MEKATRONIKA
JOURNAL OF MECHATRONICS AND INTELLIGENT MANUFACTURING
E-ISSN: 2637-0883
VOLUME 6, ISSUE 2, JULY, 2024, 39 – 51
DOI: https://doi.org/10.15282/mekatronika.v6i2.11332

RESEARCH ARTICLE

Tracking Humans and Objects in Video Surveillance System Using Feature-Based Method

Odai Saeed Baqalaql, Intiaz Mohammad Abir, Azhar Mohd Ibrahim* , Amir Akramin Shafie

Advanced Multi Agent System Lab, Department of Mechatronics Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia.

ABSTRACT - In recent years, video surveillance system has emerged as one of the active research area in machine vision community. This research intends to integrate machine vision into video surveillance system in order to enhance the accurateness and robustness of video surveillance system. To realize more robust and secure video surveillance system, an automated system is needed which can detect, classify and track human and objects even when the occlusion occurs. Object tracking is one of the most crucial parts of a automated surveillance system Hence, we proposed a tracking system which includes tracking of human and vehicles in real-time surveillance system and also in solving the problem of partially occluded human by utilizing fast-computation techniques without compromising the accuracy and performance of that particular surveillance system. In this research, we track the classified human and objects using feature-based tracking for five states, which are: entering, leaving, normal, merging, and splitting. The developed system can track the human even if occlusion occurs since we used merging and splitting cases in our tracking algorithm. The overall accuracy for our proposed system in tracking human and car is fine which is at 94.74%.

ARTICLE HISTORY
Received : 25th Sept 2024
Revised : 13th Oct 2024
Accepted : 29th Oct 2024
Published : 5th Nov 2024

KEYWORDS
Video Surveillance
Human Tracking
Object Tracking
Image Processing

1.0 INTRODUCTION

Video surveillance system is significant because of the demand for increasing security concern, and automated schemes to analyse surveillance video. In a conventional setting, a video surveillance system consists of one or several video cameras and video streams are displayed on one or several video monitors and recorded. Human operator(s) observes the continuous video to find out if there is activity that demands a response. However, there are a lot of associated problems using human operator for surveillance. For example, it is expensive and requires many human resources [1]. Moreover, human are prone to error. Hence, it is essential to make use of automated video analysis technologies for developing surveillance systems which can assist the human operator in both detecting and reacting to potential threats [2]. Automated video surveillance aims to automatically identify specific objects for various purposes. Monitoring Human and vehicles are among the applications of surveillance system [3], [4].

To realize a robust and secure video surveillance system, an automated system is needed which can detect, classify and track humans and objects (e.g., vehicles) even when the partial occlusion occurs. Hence, objects' tracking in videos have become a significant research area because of its possible functions in video surveillance system [5], [6]. To achieve this purpose, many problems should be well addressed (e.g., alteration in appearance of the objects with viewpoint and partially occluded human [7]). The automated human and objects' tracking system also should be able to analyse video streams by utilizing fast-computation techniques without compromising the accuracy and performance of that particular surveillance system.

Detection of moving objects in video streams is the first relevant step of information extraction in video surveillance. When the objects are successfully detected in the video surveillance system, the next step is to successfully classify the moving humans and objects. When the objects are classified in the video surveillance system, the next step is to successfully track the moving humans and objects from one frame to another in following scenes [6], [8]. Moving object tracking is essential in machine vision systems for surveillance and monitoring systems. Tracking each individual object becomes a difficult job especially when multiple tracked objects merge into groups with different complexities of occlusion. One of the most crucial criterion for intelligent surveillance systems is to track multiple objects over time in occluded scenes and to keep a consistent identity for each target object. [9] suggested that there is an increasing demand for more robust real-time tracking systems which can automatically detect and track multiple targets of interest.

*CORRESPONDING AUTHOR | Azhar Mohd Ibrahim | ✉ azhar_ibrahim@iium.edu.my

Human tracking under partially and fully occluded scenes especially in real-time system is an important research area. In this research, we focus on implementing the tracking of humans and objects especially vehicles in real-time surveillance system and also in solving the problem of partially occluded human.

The contents of this paper aside from the preliminaries are organized as follows: Section one introduces this research with some literature review, where the previous works on tracking of human and objects in video surveillance system are discussed. Section two presents the details of the proposed tracking system. Section three demonstrates the results and analyses the accuracy and performance of the proposed system. Section four concludes this research.

Tracking of human and objects in videos is an important research topic because of its potential applications in video surveillance system. Moving object tracking is a vital task in computer vision systems for surveillance and monitoring. One of the most crucial criterion for automated surveillance systems is to track objects over time in occluded and nonoccluded scenes [10]. In order to achieve real-time tracking of the moving objects, the exact position of moving objects must also be confirmed.

There are many challenges in tracking the moving objects such as associate target objects in consecutive video frames. This is difficult when the objects are moving fast relative to the frame rate. There are also other problems in tracking the moving objects such as when the features in the background images occluded to the features of the foreground objects. This can happen in the case where there are many people in the scene or in the occluded scenes. According to [11], occlusion itself can be divided into self-occlusion, inter-object occlusion, and also background occlusion. Self-occlusion can occur when tracking articulated objects and it refers to the feature of an object occludes to another. Meanwhile, inter-object occlusion happened when handling the tracking of multiple objects that are occluding each other. Similarly, occlusion by the background occurs when the features in the background images occluded to the features of the tracked objects.

There are many tracking methods which can be divided into five categories in general: region-based tracking [12], [13], [14], active contour or snakes based tracking [15], [16], feature based tracking [17], [18], [19], [20], model-based tracking [21], [22], [23], and hybrid tracking [24], [25]. In this research we have utilized feature-based tracking.

