iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 Issue: IV Month of publication: April 2015

DOI:

www.ijraset.com Volume 3 Issue IV, April 2015
IC Value: 13.98 ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Video Detection and Tracking Using Extended Kalman Filter

Mr. Profun C.J[1], S.Kavitha[2]
[1]M.E, Assistant Professor, Department of ECE, , [2]PG Student (M.E), Department of ECE,
DMI college of Engineering

Abstract--* An efficient moving object segmentation algorithm suitable for real-time content-based multimedia communication systems is proposed in this paper. First, a background registration technique is used to construct a reliable background image from the accumulated frame difference information. The moving object region is then separated from the background region by comparing the current frame with the constructed background image. Finally, a post-processing step is applied on the obtained object mask to remove noise regions and to smooth the object boundary. Surveillance system can be used to detect and track the moving objects. First phase of the system is to detect the moving objects in the video and track the detected object. Second phase of the system detected different abnormal activities like crimes and robbery in ATM. In this paper, detection of the moving object has been done using simple background subtraction and tracking of single moving object has been done using Extended Kalman filter. Detection of abnormal activities can be done by using HOG (Histogram of Gradient) and IM (Illumination Mapping). The algorithm has been applied successfully on standard surveillance video datasets. The proposed method will uses multiple object detection method and event recognition techniques of computer vision.

***Index terms-* Extended kalman filter, K-means clustering, OTSU's Threshold, HOG, IM.**

I. INTRODUCTION

Video surveillance of human activity usually requires people to be tracked. It is important to security purpose and traffic control which is also used to take necessary step for avoids undesired interaction. We present our system for single moving object detection and tracking using a static webcam mounted inside a building that monitors a typical open work area. Since the camera is not in motion here, the inherent ambiguities of ego-motion are not considered and the scene structure can be rejected at the very beginning. The motivation behind this work is to develop software for tracking which has the major application in security, surveillance and vision analysis. This paper mainly focussed on developing an algorithm for tracking of an object.

An automated teller machine (ATM) is a computerized telecommunications device that provides the customers of a financial institution with access to financial transactions in a public space. With the use of an ATM, customers can access their bank accounts for cash withdrawals and check their account balances. Nowadays Automated Teller Machines is considered as very common technology for dispensing notes to cash-holders. Many ATM attacks seek to obtain a consumer's personal information, such as their card number and personal identification number (PIN). As the number of ATMs in use increases, so do the frequency and sophistication of security threats, making the development of fraud prevention measures a top priority for financial institutions (FIs) and ATM manufacturers. ATM fraud is not confined to particular regions of the world. To further complicate matters, perpetrators and victims are often on different continents, and the problems of one region can quickly become the problems of another.

All the evaluation has been performed on a windows PC running MATLAB R2010a. MATLAB has an Image Processing Toolbox for handling images and videos efficiently. Algorithms can be extended for real time applications. The implementation of this object tracking is explained step by step as follows. In section II, reviews the related works and brief overview of the system, and detailed explanations for each stage in Section III. In section IV, gives an introduction about proposed tracking algorithm. Section V reports the experimental results. The paper is concluded and future work explained in section VI.

II. RELATED WORK

There have been numerous tracking algorithms in the literature. Tracking can be casted as an optimization problem. Sum of Squared Difference (SSD) is widely used in many tracking algorithms, e.g., optical flow approaches[2] .An alternative cost function derived from the color histogram is utilized in the popular mean shift tracking[3].The model update, based on incremental algorithms for principal component analysis, includes two important features: a method for correctly updating the sample mean, and a for- getting factor to ensure less modeling power is expended fitting older observations. Both of these features contribute measurably to improving overall tracking performance[4]. Numerous experiments demonstrate the effectiveness of the proposed tracking algorithm in indoor and outdoor environments where the target objects undergo large changes in pose, scale, and illumination. Disadvantage of this approach is didn't directly address the partial occlusion problem and Cost increases when numbers of images are increased. As an essential component in all tracking methods, object representation recently gets more and more attention because it plays a key role in determining the tracking performance. A good representation should have strong description or discrimination power to distinguish the target from the background. To account for the appearance variations of the target during tracking, many sophisticated object representation methods have been developed, including both generative and discriminative methods.

