

Dynamic Scene Understanding in Surveillance Systems Using Optical Flow

Dr. Moiz Abdul Hussain¹

¹Associate Professor, Electrical Engineering Deptt. Dr. V. B. Kolte College of Engineering Malkapur, M.S., India

DOI: 10.5281/zenodo.19289387

ABSTRACT

With continuous progress in video analysis technologies, automated visual surveillance has become increasingly prominent in consumer and security-oriented applications. Human motion analysis primarily involves the detection, tracking, and identification of individuals in video sequences. This paper proposes a unified framework for reliable detection and tracking of moving persons in surveillance environments. Initially, foreground objects are detected using an optical flow-based motion estimation technique. Subsequently, an illumination-invariant template matching approach is integrated to ensure accurate and robust object tracking under varying lighting conditions. Experimental evaluations conducted on the Weizmann and self-developed datasets demonstrate that the proposed method offers superior discriminative performance, enhanced robustness, and reduced computational complexity when compared with conventional motion detection techniques. Furthermore, the simplicity of implementation makes the proposed system well suited for real-time and online surveillance applications.

Keywords: - Video surveillance, Movement Detection, Optical Flow, Template Matching.

1. INTRODUCTION

In recent years, the importance of human location and tracking has increased significantly due to its wide range of applications in security, defense, human-computer interaction, supervision, and intelligent monitoring systems. Human detection remains a challenging problem in high-level computer vision applications. The large intra-class variability of humans, arising from differences in posture, gestures, expressions, and body shapes, makes reliable detection difficult [1, 2]. Furthermore, background clutter, varying illumination, and changing weather conditions further complicate the detection process.

Object tracking involves associating a detected target across consecutive frames in a video sequence. Typically, the tracking process begins by detecting a moving object in the initial frame and subsequently maintaining its identity throughout the following frames. Visual tracking is inherently complex due to factors such as sudden object motion, partial or full occlusion, camera movement, and illumination variations [5]. Visual surveillance in dynamic environments has numerous applications, including security monitoring of residential areas and critical infrastructure, as well as traffic surveillance on urban roads and highways. A conventional surveillance system generally consists of three core components: motion detection, object tracking, and behavior analysis. The primary goal of such systems is to accurately detect, track, and classify objects of interest. Existing visual surveillance techniques for object detection and tracking can be broadly categorized into region-based, boundary-based, and feature point-based approaches [5, 7, 8]. Among these approaches, optical flow-based methods are widely used in automated surveillance systems for motion detection.

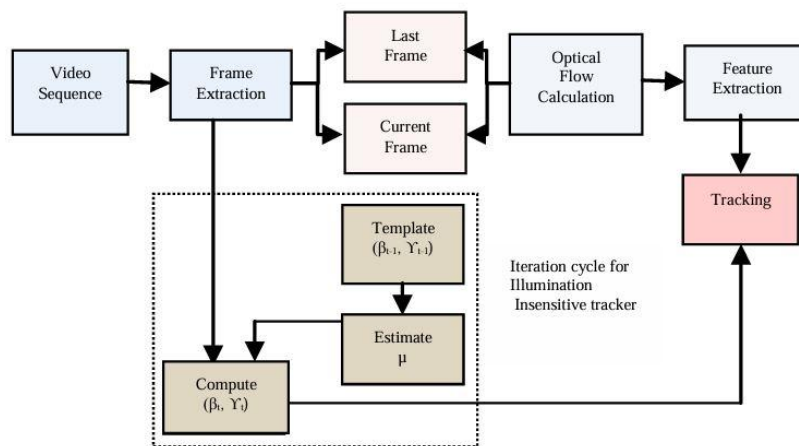


Fig 1: System Overview

Optical flow [10, 11] estimates motion by analyzing intensity variations between consecutive frames. However, optical flow cannot be computed locally using a single constraint, as motion has two components while only one independent equation is available per pixel; therefore, additional constraints are required. An effective intelligent tracking system must be capable of detecting genuine object motion while suppressing false detections caused by dynamic background elements such as moving clouds, rain, or waving objects. Illumination variations significantly affect both indoor and outdoor video sequences. In indoor environments, artificial lighting may cause rapid radiance fluctuations, occurring several times per minute, whereas cameras typically capture images at 25–30 frames per second. As a result, illumination changes influence the captured frames. Similarly, outdoor scenes are affected by rapid and unpredictable illumination changes due to sunlight variations, making robust surveillance in outdoor environments more challenging.