In feature-based tracking the idea is to abandon tracking objects as a whole but instead tracks sub-features such as distinguishable points or lines on the objects. Here, the tracking algorithm is done by extracting elements, clustering them into higher level features and then matching the feature between images. However, there is a problem in grouping similar set of features that belongs to the same object such as in the case of vehicles. In the case of vehicle tracking, [17] used vehicle sub-features as a tracking feature of the vehicle to make the system robust to partial occlusion. The constraint of common motion is used in order to group together sub-features that come from the same vehicle. Choi et al, [18] developed an image-based system that detects and tracks multiple moving vehicles from the sequences of images using the Scale-invariant Feature Transform (SIFT) and quadtree segmentation to get better performance which is robust to changes of the scale, intensity and shape of object caused by movement. Feature-based methods try to extract salient features such as edges and corners in order to track the moving vehicle.

Besides that, there are many algorithms or mathematical tools which are being used by researchers for filtering and also for data association issue (to solve occluded objects). For example, Kalman Filter [26], [27], [28], [29], [30], Particle Filter [31], [32], [33], [34], [30], [35], [36] and Euclidean Distance [37].

Gao et al, [37] utilized scale invariant feature transform (SIFT) to acquire significant points in the object region. Then, for image matching and recognition, first of all, SIFT features are extracted from a set of reference images and stored in a database. An image in the previous database is compared to the new image individually and finding candidate matching features are based on Euclidean distance of their feature vectors respectively. This Euclidean distance technique is adopted in this research together with feature-based tracking for five states of tracking and this will be discussed further in the following section.

2.0 PROPOSED VIDEO SURVEILLANCE SYSTEM

This section discusses the system architecture of the developed tracking system in video surveillance system using techniques discussed in previous section. For this research, video sequences were used that are taken from (1) CAVIAR Dataset [38] (2) SPEVI datasets [39] (Courtesy of EPSRC funded MOTINAS project (EP/D033772/1)).

Ting et al, [40] presented multiple object tracking using region based method by separating motion of moving objects into five states: entering, leaving, merging, splitting and normal. Region corresponding is used to detect these 5 states. The idea in region based method is to identify a connected region in the image associated with each moving objects and then track it over time. In contrast, in this research we utilized feature based tracking for five states which are mentioned earlier: entering, leaving, merging, splitting and normal. We used the position of middle top vertex of the classified objects as a feature for tracking in the four states of entering, leaving, normal and merging as shown in figure 1.

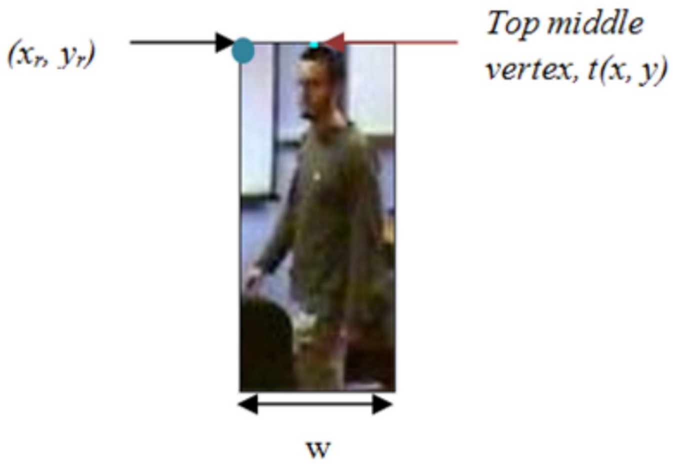

Figure 1. Top Middle Vertex Position of the Extracted Human

The coordinate of t(x,y) is as in 1. The coordinate position of y is added with zero because this position is at the top of the extracted blob.

$$x = x_r + \frac{w}{2} \; ; \; y = y_r + 0 \tag{1}$$

Successful tracking will give the tracking result which is either identified or unidentified. Identified means an object is successfully tracked as belong to an object detected in earlier scene as indicated using same colour bounding box. While unidentified means an object is just entering or leaving the scene. Colour of bounding box is used in this research as an indicator to display the tracked objects. The same human or object should have the identical colour of bounding box throughout the scene.

Here, the identified human or objects in the previous frame, fr_{c-1} have the property of top middle vertex $t_i(x_i,y_i)$. Meanwhile, the unidentified or untracked human or objects in the current frame, fr_c have the property of top middle vertex $t_n(x_n,y_n)$. The algorithm that we used in this research in order to track the objects is by using Euclidean Distance formula, $d_t(t_n,t_i)$ as in (2). The unidentified or untracked human or objects in the fr_c is matched with previous ones if the distance, d_t between $t_n(x_n,y_n)$ and $t_i(x_i,y_i)$ is less than threshold value, T_d.

$$d_t(t_n, t_i) = \sqrt{(|xn - xi|^2 + |yn - yi|^2)} \tag{2}$$

Meanwhile, for splitting states, we used colour of classified foreground objects to match the objects with the previous tracked objects before the merging of the objects occur. We have five different cases in order to perform the tracking in this research which are as follows:

1) Entering Case
2) Leaving Case
3) Normal Case
4) Merging Case
5) Splitting Case

A. Entering Case

A new human or a new object will have the new colour bounding box. Entering case happen when an unidentified human or object in the current frame fr_c is regarded as a new human or new object. This would happen if it satisfied one of the following properties:

1) There is at least one of the identified human or object in the current frame fr_c, and no human or object is detected in the previous frame fr_{c-1} as illustrated in figure 2.

