[Moving Object Detection Based on Background Subtraction using Energy Function]

REPORT OF PROJECT SUBMITTED FOR PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF COMPUTER APPLICATION

By

Priyanka Dutta, Samanwaya Bhattacharya, Suman Naskar, Adrija Mukherjee

REGISTRATION NO – 221170510001,221170510043, 221170510030, 221170510025

UNIVERSITY ROLL NO –11771022003, 11771022050, 11771022021,11771022022

UNDER THE SUPERVISION OF
Mr. Manas Ghosh
[Assistant Professor at RCCIIT]

[PUT COLLEGE LOGO HERE]

AT
RCC INSTITUTE OF INFORMATION TECHNOLOGY
[Affiliated to West Bengal University of Technology]
CANAL SOUTH ROAD, BELIAGHATA, KOLKATA – 700 015
DECEMBER – 2023

RCC INSTITUTE OF INFORMATION TECHNOLOGY

KOLKATA – 7OOO15, INDIA

CERTIFICATE

The report of the Project titled Moving Object Detection Based on Background Subtraction using Energy Function submitted by Priyanka Dutta, Suman Naskar, Adrija Mukherjee, Samanwaya Bhattacharya (Roll No.: 11771022003, 11771022050, 11771022021, 11771022022 of MCA 3rd Semester of 2023) has been prepared under our supervision for the partial fulfillment of the requirements for MCA degree in MAKAUT. The report is hereby forwarded.

OPTIONAL IN CASE

[Manas Ghosh]	[Manas Ghosh]
Assistant Professor	Dept. of MCA
RCCIIT	RCCIIT, Kolkata
(External Supervisor)	(Internal Supervisor)

Countersigned by

…………………………………………. Name of
Head
Department MCA
RCC Institute of Information Technology, Kolkata – 700 015, India

ACKNOWLEDGEMENT

We express our sincere gratitude to Mr. Manas Ghosh Sir, Assistant Professor of Department of MCA, RCCIIT and for extending his valuable times for us to take up this problem as a Project.

We are also indebted to him for his unconditional help and inspiration whenever needed.

Date: 04.12.2023

MCA – 3rd Semester, 2023, RCCIIT

Name	Roll No	Regn No
Suman Naskar	11771022050	221170510043
Samanwaya Bhattacharya	11771022021	221170510030
Priyanka Dutta	11771022022	221170510025
Adrija Mukherjee	11771022003	221170510001

RCC INSTITUTE OF INFORMATION TECHNOLOGY

KOLKATA – 7OOO15, INDIA

CERTIFICATE of ACCEPTANCE

The report of the Project titled Moving Object Detection Based on Background Subtraction using Energy Function Priyanka Dutta, Suman Naskar, Adrija Mukherjee, Samanwaya Bhattacharya (Roll No.: 11771022003, 11771022050, 11771022021,11771022022 of MCA 3rd Semester of 2023) is hereby recommended to be accepted for the partial fulfillment of the requirements for MCA degree in MAKAUT.

Name of the Examiner	**Signature with Date**
1. …………………………………….	…………………………….
2. ……………………………………..	……………………………..
3. ………………………………………	………………………………
4. ……………………………………….	……………………………

TABLE OF CONTENTS

Moving Object Detection by Using Background Detection

1. Introduction

Human vision has played a crucial role in shaping the field of computer vision. Our understanding of how the human eye perceives and processes visual information has served as a foundation for developing algorithms that enable computers to mimic this ability. The study of human vision has provided insights into the complexities of visual perception, such as pattern recognition, object detection, and depth perception. These insights have been instrumental in developing computer vision systems that can perform tasks such as image recognition, object tracking, and autonomous navigation. The intricate mechanisms of human vision have inspired researchers to create computational models that replicate these processes. These models, often based on artificial neural networks, have been trained on vast amounts of visual data, allowing them to learn and adapt to new stimuli. This ability to learn from data has been essential for the development of robust and versatile computer vision systems.

Furthermore, the study of human vision has provided valuable guidance in designing user interfaces and interacting with computer systems. By understanding how humans perceive and interact with visual information, researchers can create interfaces that are intuitive, efficient, and accessible. This has been particularly important in the development of augmented reality and virtual reality technologies, where the seamless integration of computer-generated imagery into the real world requires a deep understanding of human visual perception.

In essence, human vision has served as a compass guiding the development of computer vision. Our understanding of how we see has inspired algorithms, shaped methodologies, and influenced the design of interactive systems. As computer vision continues to evolve, the lessons learned from human vision will remain invaluable in creating machines that can truly understand and interact with the visual world.