For generative appearance modeling methods, Black et al. [6] propose a tracking algorithm based on the subspace constancy assumption. They construct a subspace model offline with some collected target observations, and keep the model fixed during tracking. However, for most tracking applications, it may be difficult to obtain such target observations in advance. As a result, this method has limited application domains and may fail when the target undergoes a different view from the ones used for constructing the model. Recently, many adaptive appearance models have been proposed for object tracking. Jepson et al. [7] propose a wavelet-based mixture model via an online Expectation–Maximization (EM) algorithm to account for appearance variations of the target during tracking. Incremental subspace methods based on Principal Component Analysis (PCA) or its variants have also been used for online object representation [8,9].

These two methods use target observations obtained online to learn a linear subspace for object representation. Since the appearance manifold of a target in a long-time interval may be quite nonlinear and complex, these models may not describe the appearance variations of the target well. Moreover, grossly corrupted observations often hurt the validity of the learned subspaces [10]. In addition, when an abrupt appearance variation of the target (due to changes such as lighting variation, pose variation or occlusion) occurs, the first few observations of the target will be identified as outliers and will not be fit into such linear appearance models. Therefore, these linear-subspace methods will run into the danger of losing the target as time progresses. Different from these subspace methods based on PCA, Ho et al. [11] propose a subspace learning method based on uniform L2 optimization. During tracking, they divide the most recent observations of the target into several batches and use the means of these batches for appearance modeling. However, this method only preserves the most recent appearance information of the target, and is sensitive to the chosen size of batches. When the batch size is not appropriate, the learned subspace model will be inaccurate and the resulting tracking may be in danger of drifting. For discriminative appearance modeling methods, Collins et al. [12] propose an online feature selection method to select the most discriminative color spaces for tracking. Avidan et al. [13] use online boosting method for tracking. They propose an ensemble tracking framework, in which a set of weak classifiers are trained to construct a strong classifier to distinguish the target from the background.

Parag et al. [14] also utilize the online boosting algorithm for tracking, but propose a different method for classifier updating. Babenko et al. [15] use Multiple Instance Learning (MIL) instead of traditional supervised learning to avoid the inaccuracy accumulation problem caused by self-learning. In these methods, tracking is usually treated as a binary classification problem. Different from the generative models which only model the target, the discriminative models can model both the target and background. Although these discriminative tracking methods have the capability to select good features for tracking, they need correctly labeled samples to train and update classifiers, which may not be available in many real tracking applications.

III. OVERVIEW OF THE SYSTEM

A. Object Tracking

Object tracking is the process of locating and following the moving object in sequence of video frames. Smart cameras are used as input sensors to record the video. Object tracking, the main application for security, surveillance and vision analysis, by using the code composer studio. In this, a video is recorded using digital camera. The recorded video frames are converted into individual frame and then converted into Portable Gray Map (PGM) format images. C language code is implemented to convert the pixel values in the PGM format image into Hexadecimal values and store the values in the files, which are inputs for code composer studio. Here the images in Hexadecimal format are imported into code composer studio. The objective of video tracking is to associate target objects in consecutive video frames. The association can be especially difficult when the objects are moving fast relative to the frame rate. Another situation that increases the complexity of the problem is when the tracked object changes orientation over time. For these situations video tracking [3] systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object. Associate a target objects in consecutive video frames is the objective of video tracking. It can be especially difficult when the objects are moving fast relative to the frame rate. when the tracked object changes orientation over time then it was increased a complexity. In these situations video tracking systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object.

B. Block Diagram

Video is taken as the input. The videos are converted into sequence of images. The images are stored in JPEG format. The image sequences are undergone with pre-processing techniques. P- Filter (Particle) is used to remove the blobs and to support the P-filter Kalman filter is used to for object tracking. The Kalman filter is a method of combining noisy (and possibly missing) measurements and predictions of the state of an object to achieve an estimate of its true current state based on position, velocity and size. The gray scale values are passing to the segmented image. To calculate the displacement of the pixel values square root of sigma is to be taken. Sigma is measured as gradient of the image and the impulse response coefficient. The segmented image is converted into gray foreground mapped image to found out the foreground and background pixels for object detection, because color image does not support the foreground mapping. The frame sequence contains number of objects and it is identified by high intensity values and it is compared with foreground object. Detection of moving object in video is based on the analysis of the 3D array of motion of the objects.