Figure 2. Entering Case in Property 1

2) If the distance between unidentified human or object in the current frame fr_c with all the human and objects in the previous frame fr_{c-1} is more than threshold value as expressed in (3) and as illustrated in figure 3.

$$d_t(t_n,t_i) > T_d \tag{3}$$

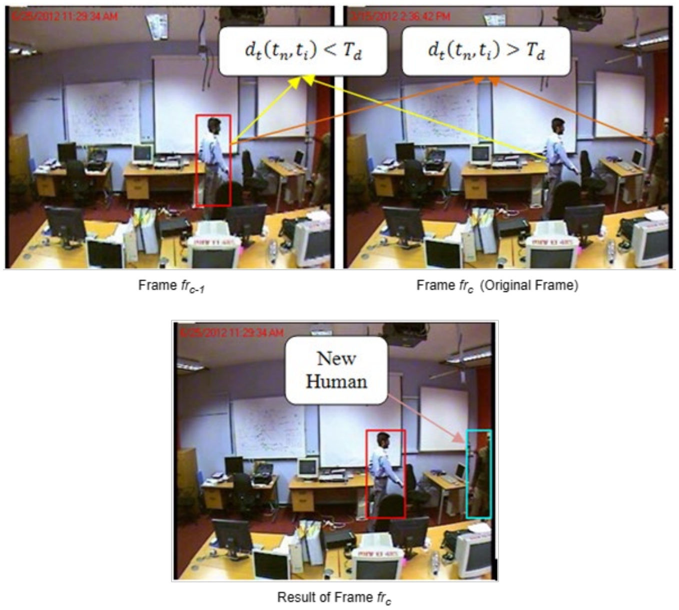

Figure 3. Entering Case in Property 2

B. Leaving Case

Leaving case happen when an identified human or object in the previous frame frc−1 is regarded as a leaving human or object. This would happen if it satisfied one of the following properties:

1) There is at least one of the identified human or object in the previous frame fr_{c-1}, and no human or object is detected in the current frame fr_c as illustrated in figure 4.

Figure 4. Leaving Case in Property 1

2) The number of identified human or object in the is more than the detected human or object which is not tracked yet in the current frame fr_c and there is at least one of the identified human or object in the previous frame fr_{c-1} is not matched or tracked in the current frame fr_c as illustrated in figure 5.

Figure 5. Leaving Case in Property 2

C. Normal Case

The tracked human or object for normal case in the current frame fr_c will inherit the same colour of bounding box from the previous frame fr_{c-1} . Normal case happens when an identified human or object in the previous frame, fr_{c-1} is matched and tracked in the current frame fr_c. This would happen if it satisfied one of the following properties:

1) If only a human or an object in the current frame fr_c which has the threshold value T_d less than 25, then that human or object will inherit the same colour of bounding box from the previous frame fr_{c-1} as illustrated in figure 6.

Figure 6. Normal Case in Property 1

2) If there are more than one identified human or object in the previous frame fr_{c-1} and in the current frame fr_c which has the threshold value T_d less than 25, then, the human or object in the current frame, fr_c which has the minimum distance,$min(d_t(t_n,t_i))$ will inherit the same colour of bounding box as illustrated in figure 7. In figure 7, in order to track the human ($O1$), we find the distance between $O1$ with the humans in previous frame, which are $H1$ and $H2$. We found out that the distance ($d1$) is less than distance ($d2$), thus, $O1$ will inherit the colour of bounding box from $H1$. The same step will be taken to track the $O2$. The result of tracking in the current frame fr_c is also shown in figure 7.

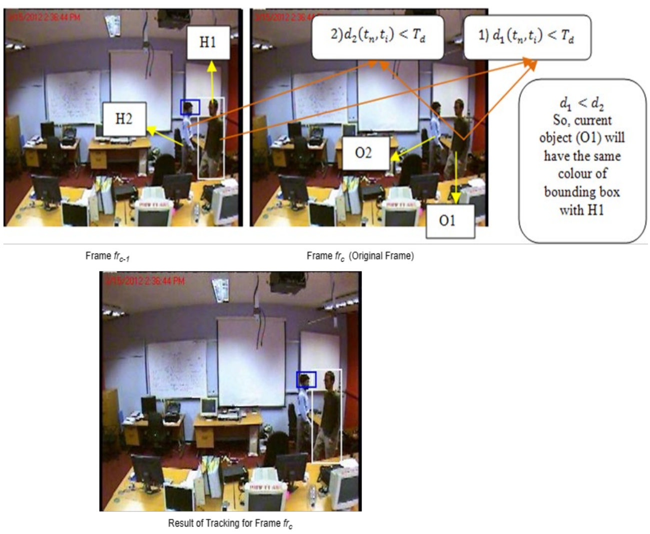

Figure 7. Normal Case in Property 2

D. Merging Case

The merged humans or objects in the current frame fr_c will have the black colour of bounding box. The merged humans or objects also will have the aspect ratio in the range of aspect ratio of single human or single object. The merging case would happen if it satisfied the following conditions:

1) The number of identified humans or objects in the previous frame fr_{c-1} is more than the detected human or object in the current frame fr_c.

2) The distance d_t of merged human or object with the identified humans or objects (at least two humans or objects) in the previous frame fr_{c-1} should be less than threshold value Td which is 25. If the distance, dt is more than threshold value, Td then, it is considered as leaving case.

The result for merging case is shown in the figure 8.

