International Journal of Advanced Intelligence
Volume 9, Number 3, pp.397-407, September, 2017.

Human Detection and Tracking Using HOG Feature and Particle Filter in Video Surveillance System

Bhuyain Mobarok Hossain

System Innovation Engineering, Faculty of Engineering, Tokushima University, 2-1, Minami-Josanjima-cho, Tokushima, 770-0811, Japan. E-mail:mdlalien001@gmail.com

Stephen Karungaru, Kenji Terada, Akinori Tsuji

System Innovation Engineering, Faculty of Engineering, Tokushima University, 2-1, Minami-Josanjima-cho, Tokushima, 770-0811, Japan. E-mail:karunga@is.tokushima-u.ac.jp

Received (15 Sep. 2017)
Revised (25 Sep. 2017)

Video surveillance system has recently attracted much attention in various fields for monitoring and ensuring security. One of its promising applications is in crowd control to maintain the general security in public places. However, the problem of video surveillance systems is the required continuous manual monitoring especially for crime deterrence. In order to assist the security monitoring the live surveillance systems, intelligent target detection and tracking techniques can send a warning signal to the monitoring officers automatically. Towards this end, in this paper, we propose an innovative method to detect and track a target person in a crowded area using the individual's features. In the proposed method, to realize automatic detection and tracking we combine Histogram of Oriented Gradient (HOG) feature detection with a particle filter. The HOG feature is applied for the description of contour detection for the person, while the particle filter is used for tracking the targets using skin and clothes color based features. We have developed the evaluation system implementing our proposed method. From the experimental results, we have achieved high accuracy detection rate and tracked the specific target precisely.

Keywords: HOG Feature, Particle Filter, Tracking people, Video Surveillance System.

1. Introduction

Human detection in crowds has been one of the major challenges over the years owing to their variable faces, shapes and environment. Moreover, tracking of human is a very important research subject in the field of computer science [1]. With the spread of applications of video systems surveillance, the interest in the field of research in image processing and computer vision in the crowded scenes is growing because of the demand for security and safety systems in public places [2]. Many researchers proposed their algorithms focusing on normal problem of tracking without specific crowded scene [3].

For those interested in safety and security systems, crowded scenes in public places such as airports, bus station, concerts, subway, religious festivals, football matches, railway stations, shopping markets and places where a lot of people gather, are a huge challenge. In these cases, we need to watch only a specific place or area like entrance and exit for monitoring. A crowded scene can be divided into structured and unstructured, fig. 1.

Figure 1. Experimental Scenes (A) Structured and (B) Unstructured.

Structured crowded scene have a fixed direction of motion the does not change much with time. On the other hand, in case of unstructured crowded scene, the direction of movement is random and overlapping movements of people is visible in all directions [3].

In the past few years, computer vision algorithms have incorporated in wide area of video surveillance systems, such as city streets and subway stations, shopping centers and other places. However, one major weakness of these systems is their inability to deal with crowded scenes. The challenge of research in crowded scenes is frequent and instantaneous occlusion of the target by other objects. The visual occlusions and ambiguities in crowded scenes are complex, making scene semantics difficult to analyze.

In this study, we have focused on the overlapping issues in the case of a target person wearing similar clothes aiming to overcome the occlusion problem.

2. Related work

For some years now, many algorithms have been proposed to detect and track a human target, and recognize a crowded scene in the field of computer vision. For target tracking, many survey papers have been published recently. Myo Thida et al. presented macroscopic and microscopic models for crowd event detection [4]. Li Teng, et al. provided survey of crowded scene analysis, which considered crowd behavior and activities [5].

Aggarwal and Cai reviewed motion analysis, tracking and recognizing human activities [6]. Wang et al. presented human detection, tracking and behavioral understanding [7]. Hu et al. also presented motion detection, tracking and behavior understanding [8]. Zhan et al. investigated crowd information extraction and crowd modeling [9]. On the other hand, using particle filter aiming for tracking people was presented with the image classification techniques to enhance the results and solve the problems of contour information in video surveillance.

