10.48047/jocaaa.2024.33.08.154

A Research and Analysis of an Optimized FPGA-Based Design for Efficient Parallel Data Computation

Ms. MONIKA RANI
Assistant Professor, Department of Electronics and Communication Engineering, Dronacharya College of Engineering, Gurugram, Haryana
monika.rani@ggnindia.dronacharya.info

Ms. AKANCHA SHEKHER
Assistant Professor, Department of Electronics and Communication Engineering, Dronacharya College of Engineering, Gurugram, Haryana
akancha.shekhar@ggnindia.dronacharya.info

Abstract

The problem of detecting moving objects is complex due to various factors, such as different detection conditions, noise intensity variations, and the need for robust algorithms. The first step is to find every possible version of the image, which is typically achieved through simulation-based methods. The second phase involves a background subtraction model based on a Gaussian mixture model mask and an entropy-based spatiotemporal mask, with differential evolution optimization used to find the optimal threshold. The third phase proposes a hybrid approach that combines a sparse matrix with a combination of Gaussian (MoG) as foreground modeling, reproductive RPCA model, and total variation L1 (TV-L1) features in low rank background subtraction modeling. The hybrid structure with TV-L1 characteristics imposes a hierarchical RPCA on the singular values of MOG sparsity indicators and the low-rank component. The proposed method achieved an accuracy of 92.9% for Highway, 86.7% for Escalator, and 95.7% for Indoor when tested on the CDnet2014 dataset. The relative reconstruction error generated by the proposed method is 0.01529 compared to traditional methods.

Keywords: Robust Principal Component Analysis, TV-L1, Mixture of Gaussians, FPGA, and Differential Evolution.

Introduction

There is no exact solution for the long-standing problem of detecting moving objects because of its peculiarities. First, the conditions for detection are different. For instance, compared to using full-color raster with smooth color transitions from one shade to the next, dealing with a binary black and white image greatly reduces effort. Every frame will have a background that is almost the same, with potential lighting differences, because the video may have been recorded using a static camera. A stationary object can also be captured by a camera attached to a moving object. There may be wide variations in noise intensity. Snow, wind, fog, and rain can all cause significant variations in a still scene. Procedures that function wonderfully in one set of conditions could therefore be completely inapplicable non another. For autonomous motion detection systems to operate in a wide range of environmental conditions, their algorithms need to be robust. Finding every possible version of the image is the first step in identifying a moving object. Mostly, simulation-based methods are used to achieve this [1].

