The Background Study for An Optical Flow Analysis Based Real Time Intelligent Video Surveillance System for People Safety

Dheeraj D[1], Abhishek V Tatachar[2], Shivanee Kondur[3], Sumukh B K[4], Vishwas K V[5]
[1]*Assistant Professor, Department of Information Science and Engineering, Global Academy of Technology*
[2,3,4,5] *Undergraduate Students, Department of Information Science and Engineering, Global Academy of Technology*

***Abstract* - The necessity of security exists in the society from a very long time and humans especially have felt the need to secure their life and property. With such a necessity there has always been an inclination towards the need of effective surveillance systems – both manual and intelligence driven. Research in this domain has been very active, given its social relevance in today's world. There have been several attempts made to detect intrusions and anomaly activities with these surveillance systems to assure security in the environment under consideration. The goal of this paper is to offer a background study on surveillance systems that used motion estimation and optical flow analysis to detect such unusual actions. The words "optical flow estimation" and "optical flow analysis" have been used interchangeably throughout this research.**

***Index Terms* - Motion Detection, Motion Estimation, Surveillance, Optical Flow, Occlusion Handling.**

1.INTRODUCTION

Safety and protection have always been a need for the human race. Safeguarding one's life and property has become the priority for people in today's world. Households, workplaces, and public places have started implementing a number of protective measures to ensure safety. One such preventive measure is the use of surveillance systems. The first video surveillance system was built to monitor the launch of missiles, in the year 1942, in Nazi occupied Germany. In the past decade the requirement for surveillance systems have increased. With the advancement of technology, and availability of resources, people have shown an increasing interest in the field of Artificial Intelligence (AI) and Computer Vision in the past few years, especially in the field of monitoring and surveillance. Research in this domain is being continuously conducted and several systems have been implemented for the purpose of surveillance. The various purposes for which the surveillance systems have been implemented include object detection and tracking [1, 2], intrusion detection [3, 4, 5], anomaly event detection [6, 7, 8], disaster management [9, 10], medical monitoring and diagnosis [11, 12], forensics [13, 14] and many more.

Even within computer vision, several approaches have been employed for surveillance. One common approach is to detect an object and then track it's motion, to classify the event as a normal event or an anomaly event. Optical flow analysis is a technique that is used to estimate motion between consecutive images. Optical flow estimation provides a pretty good approximation of the real physical motion projected onto the picture plane in many cases [15]. With the recent advancement in the technology, optical flow has been used in a wide range of applications and has proven to be a useful method for motion estimation [16, 17]. Also, better and more advanced optical flow estimation algorithms such as the Horn-Schunck algorithm [18], Lucas-Kanade [19] and Farneback Optical Flow [20] have come into light. We can view the optical flow problem as shown in figure 1.1. In the first frame the pixels are located at the position (x, y) at time t. After a time interval say dt, the pixels would have undergone a displacement of dx and dy on the x and y plane respectively. The new position of the pixels can be obtained at (x+dx, y+dy) at time d+dt.

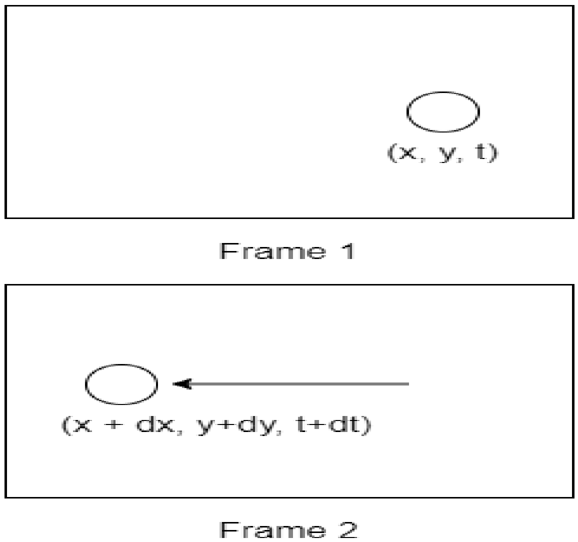

Figure 1.1 Pixels displacement over time

This paper is structured as follows. We continue form this point with the literature survey that acts as a background study for a surveillance system implemented using optical flow estimation. The we go ahead and list the findings in a tabular format to bring out the outcomes of the study. This paper aims at reviewing certain key literatures that are essential for the implementation of the system.

