Development of Video Surveillance Device

Kenji WATANABE*, Michikazu UMEMURA and Masakatsu HIGASHIKUBO

Information communication technology has been widely used for crime prevention and protection of the elderly. Currently, there is an increasing demand for video surveillance systems that can detect and report a suspicious individual or abnormal behavior. Therefore, we have developed a video surveillance device that detects and tracks persons from video sequences by using filtering logics and reports the results to the user through a network. This paper outlines the video surveillance device and its detection performance.

Keywords: video surveillance, tracking, machine learning

1. Introduction

For security improvement, i.e. crime prevention and protection of the elderly, there is an increasing demand for video surveillance systems that can detect suspicious individuals and abnormal behavior and report the detection results to the user using information communication technology (ICT). In response to this demand, we have developed a video surveillance device that can detect persons in a video sequence taken by a wide-angle camera and control another camera (such as a zoom camera) to find, track and zoom in on the target based on the information provided. This device can also send a warning to predetermined terminals.

This paper summarizes the newly developed surveillance device.

2. Outline of the Video Surveillance Device

In such a video surveillance system as illustrated in **Fig. 1**, the developed video surveillance device controls a wide-angle camera, pan-tilt-zoom (PTZ) camera, and infrared LED lighting device to detect a suspicious individual or abnormal behavior. The device uses video sequences taken by the wide-angle camera and sends warning information to predetermined terminals[(1)] through a wired/wireless communication network.

The developed video surveillance device processes wide-area images taken by the wide-angle camera to detect a target object. The PTZ camera is controlled based on the target's location to take clear images of the target while tracking its movement. When the target stops, the PTZ camera zooms up on it. During these detection/tracking procedures, the image data acquired by the wide-angle camera is used to track the movement of the person who is entering or has already entered a certain surveillance area. When a suspicious individual or the like is detected, the surveillance device sends detection results to the user's smartphone or other terminals on a real-time basis. Since the surveillance device can produce detection results in several forms including text, still image, and video, the users can select the communication mode according to their needs or network bands.

As mentioned above, this video surveillance device is expected to be used for various purposes including the prevention and investigation of crimes as well as the protection of small children and the elderly in a multiple dwelling area.

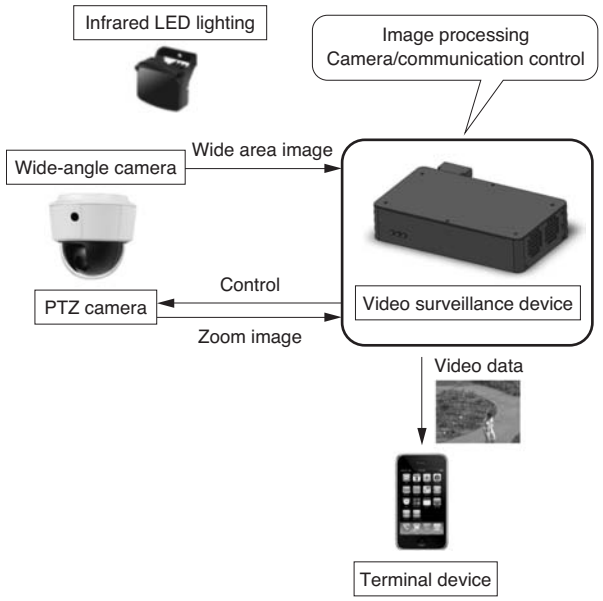

Fig. 1. Outline of video surveillance device's functions in video surveillance system

3. Detection Process

This chapter outlines the person-detection algorithm implemented in the developed video surveillance device. In particular, this algorithm is used to track the movement of a target object and take close-ups of it using a PTZ camera, and then to judge whether the object is a suspicious individual or not.

According to the person-detection algorithm implemented in the video surveillance device, a region containing a moving object (moving object region) is initially estimated by the video sequence taken by the wide-angle camera, as shown in **Fig. 2**. Then the estimated moving object region is searched for a person. If a person is detected, this region is reported.