10.48047/jocaaa.2024.33.08.154

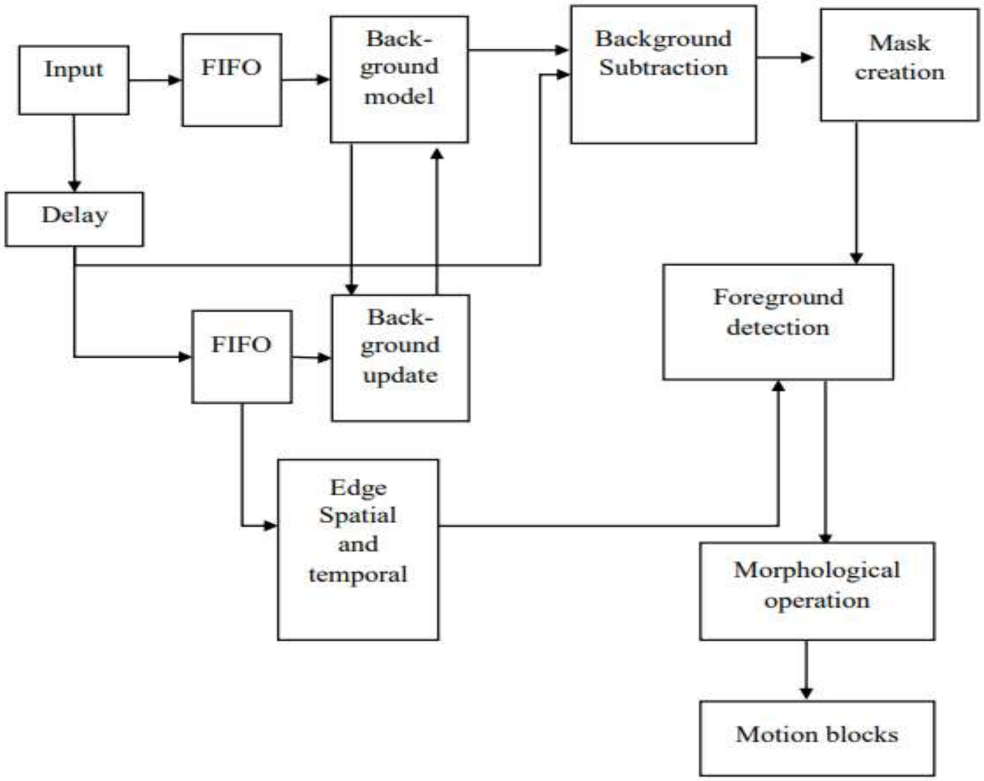

Fig.-1 Block diagram for FPGA implementation of proposed moving object detection

The produced models of a moving object are compared to certain areas of the image. The method of background removal is used to identify moving objects in videos. For every video frame, the algorithms that use this method produce a binary mask known as the foreground mask, which indicates which pixels belong to the image's backdrop and which comprise moving objects over it. Depending on the kind of object you want to detect, the time frame you want to use, and other factors, the background definition can change. For instance, it's common to include the movement of sea waves or tree leaves in the background, but not changes in the image brought on by distant pedestrians. Keeping these warnings in mind, backdrop refers to the parts of the image that remain consistent over time. The results of a subtraction algorithm are primarily determined by its background modeling. Background removal has becoming more challenging for moving items. In particular, the difficulty of implementation in a changing environment. There are three stages to this study project. A background subtraction model based on a Gaussian mixture model mask and an entropy-based spatiotemporal mask is put forth in the first phase. Additionally, differential evolution optimization is used to find the optimal entropy-based approach threshold. The area-efficient hardware implementation of a moving object detection algorithm using FPGAs is covered in the second phase of the study. The goal of the program is to remove moving object backgrounds from the CDNet dataset. Basic highlight and shadow reduction is done even if there are no occlusions found. The technique works well with pipelined technology and is optimized for high frame rates. The efficiency and robustness of hardware implementations were enhanced by the authors' novel additions, such as "background binary mask combination" or "non-linear functions" in highlight detection. The CDNet dataset was used to evaluate the approach after it was built in FPGA [2-6].