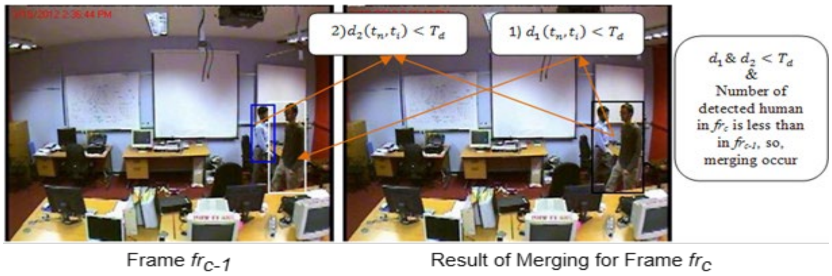

Figure 8. Result of merging

E. Splitting Case

The split humans or objects in the current frame fr_c will inherit the same colour of bounding box as humans or objects in the frame before the merging occur. The splitting case will happen if it satisfied the following conditions:

1) The number of identified humans or objects in the previous frame fr_{c-1} is less than the detected human or object in the current frame, fr_c.

2) Merging has occurred in the previous frame fr_{c-1}.

3) The colour distance dc of identified humans or objects in the frame before merging occurs and the split humans or objects in the current frame, fr_c should be less than threshold value Tc which is 25.

In order to get the colour similarity of human or objects, firstly we have to build the colour model of the identified humans or objects in the frame before merging occurs and also the colour model of the detected human or object in the current frame fr_c. Then, we check the similarity by using Euclidean distance formula as explained in the next sub section.

1) Colour Model Building: We utilized colour histogram in order to build the colour model for the detected human or object. The distribution of colours in an image is represented by a colour histogram. Thus, the distribution of pixels that span the image's colour space can be represented by a colour histogram. The colour histogram also is a statistic that could be analysed as an estimation of continuous distribution of colours' pixels values. Colour model for an image using colour histograms can be built in various colour spaces such as RGB, RG Chromaticity, HSV or any other colour space. A histogram of an image is produced by distributing the colours in the image into a number of bins and counting the number of image pixels in each bin. For colour model building in splitting case, we have used RGB colour space. Even though HSV colour space is able to take apart colour component from intensity, RGB colour space is capable of representing the distribution of colour of a blob (i.e., colour of vehicle's body and dress of a person) more correctly. Here, first of all we did the experiments for colour model RGB colour space using four and eight bins. Then, we analyzed the similarity between the images by using Euclidean Distance formula. The most accurate results that we got from the RGB colour space using four and eight bins will be utilized in our research. The details of the experiments are explained in the next sub section. Based on the experiments results, we choose to build our colour model based on RGB colour space using eight bins.

A) RGB Colour Space using Four Bins: For colour model based on RGB colour space using four bins, a histogram of an image is produced by distributing the colours in the image into four bins and counting the number of image pixels in each bin as shown in I. Rb, Gb and Bb are number of image pixels of red, green and blue colours respectively in each bin. For example, R3 means number of image pixels of red colour in Bin 3 which falls under the range of 192 – 255 from red colour pixels' values of that particular image. Table I shows the distribution of colours for the image in figure 9.

Figure 9. Extracted Human

B) RGB Colour Space using Eight Bins: Here, a histogram of an image is produced by distributing the colours in the image into eight bins and counting the number of image pixels in each bin as shown in table 2. R_b, G_b and B_b are number of image pixels of red, green and blue colours respectively in each bin. For example, R_6 means number of image pixels of red colour in Bin 6 which falls under the range of 192 – 223 from red colour pixels' values of that particular image. Table 2 shows the distribution of colours for the image in figure 9.

Table 1. Distribution of Colours for the Image in Figure 3.35 based on RGB Colour Space using Four Bins

	Bin 0	**Bin 1**	**Bin 2**	**Bin 3**
	0 – 31	32 – 63	64 – 95	96 – 127
R_b	192	63	1	0
G_b	203	48	5	0
B_b	201	53	1	0

2) Similarity between Colour Models: Similarity d_c denotes the matching degree of the colour histogram of the identified humans or objects in the frame before merging occurs and the split humans or objects in the current frame f_{rc}. We adopted the Euclidean Distance to compute the similarity between two colour histograms as in (4).

$$d_c(c_n, c_i) = \sqrt{\sum_{b=0}^{b=7} |R_{bn} - R_{bi}|^2 + |G_{bn} - G_{bi}|^2 + |B_{bn} - B_{bi}|^2} \quad (4)$$

Here, $c_n(R_bn, G_bn, B_bn)$ is the colour model of the split humans or objects in the current frame f_{rc} which is based on RGB colour space using eight bins. Meanwhile, $c_i(R_bi, G_bi, B_bi)$ is the colour model of the identified humans or objects in the frame before merging occurs which is based on RGB colour space using eight bins. The $R_bn, G_bn, B_bn, R_bi, G_bi$ and B_bi are similar to R_b,G_b and B_b in the table I and table II. Experiments showed that the different images give the colour distance of higher value using eight bins and also the similar images give the colour distance of quite lower value using eight bins. Meanwhile, using four bins, it is difficult to classify the images since the colour distance value gives almost random values for both similar and different images. Thus, we choose to build our colour model based on RGB colour space using eight bins. We choose the threshold value T_c to be 25. This means that if the d_c for image1 and image2 is more than 25, then both image is concluded as different image and if the d_c for image1 and image2 is less than 25, then both image is concluded as similar image.

Figure 10 (frame 410) shows the results of splitting case tracking using similarity colour histogram based on RGB colour space using eight bins. There are three tracking cases are included in the 10 which are normal case, merging case and splitting case. Frame 392 until frame 394 demonstrates the normal tracking case, while frame 395 until frame 409 illustrates the merging tracking case. Frame 410 illustrates the splitting case. We can observe in the frame 410 that split humans in that particular frame inherit the same colour of bounding box as before the merging occur which is at the frame 394. Frame 411 and frame 412 again demonstrates the normal tracking case. The GRAPHICAL USER INTERFACE (GUI) of the developed system is shown in figure 11.

