International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-8, June, 2019

Identification of an Object in an Image using Frame Differencing, Optical Flow and Support Vector Machine

Vikash Yadav, Rati Shukla, Vandana Dixit Kaushik

Abstract: The main aim of Identification of any object is to trace a region of interest from an image and locate its position and movement. This paper introduces a novel technique for object identification using frame differencing for background subtraction, optical flow for feature extraction and Support Vector Machine classifier for classification. Existing techniques are capable of identifying the object but unable to handle the situations of cluster of pixels in the given images. To overcome this drawback, new frameworks have been developed for object identification and have maximum accuracy. This framework is very helpful in surveillance system and security system.

Index Terms: Optical flow, Position, SVM, Object classification

I. INTRODUCTION

Video is the collection of images .Each image is called frames. In video, each frame is displayed so quickly so that our eye can feel the continuity of the content available in it. So we can also apply image processing techniques for every frame. In the field of computer vision, object identification is very popular now-a-days. This technology can be used in fields like video surveillance, traffic monitoring systems, wildlife and robotics. For example it is also used in CCTV monitoring system to identify objects and its behaviour. They are also used to differentiate between living and non-living objects in CCTV systems. Traffic monitoring system is also a good example of object identification system. They are used to identify abnormal behaviour of various vehicles and to issue challan by identifying those vehicles. Object identification can be classified into three steps. A first step is to detect the object you are going to identify. We can identify an object by using image processing techniques like frame differencing, optical flow and background subtraction. A second step consists of object classification. After detection we can classify an object into pedestrians, cars, flying birds and other moving object. This classification can be done by using shape-based, colour-based, motion-based and texture-based classification techniques. Third and final step is the identification of that object. Object identification can be done by using point-based, kernel based and silhouette-based identification techniques. Literature survey of different papers has been explained in next Section, Third Section describes the basics of optical flow, fourth Section describes the proposed work. Fifth section contains experimental work and results and finally sixth section contains conclusion.

Revised Manuscript Received on June 05, 2019

Vikash Yadav, Computer Science & Engineering Department, ABES Engineering College, Ghaziabad, India.
Rati Shukla, GIS Cell, MNNIT Allahabad, Prayagraj, India.
Vandana Dixit Kaushik, Computer Science & Engineering Department HBTU Kanpur, India.

II. LITERATURE SURVEY

Bhavna et al. describes a method for identifying moving object from a sequence of video frame [1]. They used the optical flow method in MATLAB simulink. This technology is very useful in security and surveillance, video communication, medical imaging, traffic control and video compression. Optical flow is the observations of the motions of the objects in a video obtained by the relative motion observed between the scene and the observer [2] [3]. Their techniques was based on the simulink model, contains three steps: estimation of the velocity, calculation of threshold and calculation of boundary box of the objects. The prediction of the motion of the objects is done using optical flow with the help of lucaskanade. Afterwards calculation of velocity threshold has done relative to moving object in the scene. Later, calculation has done for the boundary box of the object and its box area. In their work, boundary box was greater than the size of the object to be identified. Peng et al. incorporates hint information as a new component to a object identification framework [4].Hint information is used to train the classifier for the creation of highly efficient knowledge dataset. The framework improves the performance as compared to typical detection of object systems. They proposed the identification of object in the sequence of the frames [5].Some researchers worked on detection of face and cars [6] [7], while some takes more general approach [8].They used the facial features and viola & Jones technique to identify faces in each frame [6]. They extract the frames not having object of interest from the identification of object results. Dynamic programming is used to filter out important frames from the remaining input frames. Finally, identification of object is done by applying the classifier on the database. At this point, they used hint information to filter out the database. The feature extraction is done with the help of Local Binary Pattern (LBP) [9] and classification using K-Nearest-Neighbour (KNN) [10] to identify all frames from the database. In the object recognition phase, they extract LBP feature vector from every frame in the sequence.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

Then they compare the feature vector of every face with the feature vectors of all faces in the training data to find the nearest neighbour by a certain measure. If measure passes a threshold value, assign this face with the closest face else not recognized. Then they applied general methods for object recognition. They get the improved result as compared to framework without using hint information to their framework with hint information.