10.48047/jocaaa.2024.33.08.154

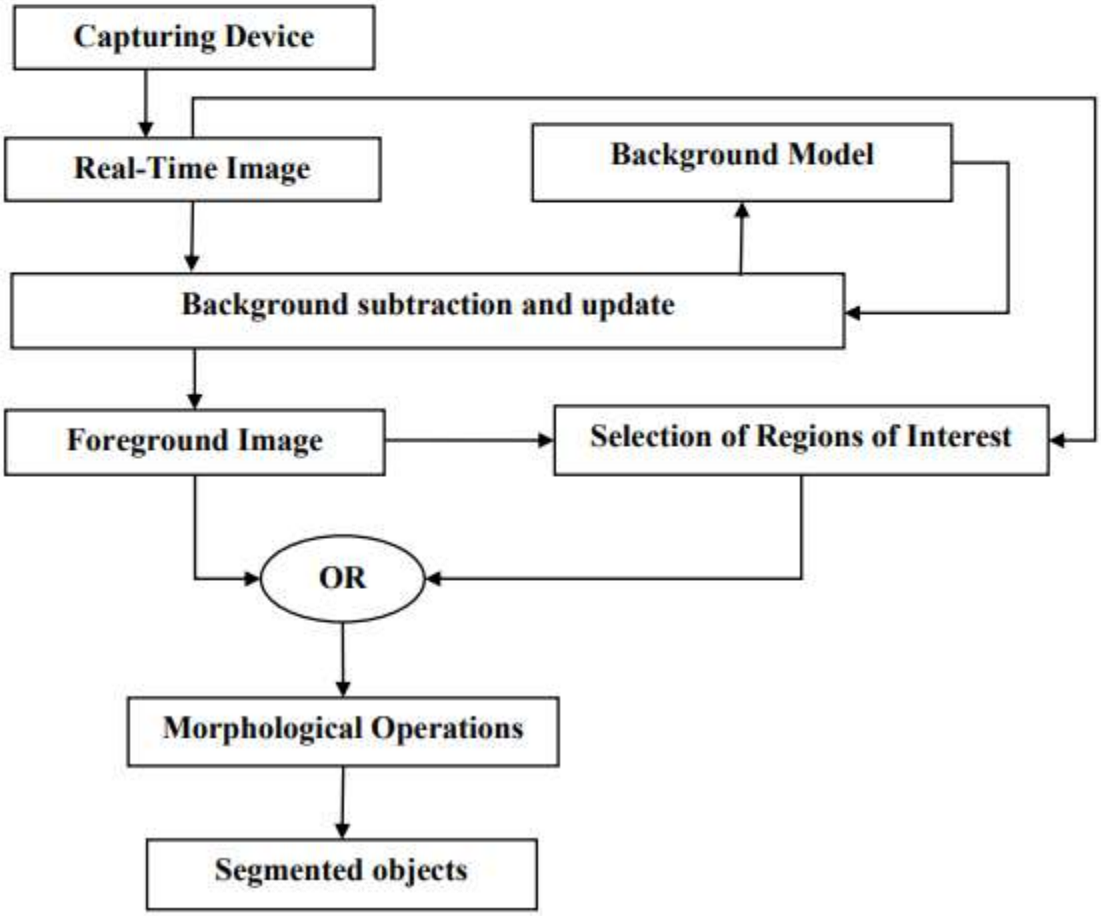

Fig.-2 Block diagram of the motion segmentation process

In image processing, presentation, and pattern recognition for dimension reduction, robust principal component analysis, or RPCA, has recently gained popularity. Recently, there has been a lot of interest in background modeling techniques that rely on low-rank sparse representations and make a few key assumptions. Meanwhile, a strong analytical framework is needed to manage foreground motions or background regions at various sizes. The third study phase suggests a hybrid approach that combines a sparse matrix with a combination of Gaussian (MoG) as foreground modeling, reproductive RPCA model, and total variation L1 (TV-L1) features in low rank background subtraction modeling. The hybrid structure with TV-L1 characteristics imposes a hierarchical RPCA on the singular values of the MOG sparsity indicators and the low-rank component. The accuracy of the proposed method was 92.9% for Highway, 86.7% for Escalator, and 95.7% for Indoor, respectively, when tested on the CDnet2014 (ChangeDetection.net) dataset [7-10].

10.48047/jocaaa.2024.33.08.154

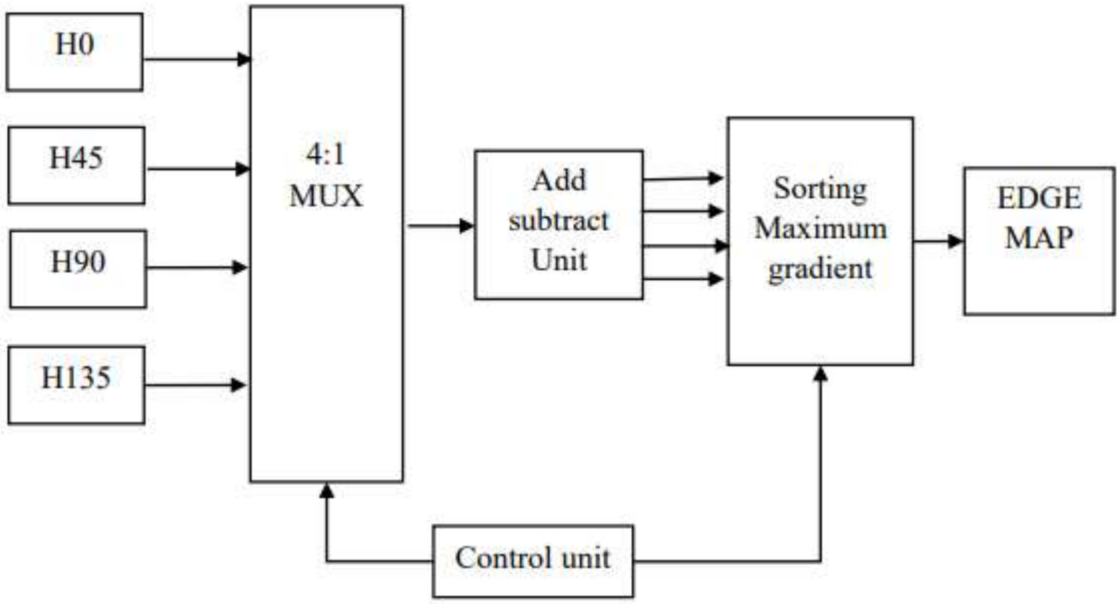

Fig.-3 Edge detection block

The relative reconstruction error generated by the proposed method is 0.01529 on the lower side when compared to traditional methods.