2. LITERATURE SURVEY

Video Surveillance in real time videos can be a complex task and it is very much possible to encounter a number of problems in the process. In [21] The authors identified five major challenges in motion detection – problems in identifying entities in a background with complex structures, anomaly event detection, problem of occlusions, problem with deformation and effective tracking, and intensity. The authors have also proposed some probable solutions to overcome these problems. Each of these problems is an hinderance in efficient tracking of motion of objects.

In [22], the authors present a method for detecting moving objects that combines background adjustment with object detection and consists of three modules: motion detection, object detection, and object matching. Using the background compensation technique, the motion detecting module will obtain a rough motion position. In this paper, the YOLOv3-Small Objects Detection (SOD) which is a better and improved version of YOLOv3, is used in the object detection module for correctly identifying the position of objects. The motion detection and object detection findings are matched in the object-matching module, and temporal and spatial forms of information from surrounding frames are employed to recollect the missing detections. In the background compensation technique, the motion parameter model is built utilizing the coordinate relationships of the matched feature points. The SIFT method was chosen after the examination of many point feature identification algorithms, by the authors. While satisfying the time limitations, using SIFT method, it was possible to extract the maximum number of feature points with the largest dispersion. This proposed technique outperformed the available methods, and it was effective in correctly identifying the moving objects in a number of settings, according to their findings.

In [23] the authors have offered a method for detecting anomalous occurrences. In this, the segments of the video with large scores for anomaly would be identified as anomalous event occurrence. Each video will be broken into an even number of segments that do not overlap. A video that comprises of an anomalous segment is considered positive, whereas one without ana anomalous segment is considered negative. A positive/negative video is used as a positive/negative bag in multiple instance learning, and the parts are used as examples. The ranking technique may be used to produce scores for each video segment, and those video segments that have high anomaly scores are deemed anomaly events.

In [24] The authors provide a tracker-independent, computationally efficient occlusion detection and management system that excels in scenarios with highly regularized motion patterns, like transportation. Similarly, to numerous previously described techniques, the proposed methodology decouples occlusion detection from occlusion management. They demonstrated that the proposed method enhances the performance of a number of popular object trackers through experimental validation. Occlusion detection and management are separated in the proposed technique. During the detection stage, the algorithm determines whether and where static occluding components are present in the scene. The direction in which a tracked object approaches an occlusion is defined at this stage. The object's pixels are then copied onto the pixels of the discovered occlusion using the direction. Throughout the handling stage, the method repeats pixel values from neighboring regions, as well as detected occlusions at frame locations when occlusions are present. As an object approaches towards an occlusion, pixels with the values representative of the item's looks are replicated onto the related pixels of the occlusion, reducing the occlusion's impact on the tracking technique. Because the proposed algorithm performs best in cases with regularized motion conditions, they achieve notable tracking performance gains at reduced computational cost by making use of assumptions on the basal patterns of movement of the objects that are being tracked as they move the scene. Because they work best in settings with regularized motion circumstances, they are well suited for transportation and surveillance applications.

In [25] the authors provided a technique in which they established an accurate motion estimation method for determining the optical flow of a moving object based on the Lucas–Kanade algorithm. The effectiveness of our precise real-time implementation was confirmed by the testing findings, which showed that it outperformed the conventional computation. Lucas and Kanade proposed an outstanding approach for evaluating the mobility of significant features by comparing two subsequent frames in their research An Iterative Image Registration Technique with a Stereo Vision Application [26]. To assess the precision of the optical flow estimation, the authors used the Middlebury Dataset to test the suggested technique. This proposed approach was tested on a Raspberry Pi 4 local treatment zone and found to be very effective. Experiments demonstrated that their accurate real-time implementation beat conventional calculation, proving its usefulness.