2. RELATED WORK

A substantial body of research exists on video surveillance systems for human motion detection and tracking. In our earlier study [12], an adaptive threshold initialization approach was introduced to segment objects from video sequences, based on the assumption that moving objects are visually distinct and do not overlap. Avidan employed a neural network-based classifier, specifically a Support Vector Machine (SVM), trained offline and later integrated into an optical flow framework for object tracking. Collins et al. utilized the inconsistency ratio between target objects and background classes to identify discriminative features suitable for tracking applications. Jepson et al. [23] modeled pixel intensities using a Gaussian mixture model and applied an online expectation-maximization (EM) algorithm to manage appearance variations of targets during tracking. Ross et al. proposed an online multiple instance learning approach that demonstrated effective object tracking in real-world scenarios, even in the presence of occlusion and illumination changes.

Hager and Belhumeur (1996) presented a tracking method that addresses appearance variations caused by lighting and pose changes using parametric models. However, since these techniques rely solely on generative representations of target objects and neglect background information, their performance degrades in cluttered environments. Existing human action recognition techniques are generally categorized into global and local representation approaches. Bao and Ravi et al. introduced methods for predicting activities such as walking, standing, running, and sitting by extracting features from raw accelerometer data and applying various machine learning algorithms. Nevertheless, these methods do not utilize the extracted information for indoor localization. Furthermore, as these approaches focus primarily on intra-image illumination variations and spatial characteristics rather than sequential inter-frame lighting changes, they represent a different category of illumination compensation techniques.

3. OBJECTIVES OF WORK

This work focuses on developing techniques for reliable motion detection and monitoring in video sequences. The proposed system is designed to:

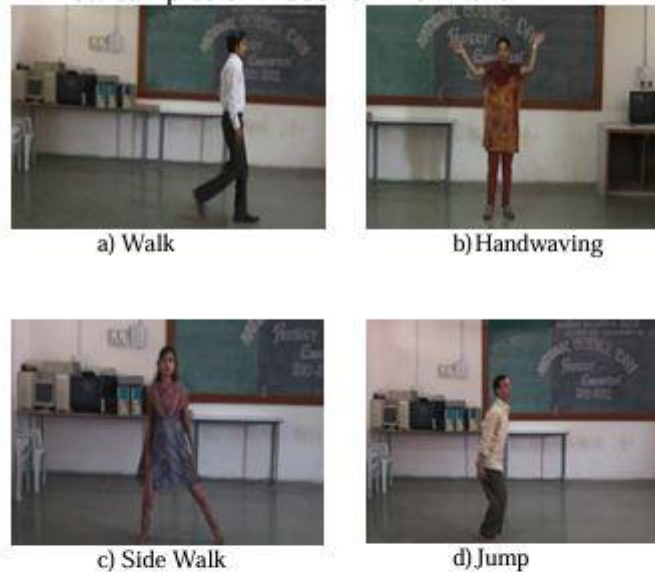
- **Automatically identify motion** within a surveillance area by employing optical flow to examine positional changes of objects across successive video frames.
- **Accurately track moving targets** in a stable and efficient manner, even under complex or dynamically changing scene conditions.
- **Improve situational awareness** by differentiating between regular and irregular movement behaviors, such as loitering or abrupt acceleration.

- **Increase surveillance reliability** by minimizing dependence on continuous human supervision, thereby reducing errors and operational costs.
- **Support real-time analysis**, allowing rapid detection and response to potential intrusions or suspicious activities.

4. DATABASE COLLECTION

In our experiment we used two different databases. One is Weizmann database which is publically available and another is our own database this database contains 86 sequences having five & four of classes of actions (Jump, Run, Walk, Side, Hand waving, Jogging) respectively performed by 19 different subjects in two different conditions d1-d2. d1 – Indoor Environment d2 - Outdoor (Playground) + illumination variations

Few samples of Indoor environment



Outdoor (Playground) + illumination variations Environment



Fig.2: Sample of Own database

5.METHODOLOGY

We present an approach based on optical flow estimation for human motion detection, combined with an illumination-invariant template matching technique for object tracking. Motion detection and tracking algorithms in video surveillance environments must address several challenges arising from real-world conditions, including gradual illumination variations, dynamic backgrounds, sensor noise, and other environmental disturbances. The primary goal of the proposed approach is to accurately detect moving objects

within video frames and reliably track them over time. The method effectively minimizes false detections that may occur due to lighting fluctuations or background motion. The overall architecture of the proposed system, illustrating its major processing stages, is shown in Fig. 1.