Table 2. Distribution of Colours for the Image in Figure 3.35 based on RGB Colour Space using Four Bins

	Bin 0	**Bin 1**	**Bin 2**	**Bin 3**	**Bin 4**	**Bin 5**	**Bin 6**	**Bin 7**
	0 – 31	32 – 63	64 – 95	96 – 127	128 – 159	160 – 191	192 – 223	224 – 255
R_b	192	63	1	0	0	0	0	0
G_b	203	48	5	0	0	0	0	0

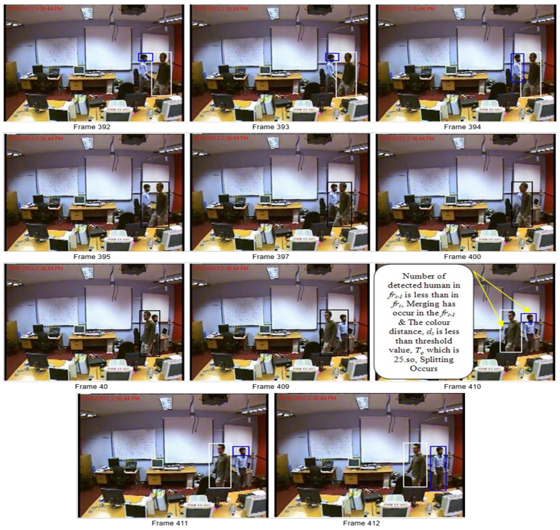

Figure 10. Results of Tracking

3.0 RESULT AND DISCUSSION

This section examines the experimental results of proposed video surveillance systems. The experiments have been conducted using public datasets and recorded videos in International Islamic University Malaysia (IIUM). The experiments were designed in order to verify the accuracy and performance of proposed video surveillance systems in various environmental conditions. These types of feasibility analysis are believed to be significant to execute the developed video surveillance system practically. Object in this section means vehicle which is car.

3.1 Results on Occluded Human

1) Results on Occluded Human with Background Objects: Our technique gives robust results by classifying human even in the presence of occlusion and in under different environments and lighting variations. We used IIUM video1 as our benchmark video to identify the effectiveness of our algorithm to extract and detect human head-shoulder in the presence of occlusion. Figure 12 shows the tracking result using our technique. We can observe from figure 12 that same human is tracked from partially occluded (human head shoulder) to not partially occluded (human full body) and this is indicated using same colour of bounding box. Thus, our tracking technique works on tracking the same human even under partially occluded or not occluded.

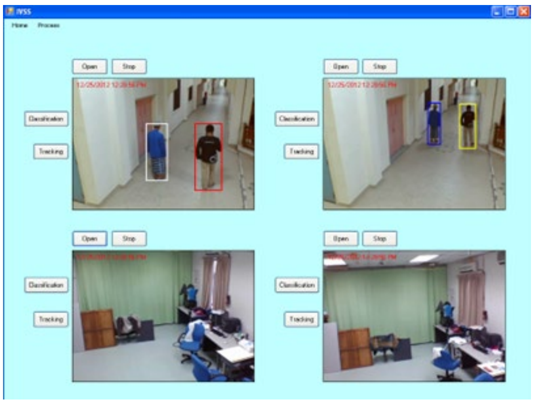

Figure 11. The GRAPHICAL USER INTERFACE (GUI) of the developed system

Figure 12. Results for Human Tracking from IIUMvideo1

2) *Results on Occluded Human with another Human:* Figure 13 shows IIUMvideo5 which consists of 623 frames. This video consists of two people entering the lab, then occluded each other and then leave the lab. This video also contains partially occluded human due to background objects which is table. Figure 13 shows the result of successful tracking of human. It can be observed from figure 13 that, two people is tracked with different colour of bounding box, but when they occluded each other, merging occurs, and this is indicated using black colour of bounding box. Then, when they split, two people is identified and tracked as before occlusion occurs and this is shown by the colour of bounding box which is same as before the occlusion occurs.

Figure 13. Results for Human Tracking from IIUMvideo5

3.2 Accuracy of Human and Objects' Classification and Tracking

We evaluate the proposed framework in two aspects which are 1- Human in indoor and outdoor 2- Car in outdoor. The proposed method used single core programming using C# programming language and it was implemented to a Laptop (Intel Core i5-2410M, 2.30GHz with 2GB of memory). For human, we use two public datasets and recorded videos in International Islamic University Malaysia (IIUM) to evaluate the algorithmic effectiveness. Meanwhile, for car, we use other recorded videos in IIUM. Public datasets that we used is adopted from CAVIAR Datasets and SPEVI Datasets. SPEVI datasets have frame rate of 25 fps and the frame size of 360 × 288 pixels, while CAVIAR datasets have frame rate of 25 fps and the frame size of 384 × 288 pixels. Recorded videos in IIUM have frame rate of 30 fps and the frame size of 320 × 240 pixels.

The results for the proposed framework are categorized into three types which are true positives, false negatives, and false positives error. For human and object's tracking, true positives happen when the human or object is successfully tracked when it supposed to be tracked, while false negatives happen when the human or object is not tracked when it supposed to be tracked. Meanwhile, false positives happen when the system tracked non-human as human or non-object as object. Here, colour of bounding box is used in this research as an indicator to display the tracked human or object. The same human or object should have the identical colour of bounding box throughout the scene. However, we utilized different thickness of bounding box to indicate the difference between tracking of human and object (car). The bounding box of object (car) will be thicker than bounding box of human as shown in figure 14.

