International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Effective Video Saliency Mapping for Object Detection using Deep Learning Neural Networks

R. Vedha Priyavadhana, G. Shanmuga Priya, N. Renee Reddiar, M. Mano Priya

Abstract: ***In this paper, we propose a new novel DNN-based video saliency prediction method.The main aim of this project is to separating foreground and background images and also finding the motion of the objects.The methodogy used in this project is to seperating the foreground images using background modelling techniques(Pixel searching algorithm).The another technique is Homography based search algorithm for block based motion estimation is used to finding the movement of the objects.Saliency object detection is also used to highlighting the particular region in the images.In our base paper[1],context subtraction algorithm is introduced.But in this paper,Convolutional neural networks is used to classifying the foreground objects from background images.We propose an deep convolutional neural network (CNN) Classifier is used to predict intra-frame salience by exploring object and object movement details .The CNN classifier is also used to separating the new objects from background. We also find from our database that a temporary correlation of human attention exists, with a smooth transition of saliency across video frames.Therefore, in our DNN-based approach we create a two-layer convolutionary long-term memory (2C-LSTM) network, using the extracted CNN features as the data.The fuzzy clustering is used to store the information as the processed images in this project.This project can be implemented in real time applications such as defence,forest and miltrary area.Saliency Mapping is also used to reduce the Zipper effect,it mainly occurs in edges.It gives High resolution too. Finally, the proposed system is developed using matlab simulation.The PSNR value is attained above 95%.***

Index Terms: Convolution Neural Network (CNN), Convolution Long Short-Term Memory (2C-LSTM) Network, Region Based Pixel Separation(Rot),Fuzzy Cluster.

I. INTRODUCTION

The main objective of this paper is to study, implement, and evaluate video object tracking algorithms able to track multiple objects in surveillance videos suitable for use by a robust surveillance system.

Manuscript received on February 10, 2020.
Revised Manuscript received on February 20, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
Dr. R. Vedhapriyavadhana, Associate Professor/ECE, Francis Xavier Engineering College, Tirunelveli, India. E-mail: vedhapriya.vadhana@gmail.com
G. Shanmuga Priya, PG Scholar/ECE, Francis Xavier Engineering College, Tirunelveli, India. E-mail:priyakailash21@gmail.com
N. Renee Reddiar, PG Scholar/ECE, Francis Xavier Engineering College, Tirunelveli, India Email:reneereddiar@gmail.com
M. Mano Priya, PG Scholar/ECE, Francis Xavier Engineering College, Tirunelveli, Indi. E-mail:manopriyam1997@gmail.com

The objective is to separate the back ground and fore ground of the moving object images.In this paper, we concentrate on extending the block Motion searching algorithms (BMA) from single reference frame to multiple reference frames.
There are many oldest technique for salience detection.The following methods have some drawbacks and they are discussed.
[1]Motion is one of the most important signs of distinguishing foreground objects in a video from context. These complexities that prevent existing context subtraction algorithms from properly identifying the artifacts in the foreground.But we propose a new approach with regular and/or infrequent movements that can identify the foreground objects. [2]Hidden Markov fields have been extensively applied mainly for segmentation and change detection.In this maximization is achieved by iterative conditional mode.This is not suitable for many un natural systems.[3]The line segments are used to analyze lane structural information and fit the two lane markings closest to it. The results demonstrate our method's Low accuracy and robustness in various complex environments.[4]High-performance hardware architecture is designed to measure the Amount of Absolute Difference in H.264 Video Coding and improve performance with a reasonable frequency in terms of speed. But block size varies with the increase in complexity. [5]Transmits high quality video information in a less bandwidth. But Time decoding in inter frame coding is complex.[6]REGIM Video coding technique is proposed and it reduces the search point's number in Motion Estimation block of video codec. But Computation is more expensive.[7]First, we estimate the scene flow using color maps and aligned depth maps in two consecutive frames, not extracting each point's feature vector from the estimated scene flow. Experiment does not provide satisfactory process results.[8]A new framework for detecting abandoned objects using dual context model subtraction is introduced in this paper.It is highly complex.Efficiency of the system is Low.[9].It is used to detect the multiple objects.But the computational time is high.[10] Sorted method algorithm is proposed.But in this paper,we reduces time and complexity.[11]A geometric motion model is proposed for the derivation of a geometric derived motion vector (GDMV). Yet challenge in choosing a suitable neighbor MV pair. But we don't have to choose a neighboring pair in ours. [12]The method of estimation and motion compensation is proposed for the recovery of lost packets in motion areas. [13] The 1-D and 2-D ME block designs are designed for VBS using the FS algorithm. Computational time is high in order to identify the best matching block.[14]A motion estimation algorithm based on a weak representation is proposed.

