Unattended and Stolen Object Detection Based On Relocating of Existing Object

Waritchana Rakumthong, Natpaphat Phetcharaladakun, Wichuda Wealveerakup and Nawat Kamnoonwatana
Faculty of Information and Communication Technology, Mahidol University
999 Phuttamonthon 4 Road, Salaya, Nakhon Pathom, 73170

Abstract **– This paper proposes a new design and implementation method in supporting a smart surveillance system that can automatically detect abandoned and stolen objects in public places such as bus stations, train stations or airports. The developing system is implemented by using image processing techniques. In the circumstance such as when suspicious events (i.e. left unattended or stolen objects) have been detected, the system will alert to people responsible for the role such as security guards or security staff. The detection process consists of four major components: 1) Video Acquisition 2) Video Processing 3) Event Detection and 4) Result Presentation. The experiment will be conducted in order to access the following qualities: 1) Usability: to verify that the system can detect the object and recognize the events, and 2) Correctness: to measure the accuracy of the system. Finally, the data sets are tested in the experimental result can measure and represent the correctness of system by percentage. The correctness of object classification is approximately 76%, and the correctness of event classification 83%.**

Key Words: Object Detection, Surveillance System, Unattended Object Detection, Stolen Object Detection

I. INTRODUCTION

A surveillance system has become a part of every day's life without even noticed. It has been used in many places such as warehouses, airports, streets, shopping malls, and even in our houses.

Traditionally, video surveillance systems do not provide the ability to analyze the contents, which leaves heavy burdens in detecting suspicious events both in real-time and offline. From this reason, this traditional system is still very much manually controlled and relied a lot on people. For example, it may requires a security team to standby at the monitor in order to detect concerned events and periodically views (e.g. every day or every week) of the recorded video to search for suspicious events that have already occurred.

In order to have a smarter surveillance system, In [1] is based on the combination of the evidences generated by three independent detectors based on different extracted features. In [2] represented the smart surveillance system related to automatic detection of abandoned object and unknown object. It automatically recognizes activities around public area to improve safety and security of those areas. For system of [3], additionally, can provide the efficient analysis of video surveillance system for detect an abandoned or removed object [2] by using video processing technique, as background subtraction.

The developing system is expected to have an ability to analyze when concerned events occur and automatically alert to the corresponding personnel. For this paper, we define concerned events to cover two scenarios 1) Unattended object: when an object is left for a period of time, and 2) Stolen objects: an object that might be left and is picked by someone. The proposed system is based on background subtraction, object detection (bags and people), and object recognition and identification.

II. BACKGROUND

The growing of technologies requires the surveillance system to give higher quality of works. However, as mentioned, only a few surveillance systems in the market provide an ability to analyze. We have noticed that CCTV is used in both public places and private organizations. There are CCTVs installed. Therefore, we can apply image processing techniques to videos that are taken by those CCTVs. We have learned from some researches about how to implement unattended and stolen object detection.

In [1], the system starts processing with Foreground Segmentation process. This process is done by applying an approach similar to the one of the semantic analysis module in [4], a background update model based on a mixture between Average and Running Average methods [5] to get less noise, the binary foreground mask that is performed using mathematical morphology and blob extraction. After that, it tracks the object by comparing the detected object in the current frame and the previous ones. It considers any position that does not change during 50 consecutive frames as a background while detect human by using the module that is structured as described in [6]. Afterwards, the unattended and stolen object detection process will be implemented. An object will be shaped by Active Contours. Two of the three cameras will be used to check the similarity of the object by High-Low gradient detector. Another camera will used mainly for color similarity detection. The result shows that an accuracy of the system in this paper is quite high in low and medium complex scenes.