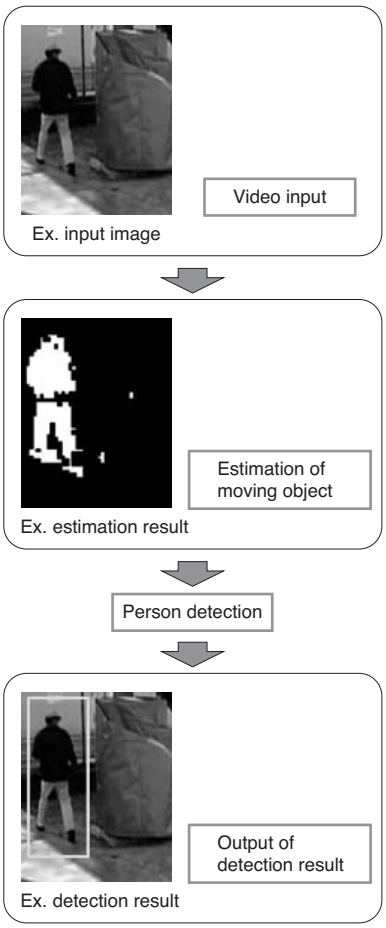

Fig. 2. Person-detection procedure
The rectangular region, given as an example of detection result, shows the moving object region (person-detection result) that is identified as a person.

3-1 Estimation of moving objects

To estimate moving objects in video sequences, the background subtraction methods and frame differential methods are generally used[2]. Here, we discuss the simple frame differential method that uses pixel values. When the absolute value of the difference between the pixel value of t-th frame $I^{(t)}$ and the pixel value of $(t-1)$-th frame $I^{(t-1)}$ exceeds a given threshold, it is regarded that the pixel value is changing remarkably in the input image. Assuming that the video sequence is being taken by a stationary camera, the pixel M whose pixel value has changed remarkably is suspected of containing a moving object. The above difference in M-pixel value is expressed as follows:

$$M = \begin{cases} 1 & if\ |I^{(t)} - I^{(t-1)}| > \text{Threshold} \\ 0 & else \end{cases} \quad \cdots\cdots\cdots\cdots (1)$$

The developed video surveillance device uses the frame differential method, as expressed by (1), to estimate a moving object region.

3-2 Classification

In an ideal measurement environment, the moving object estimated by the methods as shown in section 3-1 can be judged to be a person. However, since our surveillance system can be also used outdoors, the estimated moving object region may contain non-human moving objects as well (e.g. rustling leaves). To minimize false detections, the developed surveillance device is provided with a person-detection function that uses a supervised learning technique.

More specifically, the surveillance device initially extracts the histogram of orientation gradients (HOG) features[3] $\boldsymbol{x}_i$, which is one of the differential features having the robustness of luminance change, from the i-th estimated moving object region in the frame. In the surveillance device, the i-th extracted HOG features are mapped into a feature space constructed by the support vector machine (SVM)[4], and the i-th moving object region is classified as a person or not.

SVM is one of the popularly machine learning techniques used in the field of computer vision. When SVM is considered as a single-layer neural network, it can be defined as a technique for constructing a two-class classifier by using the simplest linear threshold element as a neuronal model.

In the leaning of SVM, the minimization problem of the objective function (2) is solved for a learning sample set under the condition of "margin maximization" to estimate the parameters of a linear threshold element (i.e. a learning model $\boldsymbol{w}$).

$$L(\boldsymbol{w}, \xi) = \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{N}\xi_i \quad \cdots\cdots\cdots\cdots (2)$$

where, ξ_i represents the slack variable relating to the i-th sample and C represents an given balancing parameter. SVM estimates the learning model $\boldsymbol{w}$ by solving the above optimization problem as a dual problem.

In the video surveillance device that classifies a person by using SVM, the classification function for input feature vector $\boldsymbol{x}_i$ is given by:

$$y_i = \text{sign}\,(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i - h) \quad \cdots\cdots\cdots\cdots (3)$$

where, h represents a balance term. From (3), the binary output value relating to the i-th input vector y_i is calculated. This model will output either 1 (the moving object is classified as a person in the case of the video surveillance device) or −1 depending on whether or not the inner product of input vector and learning model exceeds a preset threshold.

