Flow are quick calculations and the disadvantage is a lack of complete object tracking. The advantage of GMM is complete results of the operation but disadvantage long computation time with more noise.

Detailed study of both the algorithms revealed that the two methods can complement each other. This is the basic guideline for our dissertation work which leads us to the idea of assembling the two models.

Both the models are applied for object tracking side by side. Assembling the result of the two algorithms is the main tracking step. The novel algorithm proposed in this paper to track an object in video sequence was successfully tested. The results show that the performance of our proposed algorithm is better as compared to the each single model.

VIII. FUTURE SCOPE

A In the future, we can extend the work to detect the moving object with non-staticbackground, having multiple cameras which can be used in real time surveillance applications. The testing platform was an important tool at the beginning of the design of the proposed algorithm as well as in the analysis of the tracking process in general, and it becomes an important instrument to study in future works .Testing the performance of proposed algorithm with phenomena such as occlusion, scale changes, illumination changes, etc in video object tracking could also offer challenges for future work.

In future this method can be modified to differentiate different class objects in real time video. Later characteristics are extracted and applied to a Neural Network so that segmented objects are classified as vehicles and non vehicles and, in the case of vehicles, they will be classified according to the size of the vehicle as follows: large size, intermediate size, small size. With the help of Data Mining we can include a real database that could be helpful in Traffic Management.

REFERENCES

[1] J.J. Koenderink and A.J. Van Doorn,"Representation of local geometry in the visual system," Biological Cybernectics,55(6), 1987.

[2] C.Harris; M.Stephens., "A combined corner and edge detector," 4th Alvey Conference, pages 147–151, 1988.

[3] A. J. Lipton, H. Fujiyoshi, R. S. Patil, Moving Target Classification And Tracking From Real-time Video, Applications of Computer Vision,1998. WACV '98. Proceedings., Fourth IEEE Workshop on, pp. 8-14, 1998.

[4] E. Grimson, C. Stauffer, R. Romano, and L. Lee, ªUsing Adaptive Tracking to Classify and Monitoring Activities in a Site, Proc. Computer Vision and Pattern Recognition Conf., pp. 22-29, 1998.

[5] G. L. Foresti, Object Recognition And Tracking For Remote Video Surveillance, IEEE Trans. Circuits Syst. Video Technol., 9(7):1045-1062, October 1999.

[6] C.Stauffer, W.E.L. Grimson. "Adaptive Bacckground Mixture Models for Real-Time Tracking," in Proc. Computer Vision and Pattern Recognition Conf., vol. 2, Fort Collins, CO. USA, June 1999,pp.246-252.

[7] E. Grimson and C. Stauffer, Adaptive Background Mixture Models for Real Time Tracking, Proc. Computer Vision and Pattern Recognition Conf., 1999. [3] Robert T. Collins, Alan J. Lipton, Takeo K.

[8] Y. Li, A. Goshtasby, and O. Garcia, Detecting and tracking human faces in videos, Proc. ICPR'00 vol. 1, pg. 807-810 (2000).

[9] Haritaoglu, I., Harwood, D., and Davis, L., "real-time surveillance of people and their activities," IEEE Trans. Patt. Analy. Mach. Intell. 22, 8, 2000.

[10] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," IEEE Proc. on Computer Vision and Pattern Recognition, pages 673–678, 2000.

[11] A. D. Bue, D. Comaniciu, V. Ramesh, and C. Regazzoni. Smart cameras with real-time video object generation. In Proceedings of the IEEE International Conference on Image Processing, volume 3, pages 429–432, June 2002

[12] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", Int'l Journal of Computer Vision, vol. 60, pp. 91-110, 2004.

[13] Sato, K. and Aggarwal, J., "Temporal spatio-velocity transform and its application to tracking and interaction," Comput. Vision Image Understand. 96, 2, 100–128,2004.

[14] Amer, A., "Voting-based simultaneous tracking of multiple video objects," IEEE Transaction on Circuits and Systems for Video Technology, 2005.