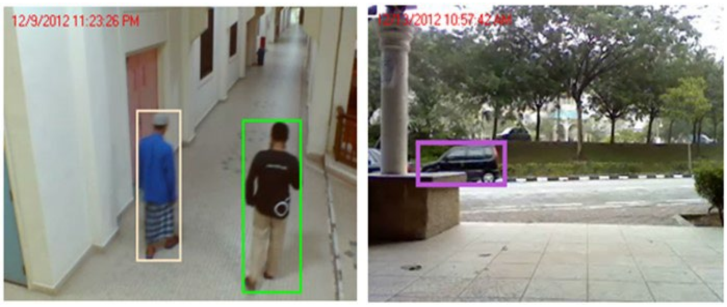

Figure 14. Different Thickness of Bounding Box to Indicate the Difference between Tracking of Human and Object (Car)

1) *Accuracy for Human Tracking:* Here, we used two videos from CAVIAR datasets, one from SPEVI datasets and six from recorded videos in IIUM to test our proposed technique. The videos that were used are in different environments and lighting variations. The videos from SPEVI datasets and recorded videos in IIUM include occlusion where the human full body cannot be classified. In the cases where partial occlusion occurs, the system identified human by extracting and classifying the human head-shoulder. During evaluation, the following criteria have been taken into account which is humans with less than 50% visibility at image boundaries and humans outside effective detection area region of the datasets were not used in the evaluation process. The effective detection area corresponds to the area in which human can be successfully detected due to preservation of human size. The accuracy for result of human tracking using our technique is calculated as shown in (5).

$$Accuracy\ (\%) = \frac{Number\ of\ True\ Positives}{Total\ of\ Humans * 100} \tag{5}$$

The graphical representation of human tracking's result for various videos are shown in figure 15. Moreover, the overall accuracy for our technique in tracking human is summarized in table 3, where the overall accuracy is very good which is at 95.83%.

Table 3: Overall Result of Human Tracking Using Our Technique

Total Humans	True Positives	False Negatives	Accuracy (%)
24	23	1	95.83

2) *Accuracy for Object Tracking:* Here, we used six recorded videos in IIUM to test our proposed technique. During evaluation, the following criteria have been taken into account which is car with less than 65% visibility at image boundaries and car outside effective detection area region of the datasets were not used in the evaluation process. The effective detection area corresponds to the area in which car can be successfully detected due to preservation of car size. The accuracy for result of car tracking using our technique is calculated as shown in (6).

$$Accuracy\ (\%) = \frac{Number\ of\ True\ Positives}{Total\ of\ Car\ * 100} \tag{6}$$

The graphical representation of car tracking's result for various videos is shown in figure 16. The overall accuracy for our technique in tracking car is summarized in table 4, where the overall accuracy is good which is at 93.94%.

Table 4. Overall Result of Car Tracking Using Our

Total Cars	True Positives	False Negatives	Accuracy (%)
33	31	2	93.94

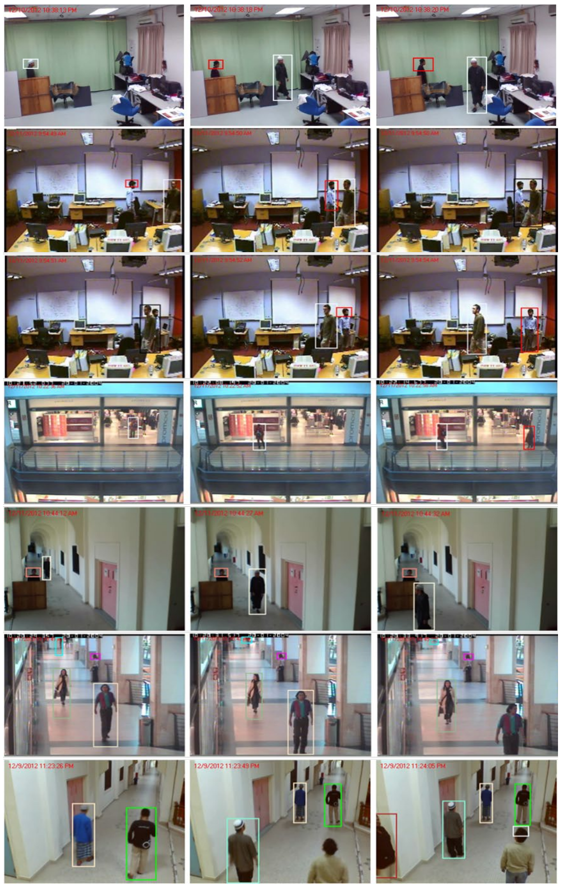

Figure 15. Results for Human Tracking from various videos

3) *Discussion:* The overall test result shows the successful implementation of our technique in tracking human and objects in various environments and also at various camera parameters which includes angle and height of camera. The overall accuracy for our technique in tracking human is also very good which is at 95.83%. The overall accuracy for our technique in tracking car is fine which is at 93.94%. The overall accuracy for our technique in tracking human and car is summarized in table 5 which is 94.75%. However, our system gives slightly lower accuracy in tracking the human and objects in cluttered environment, night surveillance and completely occluded situations.