The moving object is detected by means of motion estimation to calculate the position of the moving object in the video plane. To detect the blocks containing moving object boundaries by using the information of the motion vector field. A new algorithm of moving objects detection and description is proposed to detect and track the moving object in video. Based on the analysis of projection of the 3D array motion of objects, the information of motion field is exploited to make moving object detection more efficient. The discontinuities of motion vector field on the boundaries of moving objects enable us to detect the moving objects blocks in which the potential boundaries of the moving objects locate. A further refinement of the boundary of moving objects and efficient descriptor for moving objects are proposed as well.

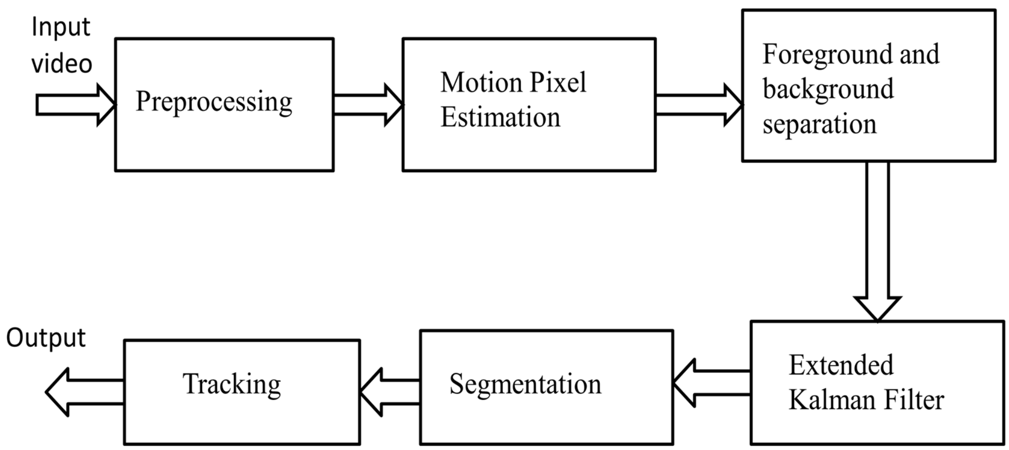


Figure 1: Block diagram of tracking system

C. Preprocessing

Extracting the points from an image that gives the best define of an object in an image namely key-points (High Intensity Pixels) is very important and valuable. These points have many applications in image processing like object detection. Before detecting the anomalies, pre-processing of the video is performed. The image sequences are extracted from the input video as in the form of frames. The three layered color (RGB) image is converted to the grey colored image in the frames.

D. Motion Pixel Estimation

Motion estimation is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence. It is an ill-posed problem as the motion is in three dimensions but the images are a projection of the 3D scene onto a 2D plane. Closely related to motion estimation is optical flow, where the vectors correspond to the perceived movement of pixels. In motion estimation an exact 1:1 correspondence of pixel positions is not a requirement. Applying the motion vectors to an image to synthesize the transformation to the next image is called motion compensation. The combination of motion estimation and motion compensation is a key part of video compression as used by MPEG 1, 2 and 4 as well as many other video codecs.

E. Algorithms

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

The methods for finding motion vectors can be categorised into

1) Pixel based methods ("direct")
2) Feature based methods ("indirect").

F. Direct Methods

1) Block-matching algorithm
2) Phase correlation and frequency domain methods
3) Pixel recursive algorithms
4) Optical flow

G. Indirect Methods

Indirect methods use features, such as corner detection, and match corresponding features between frames, usually with a statistical function applied over a local or global area. The purpose of the statistical function is to remove matches that do not correspond to the actual motion. A new algorithm of moving objects detection and description is proposed to detect and track the moving object in video. Based on the analysis of projection of the 3D array motion of objects, the information of motion field is exploited to make moving object detection more efficient. A further refinement of the boundary of moving objects and efficient descriptor for moving objects are proposed as well. As the ideal imaging model, perspective projection model is adopted in this proposed system. Calculate the first pixel value of the first frame named as p1 and the first pixel value of the second frame and named as p2.Find the mean value, add all the pixel values in all the frames. It is reasonable to assume the motion in same objects should be same.