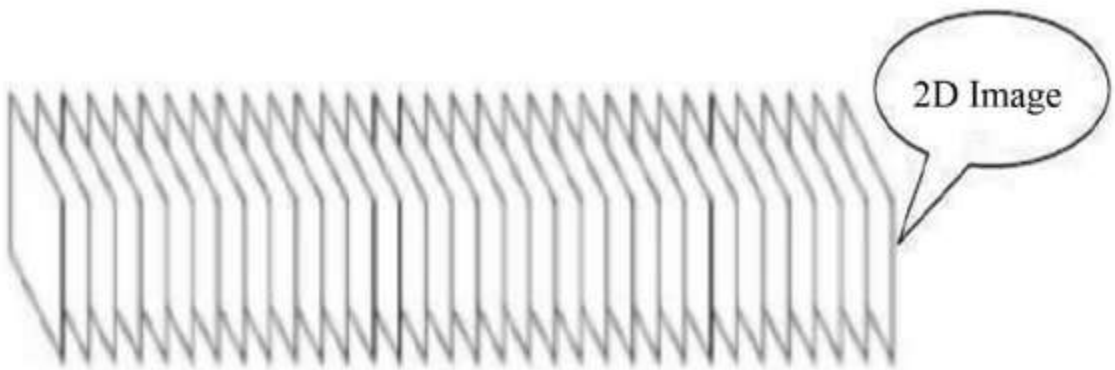

Fig.-4 Overview of 2D image

Motion detection is crucial for intelligent video surveillance systems, as it extracts pixels containing moving objects from a sequence of images observed by cameras. The increasing interest in automatic detection algorithms and object tracking has led to its use in various fields, including medical, industry, traffic flow evaluation, and sports referee assistance systems. However, setting up such systems can be challenging due to various factors such as sudden movement, changes in appearance, deformation, camera movement, illumination changes, and occlusions between objects or the scene [11-15].

10.48047/jocaaa.2024.33.08.154

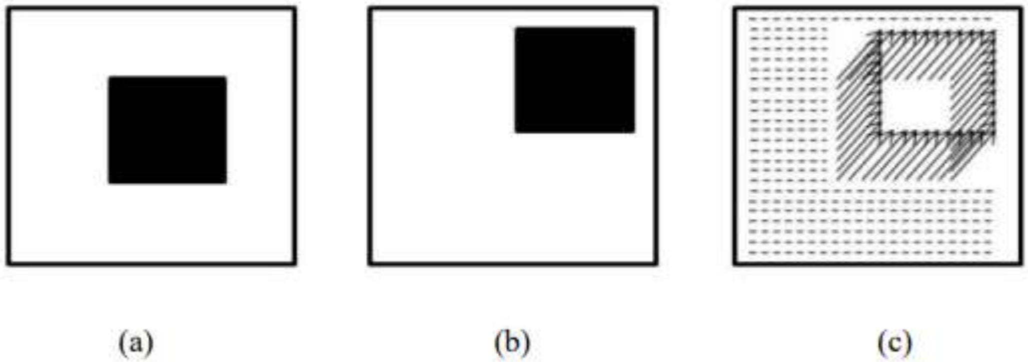

Fig.-5 position of the object at time t, (b) position of the object at time t+1, (c) resulting optical flow

Various methods have been proposed for object detection, including optical flow, time derivative, and background subtraction. The object must be represented by specific characteristics that distinguish it from other objects in the scene. Tracking a moving object requires providing the spatiotemporal location and response in a short time. This can be achieved by using a collection of points of interest, calculating the movement of the kernel in successive photos, and extracting the object's silhouette for tracking. Regardless of the difficulties encountered and the structure of the scene, follow-up must be carried out [16-20].

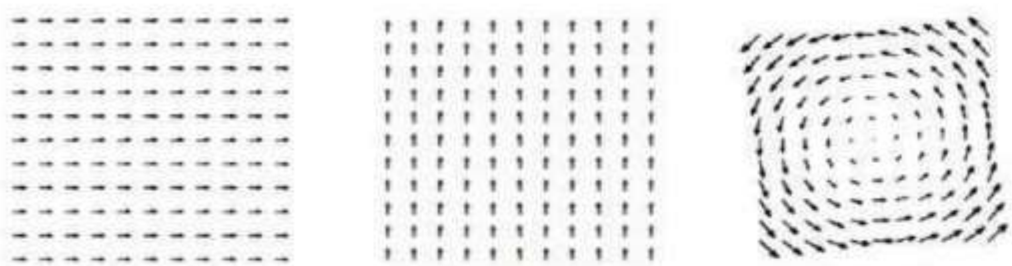

Fig. 6 Examples of moving object fields

The rapid advancement of computer vision systems has made motion detection and tracking an attractive research area, with applications in automated surveillance, vehicle tracking, traffic accident detection, and robots. Object tracking is a challenging process due to the voluminous information contained in videos and the need for complex algorithms to identify, distinguish, and track objects. This work focuses on processing necessary for detecting moving objects in video picture sequences and estimating their positions and speeds.

A video stream is a sequence of 2D images, with resolution defined by the number of pixels. There are three types of images: binary, intensity, and RGB color images. Physical objects are real-world objects that appear in the scenes observed by cameras. Context objects are static objects, while moving objects are perceived in the scenes by their movements. Abandoned objects are immovable objects that have not been in contact with a person for a certain threshold of time.