Viola et al. describes a face detection framework which performs so fast as compared to existing framework [6]. They used the three key methods for face detection. Firstly, they used the new feature called Integral Image to make the detection very fast. Secondly, they built the classifier using Adaboost learning algorithm [11] to choose the small subset of features from a huge set of features. Thirdly, they used the "Cascaded" method by processing background subtraction or foreground detection of the image so quickly. These all features make the framework so fast as compared to existing framework.

Chen et al. introduced a system which detects moving vehicles in night-time traffic scenes [12]. Their method used the pattern analysis and image segmentation, identifies vehicles by vehicles headlights and taillights. They extract the bright objects by applying multilevel thresholding that works on night-time illumination conditions. They were able to recognize motorbikes and moving cars in dense traffic image by using spatial clustering and tracking procedures from extracted bright objects. They were also incorporates this valuable technique to urban roads and traffic surveillance systems. Experimental Result shows that the proposed system was very robust and efficient for identification of vehicles in night-time environments.

Naik et al. proposed a system for the security purposes of the society [13]. Many criminal cases were solved by use of camera. To handle unauthorized activity, so many cameras are used to detect object. They proposed a distributive tracking system. They introduce a feature "field of view (FOV)", which exist on each camera. The whole system consists of "camera processing modules (CPM)". This combination of CPM makes central module, which was used to maintain coordination between the tracking tasks among the different cameras. They used the training and testing phases. Training phase allows a system to learn the relation between FOV of each camera. Testing phase belongs to the initial tracking and detection, which is performed at the single camera level. Experimentally there proposed method gives good result and easily detect multiple people or object using multiple cameras.

In this paper, they used the amalgamation of Gaussian Mixture Model (GMM) and Optical flow method [14]. The above two mentioned methods are totally opposite with respect to advantage and disadvantage. The GMM gives the complete result of operations but also gives incomplete object tracking whereas Optical flow gives the quick calculations but gives complete object tracking. These two methods combine to gives the quick and complete object tracking as compared to existing framework. This new framework can be applied to many areas such as video communication, medical imaging and video editing.

Kuralkar et al. presents an algorithm for moving object detection from a static background scene which is having shadows [15]. They take the reference frame as background model so that the foreground and background information may identified. They also solved the problem of object detection when merging of foreground object and background information takes place. Merging makes the task more complex. They also remove the shadow of object by performing the subtraction of the image from background image. Then the resultant image was converted into gray level. If this gray level image content having the value greater than threshold value then they were filled with holes in binary image. They found the removal of shadow by this technique; their algorithm gives the better result when compared to existing algorithms and efficiently removes the shadow of human object.

Zhang et al. presented an algorithm for tracking based on adaptive background subtraction from video and image or sequence of frames [16]. This paper used median filter to get the background image of the video and then remove the noise from that image. Then they used adaptive background subtraction method to detect and track the moving object in the video. They also improved the detection and tracking of the video. They showed that adaptive background subtraction is very useful in tracing moving objects from simulation in MATLAB. They also found that there algorithm runs more quickly than existing algorithm.

Ching et al. identified that the moving object tracking in a sequence of frame is very difficult task [17]. Background subtraction is very common approach in tracking moving object. Background subtraction is a method in the field of image processing and computer vision, to extract the image's foreground for further processing. They divide the background subtraction technique into four steps. These four steps include preprocessing, Background modeling, foreground detection and data validation. The preprocessing step is used to remove the noise and change the video in the required format. The background modeling is used for any background subtraction algorithm. They divide the background modeling into recursive and non-recursive. In recursive technique, they used Approximated Median Filter, Kalman Filter and Mixture of Gaussian. In non-recursive technique, they used Frame Differencing, Median Filter, Linear Predictive Filter and Non-parametric model. Further in foreground detection, they compare the input video frame with the background model and identify foreground pixels from the input frame. Data validation is the process of enhancing the candidate foreground mask on the basis of the information gained from outside the background model. After that they concluded that mixture of Gaussians produces the best result and adaptive median filtering is the simple alternative with competitive performance.