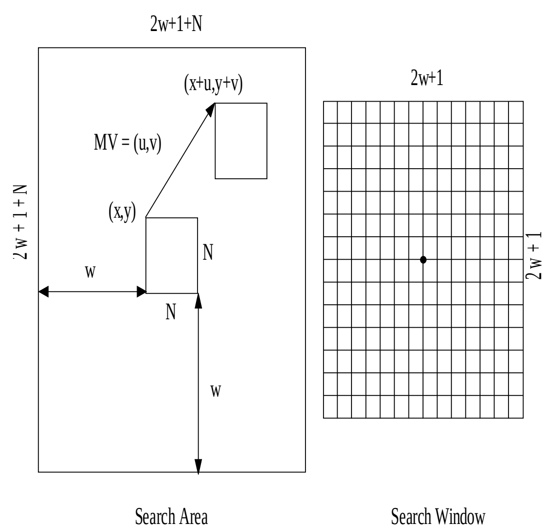


Figure 2 : Moving pixel compensation

H. Foreground and Background Separation

Background subtraction, also known as Foreground Detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further processing (object recognition etc.). Generally an image's regions of interest are objects (humans, cars, text etc.) in its foreground. After the stage of image pre-processing (which may include image denoising, post processing like morphology etc.) object localisation is required which may make use of this technique. Background subtraction is a widely used approach for detecting moving objects in videos from static cameras.

The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called “background image”, or “background model”. Background subtraction is mostly done if the image in question is a part of a video stream. Background subtraction provides important cues for numerous applications in computer vision,

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

I. Background Updating

A background occupancy map is computed on each cycle, and subtracted from the occupancy map to generate a foreground map, eliminating much of the distractions that tend to confuse a tracker. If the robot is stationary, the background is computed by taking an average over the past 10 frames, as in a fixed-viewpoint tracker.

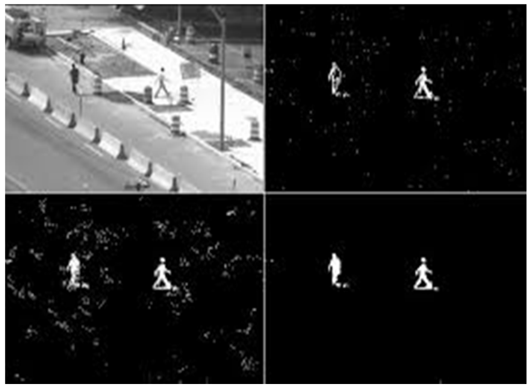


Figure 3 : Foreground and Background Separation

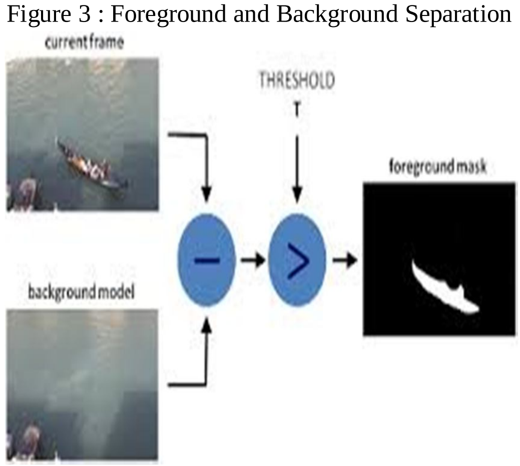


Figure 4 : Foreground mapping

J. Otsu's Thresholding

Otsu's method is used to automatically perform clustering-based image thresholding, or, the reduction of a graylevel image to a binary image. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal. The extension of the original method to multi-level thresholding is referred to as the Multi Otsu method.