The coefficients of sparse representation are calculated in real time, so there is more computational complexity.[15] Variable Block Size Motion Estimation is aimed at reducing the block effect and improving the quality.

Nonetheless, optimizing distortion efficiency in any mode would greatly increase the difficulty of coding.[16]The fixation prediction models is used to evaluating the better fixation values.It becomes feasible.[17]The saliency cut algorithm is used and it is very simple and efficient.It gives full resolution.[18]Theweighted graph model is used to provide high precision and volume data as well as video and to determine the optimum number of textural representations. But the downside is that the entire weighted graph is needed for the network.[19]Use main edge distribution with high quality, it is used to distinguish animal images from the background in natural senes. But it distinguishes animal representations only from humans, objects etc.[20]The image can be segmented by means of active contours and polygons. The accuracy is high. The topology for geometric flow of bounded images is very difficult to preserve.

II. PROPOSED SYSTEM

Security and video surveillance has developed into an increasing marketplace.There will always be robbers and criminals. The monitoring system has become a key element in security.

A.*Block diagram*:

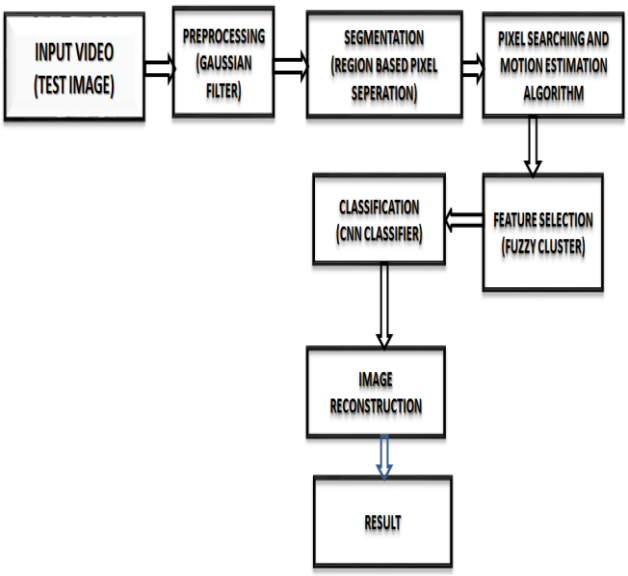


Fig 1: Proposed system Block diagram

From the above block diagram, the input video is converted into images or frames.

i)RGB Plane:

The input image is color image, the color processing steps converts the color image into grayscale image.

ii) Resizing:

The resizing can be done by gaussian filter.The rise and fall time decreases when no overshoot to a step function input.It is very faster when compared to other filters.It particurarly removes salt and pepper noise in the input.The mean is zero and Standard deviation is 99.7% for gaussian filters.

iii) Gaussian filter:

Filtering is a process of improving an image. Filtering can highlight some unique features or eliminate some unwanted features. Median filter conserves the edges of an image while minimising non specific noises.The most commonly occurring noise is the impulse noise. It introduces at the time of image acquisition and transmission of image is the impulsive noise.Image noise comes from various sources. Noise can be obtained due to communication errors and compression of images.

iv) Segmentation:

Segmentation is the mechanism by which an image is fractionated into its component pieces. Within MATLAB this can be achieved within different ways. One approach is to use a combination of morphological operations to segment objects within an image.In region based pixel separation,histogram based intensity method is used to segmenting the foreground and background images

v) Fuzzy logic Controller(Background Estimator)

Fuzzy Logic Control (FLC) has proven to be effective in dynamic, non-linear and imprecisely specified systems for which traditional model-based control methods are inefficient or unlikely.

vi) Motion Estimation:

In this paper, we use block-based motion estimation. The basic concepts of digital video including capture, color spaces, formats and quality. A video signal can be sampled as a series of complete frames or as a sequence of interlaced fields. The commonly used color spaces are RGB and YCbCr and their variations (sometimes referred to as RGB and YCbCr).We used the intermediate quarter format(QCIF) is 176. Objective quality measurements include peak signal to noise ratio(PSNR). In block movement estimation, images are normally rectangular in shape and block division can be easily done. The block sizes are etc. The goal of ME is to predict the next frame from the current frame to predict macro blocks as accurately as possible.Some of the following procedures are:

1. The macro block is (16). The movement between the consecutive frames is small, the search range is confined to this area. The best match within this area is determined based on the minimum MV displacement. The process of finding the best match, block by block is called ME(BBME) based block.
2. The motion vectors and residues between the current block and the reference blocks are measured. The residue and motion vector recovery process is known as motion compensation.
3. The best matching residues and MVs are encoded by the units TC and EC and transmitted to the decoder.
4.To recreate the original image, the process is reversed at the decoder.

vii) *Cnn Classifier*:

A Convolutionary Neural Network (CNN or ConvNet) is a class of deep neural networks in deep learning, most commonly applied to visual imaging research.