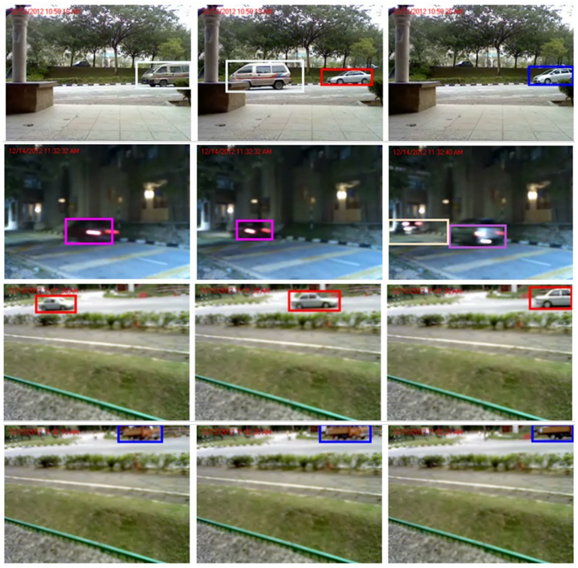

Figure 16. Results for Car Tracking from various videos

Table 5. Overall Result of Human and Car Tracking Using Our Technique

Total Humans and Cars	True Positives	False Negatives	Accuracy (%)
57	54	3	94.75

4.0 CONCLUSION

In this research, we have developed a system for human and objects tracking as this is a significant research area in both pattern recognition community and machine vision because of its possible functions in video surveillance system. We proposed a system to track humans and cars by utilizing feature-based tracking and Euclidean distance formula for four states, which are: entering, leaving, normal and merging. For splitting case, we build colour model based on RGB colour space using eight bins. Then, we find the similarity between the colour models of the objects using Euclidean distance formula to track them.

Effective performance evaluation is deemed important towards achieving successful tracking systems with higher accuracy and less error. The developed system was successfully tested in the different real environments and lighting variations. From the point of practical implementation, several experiments have been conducted in order to verify the performance of the system and the results showed that the developed system is proficient with high percentage of accuracy and also with less error.

The developed system can track the human even if occlusion occurs since we used merging and splitting case in our tracking algorithm. Moreover, the overall accuracy for our technique in tracking humans and cars is high, which is at 94.75%. However, our system also has limitations which produce slightly lower results of accuracy. Our system gives poor tracking results when many classified humans or objects are available in the scene. It also fails to track the cases where merging and splitting occur for more than three humans. These limitations of our system can be further researched in order to achieve better results.

5.0 REFERENCES

[1] K. Goya, X. Zhang, K. Kitayama, and I. Nagayama, "A method for automatic detection of crimes for public security by using motion analysis," *IIH-MSP 2009 - 2009 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 736–741, 2009.

[2] A. Hampapur, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl, S. Pankanti, A. Senior, C. F. Shu, and Y. L. Tian, “Smart video surveillance,” *IEEE Signal Processing Magazine*, vol. 22, no. 2, pp. 38–51, 2005.

[3] Z. Soleimanitaleb, M. A. Keyvanrad, and A. Jafari, “Object tracking methods: A review,” *2019 9th International Conference on Computer and Knowledge Engineering, ICCKE 2019*, pp. 282–288, oct 2019.

[4] W. Luo, J. Xing, A. Milan, X. Zhang, W. Liu, and T. K. Kim, “Multiple object tracking: A literature review,” *Artificial Intelligence*, vol. 293, p. 103448, apr 2021.

[5] A. Mangawati, Mohana, M. Leesan, and H. V. Aradhya, “Object

Tracking Algorithms for Video Surveillance Applications,” *Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing, ICCSP 2018*, pp. 667–671, nov 2018.

[6] S. Ojha and S. Sakhare, “Image processing techniques for object tracking in video surveillance- A survey,” 2015 International Conference on Pervasive Computing: Advance Communication Technology and Application for Society, ICPC 2015, apr 2015.

[7] B. Wu and R. Nevatia, “Detection and tracking of multiple, partially occluded humans by Bayesian combination of edgelet based part detectors,” *International Journal of Computer Vision*, vol. 75, no. 2, pp. 247–266, nov 2007. [Online]. Available:

https://link.springer.com/article/10.1007/s11263-006-0027-7

[8] H. K. Chavda and M. Dhamecha, “Moving object tracking using PTZ camera in video surveillance system,” *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing, ICECDS 2017*, pp. 263–266, jun 2018.

[9] J. S. Yuk, K. Y. K. Wong, R. H. Chung, F. Y. Chin, and K. P. Chow, “Real-time multiple head shape detection and tracking system with decentralized trackers,” *Proceedings - ISDA 2006: Sixth International Conference on Intelligent Systems Design and Applications*, vol. 2, pp. 384–389, 2006.

[10] L. Li, W. Huang, I. Y. H. Gu, R. Luo, and Q. Tian, “An efficient sequential approach to tracking multiple objects through crowds for real-time intelligent CCTV systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 5, pp. 1254– 1269, 2008.

[11] A. Yilmaz, O. Javed, and M. Shah, “Object tracking,” *ACM Computing Surveys (CSUR)*, vol. 38, no. 4, dec 2006. [Online]. Available: https://dl.acm.org/doi/10.1145/1177352.1177355

[12] H. Lee and D. Kim, “Salient Region-Based Online Object Tracking,” *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018*, vol. 2018-January, pp. 1170–1177, may 2018.

[13] H. Tjaden, U. Schwanecke, E. Schomer, and D. Cremers, “A RegionBased Gauss-Newton Approach to Real-Time Monocular Multiple Object Tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 8, pp. 1797–1812, aug 2019.

[14] M. Stoiber, M. Pfanne, K. H. Strobl, R. Triebel, and A. Albu-Schaffer,¨ “SRT3D: A Sparse Region-Based 3D Object Tracking Approach for the Real World,” *International Journal of Computer Vision*, vol. 130, no. 4, pp. 1008–1030, apr 2022. [Online]. Available:

https://link.springer.com/article/10.1007/s11263-022-01579-8

[15] T. Chen, “Object tracking based on active contour model by neural fuzzy network,” Proceedings - 2009 IITA International Conference on Control, Automation and Systems Engineering, CASE 2009, pp. 570– 574, 2009.