SIFT, Harris-SIFT [10], AdaBoost, and Cascaded AdaBoost are used to extract the human area from video [11]. Dalal & Triggs presented a human detection algorithm with excellent results. Their method used a dense grid of HOG feature, computed over blocks of size 16 × 16 pixels to represent a detection window. This representation is enough to classify human using a linear SVM [12]. Ruiyue Xu applied multiple human detection and tracking based on head detection for real-time video surveillance [13].

In contrast to existing methods, our proposed method have focused on the target detection and tracking based on a feature of individual person using the combination HOG feature detection and particle filter as described in the next sections.

3. Proposed Method

Tracking and detecting target object in crowded area function is essential to assist the monitoring officers in video surveillance system. In crowded scenes including many people, each of them has different figures and features like hair color, height, weight, clothes color etc. In our paper, we evaluated an innovative method for detecting and tracking only a specific person in a crowded area.

Figure 2 shows the flowchart of our proposed method. Our proposed method is divided into two parts; human detection and tracking. First, we fix the features of the specific person, which are already known by the monitoring officers. For tracking specific person, we use particle filter.

In our method, first we use Histogram of Oriented Gradient (HOG) for detecting human area. Next, we detect multi-human area in crowded scene. When we find multi-human area, we use particle filter with feature for tracking the specified person.

4. Pre-processing

In our method, we use the HSV color space in order to detect and track specific person. For this reason, first we transformed RGB color space to HSV color space, because, the RGB color space is not robust because it is affected by environmental conditions when the illumination is changed. The HSV color space is useful for detection and tracking to extract the color information of the target such as skin, clothes, hair, and shoes. In the flowchart, when setting the features of target, we set the parameters of the targets information as hue and brightness selecting from HSV color space.

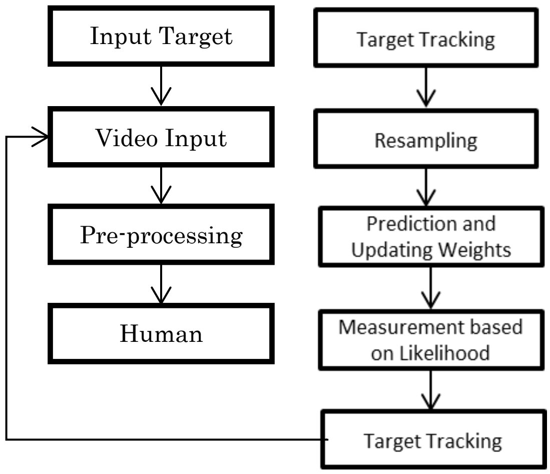


Figure 2. Proposed method flowchart

4.1 Human detection

Human detection system is divided into two main categories: Component-based methods and single detection window analysis [14]. Component-based techniques are used to detect object parts separately and check them in case of normal geometric configuration and uses hierarchical systems detection window. Body parts are detected according to priority; if one part is missing, other parts are not searched. It is slower than a single detection window and doesn't deal with multi-view and multi-pose cases. In contrast, the single detection window method is based on sequence-based detection window labeled at all possible sub-windows specific application image [14]. The most important feature is its speed, while in partial occlusion handling flaw is its limitation. In our system, we choose one frame analysis because of its speed.

The HOG-SVM detector is one of the composite and computationally heavy algorithms. The important feature of this system is calculating the image using horizontal and vertical kernels. Also, it computes the gradient magnitude and orientation image. After extracting image patch, the patch is separated into cells of equal size. For each cell, a histogram of some length L is created and each gradient pixel from the cell casts a vote proportional to the magnitude into the bin.