These are also known as artificial neural networks (SIANN) invariant shift invariants or space invariants.

They have applications in image and video recognition, suggest systems, image detection, medical image analysis and natural language processing. A CNN's main building block is the convolutionary sheet.

The layer parameters consist of a series of learnable filters (or kernels) that have a limited receptive field, but expand through the input volume's full depth. The filter is translated over the width and height of the input volume during the forward transfer, computing the dot product between the filter entries and the input, and generating a 2-dimensional activation map for that filter.

B. Flowchart of the proposed system:

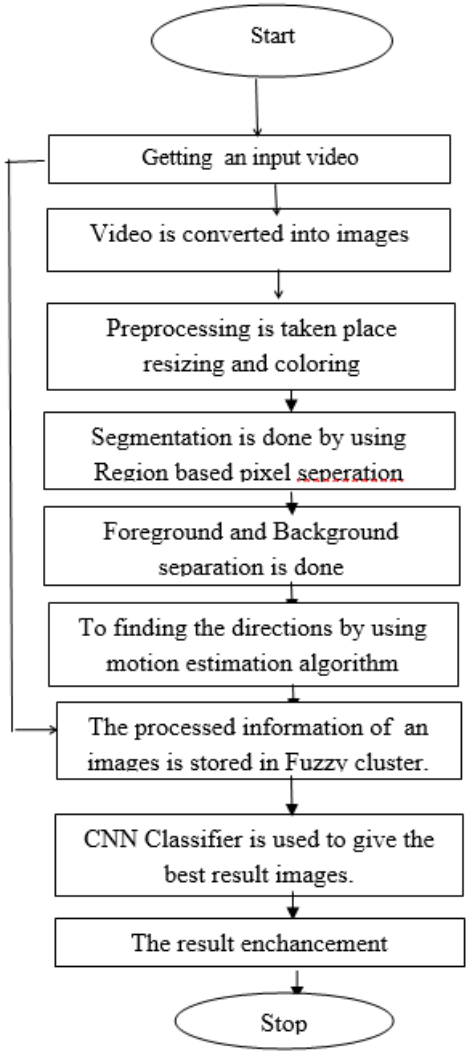


Fig 2: Flowchart of the proposed system

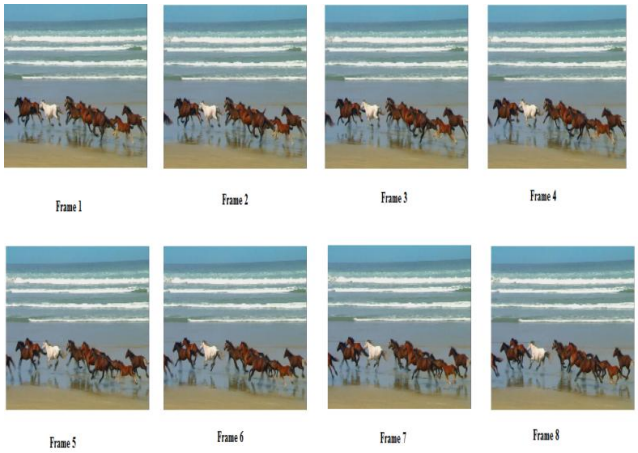
The video will be transformed into frames, then the frames will continue to check whether they are related to the background model or the foreground model, if they are related to the background model, then the frames will be retained as the background model, and the changes in the frames referred to as foreground objects from which we can detect the moving object using the proposed method.

III. RESULTS AND DISCUSSION

This project is implemented using Matlab simulation.For segmentation,2 different types of videos are applying to giving results.The first one is animal video and the second one is car video and they are converted into images,segmentation is done by giving our input images.The next is to finding the directions of the given images.

A. Video Conversion:

i)The first input video is converted into frames.



ii)The second video is converted into frames.