In [27] the authors describe an object identification and tracking system that can recognize, categorize, and track moving objects in real time. Both color and greyscale input footage from stationary and moving cameras can be used by the system. An optical flow is used in the proposed system to handle moving object detection and tracking. This method allows the tracking system to perform more accurately in both indoor and outdoor environments. The performance of the object recognition and tracking system is evaluated using several types of input videos. The suggested tracker is applied to the input videos with various classifications. There are two forms of performance analysis: qualitative analysis and quantitative analysis. In circumstances including single and multiple object tracking, as well as viewpoint adjustments, the proposed system outperformed.

[28] provides a way to detect objects based on RGB and Optical Flow Analysis. The authors have discussed four important aspects - Flow-guided Feature Aggregation, Flow-guided Partial Warp (FGPW), Label Consistency Check, Autonomous Vehicles Simulator. On numerous frames, they discovered that YOLO v3 produced incorrect bounding boxes. Using the FGPW, however, YOLO v3 produces the correct findings based on the updated aggregate feature maps. They manually label the video output and train the model to evaluate the FGPW VID technique that was recommended. Following FGPW, the improved YOLO v3 outputs remove inaccurate detections and optimise bounding box placements in three common scenarios. The addition of other warped feature maps strengthens its feature, resulting in an accurate detection. People find it difficult to prepare training data since a second video has around 25 frames, each of which must be manually labelled.

Along with these, few more literatures were reviewed to find the outcomes, the results of which are provided in the following discussion table.

Table 2.1: Literature Review and Findings

Sl. No	Citation	Year	Methodology / Algorithms used	Remarks
1	Karthika Pragadeeswari, et Al [21]	2019	Highlights five major challenges in motion estimation and initial methods to overcome these challenges.	Identified key challenges in motion estimation and provided solutions to the challenges faced.
2	Juncai Zhu, et Al [22]	2020	YOLOv3-Small Objects Detection, SIFT Algorithm.	For motion detection – a correct threshold value aids motion detection. For object detection - YOLOv3-SOD model yielded a higher precision and recall rate.
3	Wangli Hao, et Al. [23]	2020	A deep convolutional network (Convent)	To improve anomalous event detection performance, the proposed technique can make advantage of complementary information from the two streams buried in the movie.

				The fusion of the RGB stream and Focus stream provided better results.
4	Matthew Shreve, et Al [24]	2015	Occlusion detection – Gaussian Mixtures for foreground segmentation.	Higher precision and recall rate when occlusions are handled. Identified the three factors that affected performance. Average precision and recall value obtained was 0.74 and 0.6 respectively for static occlusion detection performance.
5	Anis Ammar Hana Ben Fredj. et Al [25]	2021	Lucas-Kanade, Iterative refinement,Pyramidal implementation	Algorithm was highly parallelizable, particularly for the pyramid building process.
6	Sanap, et Al [27]	2017	Object Detection, Object Classification, Object Tracking	Object detection, object segmentation, and then object tracking are all steps in the object detection and tracking process
7	Shu nyao Zhang, et Al [28]	2019	Flow-guided feature aggregation, Flow-guide partial warp, label consistency check, autonomous vehicle simulator.	They discovered that yolo v3 produced incorrect bounding boxes, thus they employed FGPW, YOLO v3, which produces correct results based on aggregating feature maps. Their method had a recall of 0.86, precesion of 0.96, and an F1 score of 0.91 in their investigation.
8	Keshav Bhandari, et Al [29]	2021	They followed three important stages to provide better model like transformation, Intermediate refinement, final refinement.	They presented two kinds of result – qualitative and quantitative results. The quantitative results explained about the relationship convergence rate and the number of layer and how well their model performed after refinement. In the qualitative results, they observe that warping of flow helps in preserving the spherical nature of the video. They also note that the algorithm did not provide expected results for certain cases.
9	K.kalirajan, et Al [30]	2015	Object identification, object localisation, and video compression	The paper produced a four part results – Performance Analysis, Qualitative Analysis, Comparative Study and Future Directions. Qualitative results provides a way to compute he occlusion rate. From the comparative study they inferred that their proposed scheme excelled under complex environments.
10	Stefano SAVIAN, et Al. [31]	2020	Variation method, Batch-based method, Refinement method. CNN-UNet architecture. Coarse-to-Fine Iterative Refinement	Increased Action Recognition accuracy
11	Prasad. D. Garje, et Al. [32]	2018	RGB to HSV conversion, emphasis on pre-processing, Gunnar Farneback algorithm. Categorisation based on threshold value.	Categorised activities as anomaly activity and normal activity based on intensity rating being greater or smaller than a threshold.
12	Muhammad Kamal Hossen, et Al [33]	2016	Motion detection, estimating optical flux, Optical flow, HornSchunck optical method	The monitoring system's motion detection circuitry is computationally quicker, with a motion detection accuracy of around 98.6%.
13	Yan Wang, et Al [34]	2019	FlowNet 2.0, FGBT, YOLO-v3, LSS:	After Flow-guided Partial Warp, the improved YOLOv3 will eliminate the incorrect identifications and optimise the positions of the bounding box in the three usual circumstances. A tight loop is formed between the simulator and the model, the simulator is used to train the model, and the output of the simulator is raised or lowered in response to the model's performance.
14	Ibrahim kajo, et Al [35]	2015	Brox method, Zach method	For many examples of crowd movement, the Zach, Brox, and BA approaches are effective, although they are less precise than the standard NL method.
15	Alwyn Mathew, et Al [36]	2017	Convolutional modules, Transfer learning, Google's inception v3 model, Principal Component Analysis and t-Distributed Stochastic Neighbour Embedding	The research exhibits a deep CNN's capability in the realm of object detection.