K. Proposed Tracking Algorithm

1) *Kalman Filter:* Once the object has been detected then it can be tracked along its path. Many standard methods are available for object tracking Wiz. Kalman filter, Particle filter, Mean-shift based kernel tracking etc.

a) *Kalman filter tracking:* Kalman filter technique is used to estimate the state of a linear system where state is assumed to be distributed by a Gaussian published his famous paper describing a recursive solution to the discrete-data linear filtering problem. Object tracking is performed by predicting the object's position from the previous information and verifying the existence of the object at the predicted position. Secondly, the observed likelihood function and motion model must be learnt by some sample of image sequences before tracking is performed.

The equations for Kalman filters fall in two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimate for the next time step. The measurement update equations are responsible for the feedback. That is used for

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate. The time update equations can also be thought of as predictor equations. The measurement update adjusts the projected estimate by an actual measurement at that time. Tracking of moving object has been done using Kalman filter. Here tracking of any object can be done by providing the frame number from which tracking has to be started. From the selected frame any object can be picked for tracking by setting the position of the mask and then the object can be tracked in subsequent frames. Following steps have been implemented for tracking a single object. Background frame has been calculated by taking average of all the pixels. Frame number has been selected from which tracking of any object has to be started. From selected frame object to be tracked has been selected by repositioning the mask. For selected object its centroid position has been found out and from centroid information all the equation of time and measurement update have been calculated. For selected frame the actual position X and error P has been calculated.

b) *Detection Of Anamoly Activities:* ATM surveillance videos capture the behavioural activities of the objects accessing the ATM system. Some behaviour is frequent sequence of events and some deviate from the known frequent sequences of events. These events are termed as anomalies and may be susceptible to criminal activities. In the past, work was based on discovering the known abnormal events. Here, the unknown abnormal activities are to be detected and alerted such that early actions are taken.

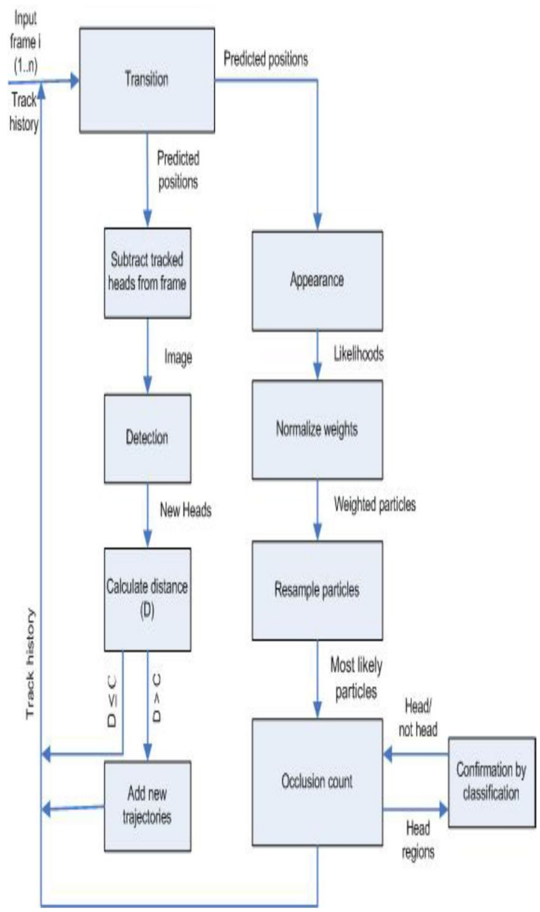


Figure 5 : Block diagram of Kalman filter

c) *HOG(Histogram of Gradient):* We compute Histograms of Oriented Gradients (HOG) to describe the distribution of the selected edge points. HOG is based on normalized local histograms of image gradient orientations in a dense grid. The HOG descriptor is constructed around each of the edge points. The neighborhood of such an edge point is called a cell. Each cell provides a local 1-D histogram of quantized gradient directions using all cell pixels. To construct the feature vector, the histograms of all cells within a spatially larger region are combined and contrast-normalized.

Volume 3 Issue IV, April 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

d) *IM(Illumination Mapping):* Estimation of Illumination from a single image or an image sequence has been widely studied in computer vision. The approach presented in this paper, introduces two new issues: 1) Illumination classification is performed rather than illumination estimation. 2) An object based approach is used for Illumination Evaluation. The input image is segmented into homogeneous regions. Per illuminant estimator, a new image is created where each region is colored with the extracted illuminant color. This resulting intermediate representation is called illuminant map (IM).

