A novel enhanced algorithm for efficient human tracking

Mehdi Gheisari¹, Zohreh Safari², Mohammad Almasi³, Amir Hossein Pourishaban Najafabadi⁴, Abel Sridharan⁵, Ragesh G. K.⁶, Yang Liu¹, Aaqif Afzaal Abbasi⁸

¹Department of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, China
²Department of Computer Science, Texas Tech University Texas, Texas, USA
³Department of Computer Science, University of Barcelona, Barcelona, Spain
⁴Department of Electrical and Computer Engineering, University of Yazd, Yazd, Iran
⁵Department of Computer Science, University of Madras, Tamil Nadu, India
⁶Indian Institute of Information Technology Kottayam, Kerala, India
⁷Department of Computer Science and Engineering, Islamic Azad University, Tehran, Iran
⁸Department of Software Engineering, Foundation University, Islamabad, Pakistan

Article Info

ABSTRACT

Tracking moving objects has been an issue in recent years of computer vision and image processing and human tracking makes it a more significant challenge. This category has various aspects and wide applications, such as autonomous deriving, human-robot interactions, and human movement analysis. One of the issues that have always made tracking algorithms difficult is their interaction with goal recognition methods, the mutable appearance of variable aims, and simultaneous tracking of multiple goals. In this paper, a method with high efficiency and higher accuracy was compared to the previous methods for tracking just objects using imaging with the fixed camera is introduced. The proposed algorithm operates in four steps in such a way as to identify a fixed background and remove noise from that. This background is used to subtract from movable objects. After that, while the image is being filtered, the shadows and noises of the filmed image are removed, and finally, using the bubble routing method, the mobile object will be separated and tracked. Experimental results indicated that the proposed model for detecting and tracking mobile objects works well and can improve the motion and trajectory estimation of objects in terms of speed and accuracy to a desirable level up to in terms of accuracy compared with previous methods.

Keywords:
Background subtraction
Bubble routing
Deep learning
Human tracking
Image filter
Movable objects
Object tracking

This is an open access article under the CC BY-SA license.

Corresponding Author:
Yang Liu
Department of Computer Science and Technology, Harbin Institute of Technology (Shenzhen)
Nanshan, Shenzhen, Guangdong, China
Email: liu.yang@hit.edu.cn

1. INTRODUCTION

This paper aims to introduce a fast method of human tracking based on background subtraction. Object tracking is a substantial task in the field of computer vision. Tracking can include the path and movement type of object. Besides, tracking can provide other information such as the direction, area perimeter or shape, speed, and acceleration of the object. But object tracking can be complicated for some reasons like loss of some part of 3D real-world information when converting to the 2D image, different types of noise in images, complex movements of the object, the non-networked nature of the object, and even more. Therefore, restrictions can be placed on object tracking in dimensions such as movement or body shape. An object's motion is almost assumed to be smooth in all tracking algorithms, without sudden direction changes. Another assumption is constant speed and acceleration for the object.
There are several methods for tracking an object. The purpose of tracking objects is to analyze the video frame by frame. The applications can be named video surveillance systems, human action recognition, and human-robot interactions. Background subtraction is an easy and crucial technique for separating foreground objects from an image. The overall accuracy of an algorithm depends on the accuracy of the object preview. The process of image background subtraction requires maintaining the image model and subtracting the foreground image from the processed frame [1]. One of the main objects in any frame to track is humans. Understanding the human from the background, different types of motion, recognizing the person with some other features like name, are just some of the main applications of human detection from a simple approach to biometric recognition.

Recent papers as [11–14] have focused on the teaching-based approach. The support vector tracker in the paper Bradski and Kaehler [5] uses an out-of-line learning support vector machine that is classified and embedded in a current-based optical tracker. In the article Allen et al. [6] the current frame is classified using a classifier learned in the previous frame. In order to track, the ratio of variance is used to measure the distinguished feature and to select the best feature from a feature set. In Ning et al. [7] poor classifiers are trained online, and pixels are labeled as target or background. Since only local spatial structures have been extracted, the resolution power decreases in crowded scenes.