3. CONCLUSION

For safety, surveillance has become a necessary, feasible and convenient option. In the recent past, extensive amount of work has been done in this field to identify more effective and efficient methods to build intelligent surveillance systems. This paper concentrates at identifying the existent research and implementation work done in the field of Artificial Intelligence and Computer Vision, especially in

Optical Flow Estimation to perform surveillance on real-time video.

It is important to apprehend the necessity to process real time videos and be open to address scenarios that might occur in the real-time. Several challenges known to occur while processing real time videos include occlusions, anomaly or intrusion activity detection, light intensity variation, distorted and disoriented images and so on. It becomes necessary to consider such situations as a part of the scenario and address them carefully.

On reviewing several articles, we made a few observations in the systems that are used for surveillance.

- Occlusions are a part of the real time videos and can be handled. The solutions include pixel replication and cutmix augmentation.
- One of the best available object detection models is YOLOv3 with Flow-guided Partial Wrap (FGPW)
- Lucas-Kanade algorithm is an ideal approach to optical flow estimation.

From our perspective, the suggested techniques would provide a better performance in real-time implementation.

CONFLICT OF INTERESTS

The authors project no conflict of interest in publishing of this paper.

REFERENCES

[1] J. S. Kim, D. H. Yeom and Y. H. Joo, "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems," in IEEE Transactions on Consumer Electronics, vol. 57, no. 3, pp. 1165-1170, August 2011, doi: 10.1109/TCE.2011.6018870.

[2] U. M. Prakash and V. G. Thamaraiselvi, "Detecting and tracking of multiple moving objects for intelligent video surveillance systems," Second International Conference on Current Trends In Engineering and Technology - ICCTET 2014, 2014, pp. 253-257, doi: 10.1109/ICCTET.2014.6966297.

[3] Kim, Seunghyun, Suchang Lim and Do Yeon Kim. “Intelligent intrusion detection system featuring a virtual fence, active intruder detection, classification, tracking, and action recognition.” Annals of Nuclear Energy 112 (2018): 845-855.

[4] Y. -L. Zhang, Z. -Q. Zhang, G. Xiao, R. -D. Wang and X. He, "Perimeter intrusion detection based on intelligent video analysis," 2015 15th International Conference on Control, Automation and Systems (ICCAS), 2015, pp. 1199-1204, doi: 10.1109/ICCAS.2015.7364811.

[5] M. T. Awad, S. M. Aldaw, S. M. Aldaw and B. A. r. Osman, "Video Security System for Intrusion Detection," 2019 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), 2019, pp. 01-04, doi: 10.1109/ICCCEEE46830.2019.9070823.