Object detection is essential in any tracking method, whether in every frame or when the item first appears in the video. Various methods and techniques have been proposed and developed, such as MATLAB's detection of a single object in a motion scene and several objects in the

10.48047/jocaaa.2024.33.08.154

scene. Detection depends on several variables, such as luminosity variation and the presence of shadows. Several methods have been proposed for motion object detection, including the use of a motion model to describe how an object's image changes with different motions and orientations. Optical flow is a method of detecting moving objects in an indoor or outdoor environment using a sequence of pictures captured by a fixed or mobile camera. It is a two-dimensional vector field represented by actual motion projected onto the picture plane. LUCAS and KANADE are popular algorithms for calculating optical flow, based on the hypopaper of conservation of light intensity. The motion vector of a point can be calculated from techniques established on the gradient. The estimation of optical flow can be expressed as an optimization problem in the neighborhood of (x) which minimizes the squared error function ε.

Time derivative detection is another method for detecting moving areas in a field of vision. It measures the change in appearance of pixels between two consecutive frames, with the instantaneous time derivative of the signal at time t. However, this method is not very robust in the face of phenomena such as slow or jerky movements, brief stops of a moving object, or redundant frames in certain video sequences. The object detection method involves applying a moving average operator to the measured data, which can be done using an accumulator matrix. The term wA weights the contribution of past measurements relative to the most recent measurement. With a low wA, the effects of temporal smoothing are not very visible, and the problems that motivated the use of this method are likely to appear all the same. With a high wA, smoothing is important.

Moving Object Detection by Background Subtraction is a famous method due to its simplicity and precision in detecting moving objects. This method involves providing or generating the background of the image to be detected. Both methods perform a scan of each pixel in an image and compare it with the background model generated or provided. However, this method has several problems, such as the inability to differentiate between detected objects and changes in light and position of the image. Background subtraction is an important step in detecting changes in the image sequence, as it converts uninteresting regions to the background, considering changes in the scene. The most used method is the mixture of Gaussian distributions, which considers nonuniform background due to sensor noise or small movements.

This paper focuses on developing a moving object detection system using background subtraction with soft computing techniques, area efficient hardware implementation using FPGAs, and a hybrid approach of moving object detection using total variation L1 (TV-L1) features and robust principal component analysis (RPCA). The research is divided into three phases, with each phase elaborated in upcoming papers. The introduction provides a basic overview of the research area and concepts of moving object detection, followed by a literature survey to improve skills in the field. The third paper presents an unsupervised segmentation and detection process based on background subtraction and entropy-based spatiotemporal with histogram modeling by mixtures of GMM models. The simulation results achieve good detection of moving objects in different situations thanks to good segmentation and differential evolution optimization.

The fourth paper discusses the area efficient hardware implementation of a moving object detection algorithm using FPGAs, which removes moving items' backgrounds from the CDNet

10.48047/jocaaa.2024.33.08.154

dataset. The authors present novel enhancements, such as background binary mask configurations or non-linear factors in highlight detection, to increase the robustness and reliability of hardware implementations. The fifth paper presents a hybrid approach using total variation L1 features and reproductive RPCA model in low rank background subtraction modeling and sparse matrix with mixture of Gaussians (MoG) as foreground modeling. The proposed work was evaluated on the CDnet2014 dataset and compared with traditional methods. The MATLAB-based simulation outcomes and comparative analysis of the developed work with previous research works are also presented.

This paper provides a comprehensive study of the literature on moving object detection in computer science. Multimedia tools have become a fundamental direction in the development of information technology, with their popularity growing rapidly. Video surveillance systems play a crucial role in monitoring and addressing security and awareness requirements. Computer vision focuses on creating algorithms capable of processing and analyzing images or sequences of images, solving problems such as face recognition, object tracking, and localization in space. Motion analysis in video has shown value in various applications, including video surveillance, medical imaging, robotics, video compression, motion sequence analysis, and human-computer interaction.

In computer vision, the notion of object is a key element, as it designates an area of the image characterized by texture, shape, color, orientation, or movement. There are two types of objects: stationary objects (stationary) and moving objects (in motion). Fast moving objects are important, while slow objects move slowly. Simple motion is the same motion in the image sequence, while complex motion has random motion in the image sequence.

In conclusion, computer vision focuses on the notion of object, which is a key element in detecting moving objects in video scenes. Object detection is a computer vision approach that helps in identifying and localizing objects in pictures or videos. It creates bounding boxes around observed objects, allowing us to identify them or determine how they move in a given scene. Object detection offers more details about the image than recognition, as it assumes where each object will be and what label will be assigned to it. The principle of object detection is to search for regions of an image that could contain an object, extract it, and classify it using an image classification model. To have a good object detection method, it is necessary to have a solid region detection algorithm and a good image classification algorithm.