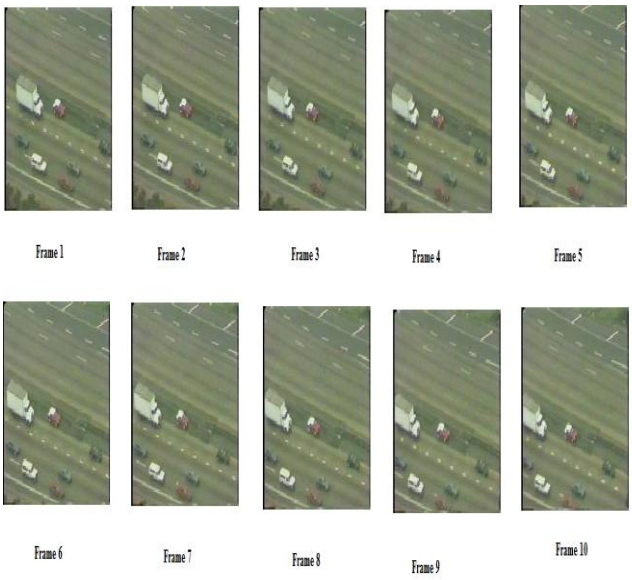
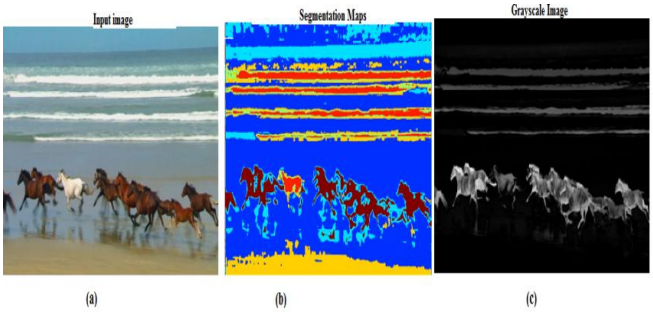
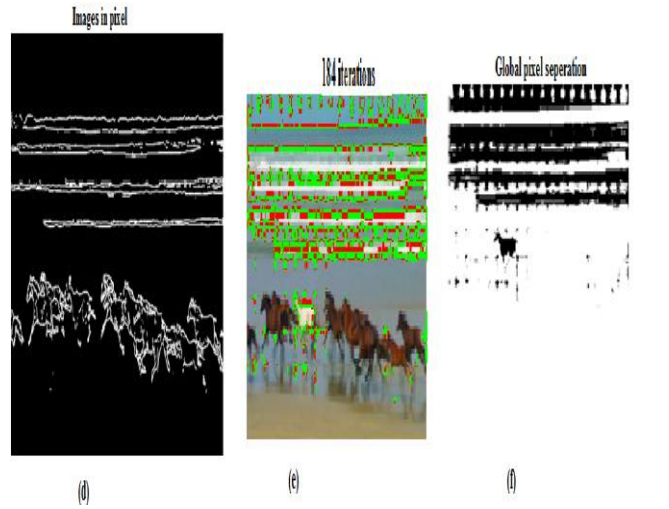


Fig 3: Video conversion into images

B. Segmentation Output:

(i)The first input image is given as horse image for segmentation.