To increase the efficiency of this algorithm, some steps were explored including parameter tweaking and the way data is acquired. The HOG descriptor counts occurrences of gradient orientation in localized portions of an image; it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization. In addition, SVM is applied with the HOG descriptor to efficiently detect human [15]. For the HOG descriptor, the important features in the scene are calculated using horizontal and vertical kernels, which compute the features of gradient magnitude and orientation. The SVM parameters are already learned especially for detecting human before starting the algorithm. It reduces the computational time for learning process [12, 15].

In this process, we use the following HOG detection parameters; block size of (16x16) and cell size of (8x8). After detecting human successfully, they are surrounded by rectangles including the target person.

4.2. Particle filter and Tracking specific target

The particle filter method was originally designed for single object tracking. To reach that goal, it tends to focus and create the particles used to estimate [16]. Particle filters are ways the state assessment systems with non-linear process and measurement models deals with noise that may be non-Gauss and multimodal [17]. We apply the particle filter to the image with human region extracted by the HOG-SVM. The processing by particle filter has several steps, resampling, prediction and updating the weights, and measurement based on likelihood function. The main aims for using particle filter is for tracking a single or multiple objects [18].

Our particle filter is implementing recursive Bayesian filter by Monte Carlo sampling; it represents the intensity of the rear set of random particles with associated weights and also calculates estimates [18]. In the case of our method, the Bayesian filter by approximating the prior and posterior density functions are used with a set of the target feature. Therefore, the generated particles move to the specific target according to the feature set of individual person. For set up of the particle filter, we use the number of particles, variance to randomize particle and scanning range for likelihood function as the parameters. When we successfully track the target, particles gather around the area of detected target.

We also defined the parameter of target features as skin color, red, blue and black color shirts. For the particle filter, the number of particle is set as 200, variance to randomize particle set as 8.0 and scanning range of likelihood set as 30 by 30 pixels.

5. Experiments

The aims of experiments are to detect and track specific target person. We assume a crowded situation where human body embolism happens due to the transfer of the person. The occlusion is a significant and difficult problem in crowded scenes as we mentioned before. We conduct the experiments outdoors creating a dataset to evaluate our method. The specification of the experiment environment is shown in Table 1.

Table 1. Specifications of experiment environment.

Computer	Intel Core i7, Memory 8GB
Camera	Logicool USB Camera
Resolution	640 x 340, 30fps
Software	Visual C++, OpenCV

5.1 Experiment Setup

In this paper, our experiments are used for human detection and tracking specific person in video surveillance system. We used four structured and unstructured crowded scene for testing our method. Each scene is different; Scene1, Scene 2, Scene 3, and Scene 4 of the recorded videos are as shown in Table 2. All dataset are crowded scenes. The background and illumination of the video is natural condition.

Table 2. Test datasets for experiments

Dataset	Total No, of frames	Time duration(sec)
Scene 1	4830 frames	241
Scene 2	1199 frames	62.0
Scene 3	1962 frames	98
Scene 4	2197 frames	110

5.2 **Experimental Results**

We validate our method on the four different data sets. Every data set is our own dataset. We discuss the experimental results of the datasets. It involves detecting and tracking specified person walking on structured and unstructured crowd scenes.

Experimental results show the people were successfully detected by the HOG feature and surrounded by a blue rectangle in the image. In the processed images, the green dots indicate the particles for tracking the target.

(A) Detecting and tracking a specified walking person in a structure crowd scene.

Each person has different features such as black, red, blue or white clothes, background, etc. In this part the target is wearing black pant and blue shirts. Other persons mostly occlude him.

Figure 3. Detecting and tracking a specified walking person in a structure crowd scene (a) Initial particle filter (b) Particle filter Tracking (c) Particle Filter Tracking specific person and HOG Detection.

(B) Detecting and tracking specified person running on structured crowd scene.

Human detecting and tracking specified person running on structured crowd scene. In this scene, people are running and we test our method when people are far from camera.