In Qiu et al. [8] the target is shown in a comparatively small subdivision that has been comparatively updated, using image tracking in previous frames. In the paper, Ganoun et al. [9] a cascade particle filter with a set of different longevity characteristics is proposed, while the probabilistic model is noisy and has many peaks without Gaussian diffusion in the sampling stage. Diagnostics-based tracking has recently been studied in articles [10], [11]. This is, in fact, a method of detection-based tracking methods. If the learner learns outline, it could be considered as a special style of machine learning-based models.

As discussed above, object detection is a key important section in computer vision. The goal is to find, label and recognize the object properly. This task can be done, easily, by machine and deep learning techniques. And human is one of the crucial objects, that needs to be detected. This task has many applications in video surveillance, action recognition, robotics, and human-computer recognition. Even on the farm, human detection could be useful in farms and Agricultural vehicles. By using the histogram of oriented gradients (HOG) algorithm to detect the person and support vector machines (SVM) to classify the motion. The assembly of parts of the human body is used to some extent to identify and trace. New feedback from the object detector (visual inclination) is used to track humans [12]. Alimardani and Almasi [13] proposed a two-step approach that first creates an apparent model of individuals and then identifies them by the model detection in any frames. These methods are based on the analysis of different parts of the human body tracking. However, the image quality of many video surveillance applications may not be clear enough on these issues because the movement of body parts is difficult to precisely track. Also, using deep learning algorithms, there is a fantastic method. As in Chahyati et al. [14], they tried to track the human body, using a convolutional neural network (CNN), they got an accuracy of 70%. In Almasi et al. [15] they tried to find the human motions in an egocentric viewpoint, the goal of this method is to find the subject motion without directly tracking the subject on the scene, and by tracking objects or background purely. Another challenging task in human detection is underwater body detection in Yasar and Kusetogullari [16] they tried to track the human body using optical flow, while in Chan [17] they tackled this issue using voxel detection, with an accuracy of 67% and 82% respectively.

It should be clarified that deep learning algorithms always need some large datasets to train, test, and validate the idea. In this regard the three best datasets in terms of human tracking are the Euro City [18] the largest dataset for humans so far, the Massachusetts Institute of Technology (MIT) ped [19] with more than 700,000 images, Institut de recherche en Informatique et en Automatique (INRIA) [20] with using histogram of a gradient. One of the applications in human detection is in robotics, as in Bachuet et al. [21] for robotics applications humans could be a good way for any detection with rehabilitation or assistive approaches [22–43]. One of the best examples of object detection especially with human detection in their feature-based algorithm has been reached almost 93% accuracy, while it can be used for faces and skin as well [37]. Regarding the new deep learning algorithms for human detection, joint detection or pose detection would be a unique application of human detection, by this approach it would be quite feasible to detect the whole body from detecting the joints. Also, from this point, many applications in health, and sports will be achievable. Their goal has been done by using standard datasets like functional living index-cancer (FLIC) with an accuracy of 97%. This approach can be a real competitor for motion capture systems these marker based systems in different model are could be expensive, but pose estimation deep neural network models can do the same job with the same accuracy [38]. Both machine learning and background subtraction methods are quite useful and powerful, while in our mixed model a new SOTA is practiced [39]. The following sections are firstly the method, which will be discussed in detail. Afterward, the results and discussion will be explained. And at the end, we sum up with a conclusion.
2. **RESEARCH METHOD**

Object tracking is the most important issue in computer vision. We are tracking objects. Background subtraction is an ordinary and useful approach for separating foreground objects from the whole image. Comprehensively, the procedure of subtracting the background image requires maintaining the image model and subtracting the foreground image from the processed frame [1], [40]. Therefore, when the frame km is compared to the reference one, the inputs of the given pixel from the incremental image indicates how many specific point intensity varies from the corresponding pixel values in the reference one. So, if R (x, y) is the reference image and K represents tk so that f (x, y, k)=f (x, y, tk). Internal changes are in the background itself, such as the tree movement water levels, and flags. t is the threshold value. Video backgrounds cannot be completely fixed. This background is called quasi-fixed background [41], [42]. Video backgrounds cannot be completely static. These types of backgrounds are referred to as background quasi-constants. After removing the shadows, the "particle filter" method has been used to distinguish the moving objects in the image from each other [43], [24] as shown in Figure 1.