(ii)The second input image is given as car image for segmentation

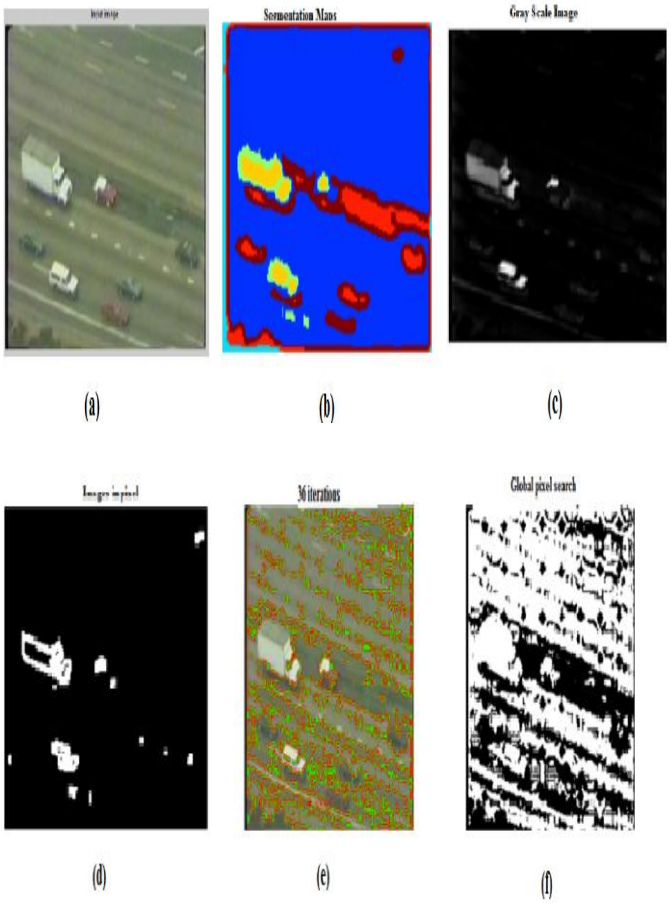


Fig 4 (a) to (f) : Segmentation using Saliency Mapping

The figure 4(a) show that the input image is given for the proposed work(b) Preprocessing is carried out to convert the input image in RGB color conversion,(c)Grayscale conversion is given for filtering using gaussian filter. (d)It shows that the images in pixel(e) CNN classifier giving foreground and background seperation within 36 iterations and 184 iterations (f)Global pixel seperation is used to seperating the background images using saliency mapping.

C. Proposed System Matlab Implementation:

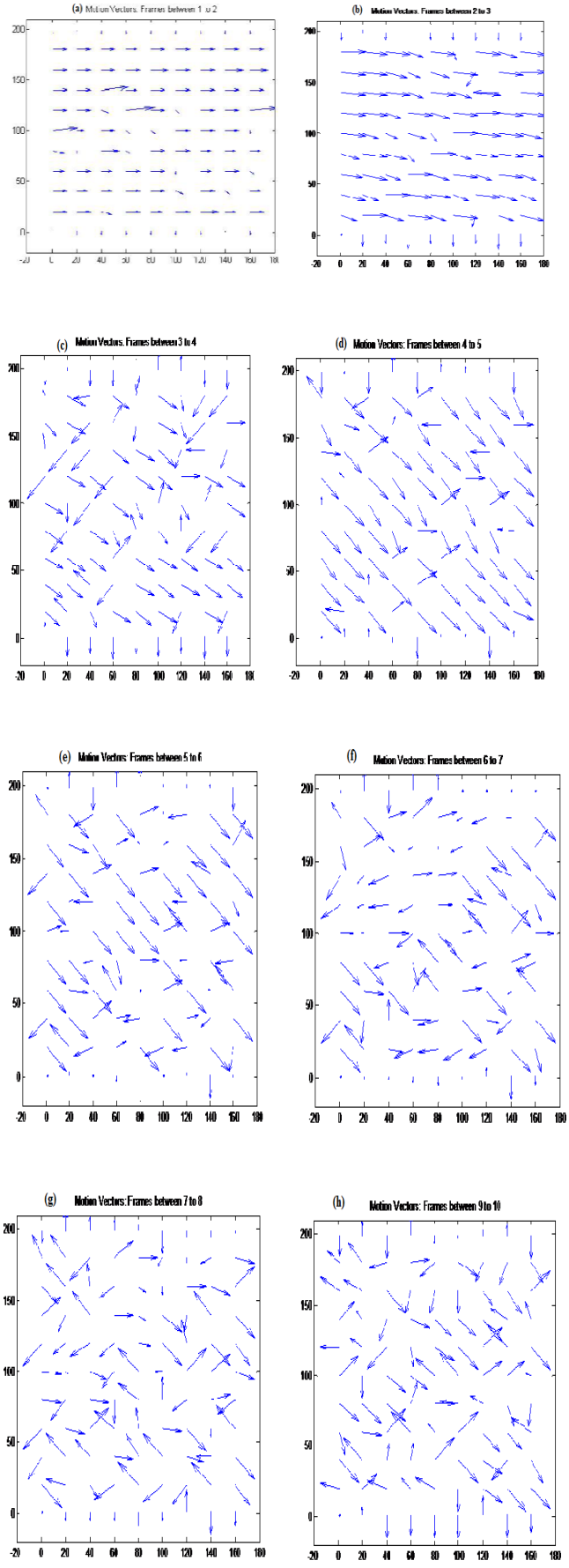
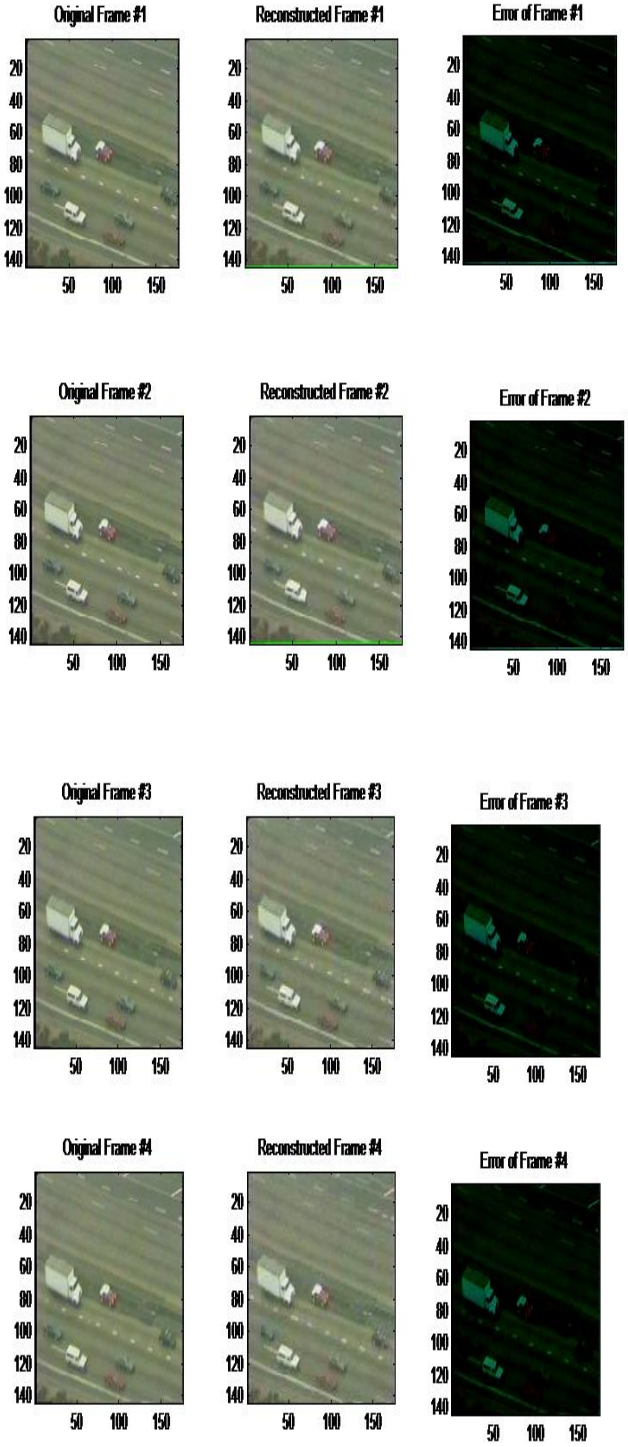