IV. SIMULATION RESULTS

The experimental results are shown in Fig. 7, 8,and 9.As our primary aim to detect and track abnormal crowd activities in the video sequences.

Figure 6: Input Frame

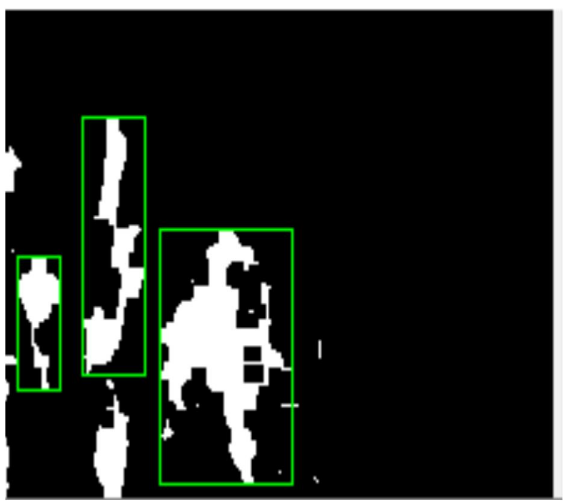


Figure 7: Background Subtracted Frame

e 

Figure 8: Moving Pixel Estimation For Crowd People

Figure 9: Tracking For Abnormal Crowd

Abnormal person and crowd are first detected in the input video. Then background subtraction processes are carried out. After that, abnormal crowd is tracked efficiently with our proposed method. Corresponding results are shown in the Figures 6, 7and 8. Each pixel in a background subtraction method's classification was determined to be: true positive (TN) for a correctly classified foreground pixel, false positive (FP) for a background pixel that was incorrectly classified as foreground, true negative (TN) for a correctly classified background pixel, and false negative (FN) for a foreground pixel that was incorrectly classified as background. By the calculation of TP, TN, FP, and FN the different methods can be evaluated. Then the sensitivity, the specificity and accuracy was calculated .Our experimental results show that the proposed method is more efficient and robust for abnormal crowd detection and tracking**.**

V. CONCLUSION

The proposed idea shows that automatically obtained person density estimates can be used to improve person localization and tracking performance in crowded scenes. It has formulated the person detection task as a minimization of a joint energy function incorporating scores of individual detections, pair-wise non-overlap constraints, and constraints imposed by the estimated person density over the scene, and also have demonstrated significant gains in detection and tracking performance on challenging videos of frames with varying density. Currently, video frames are processed individually and obtained detections are tracked in post-processing. There are hundreds of cameras deployed in public places (eg. Streets, shopping malls, airports). However, only a small fraction of the cameras outputs (<5%) is regularly observed by humans. Most of the video information is stored or destroyed without being watched or processed. This means that current surveillance systems are not able to detect abnormal events in real time. A large effort has been recently done to develop algorithms to track objects (persons, vehicles) in video sequences and to recognize activities and unusual behaviours. In the future work the gradient based algorithm is used for activity based tracking which is more efficient than the other methods. In addition to this work a accuracy of the tracking can be calculated by using precision and recall factor.

VI. FUTURE SCOPE

To detect a abnormal activities by using DSP processor is more effective in future and also GSM has been used.

REFERENCES

[1] Arulampalam.M.S, Maskell.S, Gordon. N , and Clapp.T, [2002], "particle filters for online nonlinear/non-Gaussian Bayesian tracking-tutorial," IEEE Trans Signal Process., vol. 50, no. 2, pp. 174–188.

[2] Babenko.B , Yang M.H , and Belongie.S , [2011], "Robust object tracking with online mulitiple instance learning," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 8, pp. 1619–1632.

[3]Bai and Li .Y.F, [2012], "Robust visual tracking with structured sparse representation appearance model," Pattern Recognit., vol. 45, no. 6, pp. 2390–2404.

[4] Bjoerck and Golub.G.G, [1973], "Numerical methods for computing angles between linear subspaces," Math.Comp., vol. 27, no. 123, pp. 579–594.

[5] Brand [2002], "Incremental singular value decomposition of uncertain data with missing values," in Proc. 7th Eur. Conf. Comput. Vis., Copenhagen, Demark, May 27–June 2, pp. 707–720.