The procedure used to develop the motion detection and tracking system can be summarized as follows:

- Extract individual frames from the input video sequence.
- Select the current frame and the immediately preceding frame to compute optical flow between them.
- Derive motion-related features from the optical flow output.
- Estimate the mean intensity value (μ) from the image template.
- Compute the illumination compensation parameters (β_i, γ_i) for illumination-invariant iterative processing.
- Integrate the illumination compensation parameters obtained in Step 5 with the motion features from Step 3 to track the target object throughout the video sequence.

A. Optical Flow Computation

Optical flow methods aim to estimate a vector field that represents the apparent motion of each pixel between consecutive frames over time. This vector field provides essential information for motion analysis and tracking applications. For tracking humans or objects in video sequences, the optical flow constraint equation is employed to model pixel-level motion accurately.

B. Illumination-Invariant Template Matching

To mitigate the impact of illumination changes, a compensation-based template matching approach is employed. This method estimates lighting adjustment parameters for the current template to normalize its gray-level values with respect to a reference template. To reduce computational complexity, the standard template is replaced with the current template during matrix computation, enabling the direct estimation of motion compensation parameters β and γ . Consequently, the required matrix computation is performed only once, significantly reducing processing time. To further refine the tracking accuracy, the illumination compensation parameters are combined with centroid-based motion estimation. By integrating illumination-invariant template matching with optical flow-based feature extraction, the strengths of both techniques are utilized, resulting in improved robustness and optimized object tracking performance.

6. EXPERIMENTAL RESULTS

The proposed method is compared with the simple optical flow with adaptive threshold initialization [12] for the condition of detection precision and computational time. The experimental conditions are summarized in Table I. The optical flow method uses segmentation of boundary. To improve final estimation of optical flow field borders in successive images are matched, and a set of point limitation is derived.

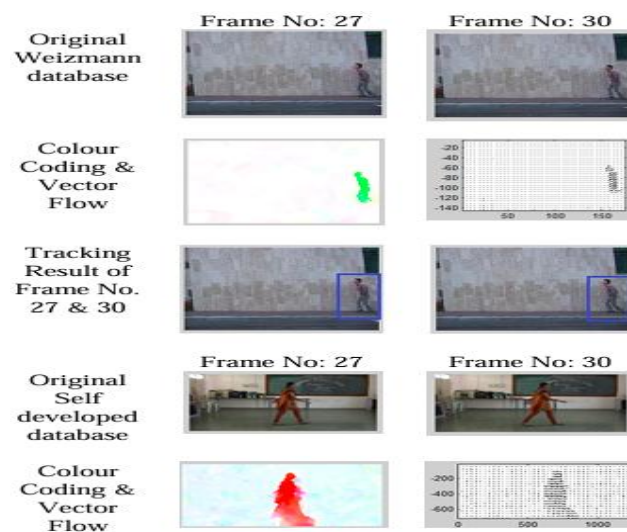


Table 1: Experiment Conditions

Sequence	1280 x720 pixels, 24 bit color 30 fps, 90 frames Motion JPEG
CPU	Corei3, 2.13GHz
RAM	4GB
OS	Windows 7.0
MATLAB	Version 7.10.0.499 (R2010a)

The proposed method can locate and track moving object and is examined for steady & moving camera. Figure shows the result for human motion detection & tracking result moving camera & dynamic background. From the result obtained it is clear that, algorithm achieves accurate target tracking not only in steady but also in dynamic conditions. The algorithm implemented has attained extremely competent & efficient results in messed up backgrounds and lighting variations. This shows the effectiveness of algorithm.