[16] R. O’Malley, E. Jones, and M. Glavin, “Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 453–462, jun 2010.

[17] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, “A real-time computer vision system for measuring traffic parameters,” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 495–501, 1997.

[18] J. Y. Choi, K. S. Sung, and Y. K. Yang, “Multiple vehicles detection and tracking based on scale-invariant feature transform,” *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 528–533, 2007.

[19] S. Guo, T. Zhang, Y. Song, and F. Qian, “Color Feature-Based Object Tracking through Particle Swarm Optimization with Improved Inertia Weight,” *Sensors 2018, Vol. 18, Page 1292*, vol. 18, no. 4, p. 1292, apr 2018. [Online]. Available: https://www.mdpi.com/14248220/18/4/1292/htm https://www.mdpi.com/1424-8220/18/4/1292

[20] N. I. H. Fauzi, Z. Musa, and F. Hujainah, “Feature-Based Object Detection and Tracking: A Systematic Literature Review,” *https://doi.org/10.1142/S0219467824500372*, feb 2023.

[21] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 34, no. 3, pp. 334–352, aug 2004.

[22] V. Lepetit and P. Fua, "Monocular Model-Based 3D Tracking of Rigid Objects: A Survey," *Foundations and Trends® in Computer Graphics and Vision*, vol. 1, no. 1, pp. 1–89, 2005. [Online]. Available: http://dx.doi.org/10.1561/0600000001

[23] C. O. Conaire, N. E. O'Connor, E. Cooke, and A. F. Smeaton,´ "Multispectral object segmentation and retrieval in surveillance video," *Proceedings - International Conference on Image Processing, ICIP*, pp. 2381–2384, 2006.

[24] C. T. Lin, Y. C. Huang, T. W. Mei, H. C. Pu, and C. T. Hong, "Multiobjects tracking system using adaptive background reconstruction technique and its application to traffic parameters extraction," *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, vol. 3, pp. 2057–2062, 2006.

[25] L. Li, W. Huang, I. Y. Gu, K. Leman, and Q. Tian, "Principal color representation for tracking persons," *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 1007–1012, 2003.

[26] S. Yang and M. Baum, "Extended Kalman filter for extended object tracking," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 4386–4390, jun 2017.

[27] P. R. Gunjal, B. R. Gunjal, H. A. Shinde, S. M. Vanam, and S. S. Aher, "Moving Object Tracking Using Kalman Filter," *2018 International Conference On Advances in Communication and Computing Technology, ICACCT 2018*, pp. 544–547, nov 2018.

[28] F. Farahi and H. S. Yazdi, "Probabilistic Kalman filter for moving object tracking," *Signal Processing: Image Communication*, vol. 82, p. 115751, mar 2020.

[29] Y. Huang, Y. Zhang, P. Shi, Z. Wu, J. Qian, and J. A. Chambers, "Robust Kalman Filters Based on Gaussian Scale Mixture Distributions with Application to Target Tracking," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 10, pp. 2082–2096, oct 2019.

[30] Y. Xu, K. Xu, J. Wan, Z. Xiong, and Y. Li, "Research on Particle Filter Tracking Method Based on Kalman Filter," Proceedings of 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference, IMCEC 2018, pp. 1564–1568, sep 2018.

[31] H. Wang, C. Liu, L. Xu, M. Tang, and X. Wu, "Multiple feature fusion for tracking of moving objects in video surveillance," *Proceedings - 2008 International Conference on Computational Intelligence and Security, CIS 2008*, vol. 1, pp. 554–559, 2008.

[32] T. Zhang, S. Liu, C. Xu, B. Liu, and M. H. Yang, "Correlation Particle Filter for Visual Tracking," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2676–2687, jun 2018.

[33] T. Zhang, C. Xu, and M. H. Yang, "Learning Multi-Task Correlation Particle Filters for Visual Tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 365–378, feb 2019.

[34] X. Zhang, Z. Yan, Y. Chen, and Y. Yuan, "A novel particle filter for extended target tracking with random hypersurface model," *Applied Mathematics and Computation*, vol. 425, p. 127081, jul 2022.

[35] T. Zhang, C. Xu, and M. H. Yang, "Multi-task correlation particle filter for robust object tracking," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017January, pp. 4819–4827, nov 2017.

[36] X. Wang, T. Li, S. Sun, and J. M. Corchado, "A Survey of Recent Advances in Particle Filters and Remaining Challenges for Multitarget Tracking," *Sensors 2017, Vol. 17, Page 2707*, vol. 17, no. 12, p.
2707, nov 2017. [Online]. Available: https://www.mdpi.com/14248220/17/12/2707/htm https://www.mdpi.com/1424-8220/17/12/2707

[37] T. Gao, G. Li, S. Lian, and J. Zhang, "Tracking video objects with feature points based particle filtering," *Multimedia Tools and Applications*, vol. 58, no. 1, pp. 1–21, may 2012. [Online]. Available: https://link.springer.com/article/10.1007/s11042-010-0676-y

[38] "CAVIAR Test Case Scenarios." [Online]. Available: https://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/

[39] "Surveillance Performance EValuation Initiative (SPEVI)." [Online]. Available: http://www.eecs.qmul.ac.uk/∼andrea/spevi.html

[40] F. Ying, W. Huiyuan, M. Shuang, and W. Xiaojuan, "Multi-object tracking based on region corresponding and improved color-histogram matching," *ISSPIT 2007 - 2007 IEEE International Symposium on Signal Processing and Information Technology*, pp. 1–4, 2007.