The detection technique implemented in this surveillance device uses two or more filtering logics in addition to the above-mentioned SVM-based process in order to ensure reliable suspicious individual detection with minimal false detection, as shown in **Fig. 3**. Although a blue sheet was mistaken for a person when no classifier was used (**Fig. 3 (a)**), the use of the classifier reduced the possibility of such erroneous classification (**Fig. 3 (b)**). This surveillance device can detect a person approximately 30 m away from the camera (**Fig. 3 (c)**). If the tree in the upper left of the picture is designated as a protector, the surveillance device will start to judge a suspicious individual from persons detected around the protector (**Figs. 3 (d), 3 (e)**). On the actual output screen, a person judged to be ordinary is surrounded by a green rectangle while a suspicious individual is surrounded by a red rectangle.

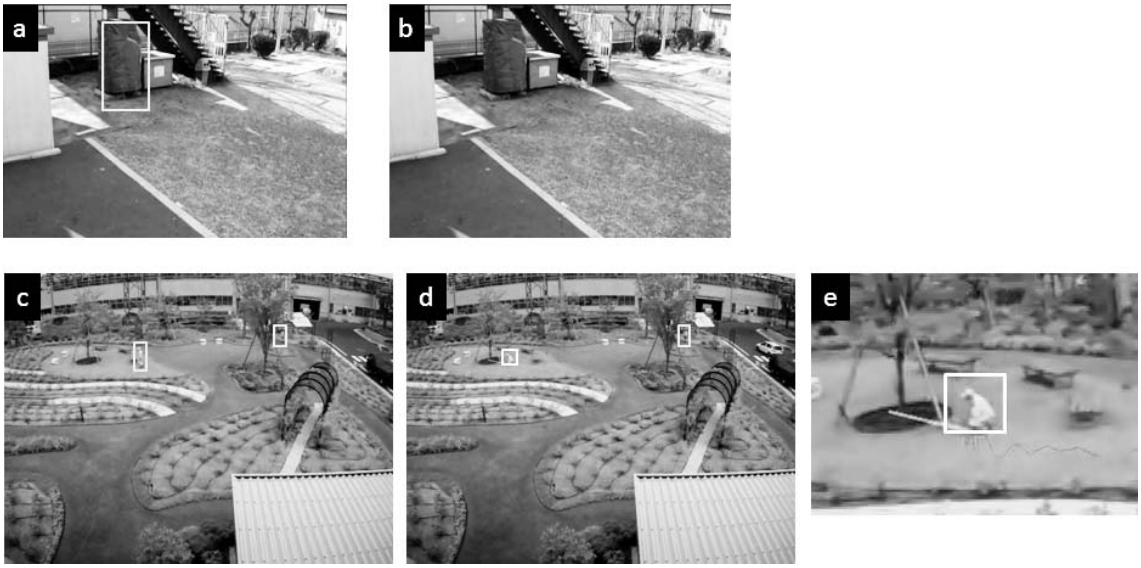

Fig. 3. Object detection result and suspicious individual detection with/without classifier
(a) object detection result without classifier; (b) object detection result with classifier; (c) distant object detection with classifier; (d) suspicious individual judgment result; (e) enlarged view of suspicious individual

4. Event Detection and Warning

This chapter describes the scheme for transmitting information (warnings).

4-1 Event detection

The developed video surveillance device performs event detection based on the person-detection processing results and preset rules. The total number of assignable events is 11 for three designated surveillance areas (rules), as shown in **Table 1**. Based on specific use, the event detection rule can be selected from "polygon," "polygonal line," and "fence." "Polygon" is used for area surveillance, while "polygonal line" and "fence" are used for spot surveillance.

The surveillance device combines a rule with an event to set detection conditions according to the video surveillance system installation environment and the object to be watched. Two or more rules can be set at the same time. Moreover, the video surveillance device monitors/controls four kinds of device operating conditions, as shown in **Table 2**.