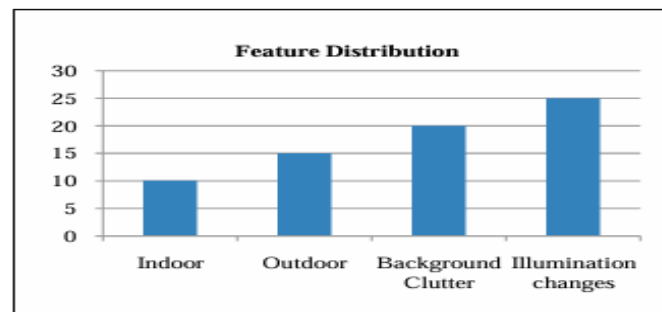


Fig.6: Range of deviation of feature distribution for different environment

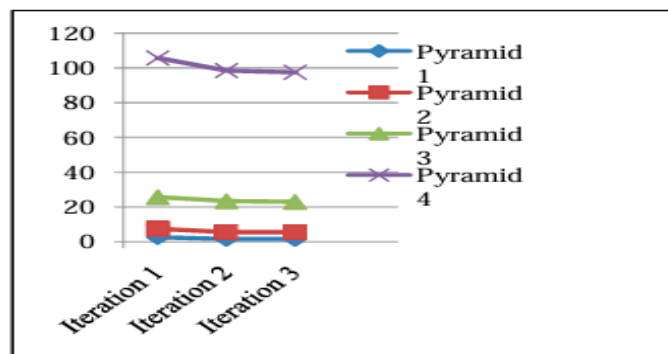


Fig.7: Norm of optical flow increment Vs No. of iterations for Optical flow calculation

From Fig 6, the number of features variation is very high if there is large change in the illumination of the image, as the pixel intensity changes to a larger extent compared to other environment conditions. Fig. 7 shows the representation of norm of optical flow for the different number of iterations.

7. CONCLUSION AND FUTURE WORK

This paper presents a robust approach for human motion localization and tracking that effectively addresses challenges such as camera instability and varying illumination conditions. The proposed system employs an efficient optical flow-based method to estimate pixel-level motion between consecutive frames. In addition, a template matching module is incorporated to mitigate the effects of illumination changes and dynamic background conditions. Experimental results demonstrate that the proposed method achieves high accuracy and fast performance across diverse environmental scenarios. Future work will focus on addressing occlusion handling issues, which are particularly relevant to social behavior analysis and realistic crowd simulation applications.

8. REFERENCES

- [1] D. Serby, E. K. Meier, and L. V. Gool, "Probabilistic Object Tracking Using Multiple Features, "IEEE Proc. Of International Conf on Pattern Recognition Intelligent Transportation Systems, Vol. 6, pp.43-53, March 2004.
- [2] L. Li, S. Ranganath, H. Weimin, and K. Sengupta, "Framework for Realtime Behavior Interpretation Form Traffic Video ", IEEE Tran. On Intelligen Transportation Systems, March 2005, Vol.6, No.1, pp.43-53.
- [3] R.T. Collins, A.J. Lipton, T. Kanade, Introduction to the special section on video surveillance, IEEE Trans. on Pattern Analysis and Machine Intelligence, 22 (8) (2000) 745-746.
- [4] S. Maybank, T. Tan, Introduction to special section on visual surveillance, International Journal of Computer Vision, 37 (2) (2000) 173-173.
- [5] N. A. OGALE, A survey of techniques for human detection from video, <http://www.cs.umd.edu/scholarlypapers/neetiPaper.pdf>, Department of Computer Science, University of Maryland.
- [6] P. Kumar, H. Weimin, I.U. Gu, and Q. Tian, "Statistical Modeling of Complex Backgrounds for Foreground Object Detection", IEEE Trans.on Image Processing, Vol.13, Mo.11, pp.43-53, November 2004.
- [7] Z. Zivkovi, "Improving the selection of feature points for tracking", In pattern Analysis and Applications, vol.7, no .2, Copyright Springer Verlag London Limited, 2004.