1.a. Object Detection

Object detection methods can be divided into two categories: traditional object detection methods and deep learning-based object detection methods [17–23]. For traditional object detection methods, Viola and Jones [10,11] propose an algorithm that realizes the real-time detection of human faces for the first time. Their algorithm is called the VJ detector. The VJ detector uses a sliding window for detection and greatly improves the detection speed by using three technologies: the integral map, feature selection, and detection cascades [9–11]. Dalal et al. propose a histogram of oriented gradient (HOG) feature descriptor [12]. HOG is an important improvement to scale-invariant feature transformation [13,14] and shape contexts [15]. The HOG detector is also an important foundation for many object detectors. Felzenszwalb et al. [16] propose the deformable parts model (DPM) algorithm, which is a milestone for traditional object detection algorithms. DPM follows the detection principle of "divide and conquer" [9]. In this principle, the training can be simply regarded as learning the correct way to decompose an object, and the inference can be regarded as a collection of detecting different object parts [9,16]. Since AlexNet won the ImageNet competition championship, more and more studies have focused on object detection by deep learning. For deep learning-based methods, Girshick et al. [20] propose regions with convolutional neural network features (RCNN) for object detection. RCNN improves the detection accuracy a lot, but the redundant feature calculation makes the detection speed very slow [9,20]. He et al. [21] propose spatial pyramid pooling networks (SPPNet). SPPNet introduces the SPP layer, which makes it faster than RCNN by more than 20 times, and the accuracy is almost unchanged [21]. Ren et al. [22] propose the Faster RCNN detector. Faster RCNN introduces the regional proposal network (RPN), making it the first near real-time deep learning detector [9,22]. Liu et al. [23] propose a single shot multibox detector (SSD). SSD is a one-stage object detection algorithm. It introduces multi-reference and multi-resolution detection technology [9], which greatly improves the detection accuracy of the one-stage detection algorithm.

Object tracking can reduce the missed detection in object detection. The main object tracking algorithms include correlation filter tracking [24–29] and non-correlation filter tracking [30–34]. For correlation filter tracking algorithms, Bolme et al. [25] propose the minimum output sum of squared error (MOSSE) filters. MOSSE changes the situation that correlation filter tracking algorithms are not suitable for online tracking [24,27]. Ma et al. [32] propose rich hierarchical convolutional features in a correlation filter (CF2) for visual tracking. Qi et al. [29] propose the hedged deep tracking (HDT) algorithm, which combines multiple weak trackers to form a strong tracker. For non-correlation filter tracking algorithms, Zhang et al. [27] propose the convolutional networks without training (CNT) algorithm, which can utilize the inner geometry and local structural information of the object. CNT can adapt to changes in the object's appearance during tracking [24,30]. The structural sparse tracking (SST) algorithm [31] uses the inherent relationship between the local object patches and the global object to learn sparse representation together. The object location is estimated based on the object dictionary template and the corresponding block with the largest similarity score from all particles [30,32]. And other object detection methods are Frame Differencing (FD), Optical Flow (OF), and (Background Subtraction).

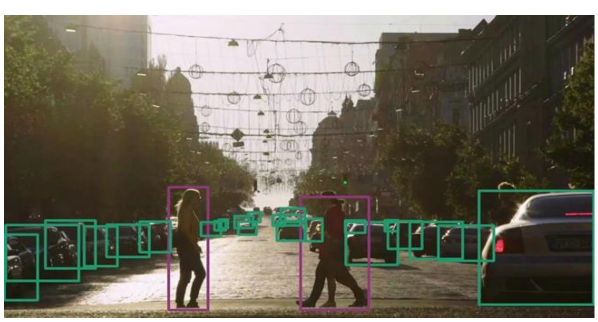


Fig: 1.1 An example of moving object detection

1.b Moving Object Detection and Its Application

Moving object detection is a computer technology related to computer vision, image processing that deals with detecting moving objects of a certain category (such as human, cars etc.) in digital image or video. Researched domains include vehicle detection, human detection. Moving object detection has many applications in the domain of computer vision, including image retrieval and video surveillance. Moving object detection [2] is the detection process of finding the location of moving objects in images or videos. Moving object detection involves locating objects in the frame of a video sequence. It’s not an easy task to do in video as the video may be taken with a static camera or may be with a dynamic / moving camera. In the first case detection is a bit easier as the background of the video is static. But in the second case the background is also dynamic. So, for the second scenario both objects and background are moving. That's why detection of only moving foreground or moving objects is a bit tough. In moving object detection various background subtraction techniques are available. Background subtraction involves the absolute difference between the current image and the reference updated background over a period of time. A good background subtraction should be able to overcome the problem of varying illumination condition, background shadows, camouflage, bootstrapping and at the same time motion segmentation of foreground objects should be done at the real time [3]. An example of moving object detection that specifies detection of moving cars on the road where everything is static apart from those cars in Figure 1.2