[15] Ashwani Aggarwal, Susmit Biswas, Sandeep Singh, Shamik Sural, and A.K. Majumdar, " Object Tracking. Using Background Subtraction and Motion Estimation in MPEG Videos", ACCV. 2006, LNCS, vol. 3852, pp. 121-130, Springer, Heidelberg (2006).

[16] Sebastian, P.;Yap Vool Voon,"Tracking using normalized cross correlation and color space,"International Conference on Intelligence and Advanced system, ICIAS, 2007.

[17] R. V. Babu, P. Perez, and P. Bouthemy, "Robust tracking with motion estimation and local kernelbased color modeling," Image and Vision Computing, vol. 25, issue 8, pp. 1205-1216, 2007.

[18] Zhi Liu; Liquan Shen; Zhongmin Han; Zhaoyang Zhang, "A Novel Video Object Tracking Approach Based on Kernel Density Estimation and Markov Random Field ," IEEE International Conference on image processing, 2007.

[19] Khan, Z.H., Gu, I.Y.-H., TieSheng Wang, Backhouse, A., "Joint anisotropic mean shift and consensus point feature correspondences for object tracking in video,"IEEE International Conference on Multimedia and Expo, 2009.

[20] Prakash Chockalingam. Non-rigid multi modal object tracking using Gaussian mixture model, The Graduate School of Clemson University, PhD thesis, 2009.
[21] Chakraborty, D.; Patra, D.,"Real time object tracking based on segmentation and Kernel based method," IEEE International Conference on Industrial and Information systems, 2010.
[22] Fuat Cogun and A. Enis Cetin "Object Tracking under Illumination Variations using 2D-Cepstrum Characteristics of the Target" IEEE 2010
[23] Mojtaba Asgarizadeh, Hossein Pourghassem, Ghazanfar Shahgholian, Hossein Soleimani "Robust and Real Time Object Tracking Using Regional Mutual Information in Surveillance and Reconnaissance Systems," IEEE 2011.
[24] Chun-Te Chu; Jenq-Neng Hwang; Hung-I Pai; Kung-Ming Lan, "Robust video object tracking based on multiple kernels with projected gradients," International Coference on Acoustics, Speech and Signal Processing (ICASSP), 2011.
[25] Chun-Te Chu, Jenq-Neng Hwang, Shen-Zheng Wang, Yi-Yuan Chen ,"Human tracking by adaptive Kalman filtering and multiple kernels tracking with projected gradients," IEEE International Conference on Distributed smart Cameras, 2011.
[26] Gabriel, P.; Hayet, J.-B.; Piater, J.; Verly, J., "Object Tracking using Color Interest Points," IEEE Conference on Advanced Video and Signal Based Surveillance, 2005. International Journal of Information Technology, Modeling and Computing (IJITMC) Vol.1, No.2, May 2013 .
[27] Khatoonabadi, S.H.; Bajic, I.V., "Video Object Tracking in the Compressed Domain Using Spatio-Temporal Markov Random Fields," IEEE Transaction on Image Processing, 2013.
[28] 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, Volume 3, Issue 4, April 2013
[29] Object Tracking: A Survey, Alper Yilmaz, Omar Javed, Mubarak Shah.
[30] Bhavana C. Bendale, Prof. Anil R. Karwankar. Moving Object Tracking in Video Using MATLAB, International Journal of Electronics, Communication & Soft Computing Science and Engineering.
[31] Qi Zang & Reinhard Klette. Parameter Analysis for Mixture of Gaussians Model, The university of Auckland.
[32] Yong-Beom Lee , Bum-Jae You, Seong-Whan Lee "A Real-time Color-based Object Tracking robust to Irregular Illumination Variations," IEEE, International Conference on Robotics & Automation Seoul
[33] Sheldon Xu and Anthony Chang "Robust Object Tracking Using Kalman Filters with Dynamic Covariance" Cornell University.