Object classification methods include shape-based classification, movement-based classification, texture-based classification, and color-based classification. Shape-based classification uses multiple descriptions for motion shape information of regions, while movement-based classification uses the periodic aspect of non-rigid object motion. Texture-based classification counts occurrences of gradient orientation in localized regions of an image for improved accuracy. Color-based classification remains generally consistent throughout perspective shifts and is easier to detect.

There are two major approaches in the field of object detection: the classical detection approach based on visual features extracted from the object, often by local detectors, and the deep learning detection approach. Detection using feature descriptors is divided into three models: detection by

10.48047/jocaaa.2024.33.08.154

points of interest, background subtraction, and image segmentation.

Detection by points of interest is crucial for computer vision processes, as it allows for the identification of points of interest in images. Point detectors have advantages such as insensitivity to lighting, pose, and scale variation. The Harris Corner Detector is a corner detector that looks for the difference in intensity for a displacement of (u, v) in all directions in the image. It is sensitive to noise due to the use of gradient information. The LK Detector is an improvement of the Harris detector, based on the Harris autocorrelation matrix. It uses this matrix to extract points of interest, but it is sensitive to changes in scale and illumination. The SIFT Detector was introduced to remedy problems caused by changes in scale in images and geometric variations, such as angle of view and illumination. It describes a point from the local orientations of the gradient and produces a large quantity of points of interest due to the extraction of points of interest at different scales and resolutions. The SURF Detector is a faster version of the SIFT detector, which relies on the Hessian Matrix determinant approximation using the convolution between the Gaussian Kernel LoG and the Box-Filter. This makes it invariant to rotations and allows for faster calculations.

In summary, detection methods based on the calculation of points of interest are essential for computer vision processes. Background subtraction is a crucial pre-processing step in vision applications, allowing for the detection of objects by eliminating them from an image. It involves constructing a background model representing the scene and monitoring any movement or change relative to that model. Several background subtraction methods have been developed, including Gaussian models, which model the color intensity of pixels by a single Gaussian distribution, and models based on learning subspaces, which use data analysis methods to represent pixels in a global approach. Principal components analysis (PCA) is another method used to construct the average image, which is less vulnerable to changes in illumination.

Segmentation is another widely used technique in computer vision, used to detect and identify objects by partitioning the image into meaningful regions. Popular segmentation techniques include MeanShift algorithms, GraphCut algorithm, and edge detection. These methods help to reduce memory space and processor time, while also being more suitable for outdoor scenes and dynamic scenes. MeanShift is a robust segmentation technique used in imaging and computer vision to estimate local gradients of similar pixel densities in an image. Iteratively, iterative estimates are performed, allowing for the identification of similar pixels in the corresponding image. MeanShift is particularly useful for image segmentation into regions, as it combines spatial information and colorimetric information.

Segmentation by Graph Slice is another technique that divides images into subgraphs representing distinct areas. The most popular method is the method by graph cut or normalized GraphCut, which treats each pixel as a node, builds a graph from these pixels, and adds two additional nodes, an object-node and a background-node. The weights between the edges are determined by the edge information indicating the similarity of the pixels. Segmentation by Contours is a simple technique that detects and traces a contour in an image, tracing it towards the boundaries of the object until the outline encompasses the entire region. An energy function controls the evolution of the contour by determining its limits on the object's region.

10.48047/jocaaa.2024.33.08.154

Level set is another edge-based segmentation technique used in video to detect dynamic objects. However, it requires previous knowledge of the item's position, making it difficult to start the contour from a single point in the image. Video segmentation involves identifying and dividing video frames based on their image properties. Two well-known methodologies are spatial segmentation and temporal segmentation. Spatial segmentation partitions individual video frames based on their image properties, while temporal segmentation groups frames from the same scenes by detecting delimiting frames. The combination of these two methodologies led to spatiotemporal segmentation, which follows the behavior of regions segmented in space over time. Space-time segmentation detects contours of parts in a video frame and identifies their location in the following frames. Segmentation methods using tunnels are considered three-dimensional, as they analyze the video in both spatial and temporal domains. The scope of motion-based video segmentation research has expanded, with more applications such as automatic event detection, video comprehension, and object-based video coding.

The standard approach to video segmentation is video segmentation between static and moving regions, with a methodology using level sets to obtain a closed curve with the highest probability of delimiting the moving region in each frame. Two extensions to this approach include pixel intensity and identifying numerous moving regions. Another approach is the coherent movement of particles, which are adaptively sampled points in video frames. Their movement between frames is estimated using a modified version of the trajectory estimation method, which includes adding particles, propagating particles, removing particles, and optimizing their location.