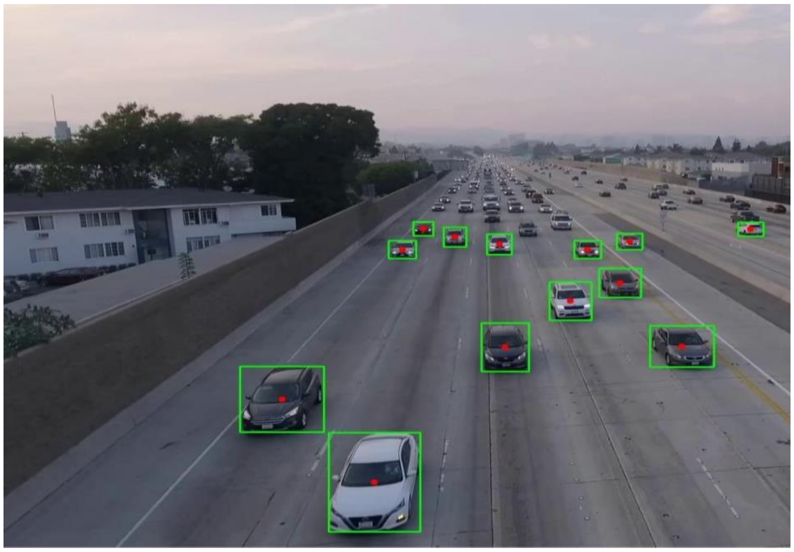


Fig: 1.2 Moving object detection

The moving object tracking in video pictures has attracted a great deal of interest in computer vision. Object detection is the first step in surveillance systems, navigation systems and object recognition [4]. CCTV is one of the main reasons for the growing interest and use of video in security systems. Moving object detection in a video stream is an essential step in video surveillance applications. In some algorithms, the moving objects may become part of the scene when they come to a stop. Also, the scene may be affected by changes in the light, leaves swaying, cameras shaking, etc. Many algorithms of moving object detection have been proposed in recent years. These involve background subtraction, optical flow, temporal difference and many other algorithms for detecting moving objects. From these, the most widely used algorithm is background subtraction, which has many algorithms such as frame difference, approximate median, Gaussian mixture. Moving Object detection has huge significance in a real time environment. As it has several important applications such as it is used to provide a better sense of security using visual information. It's also used to give security and surveillance to recognize people, to analyze shopping behavior of customers in retail space, video abstraction to obtain automatic annotation of videos, to generate object-based summaries, traffic management to analyze flow, to detect accidents, video editing to eliminate human operator interaction.

2. Problem Definition

The system should excel in accurately detecting and tracking moving objects against a non-changing backdrop, with a focus on the key challenges-Background Subtraction, Noise reduction and object tracking. This project is designed to implement a moving object detection system specifically tailored for scenarios where the background remains static.

3. Review of the Literature

Li Y, Bao H, Ge Z, Yang J, Sun J, Li Z. Bevstereo: Enhancing depth estimation in multi-view 3d object detection with temporal stereo (2023) [1]. In this paper, a novel multi-view object detector is proposed, namely BEVStereo. BEVStereo improves performance without significantly increasing memory usage by applying dynamic temporal stereo technique to create temporal stereo. Some complex scenarios that other stereo-based approaches cannot handle can be resolved by our method. In addition, we propose size-aware circle NMS, which takes the size of boxes into account while avoiding the laborious computation of rotated IoU. Under both class-aware and class-agnostic circumstances, our size-aware circle NMS performs satisfactorily. Last but not least, we present Efficient Voxel Pooling v2, which speeds up voxel pooling by improving the efficiency of memory accesses.

Zhou, Zixiang, et al. "Centerformer: Center-based transformer for 3d object detection." European Conference on Computer Vision. Cham: Springer Nature Switzerland, (2022) [2]. In this paper, we propose a novel center-based transformer for 3D object detection. Our method provides a solution to improve the anchor-free 3D object detection network through object-level attention learning. Compared to the DETR-style transformer networks, we use the center feature as the initial query embedding in the transformer decoder to speed up the convergence. We also avoid high computational complexity by focusing the crossattention learning of each query in a small multi-scale window or a deformable region. Results show that the proposed method outperforms the strong baseline in the Waymo Open Dataset, and reaches state-of-the-art performance when extended to multi-frame. We hope our design will inspire more future work in query-based transformers for LiDAR point cloud analysis.

Balasubramaniam, A., & Pasricha, S. (2022) [3]. Object detection in autonomous vehicles: In this article, they discussed the landscape of various object detectors being considered and deployed in emerging AVs, the challenges involved in using these object detectors in AVs, and how the object detectors can be optimized for lower computational complexity and faster inference during real-time perception.

Sharma, T., Debaque, B., Duclos, N., Chehri, A., Kinder, B., & Fortier, P. (2022). [4] Deep learning-based object detection and scene perception under bad weather conditions. They implemented the proposed strategy to enhance vehicle recognition in rainy and regular weather conditions by utilizing an open dataset accessible at Roboflows. Rain and snow are challenging conditions for self-driving cars and often human drivers to deal with. Snow and rain impact the sensors and algorithms that control an autonomous vehicle.

Alotaibi, M. F., Omri, M., Abdel-Khalek, S., Khalil, E., & Mansour, R. F. (2022) [5]. Computational intelligence-based Harmony search algorithm for real-time object detection and tracking in video surveillance systems. In this study, a new CIHSA-RTODT technique was developed to detect and track objects on video surveillance systems. The CIHSA-RTODT technique designed an Adagrad with an improved RefineDet-based object detector which can effectively recognize multiple objects in the video frame. The HSA with TWSVM model was also applied to properly categorize the existence of objects in the video frame and thereby lead to improved classification results.