In another paper [3] used matching the boundaries of static foreground regions to classify the object is abandoned or removed. The author purposed to present a new method for analysis the sequence of video surveillance. First step is subtracting between current frame and background frame. Mostly researches used the technique of Gaussian mixture models [7] to subtract the foreground object from the background. The next step is a shadow removing algorithm by

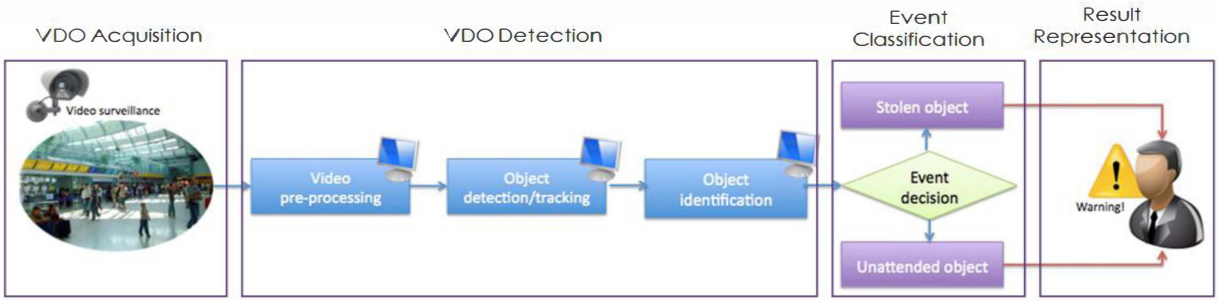

Figure 1. Overview of System Architecture

extracting the real shape of detected object. Finally is event classification as abandoned or removed object.

III. Proposed Algorithm

The system architecture is shown in Fig 1. It represents the process of the system. The system obtains a video from a video surveillance camera (e.g. CCTV) or a video file. Then, detect objects using image processing techniques. The output of the system is an event classification result that is acquired from a decision-making. The result of the system processing can be viewed via a user interface, which is a TV screen or a computer screen.

A. Video Acquisition

This processing unit is the process of importing the video from a video stream and capture into sequence frames.

1) *Video Stream* – This method receives a streaming video from a file or a CCTV camera. Currently, the following video file formats are supported, mp4, avi, bmp and others. By far, this paper works with only one video from one camera or one video file at a time.
2) *Sequence Frame* – After the program reads the video file, it takes and processes each image by querying frames from the video file.
3) *Capture Image Displaying* – Creating a window in which the captured images from camera will be shown on that window.

B. Video Detection

This processing unit uses video processing techniques to detect objects in the scene, such as Gaussian blur, and Background Subtraction.

1) *Gaussian blur* – it is an ideal form of white noise, which is caused by random fluctuations in the signal such as when we are watching the TV. Image Averaging, Average Filtering, and Adaptive Filtering are the techniques to clean Gaussian noise. There are many options to reduce noise from the video. In this paper, a video is blurred by applying a filter to each frame.
2) *Background subtraction* – Performing a subtraction between the current frame and the background model. In detecting moving object in videos, the approach that mostly used is foreground segmentation or it is known in term of background subtraction. It is used for extracting current frame or interested object from background image. This is the first step to identify moving objects in a video stream.
3) *Threshold* – it is used widely to generate a foreground mask of a video that is taken by a static camera. It is the simplest segmentation method.
4) *Contour process* – it is finding and drawing edges around the moving objects in a circle or rectangle shape

C. Event Classification

This processing unit is used to classify type of objects, and to identify if that object is unattended or stolen object.

1.) *Classification (Object/People)* – When object is tracked, it can be classified the type of object as a people or static object. Firstly use the CascadeClassifier class, by loading an .xml classifier file and apply it with a detectMultiScale class to perform the detection. This class has function is used for detect people body. After CascadeClassifier is used, the object is detected can classify as people, so we can know that the object that cannot detect by this function is a static object. (see in Fig.2 shows the example of object classification)
2.) *Identification (Unattended/Stolen)* – After the object is classified in current frame, its area in both background and first frame will be cropped out from the frame. Background frame is an image that is taken when the scene does not contains any people or static object. First frame is an image that is captured when the video starts. Then, applying background subtraction method and contours to those cropped frames again to find out whether the object is in the first frame or not. If there is no object in the first frame, it means that the event is likely to be unattended object because there is an object added to the current frame. On the other hand, the stolen object event after the processing has an object shows up at the first frame, and disappear or moved from the current frame. (see in Fig.3 shown the example and Fig.4 shown algorithm of object classification)