The text discusses various methods for video segmentation, including particle detection, meta-clustering, and motion-based segmentation. The first method uses the gPb edge detection algorithm for individual frame segmentation, followed by optical flow and spectral clustering algorithms to determine the affinity of pixels between frames. A confidence measure is calculated to identify points where the optical flow calculation may present an incorrect result.

The second method introduces a motion-based segmentation approach for long videos, which divides the video into N time windows. Particles are identified within each window, and an affine transformation model is used to obtain the complete trajectory of each particle within the window. Particles are grouped in each window according to their movement using the residual ordered core algorithm.

The third method explores the difference in behavior over time when using segmentation based on color and texture or motion. Motion-based segmentation identifies denser and more coherent regions across multiple frames, using color, texture, and movement for segmentation. Particle movement is estimated using the optical flow algorithm, and the colors are used to segment the video considering a maximum number of eight dominant colors in the lab color system, grouped using a k-means algorithm. The particles are grouped considering the distance between their trajectories and color regions.

Analysis Report

10.48047/jocaaa.2024.33.08.154

Table-1 Frame-wise comparison on highway video with ground truth and segmented outcome

Highway	Accuracy	Precision	Sensitivity	F-Score
Frame 50	92.9%	87.5%	100%	93.3%
Frame 100	92.3%	86.7%	100%	92.9%
Frame 150	94.4%	90%	100%	94.7%
Frame 200	95.2%	91.3%	100%	95.5%

Table-2 Frame-wise comparison on Escalator video with ground truth and segmented outcome

Escalator	Accuracy	Precision	Sensitivity	F-Score
Frame 50	86.7%	84.2%	88.9%	86.5%
Frame 100	90%	91.4%	88.9%	90.1%
Frame 150	91.5%	94.4%	89.3%	91.8%
Frame 200	90.39%	100%	84.78%	91.8%

Table-3 Frame-wise comparison on Indoor video with ground truth and segmented outcome

Indoor	Accuracy	Precision	Sensitivity	F-Score
Frame 50	95.7%	100 %	92.9%	96.3%
Frame 100	96.1%	100%	95.7%	96.2%
Frame 150	92.2%	96.5%	89.9%	92.7%
Frame 200	89.4%	100%	82.1%	90.2%

The binary partition tree (BPT) is a hierarchical structure built after segmenting a video, where regions detected are fused at each level. This structure aims to be a practical way to analyze

10.48047/jocaaa.2024.33.08.154

videos later, as it can be cut into regions that meet certain characteristics. It requires 1,000 seconds and 20GB of RAM memory for a video of three million voxels.

Works using motion-based segmentation for applications other than video preprocessing can also be found in the literature. For example, particle trajectory tracking is proposed to monitor the behavior of a specific object throughout the video. The GrabCut algorithm is used to determine the outline of the object within the rectangular region.

There are two basic types of foreground extraction strategies in the literature: background modelling and simultaneous background/foreground modelling. Background modelling techniques are noise resilient and can change to accommodate new background objects, but they are susceptible to environmental phenomena such as shadows, global light changes, and camouflage. Many shadow detecting approaches rely on pixel-level local information, which can cause false negatives and false positives, and camouflage. Many options have been presented to address these shortcomings by modeling both the foreground and backdrop. In this situation, the extraction of the foreground is performed by combining the two. The study compares a proposed approach to various cutting-edge algorithms on www.changedetection.net, revealing that the top-ranked algorithms are FR-CNN, FBS-ABL, SDMBS, and BSPVBGANs. These deep learning-based supervised approaches outperform the proposed hybrid algorithm, which has an F-Measure of 97.80%. However, the proposed framework outperforms some deep learning-based methods, including BSPVBGANs, and is highly flexible, allowing for easy component changes.

Conclusion

Computer vision techniques, such as video surveillance and human movement analysis, rely on the detection of moving objects. Background subtraction is a technique used in video processing to separate pixels containing information from static objects, obtaining information-rich pixels for various techniques. This technique has been widely adopted in areas such as motion detection, multimedia applications, and surveillance by means of videos. Background extraction, also known as foreground detection or background subtraction, is a well-known approach in image processing and computer vision that involves extracting the foreground of an image for on-line or off-line analysis. This problem can be tackled using relatively simple algorithms, but movies often contain unsettling features that make it difficult to classify a pixel as an object or a background. Advanced modeling approaches have been developed to handle such complicated videos, such as the Gaussian mixture model (GMM). This research work is presented into three phases: an unsupervised segmentation and detection procedure based on background removal and entropy-based spatio-temporal using Gaussian mixture model histogram modelling, an integrated technique for extracting moving objects from a real-time video stream, and noise modelling of the novel RPCA approach as a MoG distribution in a Bayesian framework. The proposed hybrid method achieves a moderate accuracy of 93% on multiple videos and has a lower relative error of recovery than existing methods.