Jiménez-Bravo, D. M., Lozano Murciego, Á., Sales Mendes, A., Sánchez San Blás, H., & Bajo, J. (2022) [6]. Multi-object tracking in traffic environments: The research covers the techniques used in MOT in traffic environments and explains the main structure of this kind of system. We focus our explanation on the two main tasks of a MOT, detection, and tracking. We explained that the most used technique for detection is based on CNNs architectures and that the effectiveness of the several techniques available for tracking is based on the detection and identification outcome. Moreover, we make a distinction between different types of MOTs.

4. Proposed solution

Proposed solution leverages the background subtraction techniques to detect moving objects from static background. A video capture object is initialized to read frames from the input video. The first frame is used to initialize the background model for subsequent frame differencing. Absolute difference between consecutive frames is then calculated to create an energy function highlighting change. Then the dilation has been applied to enhance features in the binary images. A binary mask is generated by thresholding the blurred energy function to isolate regions with significant changes. Identify contours within the binary mask to locate areas of motion. Draw rectangles around detected objects based on contours. Add a descriptive label ("motion detected") to each bounding box. Show the original frame in real-time with overlaid rectangles around moving objects.

5. Implementation details

The implementation process can be outlined as follows:

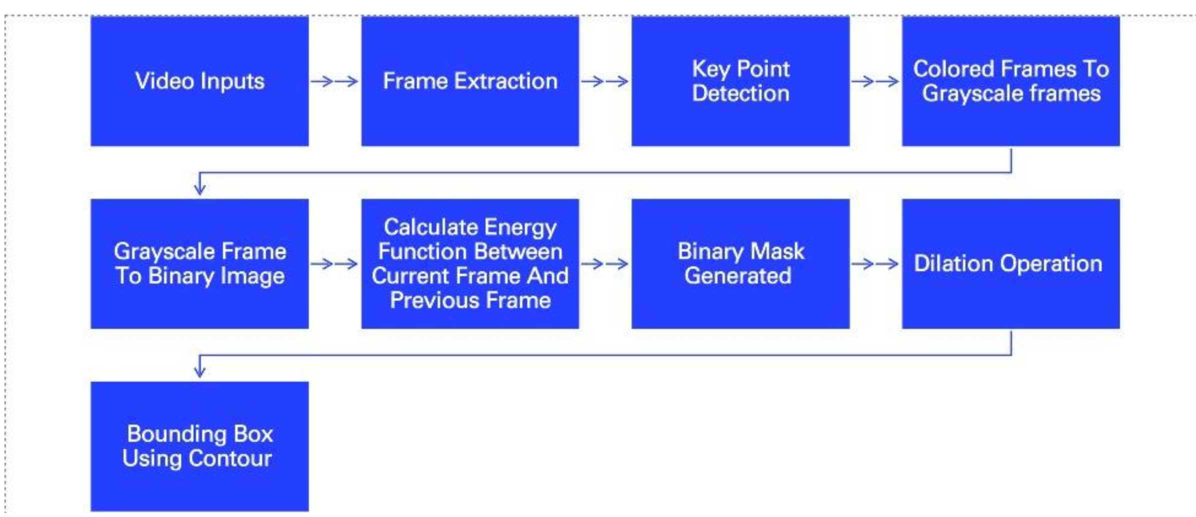


Figure: 1.3 Workflow diagram

- **Step 1 –**

The system initializes a video capture object, reading frames from a specified video file [1]. Video is a collection of multiple images that are constantly displayed to create motion effects. Thus, object detection in video files is actually detection of objects in image files (cause as mentioned earlier that video is made of multiple images). We need to perform the conversion of video into frames to determine where the areas of moving objects are, and where the area of the background is. Usually, one second of video has 24 (or maybe more) still images and 24 or more photos. So, the video which is taken as input is converted into frames at a rate of 4 frames per second. Let after conversion the video sequence V has n frames denoted as, where one frame denoted as $f\ (x,\ y)$. A frame is nothing but an image. An image is basically a 2D array where x is number of rows and y is number of columns shown in fig: 4. 1