III. BASICS OF OPTICAL FLOW

The Sequence of ordered images allows the estimation of motion as either instantaneous image velocities or discrete image displacements [4]. It focuses on depth of measurements and accuracy.

The optical flow is very efficient method to compute the motion between two consecutive frames in the video at time *t* and t+Δt. They are based on Taylor series with partial derivatives. The brightness constancy constraint can be given:

$$I(x, y, t) = I(x+\Delta x, y+\Delta y, t+\Delta t) \quad (1)$$

Applying Taylor series to equation (1)

$$I(x+\Delta x, y+\Delta y, t+\Delta t) = I(x ,y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t + \text{Higher terms} \quad (2)$$

From above equations, we have:

$$\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t = 0 \quad (3)$$

or

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t}\frac{\Delta t}{\Delta t} = 0 \quad (4)$$

Which give:

$$\frac{\partial I}{\partial x}V_x + \frac{\partial I}{\partial y}V_y + \frac{\partial I}{\partial t} = 0 \quad (5)$$

Where V_x, V_y are velocity components

Optical flow of I(x, y, t) and $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$ are the derivatives of the image at I(x, y, t) in the directions. $I_x, I_y \text{ and } I_t$ can be written:

Therefore:

$$I_x . V_x + I_y . V_y = -I_t \quad (6)$$

or

$$\nabla I^T . \vec{V} = -I_t \quad (7)$$

Equations having unknown, which cannot be easily solved. This is the example of optical flow technique. More equations are needed to solve the optical flow. All optical flow methods require some more information.

IV. PROPOSED SYSTEM

The proposed system consists of training and testing segment. The different sections of the system have been explained below:

(A) Background Subtraction: Background subtraction is very important step in every detection systems. Background subtraction reduces the dimensions of the images and hence able to maximize the performance of the results. Frame differencing technique has been used for this purpose. First we have to take an image background and frame. The background is denoted by A(t) and image denoted by B. We follow simple method of subtraction to get the foreground of the images. It is represented as:

$$P[F(t)] = P[A(t)] – P[B] \quad (8)$$

This technique is used for dynamic foreground pixels and static background pixels.

(B) Feature Extraction: Feature extraction is used to find out the local features from an image. Edge detection and colour detection is done with the help of canny edge detector, optical flow with SVM respectively. The canny edge detector has two threshold values bel (low) and abo (high) [11]. The adaptive threshold also used in various images called Tada (adaptive). Adaptive value can be formed by various formulas given below for the image p(x,y). The T_i moment of image is given by:

$$T_i = (1/n) \sum_j Nj(Zj)i \qquad i=1, 2, 3, 4 \quad (9)$$

Different moments are calculated with the help of above equations. After that lower threshold bel and higher threshold abo can be calculated easily.

The colour classification can be done by using work by Tsai et. al.[7]. They proposed very efficient technique for classification with the help of colour. Their proposed system first converts the RGB into colour.

$$U_p = (2Zp-B_p-G_p) / Z_p \quad (10)$$

$$V_p = Max\ ((B_p-G_p) / Z_p) , (R_p-Z_p) / B_p)) \quad (11)$$

Where Rp, Bp, Gp is the colour component of the pixel p and Zp is the mean of all colour components. It also known that colour of vehicle is easily separated from the colour of the image. They are first converted into UV colour combination then optical flow with SVM is applied to differentiate between vehicle colour and non-vehicle colour.