- [8] J. Lou, T. Tan, W. Hu, H. Yang, and S. H. Maybank, "3D Model based Vehicle Tracking", IEEE Trans. on Image Processing, Vol. 14, pp. 1561-1569, October 2005.
- [9] Jun Yang, Tusheng Lin, Bi Li, "Dual Frame Differences Based Background Extraction Algorithm", ICCP 2011 Proceedings, pp. 44-47, 2011.
- [10] I. Austvoll, A Study of the Yosemite Sequence Used as a Test Sequence for Estimation of Optical Flow, pages 659– 668. Lecture Notes in Computer Science. Springer Berlin/ Heidelberg, 2005.
- [11] L. Alvarez, R. Deriche, T. Papadopoulos, and J. Sanchez. Symmetrical dense optical flow estimation with occlusions detection. In Proc. ECCV, pages 721–735, 2002.
- [12] Moiz A. Hussain, Dr. G.U. Kharat, "Real Time Human Motion detection System" international conference on information, knowledge & research in Engineering, technology & sciences- icikr-ets-2012 .
- [13] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, pp. 62-66, 1979.
- [14] Rachel Kleinbauer, "Kalman Filtering Implementation with MatLab", University Stuttgart Institute of Geodesy.
- [15] Jun Yang, Tusheng Lin, Bi Li, Member, IEEE, "Dual Frame Differences Based Background Extraction Algorithm", ICCP Proceedings, 2011.
- [16] D. Serby, E. K. Meier, and L. V. Gool, "Probabilistic Object Tracking Using Multiple Features," IEEE Proc. Of International Conf on Pattern Recognition Intelligent Transportation Systems, vol. 6, pp. 43-53, March 2004.
- [17] E. Durucan and T. Ebrahimi. Moving object detection between multiple and color images. In IEEE Conf. Advanced Video and Signal Based Surveillance, pages 243–251, 2003.
- [18] Zalili Musa & Junzo Watada; "Multi-camera Tracking System in Large Area Case," in IEEE at Korea 2009.
- [19] Y. Dhome, N. Tronson, A. Vacavant, T. Chateau, C. Gabard, Y. Goyat, and D. Gruyer. A Benchmark for Background Subtraction Algorithms in Monocular Vision: a Comparative Study. In International Conference on Image Processing Theory, Tools and Applications (IPTA 2010). 7–10 July 2010.
- [20] Li Ying-hong, Pang Yi-gui, Li Zheng-xi and Liu Ya-li, "An intelligent tracking technology based on kalman and mean shift algorithm," IEEE Second Intl. Conf. on Computer Modelling and Simulation 2010.
- [21] Z. Zivkovi, "Improving the selection of feature points for tracking", In pattern Analysis and Applications, vol. 7, no. 2, Copyright Springer Verlag London Limited, 2004.
- [22] P. Perez, C. Hue, J. Vermaak, and M. Gangnet. "Color-Based Probabilistic Tracking", in Proceedings of the European Conference on Computer Vision (ECCV), 2002.
- [23] A. D. Jepson, D. J. Fleet, and T. F. El-Maraghi. Robust online appearance models for visual tracking.
- [24] Xiaoqin Zhang, Weiming Hu, Zixiang Zhao, Yan-guo Wang, Xi Li and Qingdi Wei, "SVD based kalman particle filter for robust visual tracking," IEEE 2008, 978-1-4244-2175-6.
- [25] R. T. Collins, A. J. Lipton, and T. Kanade, "Introduction to the special section on video surveillance," IEEE Trans. Pattern Anal. Machine Intell., vol. 22, pp. 745–746, Aug. 2000.
- [26] J. Steffens, E. Elagin, and H. Neven, "Person spotter-fast and robust system for human detection, tracking and recognition," in Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, 1998, pp. 516–521.
- [27] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L. Wixson, "A system for video surveillance and monitoring," Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep., CMU-RI TR-00-12, 2000.
- [28] I. Haritaoglu, D. Harwood, and L. S. Davis, "W: Real-time surveillance of people and their activities," IEEE Trans. Pattern Anal. Machine Intell., vol. 22, pp. 809–830, Aug. 2000.
- [29] J. Fernyhough, A. G. Cohn, and D. C. Hogg, "Constructing qualitative event models automatically from video input," Image Vis. Comput., vol. 18, no. 9, pp. 81–103, 2000.
- [30] Roland Kehl, Luc Van Gool, Markerless tracking of complex human motions from multiple views, Computer Vision and Image Understanding (CVIU) 104 (2–3) (2006) 190–209.
- [31] Thomas B. Moeslund, Erik Granum, A survey of computer vision based human motion capture, Computer Vision and Image Understanding (CVIU) 81 (3) (2001) 231–268.
- [32] Vladimir I. Pavlovic, James M. Rehg, Tat-Jen Cham, Kevin P. Murphy, A dynamic Bayesian network approach to figure tracking using learned dynamic models, in: Vision (ICCV'99), Greece, September 1999, pp. 94–101.
- [33] J.Y. Zou, Q. Pan, Y. Liang, H.C. Zhang, Y.M. Cheng, "Model Switching Based Adaptive Background Modeling Approach", Journal of Acta Automatica Sinica, 2007 May, 33(5): pp. 467-473.
- [34] N. Friedman, S. Russell, "Image segmentation in video sequences: a probabilistic approach", In: Proceeding of Thirteenth Conference on Uncertainty in Artificial Intelligence, Providence, Rhode Island, USA, Morgan Kaufmann Publishers, 1997, pp. 175-181.