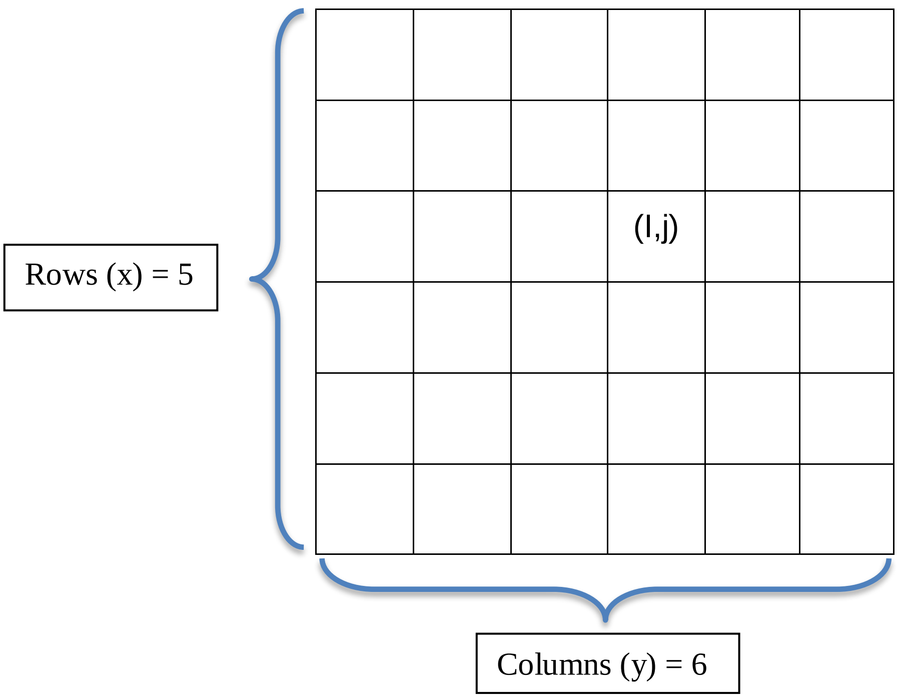


Figure: 4.1 A 2D image representation

A video consists of images which have the same dimension. That means number of rows x and number of columns y is same. $V = \{f1\ (x,\ y),\ f2\ (x,\ y),\ f3\ (x,\ y),\ \dots\ \dots,\ fn\ (x,\ y)\}$

$$or, V = \sum f_i(x, y)$$

- **Step 2 -** Frames are extracted from the video.

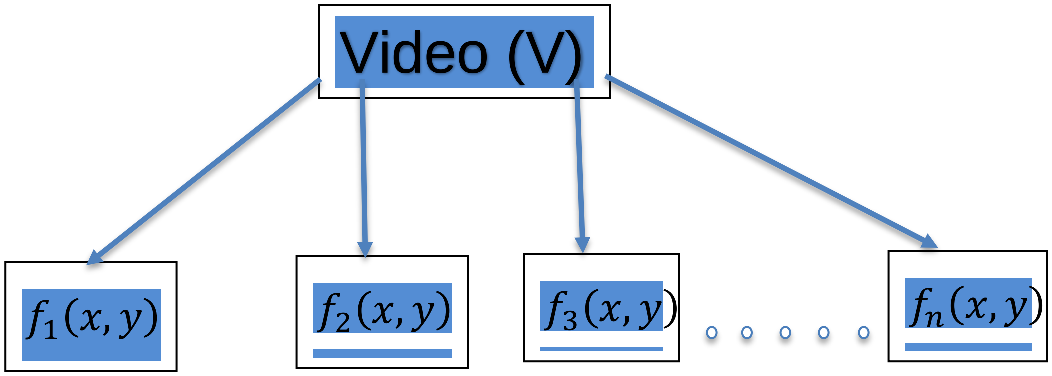


Figure 4.2: Video V is converted into n consecutive frame

- **Step 3 –**

A video contains many frames and, in each frame, there are so many pixels. Suppose a single frame has 100 rows and 100 columns so it has 10000 pixels and a video has 10 sec videos with 4fps. So, the total pixels for that will be 4,00,000 pixels. Working with these huge numbers of pixels will increase the computation time. To avoid this unnecessary computation important pixels or key points have been detected with the feature detector algorithm. A feature-detector is an algorithm that detects feature-points (also called interest points or key-points) in an image. During this step, important key points are searched in an image. AKAZE [31] feature detection algorithm has been used in this detection algorithm.

AKAZE uses the determinant of the Hessian (DoH) matrix [31]. DoH images are computed in each image. The Hessian matrix H is of second-order partial derivatives of P:

$$H = \begin{bmatrix} l_{xx} & l_{xy} \\ l_{xy} & l_{yy} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 P_i(x_m, y_m)}{\partial x^2} & \frac{\partial^2 P_i(x_m, y_m)}{\partial x\, \partial y} \\ \frac{\partial^2 P_i(x_m, y_m)}{\partial y \partial x} & \frac{\partial^2 P_i(x_m, y_m)}{\partial y^2} \end{bmatrix} \tag{2}$$

As $\frac{\partial^2 P_i(x_m, y_m)}{\partial x\, \partial y}, \frac{\partial^2 P_i(x_m, y_m)}{\partial y \partial x}$ both are same so both are denoted as l_{xy}.So the determinant of the Hessian matrix (H) is DoH as follows:

$$DoH = det(H) = \left(l_{xx} l_{yy} - l_{xy^2}\right) \tag{3}$$

The presented solution is designed to detect moving objects in a video stream using computer vision techniques. The key components of the implementation include the AKAZE feature detector and background subtraction methods.

Figure: 4.3 (Key points detected using AKAZE)

- **Step 4 -** Colored frames are converted to grayscale frames.

Figure:4.4 (Colored frames to grayscale)

- **Step 5 –**

Grayscale frames are converted into binary images. Binary image obtained pixels with values below the specified threshold will be set to 0, and pixels equal to or above the threshold will be set to 255.