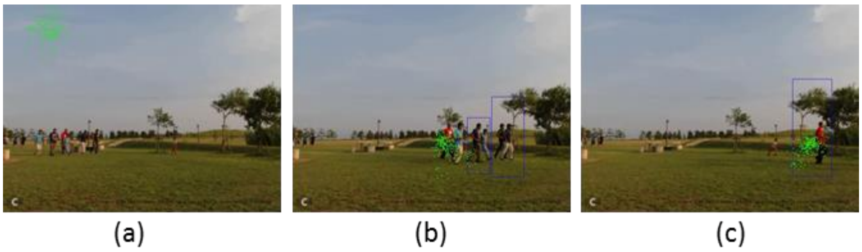
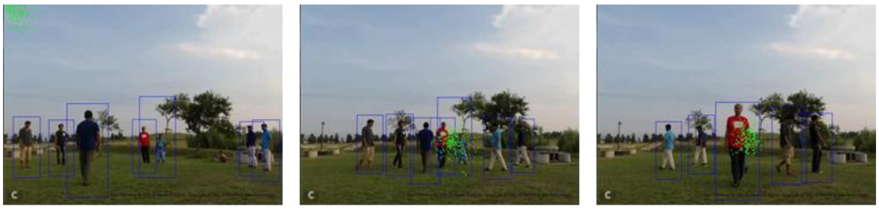


Figure 4. Detecting and Tracking Specified person running on structured crowd scene experimental (a) Initial particle filter (b) Particle filter Tracking (c) Particle Filter Tracking.

(C) Detecting and tracking specified person walking on unstructured crowd scene.

In these unstructured crowded scenes, we use seven people for detection and track a specified person. Each people have different feature but our main motivation is detect all person and tracking a specific person. In these scenes one person is moving very far from camera. This specific person wears black pants and a red shirt. To track him, first, we use Histogram of Oriented Gradient (HOG) to detect human area after we use particle filter for tracking specific person. In figure 5, we see the Histogram of Oriented Gradient (HOG) detected human area perfectly and particle filter tracking specific person accurately.



(a) (b) (c)

Figure 5. Detecting and tracking specified person walking on unstructured crowd scene experimental (a) Initial particle filter (b) Particle filter tracking (c) Particle Filter Tracking and HOG Detection.

(D) Detecting and tracking specified person running on unstructured crowd scene.

Tracking and detecting specific person using particle filter and Histogram of oriented Gradient (HOG) when people are running on the system of unstructured crowd scene. In these unstructured crowd scene we track a person wearing a black color shirt and white color pant.

(a) (b) (c)

Figure 6. Detecting and tracking specified person running on unstructured crowd scene experimental. (a) Initial particle filter (b) Particle filter tracking (c) Particle Filter Tracking and HOG Detection.

Table 3 shows the experimental results of success and failure rate for detection and tracking the target person in crowded areas.

Table 3. Evaluation of proposed method.

Dataset	HOG feature detection [%]		Particle Filter [%]	
	Success	Failure	Success	Failure
Scene 1	81.0	19.0	99.5	0.5
Scene 2	92.0	8.0	98.7	1.3
Scene 3	90.0	10.0	98.1	1.9
Scene 4	90.0	10.0	99.9	0.1

In the experimental results, our system achieved more than 90% of human detection and for particle filter tracking specific person around 99% of the specific persons in all dataset. In scene 1 and 2, although the HOG detection's failure rate was high caused by occlusion, the particle filter tracked the specific person perfectly even when the people are moving. In scene 3, the HOG detection and particle filter worked perfectly. In scene 4, the particle filter also tracked the specific person accurately. From these results, we have confirmed our proposed method that combined the HOG-SVM and particle filter with the target feature is robust toward the occlusion problem.

6. Conclusion

In this paper, we proposed the innovative method combining the HOG feature detection and particle filter for tracking the specific target. From the results of evaluation using the dataset including occlusion problem, we confirmed the effectiveness of our proposed method in such complex scenes.

In future work, we will improve this method, apply it towards to a more crowded scene, and implement the tracking of multiple targets simultaneously.