![Figure 1. The algorithm architecture of the method](image1)

In an attempt to determine a dynamic object in a video is the background subtraction method. Each video frame is subtracted by analyzing with a previously extracted constant image, and a motion preview image of the object will be the final result of the process [25], [26]. This image is then easily converted to binary color space after image processing. In this image, however, there is a lot of noise due to different radiation. The final image may include shadows or small non-fixed objects in addition to a reference object that is necessary to be removed [27], [28]. In order to remove unwanted objects, a set of morphological operations have been performed by deleting the erosion, removing the small noises that have been created unintentionally [29], [30].

3. **RESULTS AND DISCUSSION**

The proposed algorithm is implemented in detecting and tracking humans, human detection has lots of applications. The first step as shown in Figure 2 is background removal, in this level we tried to separate with high accuracy be foregrounded the person from the background. In a very wide color spectrum, applying different functions, filters and deciding on different image pixels will be complex and difficult. Thus, turning the image to gray or 1D black and white as shown in Figure 3 is a feasible and desirable step.

![Figure 2. The basic black and white detected person](image2)

![Figure 3. Background removal](image3)

Figure 3 is showing a walking human in Figure 4, the aim is to track the image of a person as a moving object, which can be seen in the image at the beginning of its movement. There is also a dot on the person's image, which is displayed as a bubble to represent each object. This point shows the local center
where the object is identified as a moving object in the image. A quadrilateral with a tracking square for it can also be seen around it. Each object is identified and tracked in the object tracking algorithm and in its implementation with a quadrilateral around it. Figure 5 shows that the person is tracked at the beginning of his movement as a moving object with high accuracy in the image. Of course, this tracking shows a little more than the person's range in the image due to the hand movement [31], [32]. In some places of the object, pixels close to the background color are also detected in motion due to the motion type or the speed. In the proposed method, the better this diagnosis, the more accurate and efficient the filters used in image processing.

As shown in Figure 5, the full range of the tracked object is not identified in this image. Also, it is related to the velocity and point of the person. In addition, since the new object is detected at this time, the start of any movement has a significant effect on determining the range of the moving object. Figure 5 reveals the effect of this problem. As can be seen in the Figure, the quadrilateral tracking in the Figure is much more limited than before. The detection of the lower body but using filters other than black and white filters can help in this regard [33], [44].

As mentioned above, after subtracting the background and applying the filter, gaps and small spaces on the image line are connected and smoothed. This is done before the noise is removed, and the result may change the scope of the object slightly. Detection of the scope change and inaccurate general estimation of the object's scope, although is not the main goal in this study, has been tried to be considered in the proposed method to a large extent to introduce a method with better performance and efficiency. The proposed algorithm can be used to detect the motion of several objects as well as faster objects. The test results indicate, the suggested system provides accurate and very reliable results from the tested samples.

Figure 5. Binary image with detection motion and human

4. CONCLUSION

This research introduced a cost-effective, accurate, and efficient method for detecting a moving object and tracking it using advanced image processing techniques. It has been shown that the proposed system for tracking moving objects with background algorithm, noise cancellation, filtering, and bubble routing, provides full automation by evaluation and dynamic object detection on images. More importantly, system accuracy in determining a moving object in road images is in accordance with the standards set by the decision-making authorities for road control. The most important applications can be named human action recognition and behavior analysis, video surveillance, and autonomous driving.

ACKNOWLEDGEMENTS

This work has been supported by the Shenzhen Basic Research (General Project) (No. JCYJ20190806142601687) and Shenzhen Stable Supporting Program (General Project) (No. GXWD20201230155427003-20200821160539001).