Figure:4.5 (Grayscale image converted to binary image)

- **Step 6 –**

The absolute difference between consecutive frames is computed pixel-wise to create an energy function that highlights changes in the scene. The energy is a measure of the localized change of the image. It's the rate of change in the color/brightness/magnitude of the pixels over local areas [32]. To detect the moving object key points or pixels in each frame Energy function is used by calculating Entropy. Entropy is the measurement of the dissimilarity of the pixels [33]. In this method entropy defines that the key point belongs to the foreground or background. Entropy is formulated as follows:

$$-\sum p(x) * \log p(x) \qquad (4)$$

Where $p(x)$ is the probability of a given symbol/even

- **Step 7 –**

A binary mask is generated by thresholding the energy function. This binary mask isolates regions with significant changes, indicative of moving objects.

- **Step 8 –**

The dilation operation has been applied iteratively. In our case, the dilation operation on a 3x3 rectangular structuring element has been applied six times. Dilation is often used to expand or enhance features in binary images, and the number of iterations controls the extent of this expansion.

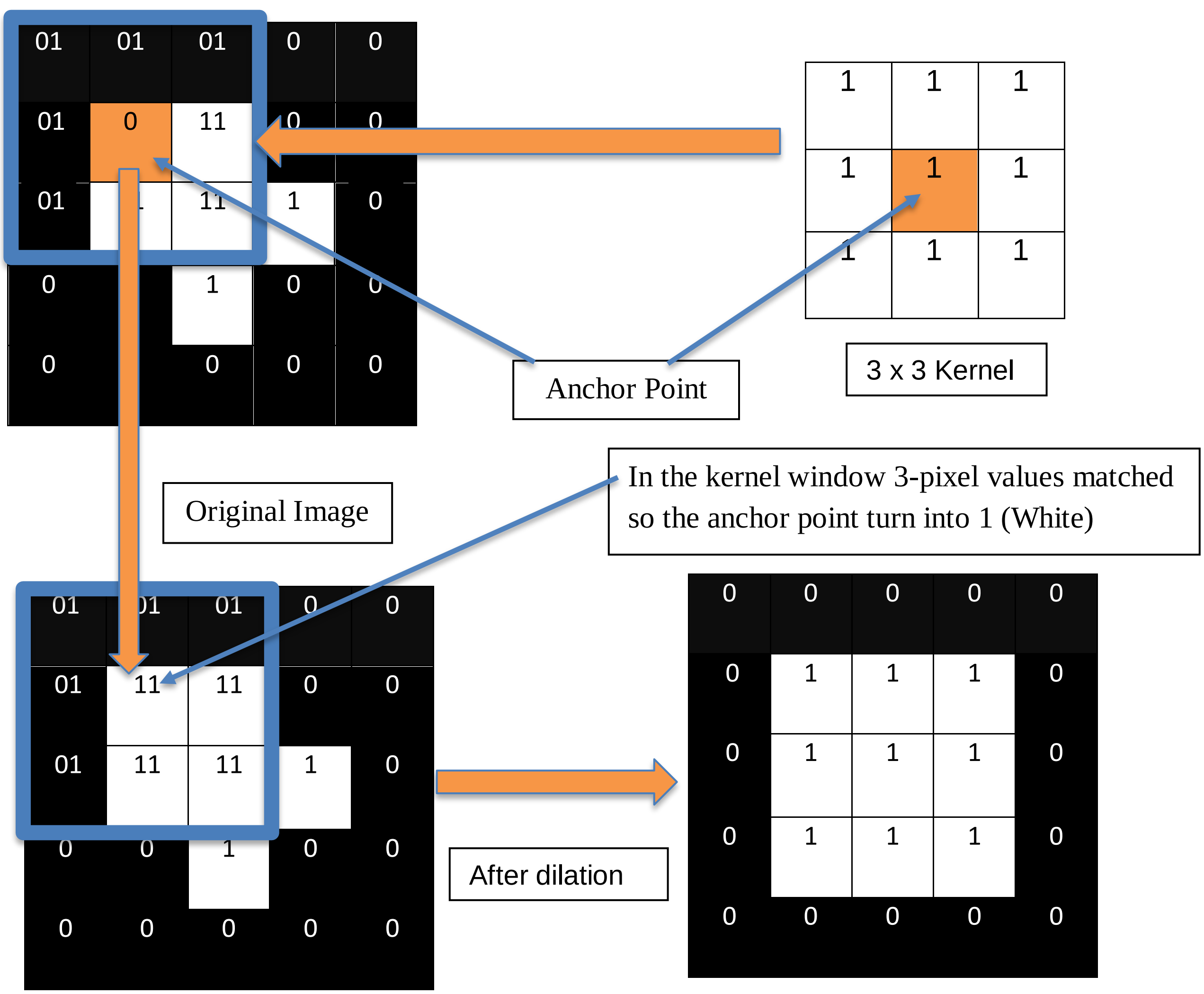


Figure: 4.6 The Process of dilation

- **Step 9 –**

Contours are identified within the binary mask, representing distinct areas of motion.

- **Step 10 –**

Rectangles are drawn around detected objects based on contours, and a descriptive label ("motion detected") is added to each bounding box.