References

1. Saad Ali, Mubarak Shah, "Floor Fields for Tracking in High Density Crowd Scenes" The 10th European Conference on Computer Vision (ECCV), 2008.
2. Saad Ali, Mubarak Shah, ALagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2007.
3. Mikel Rodriguez; Saad Ali; Takeo Kanade." Tracking in unstructured crowded scenes", 2009 IEEE 12th International Conference on Computer Vision, Pages: 1389 – 1396, 2009.
4. Thida, Myo, et al. "A literature review on video analytics of crowded scenes." Intelligent Multimedia Surveillance.Springer Berlin Heidelberg, 17-36. 2013.

5. Li, Teng, et al. "Crowded scene analysis: A survey." IEEE transactions on circuits and systems for video technology 25.3: 367-386 , 2015.

6. Aggarwal, J.K., Cai, Q.: Human motion analysis: a review. Computer Vision and Image Understanding. 73(3), 428–440 , 1999.

7. Wang, L., Hu, W., Tan, T.: Recent developments in human motion analysis. Pattern Recognit. 36(3), 585–601 ,2003.

8. Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviors. IEEE Trans Rev. 34(3), 334–352 , 2004.

9. Zhan, B., et al Crowd analysis: a survey. Mach. Vis. Appl. 19(5–6), 345–357, 2008.

10. Qi, Z., Ting, R., Husheng, F., Jinlin, Z.: Particle filter object tracking based on Harris-SIFT feature matching. Proc. Eng. 29, 924–929, 2012.

11. Marcin Wojnarski, Absolute contrasts in face detection with AdaBoost cascade https://www.mimuw.edu.pl/~mwojnars/papers/07jrs-adaboost.pdf.

12. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. Conference on Computer Vision and Pattern Recognitiom(CVPR), 2005.

13. R. Xu, Y. Guan, and Y. Huang, "Multiple human detection and tracking based on head detection for real-time video surveillance," Multimedia Tools and Applications, Vol. 74- 3, pp.729–742, 2015.

14. Smita Kakade, Dr. N.J. Uke, Niraja Jain, Real Time Human Detection using HOG Feature as Human Descriptor and Variable size Sliding Window.

15. Haofu Liao, Human detection based on Histograms Oriented Gradients and SVM.EECE, SPRING, 2013.

15. Jamal Saboune and Robert Laganiere:People detection and tracking using the Explorative Particle Filtering.Ontario Centre of Excellenceunder the project CR-10030-08 and by the Ontario.

16. Ant´onio Almeida, Jorge Almeida and RuiAra´ujo:Real-Time Tracking of Moving Objects UsingParticle Filters.IEEE ISIE 2005, June 20-23, 2005, Dubrovnik, Croatia.

17. Saad Ali and Mubarak Shah, A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2007.

18. ARNAUD DOUCET, SIMON GODSILL and CHRISTOPHE ANDRIEU, On sequential Monte Carlo sampling methods for Bayesian filtering, Statistics and Computing 10, 197–208, 2000

Bhuyain Mobarok Hossain

He received the Master degree in 2012 from Faculty of Engineering, the University of Tokushima, Japan. He is currently Ph.D. student Faculty of Engineering at the University of Tokushima. His research interests Image processing.

Stephen Karungaru

He received a PhD in Information System Design from the Department of information science and Intelligent Systems, University of Tokushima in 2004. He is currently an associate professor in the same department. His research interests include pattern recognition, neural networks, evolutionary computation, image processing and robotics.

Kenji Terada

He received the doctor degree from Keio University in 1995. In 2009, he became a Professor in the Department of Information Science and Intelligent Systems, University of Tokushima department. His research interests are in computer vision and image processing. He is a member of the IEEE, IEEJ, and IEICE.

.

Akinori Tsuji

He received the Ph.D. degree in 2013 from Faculty of Engineering, the University of Tokushima, Japan. He is currently an engineer at Tokushima University. His research interests include image processing, signal processing and embedded systems. He is a member of IEEE and IEICE.