[6]Chen and Li.Y.F , [2009], "Enhanced particles with pseudolikelihoods for three-dimensional tracking," IEEE Trans. Ind. Electron., vol. 56, no. 8, pp. 2992–2997.

[7] Chen,Wang.Q , Wang.S, Zhang.W, and Xu.W, [2011], "Object tracking via appearance modeling and sparse representation," Image Vis. Comput., vol. 29, no. 11, pp. 787–796.

[8] Comaniciu, Ramesh.V and Meer.P, [2003], "Kernel-based object tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 5, pp. 564–577.

[9] Eldar, Kuppinger.P and Bolcskei.H, [2010], "Block-sparse signals: Uncertainty relations and efficient recovery," IEEE Trans. Sig. Proc., vol. 58, no. 6, pp. 3042–3054.

[10] Everingham, Van Gool.L , Williams C.K.I , Winn.J, and Zisserman.A, [2009], "The pascal visual object classes (VOC) challenge," Int. J. Comput. Vis., vol. 88, no. 2, pp. 303–308.

[11]Hall,Marshall.D, and Martin.R, [2000], "Merging and splitting eigenspace models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 9, pp. 1042–1049.

[12] Han, Jiao.J, Zhang.B, Ye.Q, and Liu.J ,[2011], "Visual object tracking via sample-based adaptive sparse representation (AdaSR)," Pattern Recognit., vol. 44, no. 9, pp. 2170–218.

[13] Kwak, W. Nam, B. Han, and J. Han, [2011], "Learning occlusion with likelihoods for visual tracking," in Proc. IEEE Conf. Comput. Vis., Barcelona, Spain, Nov. 6–13.

[14] Kwon and K. M. Lee, [2010], "Visual tracking decomposition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., San Francisco, CA, USA, June 13–18, pp. 1269–1276.

[15] Lee, J. Ho, M. H. Yang, and D. Kriegman, [2005], "Visual tracking and recognition using probabilistic appearance manifolds," Comput. Vis. Image Understanding, vol. 99, no. 3, pp. 303–331.

[16] J. Lim, R. S. Lin, and M.-H. Yang, [2008], "Incremental learningfor robust visual tracking," Int. J. Comput. Vis., vol. 77, no. 1–3, pp. 125–141.

[17] Levy and M. Lindenbaum, [2000], "Sequential Karhunen-Loeve basis extraction and its application to images," IEEE Trans. Image Process., vol. 9, no. 8, pp. 1371–1374.

[18]Liu, L. Yang, J. Huang, P.Meer, L. Gong, and C.Kulikowski, [2010], "Robust and fast collaborative tracking with two stage sparse optimization," in Proc. 11th Eur. Conf. Comput. Vis., Hersonissos, Greece, Sep. 5–11, pp. 624–637.

[19] Mei and H. Ling, [2011] "Robust visual tracking and vehicle classification via sparse representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 11, pp. 2259–2272.

[20]Stalder, H. Grabner, and L. Van Gool, [2010], "Cascaded confidence filtering for improved tracking-by-detection," in Proc. 11th Eur. Conf. Computer Vision, Hersonissos, Greece, Sep. 5–11, pp. 624–637.