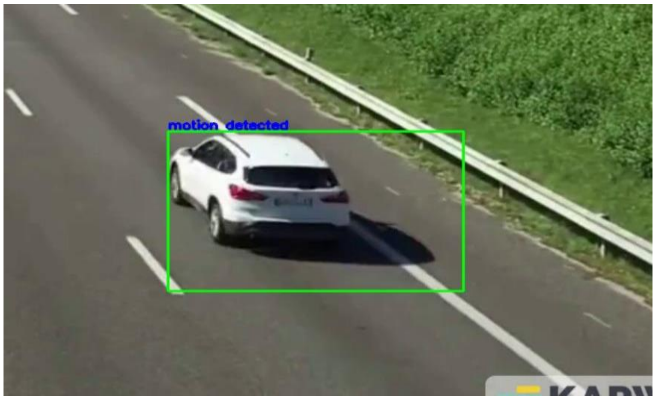
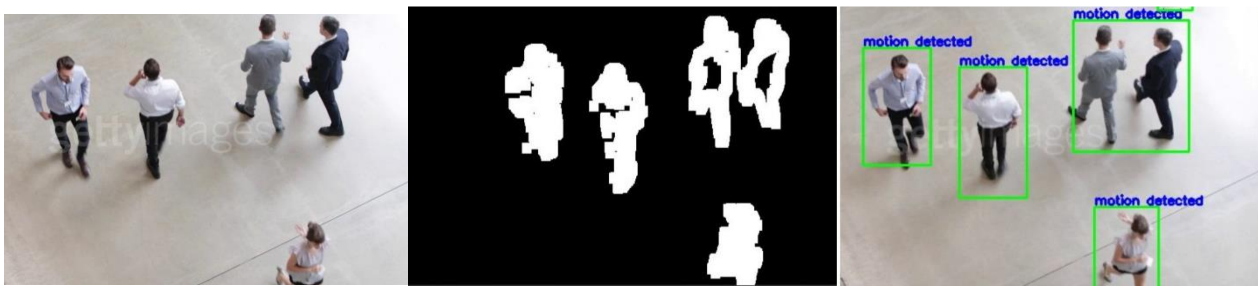


Figure:4.7 Detected moving object with bounding box

This solution provides a foundational framework for basic moving object detection, suitable for applications such as surveillance, security, and traffic monitoring. The effectiveness of the system can be further tuned by adjusting parameters, such as the AKAZE feature detection settings, blur parameters, and contour area threshold.

6. Experimental results

The test results are provided below…



Original Image | Binary Image | Detected Objects

Figure: 4.8 Testing on Human video (video Links-[2])

Original Image | Binary Image | Detected Object

Figure: 4.9 Testing on car video (video Links- [3])

In the above results of testing of the proposed moving object detection method is shown. This method is very simple and has very less time complexity as it does not use very high complex algorithm.

Overall, this method gives descent detection result with simple indoor-outdoor videos which recorded through a static camera.

7. Conclusion

Moving object detection is a crucial task in video surveillance systems. The process involves detecting and tracking moving objects in video sequences captured by static or CCTV cameras. The algorithm typically involves steps such as sampling the video into frames, converting RGB frames to grayscale, applying entropy values, background subtraction, and morphology-based object detection. The detection of moving objects is essential for video surveillance and has applications in various domains such as traffic monitoring and indoor security. While the algorithm may not be highly accurate, it provides satisfactory results within a short time frame.

This method has two shortcomings, one it works only with a static background. If the background is also dynamic it could not detect, as it cannot differentiate between foreground and background. Mostly it will consider the whole frame as one moving object. Another one is it cannot work properly on those videos, which have highly congested moving objects, meaning too many moving objects at very short places. It will consider objects as a single object with a single boundary box so the accuracy will drastically decrease.

8. References

1. Li, Y., Bao, H., Ge, Z., Yang, J., Sun, J., & Li, Z. (2023). BEVStereo: Enhancing depth estimation in multi-view 3D object detection with temporal stereo. Proceedings of the ... AAAI Conference on Artificial Intelligence. AAAI Conference on Artificial Intelligence, 37(2), 1486–1494. https://doi.org/10.1609/aaai.v37i2.25234

2. Zhou, Z., Zhao, X., Wang, Y., Wang, P., & Foroosh, H. (2022). CenterFormer: Center-based transformer for 3D object detection. In Lecture Notes in Computer Science (pp. 496–513). Springer Nature Switzerland.

3. Balasubramaniam, A., & Pasricha, S. (2022). Object detection in autonomous vehicles: Status and open challenges. In arXiv [cs.CV].

4. Sharma, T., Debaque, B., Duclos, N., Chehri, A., Kinder, B., & Fortier, P. (2022). Deep learning-based object detection and scene perception under bad weather conditions.

5. Alotaibi, M. F., Omri, M., Abdel-Khalek, S., Khalil, E., & Mansour, R. F. (2022). Computational intelligence-based Harmony search algorithm for real-time object detection and tracking in video surveillance systems.