D. Result Presentation

This processing unit is to indicate the output result to users. Result is presented in two parts, screen and console. On the screen, it shows labels of each person and object. A rectangle will be drawn around an object after it is classified as an unattended or a stolen object. Time, object status and other details are printed real-time on console. These details could be saved as a log file.

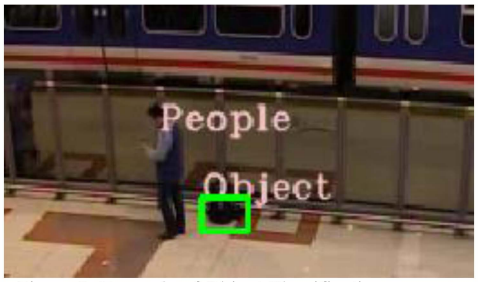

Figure 2. Example of Object Classification

(a) **(b)**

Figure 3. Example of (a) Unattended Object Event
(b) Stolen Object Event

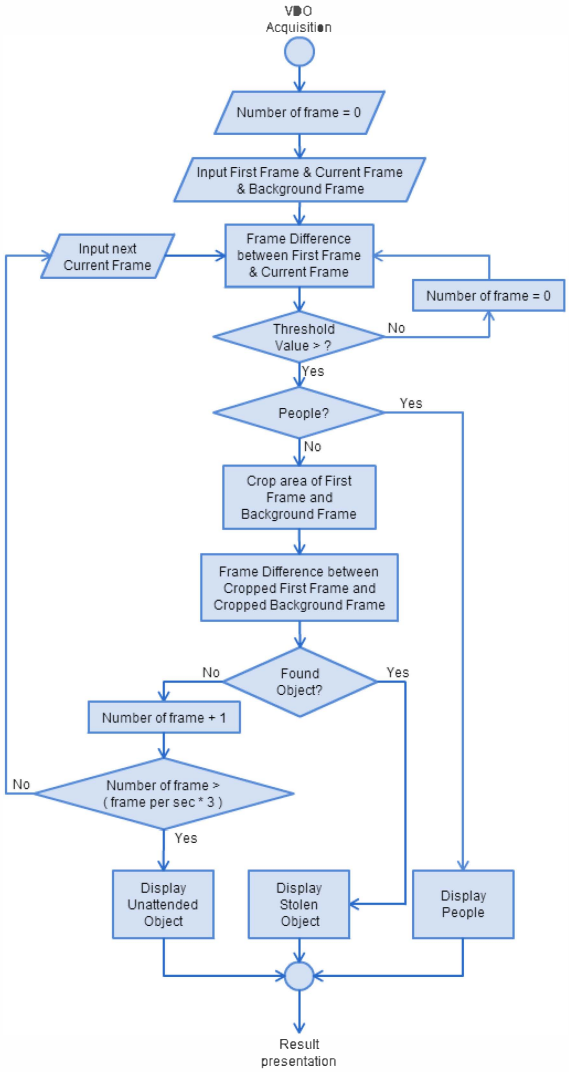

Figure 4. Algorithm of Event Classification

III. EXPERIMENTAL RESULT

In experimental result phase, algorithms are implemented by OpenCV libraries. The OpenCV library [8] is coded by C++ on Microsoft Visual Studio 2012. The system is tested with videos several datasets. The first dataset is the PETS 2006 dataset [9], the surveillance video in train station, is used for test as open environment. The second dataset is simulation dataset that is used for test as simple environment.

In the Unattended and Stolen Object Detection System, before going beyond to event classification , it can classify the type of moving object whether it is an static object or people as shown in the Fig. 2 by object classification.

The first experiment is shown in the Fig. 5, experiment: a bag is unattended. In the first two images in the first row, the background and current image are subtracted to get the result of moving object (foreground), and presented in the second row of different images. After it takes processes of Morphology and Contour techniques, then the output shows mask image as in the second figure in the second row. Finally, the result of object detection will be shown due to timer set greater than specific time on the system, and alerts an event status on the screen output as in figure of event notification.

As we can see Fig. 6, the second experiment of a bag is stolen. The background, current, different, mask is processed similar as an unattended object event, due to the event notification that alerts the event status immediately after the object is stolen or collected by anyone.

Along to the classification in the situations of unattended and stolen object in all experiments. Finally, as you can see in the TABLE I., the correctness of object/people classification is about 76 %. The divide into object classification 97% and people classification 52%. We note that the people classification is very low performance; this is caused by the CascadeClassifier class in object classification process, which cannot detect people in any frame that people change their posture to difference style. In the TABLE II, The correctness of event classification can be separated into three situations (None, Unattended, and Stolen). The nonevent is about 78%, unattended event is about 84%, and the stolen classification is about 87%. The efficiency and correctness of results is summarized by tested video dataset.

TABLE I. OBJECT CLASSIFICATION ACCURACY

Dataset	**Object Classification**		**Average**
	Object	People	
Simple	93%	82%	88%
Open	100%	29%	65%
Average	97%	52%	76%

TABLE II.EVENT CLASSIFICATION ACCURACY

Dataset	**Event Classification**			**Average**
	None	Unattended	Stolen	
Simple	86%	91%	90%	89%
Open	70%	77%	83%	77%
Average	78%	84%	87%	83%

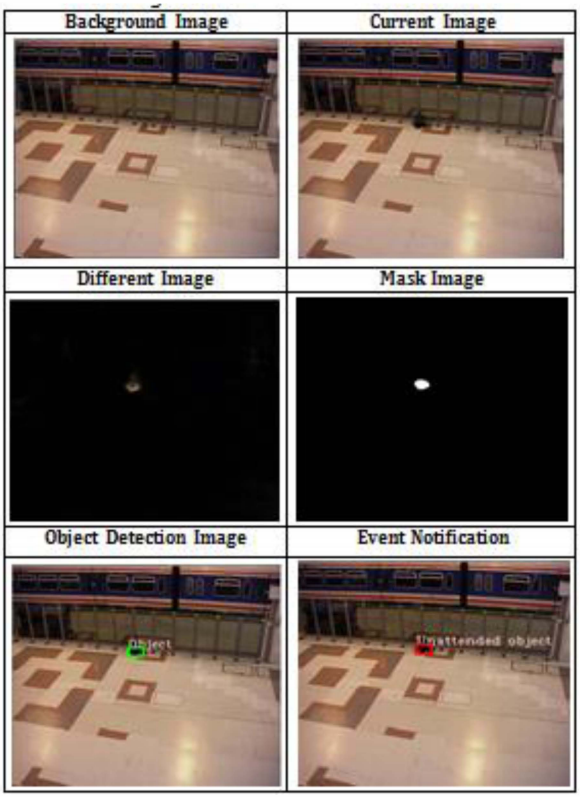

Figure 5. A bag is unattended

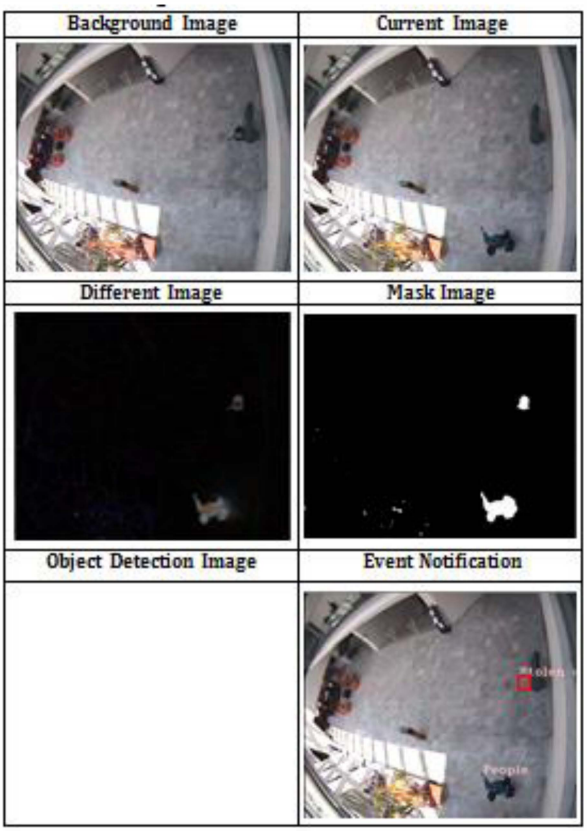

Figure 6. A bag is stolen

IV. CONCLUSION

In this paper, a method of the unattended and stolen object detection is presented. As started by a first step, video acquisition is used to import the video to process in the next step. The second step is the video detection process to transform normal image to gray scale image and distinguish the moving object by using subtraction technique with background and current image. Next step, we separate between people or object by using the Haarcascade function, normally used for people detection, and discriminate between unattended or stolen events by analyzing the boundaries of static foreground regions. Finally, the result of detection will alert on the output screen. Although all process can detect and classify types of objects (people/object) and events (unattended/stolen), the system still needs efficient and correctness for more reliable detection.

V. FUTURE WORK

We are suggesting some possible future works that could be applied to improve the system. Since a video quality could not be improved much, setting a camera in a proper angle helps getting more accuracy system.

Tracking feature is demanded in a robust surveillance system to provide an ability to recognize a person in a scene. There are cameras located along the way people walk, which coordinate together.

Developers could invent new algorithms and apply them to the given OpenCV libraries in order to get better performance. Moreover, more information about a current scene might be added to the user interface. For example, a number of people walking in and out of that area, a graph showing a number of people in each period of the day.

REFERENCES

[1] San Miguel, J.C.; Martinez, J.M. "Robust Unattended and Stolen Object Detection by Fusing Simple Algorithms", *Advanced Video and Signal Based Surveillance, 2008. AVSS '08. IEEE Fifth International Conference on,* On page(s): 18 – 25.

[2] B. Atul, M. S. Nagmode, P. Dhaval. "REAL TIME ABANDONED BAG DETECTION USING OPENCV", International Journal of Scientific & Engineering Research, Vol. 4 Issue 11,:2229-5518,2013.

[3] P. Spagnolo et al. "An Abandoned/Removed Objects Detection Algorithm and Its Evaluation on PETS Datasets". Proc. of IEEE Conf. on Video and Signal BasedSurveillance, pp. 17-21, 2006.

[4] A. Cavallaro, A. Steiger and O. Ebrahimi. "Semantic Video Analysis for Adaptive Content Delivery and Automatic Description", IEEE Trans. on Circuits and Systems for Video Technology, 15(10):1200-1209, 2005.

[5] R. Cucchiara, et al. "Detecting Moving Objects, Ghosts, and Shadows in Video Streams", IEEE Trans. on Pattern Analysis and Machine Intelligence, 25(10):1337-1342, 2003.

[6] V. Fernandez-Carbajales et al. "Robust People Detection by Fusion of Evidence from Multiple Methods", Proc. of Int. Workshop on Image Analysis for Multimedia Interactive Services, pp. 55–58, 2008.

[7] Piccardi M. "Background subtraction techniques: a review", Man *and Cybernetics, 2004. IEEE International Conference on Systems.*

[8] Itseez . (2014). OpenCV. *OpenCV Computer Vision*. Retrieved January 3, 2014, from opencv.org

[9] The PETS 2006 dataset. Retrieved September 15, 2013, htttp://www.cvg.rdg.ac.uk/PETS2006/data.html