Fig 5 (a) to (h): Motion Estimation

The figure 5 (a) to (h) shows the direction of contour between frames to frames, this gives exact directions of the object.

Motion estimation algorithm provides direction vectors in the proposed system.



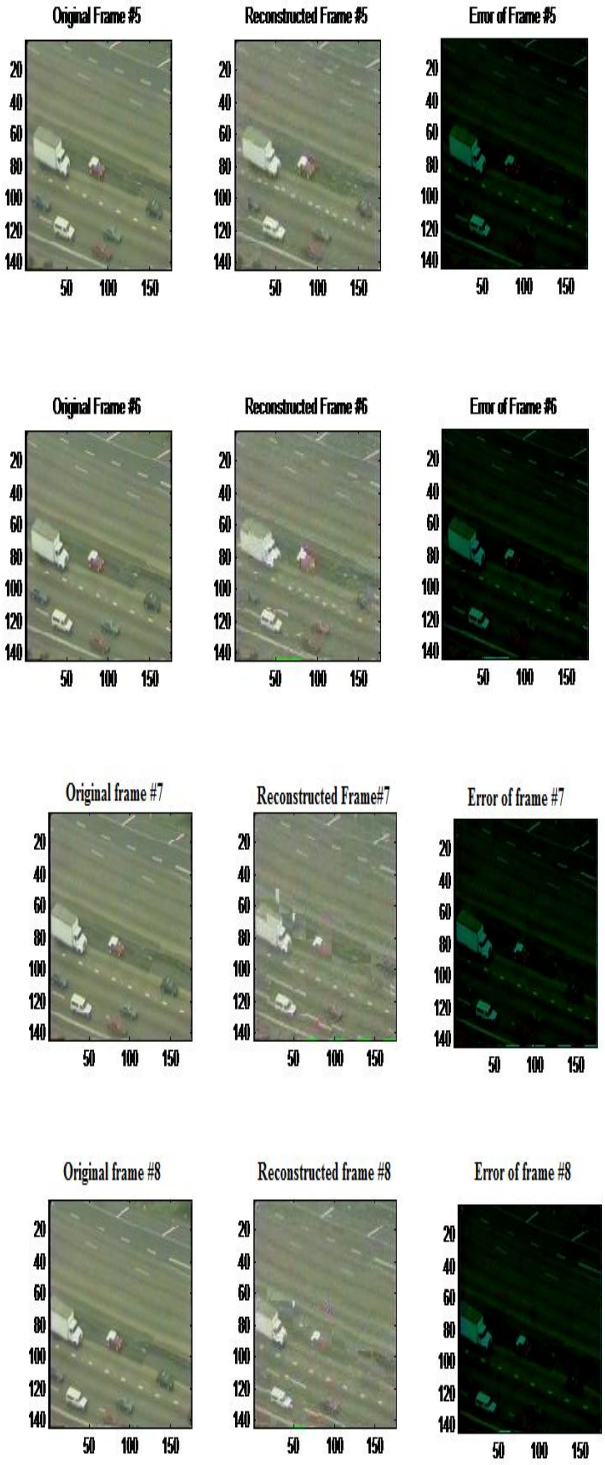
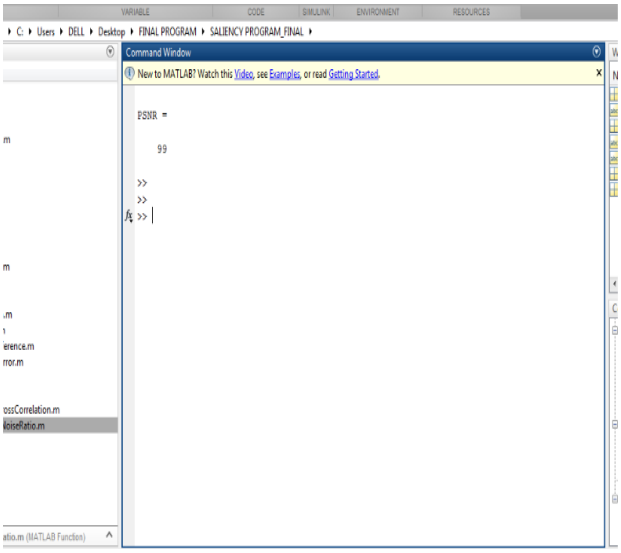