6. Jiménez-Bravo, D. M., Lozano Murciego, Á., Sales Mendes, A., Sánchez San Blás, H., & Bajo, J. (2022). Multi-object tracking in traffic environments: A systematic literature review. Neurocomputing, 494, 43–55. https://doi.org/10.1016/j.neucom.2022.04.087

7. Viola, P.; Jones, M.J. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 8–14 December 2001; Volume 1, pp. I-511–I-518

8. Viola, P.; Jones, M.J. Robust real-time face detection. Int. J. Comput. Vis. 2004, 57, 137–154

9. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. comput. Vis. Pattern Recognit. 2005, 1, 886–893

10. Lowe, D.G. Object recognition from local scale-invariant features. Comput. Vis. 1999, 2, 1150–1157.

11. Lowe, D.G. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 2004, 60, 91–110.

12. Belongie, S.; Malik, J.; Puzicha, J. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. 2002, 24, 509–522.

13. Felzenszwalb, P.; McAllester, D.; Ramanan, D. A discriminatively trained, multiscale, deformable part model. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.

14. Zhao, Z.Q.; Zheng, P.; Xu, S. Object detection with deep learning: A review. IEEE Trans. Neural Netw. Learn. Syst. 2019, 30, 3212–3232. [CrossRef]

15. Wu, X.; Sahoo, D.; Hoi, S.C.H. Recent advances in deep learning for object detection. Neurocomputing 2020, 396, 39–64.

16. Xiao, Y.; Tian, Z.; Yu, J. A review of object detection based on deep learning. Multimed. Tools Appl. 2020, 79, 23729–23791.

17. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.

18. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. Eur. Conf. Comput. Vis. 2014, 346–361.

19. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Processing Syst. 2015, 28, 91–99.

20. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. Ssd: Single shot multibox detector. In European Conference on Computer Vision; Springer: Cham, Switzerland, 2016; pp. 21–37

21. Liu, S.; Liu, D.; Srivastava, G. Overview and methods of correlation filter algorithms in object tracking. Complex Intell. Syst. 2021, 7, 1895–1917.

22. Bolme, D.S.; Beveridge, J.R.; Draper, B.A.; Lui, Y.M. Visual object tracking using adaptive correlation filters. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 2544–2550

23. Qi, Y.; Zhang, S.; Qin, L.; Yao, H.; Huang, Q.; Lim, J.; Yang, M.-H. Hedged deep tracking. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.

24. Marvasti-Zadeh, S.M.; Cheng, L.; Ghanei-Yakhdan, H. Deep learning for visual tracking: A comprehensive survey. IEEE Trans. Intell. Transp. Syst. 2021;

25. Zhang, D.; Guo, J.; Lei, X.; Zhu, C. A High-Speed Vision-Based Sensor for Dynamic Vibration Analysis Using Fast Motion Extraction Algorithms. Sensors 2016, 16, 572

26. Guo, W.; Gao, J.; Tian, Y.; Yu, F.; Feng, Z. SAFS: Object Tracking Algorithm Based on Self-Adaptive Feature Selection. Sensors 2021, 21, 4030.

27. Zhang, K.; Liu, Q.; Wu, Y.; Yang, M.-H. Robust visual tracking via convolutional networks without training. IEEE Trans. Image Processing 2016, 25, 1779–1792

28. Zhang, T.; Liu, S.; Xu, C.; Yan, S.; Ghanem, B.; Ahuja, N.; Yang, M.-H. Structural sparse tracking. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 150–158.

29. Ma, C.; Huang, J.-B.; Yang, X.; Yang, M.-H. Hierarchical convolutional features for visual tracking. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 3074–3082.

30. S. A. K. Tareen and Z. Saleem, "A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK," 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, 2018, pp. 1-10, doi: 10.1109/ICOMET.2018.8346440

31. H. K. Ryu, J. K. Lee, E. T. Hwang, L. Jing, H. H. Lee and W. H. Choi, "A new corner detection method of gray-level image using Hessian matrix," 2007 International Forum on Strategic Technology, Ulaanbaatar, 2007, pp. 537-540, doi: 10.1109/IFOST.2007.4798654.

32. T. L. Mengko, T. Adiono, H. Setyawan and R. Setiadarma, "Design and implementation of object detection and classification system based on deformable template algorithm," IEEE. APCCAS 1998. 1998 IEEE Asia-Pacific Conference on Circuits and Systems. Microelectronics and Integrating Systems. Proceedings (Cat. No.98EX242), Chiangmai, Thailand, 1998, pp. 311-314, doi: 10.1109/APCCAS.1998.743755.

33. M. Bhowmik, A. Sarkar and R. Das, "Shannon entropy based fuzzy distance norm for pixel classification in remote sensing imagery," Proceedings of the 2015 Third International Conference on Computer, Communication, Control and Information Technology (C3IT), Hooghly, 2015, pp. 1-6, doi: 10.1109/C3IT.2015.7060200.

Video Links

1. Video 1 - Haque, M. Z. U. (n.d.). Simple-Motion-Detection-Algorithm: An OpenCV based implementation of Motion Detection Algorithm in python.

2. Video - 2 ReeldealHD Ltd. (n.d.). Overhead shot of busy office reception. Getty Images. Retrieved November 30, 2023, from gettyimages.

3. Video – 3 ReeldealHD Ltd. (n.d.). Traffic CCTV cars footage.Getty Images. Retrieved November 30, 2023, from gettyimages.