[6] Y. Wang, C. Fan and J. Chen, "Traffic Camera Anomaly Detection," 2014 22nd International Conference on Pattern Recognition, 2014, pp. 4642-4647, doi: 10.1109/ICPR.2014.794.

[7] J. Henrio and T. Nakashima, "Anomaly Detection in Videos Recorded by Drones in a Surveillance Context," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2018, pp. 2503-2508, doi: 10.1109/SMC.2018.00429.

[8] R. J. Franklin, Mohana and V. Dabbagol, "Anomaly Detection in Videos for Video Surveillance Applications using Neural Networks," 2020 Fourth International Conference on Inventive Systems and Control (ICISC), 2020, pp. 632-637, doi: 10.1109/I CISC4 7916 .2020 .9171212.

[9] A. Sehrawat, T. A. Choudhury and G. Raj, "Surveillance drone for disaster management and military security," 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017, pp. 470-475, doi: 10.1109/CCAA.2017.8229846.

[10] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho and S. W. Baik, "Convolutional Neural Networks Based Fire Detection in Surveillance Videos," in IEEE Access, vol. 6, pp. 18174-18183, 2018, doi: 10.1109/ACCESS.2018.2812835.

[11] David E. Hall, Laura Eubanks, LCSW*; Swarnalatha Meyyazhagan, Richard D. Kenney, Sherry Cochran Johnson; Evaluation of Covert Video Surveillance in the Diagnosis of Munchausen Syndrome by Proxy: Lessons From 41 Cases. Pediatrics June 2000; 105 (6): 1305–1312. 10.1542/peds.105.6.1305

[12] H. Foroughi, B. S. Aski and H. Pourreza, "Intelligent video surveillance for monitoring fall detection of elderly in home environments," 2008 11th International Conference on Computer and Information Technology, 2008, pp. 219-224, doi: 10.1109/ICCITECHN.2008.4803020.

[13] J. Xiao, S. Li and Q. Xu, "Video-Based Evidence Analysis and Extraction in Digital Forensic Investigation," in IEEE Access, vol. 7, pp. 55432-55442, 2019, doi: 10.1109/ACCESS.2019.2913648.

[14] B. Yogameena and K. Sindhu Priya, "Synoptic video based human crowd behavior analysis for forensic video surveillance," 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR), 2015, pp. 1-6, doi: 10.1109/ICAPR.2015.7050662.

[15] Turaga, P., Chellappa, R., & Veeraraghavan, A. (2010). Advances in Video-Based Human Activity Analysis. Advances in Computers, 237–290. doi:10.1016/s0065-2458(10)80007-5

[16] Kai-Tai Song and Jui-Hsiang Huang, "Fast optical flow estimation and its application to real-time obstacle avoidance," Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164), 2001, pp. 2891-2896 vol.3, doi: 10.1109/ROBOT.2001.933060.

[17] C. -H. Wu, J. -W. Hsieh, C. -Y. Wang and C. -H. Ho, "Marine Pollution Detection based on Deep Learning and Optical Flow," 2020 International Computer Symposium (ICS), 2020, pp. 376-381, doi: 10.1109/ICS51289.2020.00081.

[18] J. a/l Kanawathi, S. S. Mokri, N. Ibrahim, A. Hussain and M. M. Mustafa, "Motion detection using Horn Schunck algorithm and implementation," 2009 International Conference on Electrical Engineering and Informatics, 2009, pp. 83-87, doi: 10.1109/ICEEI.2009.5254812.

[19] L. Y. Siong, S. S. Mokri, A. Hussain, N. Ibrahim and M. M. Mustafa, "Motion detection using Lucas Kanade algorithm and application enhancement," 2009 International Conference on Electrical Engineering and Informatics, 2009, pp. 537-542, doi: 10.1109/ICEEI.2009.5254757.

[20] Z. Pan, Y. Jin, X. Jiang and J. Wu, "An FPGA-Optimized Architecture of Real-time Farneback Optical Flow," 2020 IEEE 28th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM), 2020, pp. 223-223, doi: 10.1109/FCCM48280.2020.00054.