REFERENCES


A novel enhanced algorithm for efficient human tracking (Mehdi Gheisari)
BIOGRAPHIES OF AUTHORS

Mehdi Gheisari is a Ph.D. holder in computer science from Guangzhou University, China. His primary research is about the Privacy-preservation of IoT devices data, including remote sensing data in smart cities. Prior to that, he was with the Islamic Azad University, where he served in the capacity of lecturer in the department of computer science. There, he worked on Wireless Sensor Networks. Furthermore, he has published papers in several domains in highly ranked journals and ranked conferences. He has been serving in academic works as well, such as reviewing papers from several well-known venues such as IEEE communication magazine, and TPC of several conferences. He can be contacted at email: mehdi.gheisari61@gmail.com

Zohreh Safari will receive her Ph.D in August 2022 in Computer Science at Texas Tech University, her research focuses on Clustering High dimensional Big Data to optimize clustering algorithm for data sets with high dimensions, and perform efficiently for retrieval information. She is interested in developing machine learning algorithms, Optimizations, AI in healthcare applications. Has MSc and BSc in Software Engineering. She can be contacted at email: zohreh.safari@ttu.edu.

Mohammad Almasi started his Ph.D 2021 in Computer Science at the University of Barcelona UB and Computer vision center at UAB. His works is on application of Computer Vision and Deep Learning on Sports and Biomechanics. He has received his MSc. From University Polytechnical of Catalonia in Mathematical Engineering. Please kindly follow his updates on ResearchGate and Scopus. For more information and contact him use the personal email: mohamadalmas2@gmail.com
Amir Hossein Pourishaban Najafabadi received his Bachelor's degree in Computer Engineering from Yazd University in February 2021. His research focuses on Gradient Descent Algorithms (Optimization Algorithms). He is interested in Machine Learning and Neural Network Algorithms. He can be contacted at personal email: Amirhosseinpurishaban@gmail.com.

Abel Sridharan is a Senior Manager–Digital Engineering that has been graduated from the University of Madras, India. He can be contacted at email: abel.sridharan@gmail.com.

Ragesh G K secured a Ph. D. from Anna University, Chennai, India (Full-time Ph. D. Scholar, Department of Computer Science and Engineering, Government College of Technology, Coimbatore, Tamilnadu, India from July 2011 to July 2013) and B.E and M.E. also from Anna University, Chennai, India. He is currently working as Assistant Professor and Faculty-In-Charge for IPR and IIC at Indian Institute of Information Technology (IIIT) Kottayam (an Institute of National Importance), Kerala, India since August 2021. His areas of interest include Data Security and Privacy, Cloud Security, the Internet of Things, and Wireless Communication and Networks. He can be contacted at email: rageshwarrier@gmail.com.

Yang Liu received his D. Phil (Ph. D) degree in Computer Science from University of Oxford (advised by Professor Andrew Simpson) in July 2018. He is currently an Assistant Professor and an Associate Research Fellow in Department of Computer Science and Technology of Harbin Institute of Technology, Shenzhen. He is interested in security and privacy problems and, in particular, the privacy issues on mobile devices. Prior to joining Oxford, He received an MSc in Software Engineering from Peking University and a B. Eng in Computer Science from Ocean University of China. He can be contacted at email: liu.yang@hit.edu.cn.

Aaqif Afzaal Abbasi received the Ph.D degree in computer engineering from the School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China. He is currently serving as Associate Professor with the Department of Software Engineering, Foundation University, Islamabad, Pakistan. He was a visiting researcher at University of Essex, UK in 2014 and served as Assistant Professor at Wuhan University, China between 2017 and 2018. His general research interests cover the broad area of computer communications with special emphasis on resource management. Specifically, his research is focused on designing new resource allocation algorithms, analyzing their performance, and prototyping those that find practical applications in large-scale computing systems. He is also interested in areas of application-awareness in high-performance distributed computing. Abbasi published numerous research papers in impact factor journals and hosted several special issues on topics of his interest. He is a member of the IEEE and ACM. He can be contacted at email: aaqif.afzaal@yahoo.com.