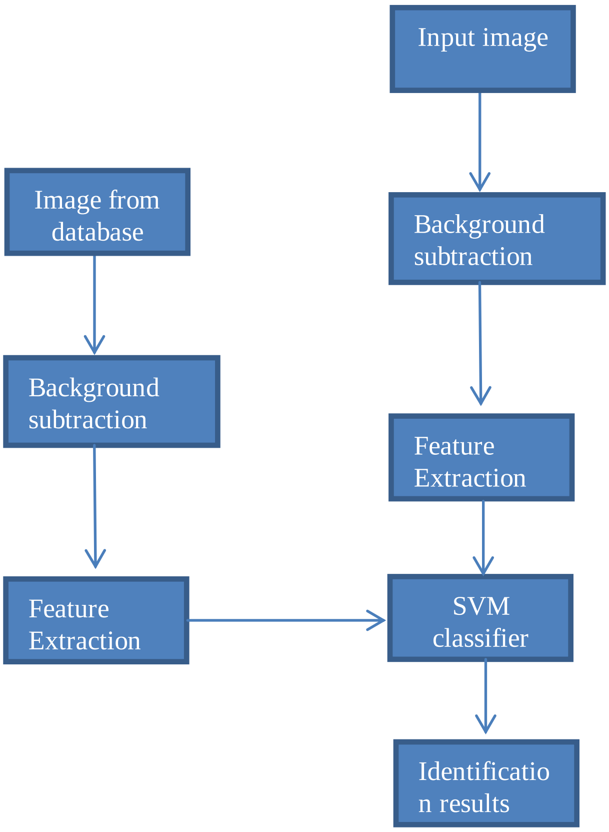

Figure 1: Framework for object identification

(C) Vehicle detection using SVM: Support Vector Machine is used to depict the model in point space where region is divided into clear gap which is wide as possible into two categories: true or false, either it is true or false depending on the condition. This new points are mapped into this trained SVM and predicted to belong to a category based on which side of the gap they fall on. The feature extracted images are used to train the SVM and later on remaining images are used to detect the vehicle.

V. EXPERIMENTAL RESULTS

To analyse the results, images are selected from different scenes and from various altitudes and angle have been selected. Frame differencing technique has been used in this framework for background subtraction, which is more efficient technique used so far.

The result has been developed using removal of highest colours from the image and SVM used to differentiate between colours. The result from the background subtraction step is shown in figure 2 and 3.

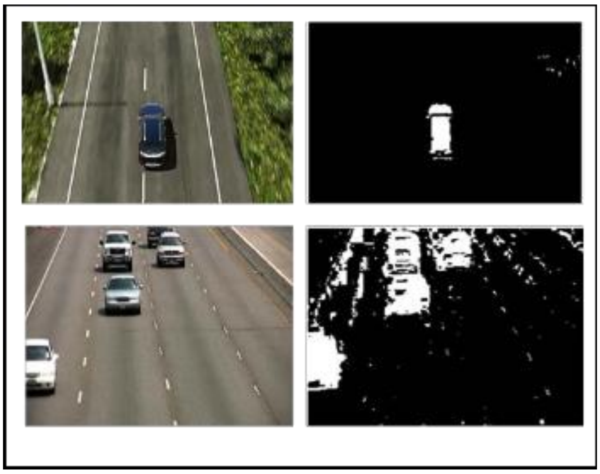

Figure 2: Background subtraction

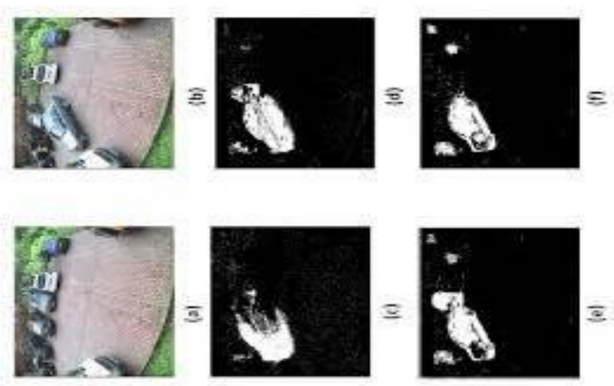

Figure 3: Background subtraction

The results using SVM is shown in figure 4 and 5.

Figure 4: Car detection in an image

Figure 5: Detection of more than one car in an image

VI. CONCLUSION

This paper proposes an efficient framework for detection of car in the traffic which can't depends on aerial distance, size of the car and aspect ratio. We have used frame differencing, Optical Flow and Support Vector Machine to detect efficiently car in the running traffic. The use of these techniques enhanced the performance of the detection. The experimental results show the efficient and robust nature of framework and works well despite of aerial height, color and dense traffic.