Fig 6 (a) to (h): Reconstructed images for Motion Estimation

The figure 6(a) to (h) shows the reconstructed frames from the input videos, the segmentation process exactly separates the background and foreground of the images.

D. Measuring PSNR Value:



The quality of motion estimation is taken by PSNR value.The value 99 is obtained by using PSNR formula.

IV. CONCLUSION

This paper used a context subtraction to detect objects moving through an SWT-based adaptive threshold technique to develop our algorithm for detecting and monitoring moving targets.Compared to other algorithms, the experimental results prove that the proposed approach can detect and track the moving objects efficiently and with robustness. Moreover, the simplicity of the proposed method indicates that the approach can be implemented in different intelligent systems.In this project,we propose pixel searching algorithm is used for segmenting the foreground and background images gives high resolution and easily seperated the foreground objects by saliency mapping.The homography based search algorithm for block motion estimation in coding the sequences captured by PTZ Cameras.The homography between consecutive frames is directly computed from coressponding points on images.The experimental results have proved that the HBS algorithms highly efficient and robust for block motion estimation in coding the sequences captured by PTZ cameras.

EXTENSION WORKS

- The K-means cluster can be replaced with fuzzy c-means logic, which will increase the speed of separation.
- The CNN Classifier can be replaced with RBFNN Classifier.The RBFNN classifier giving the best results when compared to CNN classifier.
- Fuzzy cluster is used in feature selection can be replaced with genetic algorithm.Genetic algorithm having more accuracy than the fuzzy cluster.The memory space is also high.
- In my research work,the performance characteristics and comparison is also carried out between the CNN and RBFNN classifier.