References

1. AbdelTawab, A.M., Abdelhalim, M.B. and Habib, S.E.D., 2022. Moving Object Detection over Wireless Visual Sensor Networks using Spectral Dual Mode Background Subtraction. International Journal of Advanced Computer Science and Applications, 13(1).

10.48047/jocaaa.2024.33.08.154

2. Akilan, T., Wu, Q.J. and Yang, Y., 2018. Fusion-based foreground enhancement for background subtraction using multivariate multi-model Gaussian distribution. Information Sciences, 430, pp.414-431.
3. Alipour, P. and Shahbahrami, A., 2022, February. An adaptive background subtraction approach based on frame differences in video surveillance. In 2022 International Conference on Machine Vision and Image Processing (MVIP) (pp. 1- 5). IEEE.
4. Aung, S.S. and War, N., 2019, November. Foreground objects segmentation in videos with improved codebook model. In 2019 International Conference on Advanced Information Technologies (ICAIT) (pp. 161-166). IEEE.
5. Barnard, K., 2022. Background Subtraction: Theory and Practice. Springer Nature.
6. Belagiannis, V., Schubert, F., Navab, N. and Ilic, S., 2020, October. Segmentation based particle filtering for real-time 2d object tracking. In European Conference on Computer Vision (pp. 842-855). Springer, Berlin, Heidelberg.
7. Bourouis, S., Pawar, Y. and Bouguila, N., 2021. Entropy-based variational scheme with component splitting for the efficient learning of gamma mixtures. Sensors, 22(1), p.186.
8. Çelik, Y., Altun, M. and Güneş, M., 2017, September. Color based moving object tracking with an active camera using motion information. In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP) (pp. 1-4). IEEE.
9. Chandrakar, R., Raja, R., Miri, R., Sinha, U., Kushwaha, A.K.S. and Raja, H., 2022. Enhanced the moving object detection and object tracking for traffic surveillance using RBF-FDLNN and CBF algorithm. Expert Systems with Applications, 191, p.116306.
10. Chen, F., Wang, X., Zhao, Y., Lv, S. and Niu, X., 2022. Visual object tracking: A survey. Computer Vision and Image Understanding, 222, p.103508.
11. Chen, Z., Wang, R., Zhang, Z., Wang, H. and Xu, L., 2019. Background–foreground interaction for moving object detection in dynamic scenes. Information Sciences, 483, pp.65-81.
12. Comaniciu, D. and Meer, P., 2002. Mean shift: A robust approach toward feature space analysis. IEEE Transactions on pattern analysis and machine intelligence, 24(5), pp.603-619.
13. Dong, Y. and DeSouza, G.N., 2011. Adaptive learning of multi-subspace for foreground detection under illumination changes. Computer Vision and Image Understanding, 115(1), pp.31-49.
14. Fattahi, S. and Sojoudi, S., 2020. Exact guarantees on the absence of spurious local minima for non-negative rank-1 robust principal component analysis. Journal of Machine Learning Research, arXiv preprint arXiv:1812.11466.
15. Garcia-Garcia, B., Bouwmans, T., Sehairi, K. and Zahzah, E.H., 2022. Real-Time Implementations of Background Subtraction for IoT Applications. In Computer Vision and Internet of Things (pp. 263-286). Chapman and Hall/CRC.
16. Gemignani, G. and Rozza, A., 2016. A robust approach for the background subtraction based on multi-layered self-organizing maps. IEEE transactions on image processing, 25(11), pp.5239-5251.
17. Goyette, N., Jodoin, P.M., Porikli, F., Konrad, J. and Ishwar, P., 2012, June. Changedetection. net: A new change detection benchmark dataset. In 2012 IEEE computer society conference on computer vision and pattern recognition workshops (pp. 1-8). IEEE.
18. Han, G., Wang, J. and Cai, X., 2017. Background subtraction based on modified online robust principal component analysis. International Journal of Machine Learning and Cybernetics, 8(6), pp.1839-1852.
19. Huang, R., Pedoeem, J. and Chen, C., 2018, December. YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 2503-2510). IEEE.
20. Keuper, M., Tang, S., Andres, B., Brox, T. and Schiele, B., 2018. Motion segmentation & multiple object tracking by correlation co-clustering. IEEE transactions on pattern analysis and machine intelligence, 42(1), pp.140-153.