REFERENCES

1. Bhavana C. Bendale, Prof. Anil R. Karwankar," Moving Object Tracking in Video Using MATLAB" International Journal of Electronics, Communication & Soft Computing Science and Engineering, ISSN: 2277-9477, Volume 2, Issue 1.
   Object Tracking: A Survey, Alper Yilmaz, Omar Javed, Mubarak Shah
2. Brown, L. M. View independent vehicle/person classification. Technical report, Proceedings of the ACM 2nd international workshop on Video surveillance & sensor networks, pages 114-123, 2004.
3. Liang Peng, Yimin Yang, Xiaojun Qi, and Haohong Wang, "HIGHLY ACCURATE VIDEO OBJECT IDENTIFICATION UTILIZING HINT INFORMATION"
4. Rainer Lienhart and Jochen Maydt, "An extended set of haar-like features for rapid object detection," in International Conference on Image Processing (ICIP). IEEE, 2002, vol. 1, pp. 900–903.
5. Paul Viola and Michael J Jones, "Robust real-time face detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137–154, 2004.
6. Cheng-Hao Kuo and Ramakant Nevatia, "Robust multi-view car detection using unsupervised sub-categorization," in Workshop on Applications of Computer Vision (WACV). IEEE, 2009, pp. 1–8.
7. Hae Jong Seo and Peyman Milanfar, "Training-free, generic object detection using locally adaptive regression kernels," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, pp. 1688–1704, 2010.
8. Timo Ahonen, Abdenour Hadid, and Matti Pietikainen, "Face description with local binary patterns: Application to face recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037–2041, 2006.
9. Daniel T Larose, "k-nearest neighbor algorithm," Discovering Knowledge in Data: An Introduction to Data Mining, pp. 90–106, 2005.
10. Freund, Y. and Schapire, RE 1995. A decision-theoretic generalization of on-line learning and an application to boosting. In Computational Learning Theory: Eurocolt 95,Springer-Verlag, pp.23-37.

11. Y. Chen, B.Wu, H. Huang, and C. Fan, "A real-time vision system for nighttime vehicle detection and traffic surveillance," IEEE Trans. Ind. Electron., vol. 58, no. 5, pp. 2030–2044, May 2011.
12. R. Naik, "Object Detection from Multiple Cameras View for Smart Video Surveillance: A Review" e-Reflection vol. 2, issue 2, ISSN 2278-120X, Mar-April 2013.
13. A. K. 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.
14. P.P.kuralkar, V.T. Gaikwad, "Human object tracking using Background Substraction and Shadow Removal Techniques", International journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 3,March 2012.
15. Ruolin Zhang, Jian Ding, "Object Tracking and Detecting Based on Adaptive Background Subtraction", International Workshop on Information and Electronics Engineering, 2012, 1351-1355.
16. Sen-Ching S. Cheung and Chandrika Kamath, "Robust techniques for background subtraction in urban traffic video".

AUTHORS PROFILE

Dr. Vikash Yadav received his B.Tech (Computer Science & Engineering) degree in 2009 from Dr. Ambedkar Institute of Technology for Handicapped, Kanpur (U.P. India), M.Tech (Software Engineering) degree in 2013 from Motilal Nehru National Institute of Technology, Allahabad (U.P. India) and Ph.D (Computer Science & Engineering) degree from Dr. A.P.J Abdul Kalam University (Formerly U. P. Technical University) Lucknow, (U.P. India) in 2017. He is currently working as an Assistant Professor in the Department of Computer Science & Engineering, ABES Engineering College, Ghaziabad, India.

Rati Shukla received her B.Sc in 2006, MCA degree in 2009 from U.P. Technical University Lucknow and M.Tech (Computer Science and Engineering) degree in 2014 from Motilal Nehru National Institute of Technology, Allahabad (U.P. India). She is pursuing her Ph.D at GIS Cell Motilal Nehru National Institute of Technology, Allahabad (U.P. India).

Vandana Dixit Kaushik received her B.E (Computer Science) degree in 2000 from Pune University and Ph.D (Computer Science & Engineering) in 2011 from U. P. Technical University, Lucknow. She is currently working as an Assistant Professor in the Department of Computer Science & Engineering, Harcourt Butler Technical University, Kanpur and more than 16 years of teaching experience.