[21] K. P. C. and Y. G., "Challenges and Solutions in Motion Surveillance - A Survey," 2019 International Conference on Communication and Signal Processing (ICCSP), 2019, pp. 0960-0964, doi: 10.1109/ICCSP.2019.8698046.

[22] Zhu, J., Wang, Z., Wang, S., & Chen, S. (2020). Moving Object Detection Based on Background Compensation and Deep Learning. Symmetry, 12(12), 1965. doi:10.3390/sym12121965

[23] Wangli Hao, Ruixian Zhang, Shancang Li, Junyu Li, Fuzhong Li, Shanshan Zhao, Wuping Zhang, "Anomaly Event Detection in Security Surveillance Using Two-Stream Based Model", Security and Communication Networks, vol. 2020, Article ID 8876056, 15 pages, 2020. https://doi.org/10.1155/2020/8876056

[24] M. Shreve, E. A. Bernal, Q. Li and R. P. Loce, "Static Occlusion Detection and Handling in Transportation Videos," 2015 IEEE 18th International Conference on Intelligent Transportation Systems, 2015, pp. 638-643, doi: 10.1109/ITSC.2015.110.

[25] Ammar, A.; Fredj, H.B.; Souani, C. Accurate Realtime Motion Estimation Using Optical Flow on an Embedded System. Electronics 2021, 10, 2164. https://doi.org/10.3390/electronics10172164

[26] Bruce D. Lucas and Takeo Kanade. 1981. An iterative image registration technique with an application to stereo vision. In <i>Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2</i> (<i>IJCAI'81</i>). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 674–679.

[27] Moving Object Detection By Using Optical Flow, Sanap Vaishali Arjun, et Al, 2017

[28] Shunyao Zhang, Tian Wang, Chuanyun Wang, Yan Wang, Guangcun Shan, et al. Video Object Detection Base on RGB and Optical Flow Analysis. 2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI), Sep 2019, Xi'an, China. ⟨10.1109/CCHI.2019.8901921⟩. ⟨hal-02373483⟩

[29] K. Bhandari, Z. Zong and Y. Yan, "Revisiting Optical Flow Estimation in 360 Videos," 2020 25th International Conference on Pattern

Recognition (ICPR), 2021, pp. 8196-8203, doi: 10.1109/ICPR48806.2021.9412035.

[30] Kalirajan, K. & Sudha, M. (2015). Moving Object Detection for Video Surveillance. The Scientific World Journal. 2015.1-10.10.1155/2015/9074 69.

[31] Savian, Stefano, Mehdi Elahi and Tammam Tillo. "Optical Flow Estimation with Deep Learning, a Survey on Recent Advances." Deep Biometrics (2020): n. pag.

[32] P. D. Garje, M. S. Nagmode and K. C. Davakhar, "Optical Flow Based Violence Detection in Video Surveillance," 2018 International Conference On Advances in Communication and Computing Technology (ICACCT), 2018, pp. 208-212, doi: 10.1109/ICACCT.2018.8529501.

[33] M. K. Hossen and S. H. Tuli, "A surveillance system based on motion detection and motion estimation using optical flow," 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), 2016, pp. 646-651, doi: 10.1109/ICIEV.2016.7760081.

[34] S. Zhang, T. Wang, C. Wang, Y. Wang, G. Shan and H. Snoussi, "Video Object Detection Base on RGB and Optical Flow Analysis," 2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI), 2019, pp. 280-284, doi: 10.1109/CCHI.2019.8901921.

[35] I. Kajo, A. S. Malik and N. Kamel, "Motion estimation of crowd flow using optical flow techniques: A review," 2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS), 2015, pp. 1-9, doi: 10.1109/ICSPCS.2015.7391778.

[36] Mathew, Alwyn & Mathew, Jimson & Govind, Mahesh & Mooppan, Asif. (2017). An Improved Transfer learning Approach for Intrusion Detection. Procedia Computer Science. 115. 251-257. 10.1016/j.procs.2017.09.132.