REFERENCES

1. Yuewei lin ,Yu cao,Youjie zhou,"Visual-Attention Based Background Modeling for Detecting Infrequently Moving Objects",IEEE Transcations on circuits and systems for video Technology,vol 26,2014.
2. Mohamed El Yazid Boudaren," Unsupervised Segmentation of SAR Images Using Gaussian Mixture-Hidden Evidential Markov Field",IEEE Geoscience and Remote Sensing Letters ,December 2016.
3. ChangzhengHou, Jin Hou, Chaochao Yu ," An efficient lane markings detection and tracking method based on vanishing point constraints",35th Chinese Control Conference (CCC) ,IEEE,ISSN: 1934-1768,29 August 2016.
4. Chukka Santosh ," A SAD Architecture for Variable Block Size Motion Estimation , International conference on Microelectronic Devices, Circuits and Systems (ICMDCS),IEEE in 18 December 2017.
5. Ashika K O ," Inter Frame Coding in Advanced Video Coding Standard H.264 using Block Based Motion Compensation Technique",International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT),08 February 2018.
6. Mahmoud Ahmadi, AhlemWalha, Ali Wali and Adel M. Alimi, "A New Motion Estimation Technique for Video Coding", 15th International Conference on Intelligent Systems Design and Applications, vol.5, pp.110-115, 2015.
7. XuezhiXiang ,Motion detection based on RGB-D data and scene flow clustering,12th World Congress on Intelligent Control and Automation (WCICA),September 2016.
8. Wahyono , Alexander Filonenko , Kang-Hyun Jo ," Detecting abandoned objects in crowded scenes of surveillance videos using adaptive dual background model",8th International Conference on Human System Interaction (HSI),July 2015.
9. Andrea Manno-Kovacs,"Direction Selective Contour Detection for Salient Objects", DOI 10.1109/TCSVT.2018.2804438, IEEE Transactions on Circuits and Systems for Video Technology.
10. Yogananda Patnaik ," A New Search Method for Block Motion Estimation in Video Compression",DOI:10.1109/INDICON.2014.7030562, Annual IEEE India Conference (INDICON),2014.
11. Yucheng Sun,"Motion vector fields based video coding",DOI: 10.1109/ICIP.2015.7351170,IEEE International Conference on Image Processing (ICIP) 2015.
12. Abdou Khadre Diop , K. Tall ,S. M. Farssi ," Performances of the estimation and motion compensation for the reconstruction of motion areas in a sequence video Motion JPEG 2000,16th International Conference on Advanced Communication Technology,2014.
13. Jayakrishnan,"2-Dimensional Systolic Architecture for H.264/AVC Variable Block Size Motion Estimation",International Conference on Circuits, Communication, Control and Computing,2014.
14. Yanfei Shen ,Jintao Li , Zhenmin Zhu ," Motion Estimation for Video Coding based on Sparse Representation", IEEE International Conference on Acoustics, Speech and Signal Processing,2013.
15. Ailing Zhang, Xiaopeng Fan ,Ruiqin Xiong ,"Distributed Soft Video Broadcast with Variable Block Size Motion Estimation",Visual Communications and Image Processing (VCIP),2013.
16. Ali Borji,"What is a salient object? A dataset and a baseline model for salient object detection",IEEE Transactions on Image Processing, 2014
17. Ming-Ming Cheng, Guo-Xin Zhang,"Global Contrast based Salient Region Detection",Journal of latex class files, vol. 11, 2012.
18. Alexander Wong, Khalil Fergani, John S. Zelek,"Statistical Textural Distinctiveness for Salient Region Detection " , IEEE Conference on Computer Vision and Pattern Recognition,2013.
19. *Litian Sun and Tadashi Shibata,"*Uusupervised object extraction by contour delineation and texture based discrimination" , 20th European Signal Processing Conference (EUSIPCO 2012) Bucharest, Romania, 2012.
20. Ganesh Sundaramoorthi and Anthony Yezzi, *Senior Member,"* Global Regularizing Flows With Topology Preservation for Active Contours and Polygons ", IEEE Transcations on Image Processing,vol. 16, No. 3, March 2007.

AUTHORS PROFILE

Dr. R. Vedhapriyavadhana, received B.E degree Biomedical Instrumentation Engineering from Avinashilingam Deemed University in 2005. She received the M.Tech degree in Computer and Information Technology from Manonmaniam Sundaranar University, Tirunelveli in the year 2009 and Completed her Ph.D in Information and Communications in Anna University Chennai by July 2019 .

At Present she is working as Associate Professor in Electronics and Communication Engineering Department of Francis Xavier Engineering College, Tirunelveli Tamil Nadu, India.
She has authored / coauthored over 30 research articles in various Journals and Conferences in the areas of Image and Video Processing, Embedded System and Bio Medical Instrumentation.

Mrs. Shanmuga Priya.G, received B.E degree in Electronics and communication engineering from Dr.Sivanthi Aditanar College of Engineering in 2016.I am pursuing my M.E (communication systems) in Francis Xavier Engineering College,Tirunelveli in the year 2020.I attended the national and international conferences in the area of embedded ,wireless and robotics from various colleges.

N. Renee Segrid Reddiar, completed BE in the year 2013-2017 and presently doing my master degree in the stream of communication systems attended National Conference on the topic Innovation and Sustainability in Rural and Agro Development in Digital Era on Jan 24,2020 and also published my paper in UGC Journal.

M. Mano Priya, completed B.E in National Engineering college in the year 2014-2018 and presently doing my master degree in the stream of communication systems attended National Conference on the topic Innovation and Sustainability in Rural and Agro Development in Digital Era on Jan 24,2020 and also published my paper in UGC Journal.