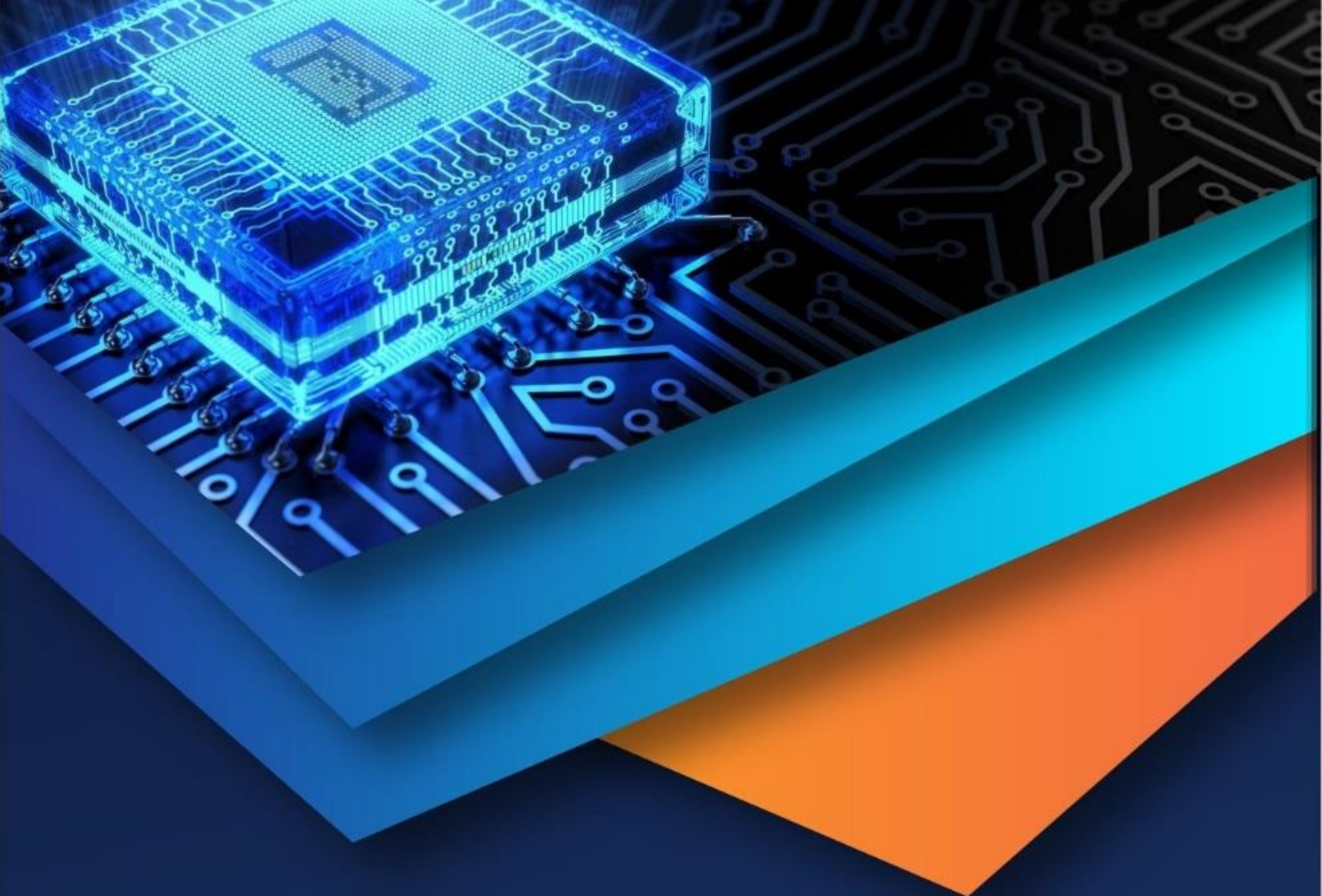
[21]Tran and M. M. Trivedi,[2012] "3-D posture and gesture recognition for interactivity in smart space," IEEE Trans. Ind. Inf., vol. 8, no. 1, pp. 178–187.

[22] Yang, Z. Fan, J. Fan, and Y. Wu, [2009], "Tracking nonstationary visual appearances by data-driven adaptation," IEEE Trans. Image Process., vol. 18, no. 7, pp. 1633–1644.

[23]Yu, T. B. Dinh, and G. Medioni, [2008], "Online tracking and reacquisition using co-trained generative and discriminative trackers," in Proc. 10th Eur. Conf. Comput. Vis.,Marseille, France, Oct. 12–18, 2008, pp. 678–691.

[24]Zhang, B. Ghanem, S. Liu, and N. Ahuja, [2012], "Robust visual tracking via multitask sparse learning," in Proc. IEEE Conf. Comput. Vis., Providence, RI, USA, Nov. 6–13, pp. 2042–2049.

[25]Zhou, Y. F. Li, and B. He, [2013] "Game-theoretical occlusion handling for multitarget visual tracking," Pattern Recognit., vol. 46, no. 10, pp. 2670–2684, Oct.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY