논문 2009-4-10

비전 기술에 기반한 위험 유기물의 자동 검출 시스템

Automatic Detection System for Dangerous Abandoned Objects Based on Vision Technology

김 원*

Won Kim

요 약 공공장소에서의 유기물은 의도적 공공테러를 목적으로 폭발물이나 화학물질 등을 포함할 수 있기 때문에 일단 가능한 위험물로 반드시 다루어져야 한다. 공항이나 기차역과 같은 대형 공공장소에서는 전체 영역을 감시하는 모든 모니터를 점검할 보안 인력을 유지하는데 있어서 비용적 측면의 한계가 있게 마련이다. 이것이 비전 기술에 기반한 위험 유기물의 자동 검사 시스템을 개발하여야 하는 기본적 동기이다. 이 연구에서는 잘 알려진 DBE 기법을 적용하여 배경 이미지를 안정적으로 추출하는 것을 보이며, HOG 알고리즘을 적용하여 물체 분류에 있어서 사람과 물건을 구분하는 기능을 구현하였다. 제안된 시스템의 유효성을 보이기 위하여 감시 지역의 한 실내 환경에 대해 금지구역 침범을 탐지하고 유기물에 대한 경보를 발생하는 실험을 수행하였다.

Abstract Abandoned objects should be treated as possibly dangerous things for public areas until they turn out to be safe because explosive material or chemical substance is intentionally contained in them for public terrors. For large public areas such as airports or train stations, there are limits in man-power for security staffs to check all the monitors for covering the entire area under surveillance. This is the basic motivation of developing the automatic detection system for dangerous abandoned objects based on vision technology. In this research, well-known DBE is applied to stably extract background images and the HOG algorithm is adapted to discriminate between human and stuff for object classification. To show the effectiveness of the proposed system, experiments are carried out in detecting intrusion for a forbidden area and alarming for abandoned objects in a room under surveillance.

Key Words : Visual detection, Abandoned object, Object classification, Background Extraction

Ⅰ. Introduction

Terror is one of modern fears for the people living in civil societies. For example, the imagination shown in action movies was realized in the 911 terror in the United States and people in the world started to feel that metro-cities are not safe places to live in. South Korea is also not a safe nation for massive terrors as shown in the tragic deaths in the subway train of Taegu metro-city by an incendiary fire. Great south gate *Soong Rye Moon*, which is the no. 1 cultural property in Korea, has been burn down in 2008 by another incendiary fire. As shown in the terrors of Taegu subway disaster and *Soong Rye Moon* fire, indirect attempt is the main type of modern terrors. For massive damage, crowed areas are certainly the main targets of terrorists such as airports, train stations, hotels, hospitals and so on. In crowded areas,

*정회원, 우송대학교 컴퓨터정보학과
접수일자 2009.7.9, 수정완료 2009.8.7

abandoned objects are dangerous because they may contain harmful materials such as toxic chemical substances, explosives or incendiary materials. The concept of a dangerous abandoned object is shown in Fig. 1, where a man carrying a briefcase put this object alone in an airport lobby after a while. This abandoned object should be classified as a dangerous object until it is identified as a safe thing. However, there are thousands of cameras in the lobbies of large airports. It is impossible for security staffs to monitor for detecting possibly dangerous abandoned objects in limited man powers. Therefore, it is important to develop the automatic detection system for dangerous abandoned objects based on vision techniques in crowded areas to prevent massive damages and tragic deaths by the indirect manner of terror using dangerous abandoned objects. This research deals with the development of automatic detection system for these abandoned objects based on vision technology. The related works on this system are summarized in Chapter Ⅱ and the system configuration is described in Chapter Ⅲ. Experimental results are explained in Chapter Ⅳ and finally conclusive remarks are summarized in Chapter Ⅴ.

Fig. 1. Illustration of dangerous abandoned object

Ⅱ. Related Works

Bird and et al. showed a real-time method for online detection of abandoned objects in public areas[1]. In their research short-term logic is configured on blob operations to make the system robust to temporary occlusion. To cover crowded scenarios, Porikli et al. proposed the dual foreground method such as short-term foreground and long-term one[2]. Haritaoglu et al. showed W^4 system which is a real time visual surveillance system for detecting and tracking multiple people and monitoring their activities[3]. Shape analysis and tracking are combined to locate people and their parts in their research. Hu et al. in [4] surveyed latest researches on visual surveillance of object motion and behaviors. In industries, advanced companies such as Creative Vistas Inc, Digi Sensory and QTCO shows their visual tracking systems of good performance. However, it is hardly found researches on the classification of detected objects into people and stuff because it is necessary to discriminate motionless people from real abandoned object. To achieve this useful function, HOG method[5] is applied to develop the entire vision system for abandoned object detection. Moreover, dynamic background extraction technique[6] is used to obtain stable foreground objects in image stream.

Ⅲ. Automatic Visual Detection System for Dangerous Abandoned Objects

The entire system of this research is shown in Fig. 2. Image streams from camera systems are the inputs for this system and the processing outputs of the automatic detection system for abandoned objects are alarms for the monitoring system of security staffs. Components of the proposed vision engine in Fig. 2 are explained specifically in the following sub-chapters.

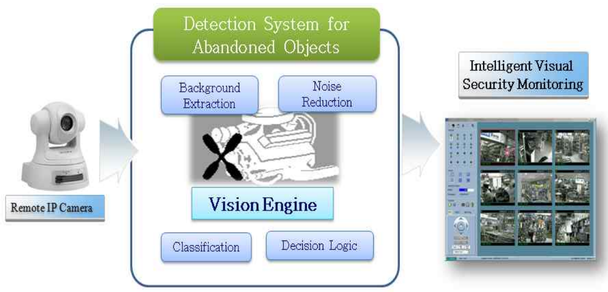

Fig. 2. Visual detection system

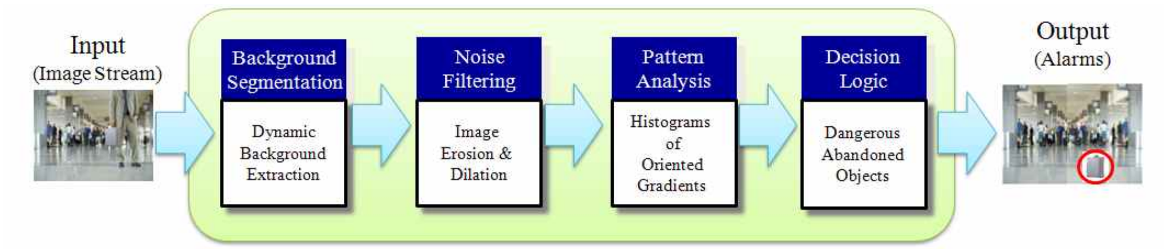

Fig. 3. Vision process for automatic detection of abandoned objects

1. System Configuration

Fig. 5 shows the process of the visual detection system in this research. Image stream from camera system as input is first processed to build up background images and the dynamic background extraction technique, which is explained in the very next sub-chapter, is applied to stably extract background images. If background image is obtained, foreground objects is easily extracted by subtracting incoming image from the background image. Since the difference areas may contain noise in their shape, erosion and dilation techniques are applied to eliminate noise in the shape and integrate segmented areas. As explained earlier, pattern analysis is needed to discriminate motionless people from real abandoned object and HOG method is applied to make the entire system intelligent. This will be explained hereafter in relevant sub-chapter.

2. Background Extraction Technique

Background extraction means separation of foreground objects and their background. While image difference method, moving average method are used in the earlier researches, the performance is not always guaranteed in real systems. To solve this problem, Kong et al. proposed the dynamic background extraction technique in their research[6], which originally targets to detect and track moving vehicles automatically on roads. In this technique, each pixel is statistically analyzed and its histograms are constructed in R, G, B channels. From the histograms, the dominant R, G, B values are selected to represent each pixel, and these selection method is robust to temporary illumination changes in background. Fig. 4 shows the application result of the dynamic background extraction method. Original image stream is shown in the left, where a person exists and the background extraction result in the right shows successful elimination of the person based on DBE technique. This background image output is used to extract foreground objects by subtraction operation.

Fig. 4. Dynamic background extraction result

3. Noise Filtering Technique

Foreground objects are extracted from the subtraction between current image frame and the background image obtained by DBE technique. However, one foreground object may split up into a few segments due to image noise or imperfect background extraction. To reduce the effects from noise, erosion and dilation techniques are applied while considering segment sizes and the length between neighboring segments. The erosion and dilation techniques are popular in dealing with image noise because they are effective and fast in calculation.

4. Object Classification Method[5]

Foreground objects may be human or stuff. Entire system should finally generate alarms for abandoned objects, not for persons who don't move for a long time because of unknown reasons. For this, pattern classification is needed to classify foreground objects into human or stuff category. However, detecting human in images is difficult due to the reasons of the diversities of clothes, postures, shapes, of occlusion and of illumination change. To solve this problem, Mohan et al. approaches human detection with Harr wavelets and SVM(Support Vector Machine) classifier[7]. Viola et al. use adaBoost classifier to detect pedestrians in streets and Mikolajczyk st al. also utilize adaBoost classifier to robustly detect human parts. In this research human detection function is implemented by the approach of Dalal and B. Triggs[5]. According to their approach automatic decision is possible whether human exists or not in specific area of image using SVM classifier, which uses the feature set encoding human shapes as input, and the features are HOG, *Histogram of Oriented Gradients*. HOG divides interesting area into *cell* units, and constructs 1-dimensional histogram of pixel's gradient orientations for each cell. Fig. 5 illustrates the process of HOG.

Fig. 5. Process of HOG method

First, input images are normalized in size and color information of R, G, B is converted into gray scale. Obtaining normalized image, the gradient of the image is calculated by using the derivative mask. Resulting gradient includes the magnitude and orientation information of a pixel. In HOG, a cell is defined by grouping neighboring pixels. Histograms on the gradient orientations in cells are constructed in the range of 0~360[degrees]. To make the system robust to illumination change, neighboring cells are grouped into their block and then blocks are normalized. After constructing histogram and normalizing gradient orientations, the histogram results are used as the input of SVM, which already learned fully the images without human and ones with human. Therefore, SVM can make outputs on whether images include human or not.

Ⅳ. Experimental Results

Experiments are carried out on the entire detection system, in which a camera, *AXIS 210A*, is used to capture image stream. Table 1 illustrates the camera characteristics. Detection software is installed in a computer with *Intel Quadcore* Q8200, 4 GB RAM, 1 Tera HDD systems. Fig. 7 shows the intrusion detection result in this experiment, which illustrates an intrusion situation. There are 4 quadrant images in the figure. 2-quadrant image shows original image stream, where a man is intruding into the room under surveillance. 1-quadrant image is background extraction result of the DBE algorithm. 3-quadrant image shows the foreground object obtained from the subtraction background image from current image. Final result is shown in the 4-quadrant, which includes object information understood in the object classification process.

Fig. 6. IP Camera

Table 1. Camera characteristics of the experiment

Item	Features
Image Sensor	1/4" sony progressive RGB CCD
Lens	4.0mm, f2.0, fixed iris, horizontal viewing angle 48°
Illumination	3.0 LUX, f2.0
Compression	Motion jpeg, mpeg4
Resolution	640×480 to 160×120, 16 step change
Frame Rate	30 frames over the all resolutions

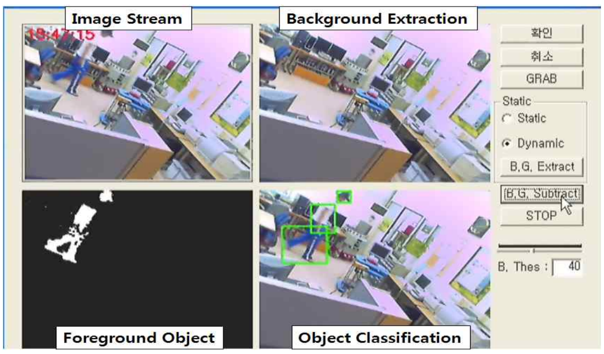

Fig. 7. Intrusion detection for a person entering the room

Fig. 8 shows the object classification ability of the detection system. In this experiment, a man with a bag enters the room and put the bag on the floor after a while. 4-quadrant image shows that the bag is classified as the first object while the man being classified as the second object.

Fig. 8. Object classification

Fig. 9 shows the experimental result on alarming performance of the system. Because the system identifies object types according to object classification algorithm, alarms are generated if objects, which are not persons, are abandoned for a predefined time, which is 30 seconds in this experiment. 4-quadrant image illustrates the experimental result on the alarm for an abandoned bag in the room in surveillance.

Fig. 9. Alarming for an abandoned object

Detection system performance is summarized in Table 2. First, PED(Percent Events Detected) shows 90%. PED means the performance index as false positive situations and is also calculated by the ratio of real detected alarms to total real alarms. For the case of FPS(Frames Per Second), it measures 15 for 640×480 image size, which means the processing speed of the entire system.

Table 2. System performance

Items	Unit	Performance
PED (Percent Events Detected)	%	90
FPS (Frames Per Second)	Frames	15
Image Size	Pixels	640×480

Ⅴ. Conclusions

This research deals with the development of automatic detection system for dangerous abandoned objects based on vision technology. Abandoned objects in public areas should be first classified into possibly dangerous objects until they are checked even though they may contain safe things. To check all the monitors linked to camera by security staffs costs a lot in man powers for large areas such as airports, train stations and so on. Because of the reason, automatic detection is necessary to detect abandoned objects based on vision technology. To enhance background extraction performance, dynamic background extraction algorithm is used in the fore part of the system as a statistical approach. After constructing background image, foreground image is obtained from subtraction operation, and erosion and dilation techniques are applied to the foreground results to reduce image noise. To discriminate between real abandoned objects and unmoved people, HOG algorithm is applied to the foreground objects. To show effectiveness of the system, experiments are carried out on real image stream captured by *AXIS 210A*. Experiment results show the basic functions are good such as stable background extraction, intrusion detection, object discrimination and alarming for abandoned objects. As an important performance index, PED is estimated at 90%, which is the ratio of real detected alarms to total real alarms. It means only 10 % are missed for total real alarms and this performance is good in considering current vision technology. The detection system can be applied to conduct visual surveillance for public areas for detecting abandoned objects automatically. The system will contribute to prevent indirect terror and reduce man-power in surveillance of large public areas.

References

[1] N. Bird, S. Atev, N. Caramelli, R. Martin, O. Masoud, and N. Papanikolopulos, "Real Time, Online Detection of Abandoned Objects in Public Areas," Proceedings of the IEEE International Conference on Robotics and Automation, Orlando, Florida, May 2006.

[2] F. Porikli, Y. Lvanov and T. Haga, "Robust Abandoned Object Detection Using Dual Fourgrounds," EURASIP Journal on Advanced in Signal Processing, 2008.

[3] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4 : Real-Time Surveillance of People and Their Activities," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, Aug. 2000.

[4] W. Hu, T. Tan, L. Wang, and S. Maybank, "A Survey on Visual Surveillance of Object Motion and Behaviors," IEEE Trans. on Systems, Man, and Cybernetics-Part C, vol. 34, no. 3, Aug. 2004.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *CVPR 2005*, vol. 1, pp. 886-893, 2005.

[6] J. Kong, Y. Zheng, Y. Lu, and B. Zhang, "A Novel Background Extraction and Updating Algorithm for Vehicle Detection and Tracking," *Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, 2007.

[7] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," *IEEE PAMI*, pp. 349 - .361, April, 2001.

[8] P. Viola, M. J. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance" *The 9th ICCV*, vol. 1, pp. 734-741, 2003.

[9] K. Mikolajczyk, C. Schmid, and A. Zisserman, "Human detection based on a probabilistic assembly of robust part detectors," *The 8th ECCV*, vol. I, pp. 69-81, 2004.

저자 소개

Won Kim(정회원)

• 1990 : B.S. degrees in Department of Korea Aerospace University.
• 1999 : M.S. degrees in Department of Electrical Engineering, KAIST.
• 2007 : Ph.D. degrees in Department of Computer Science and Electronic Engineering, KAIST.
• 1990~1997 : Researcher, Agency for Defense Development
• 2000~2007: Professor, Department of Electronic Engineering, Woosong Information College
2007~Now : Professor, Department of Computer Science, Woosong University
<Pricipal Scholarly Interests : Internet, Security, Image, Vision, Classification>