Figure 4 shows an example of setting video surveillance device rules for a parking area. In addition to watching a suspicious individual entering the parking area, the surveillance device can detect the movement loci of persons in the surveillance zone and observe them also.

Table 1. Event detection rules

Use	Rule	Event	Description
Area surveillance	Polygon	Presence	A person is present inside the polygon.
		Entry	A person is entering the polygon.
		Exit	A person is leaving the polygon.
		Leaving/ carrying away	An object is being left behind/carried away from the polygon.
		Approaching	A person is approaching a surveillance spot designated inside the polygon.
		Crouching down	A person is crouching down inside the polygon.
Spot surveillance	Polygonal line	Presence	A person is present on a diagonal line.
		Movement	A person is moving along a polygonal line in the preset direction.
	Fence	Presence	A person is present on a fence.
		Climbing up	A person is climbing up a fence.
		Moving down	A person is hopping down a fence.

Table 2. Device checks

Use	Event	Description
Device check	Digital input	Detection of digital signal sent to device.
	Abnormality in camera image	Abnormality discovered in input camera image. * Image interruption, fogging, camera failure
	Abnormality in device	Abnormality discovered in video surveillance device. * Overheating, abnormal voltage
	Abnormality in network	Abnormality discovered in network connections.

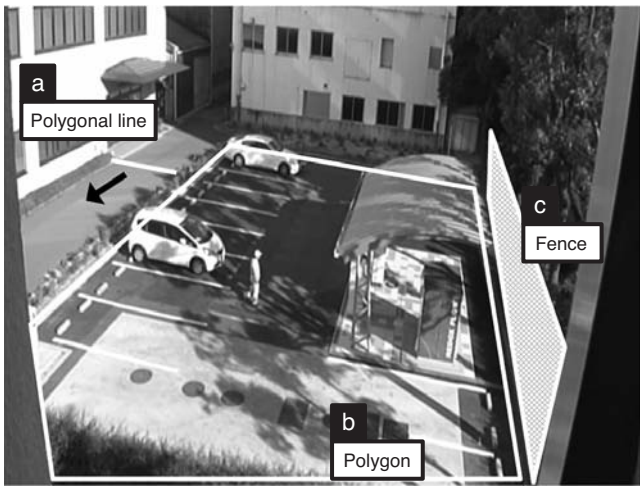

Fig. 4. Example of rule settings for a parking area
(a) Use of polygonal line to watch a person moving on a road
(b) Use of a polygon to watch a suspicious individual or person crouching down in a parking area
(c) Watching a suspicious individual climbing over a fence

4-2 Warning operation

When the video surveillance device detects an event, it conducts warning operations under the warning conditions described in section 4-1. Practical warning operations and their descriptions are given in **Table 3**. A video file, one type of data that can be attached to an e-mail, is encoded in H.264 so that the files can be watched on general smartphones. When linked to a PTZ camera, the video surveillance device can send zoomed-in images of the object, making it possible to check more detailed information remotely.

Table 3. Warning operations

Warning operation	Description
Digital signal output	Outputting digital signals Possible to interconnect with other devices
Event recording	Recording specified file size for specified time periods before/after event detection
Sending e-mail	Transmitting details of detected event through e-mail Possible to attach video and still images to literal information
PTZ camera control	Controlling PTZ camera based on information on detecting position

Table 4. Example of surveillance condition settings for Fig. 4

Setting rule - event	Warning operation
Polygonal line a – directional movement (left-oblique downward)	Digital output ch0
Polygon b – entering	Digital output ch1
Polygon b – presence	Tracking by PTZ camera
Polygon b – crouching down	Sending e-mail (with video attached)
Fence - presence	Video recording of event

Table 4 shows an example of the surveillance conditions set for **Fig. 4**. Video surveillance device users turn on the digital output modes when they set the object detection rules. Thus they can link the surveillance device to a video recorder or can operate a rotary beacon light to use the device as an automatic/manual surveillance support device.

5. Conclusion

This paper described our video surveillance device useful for detecting and reporting suspicious individuals and abnormal behavior. This device incorporates a machine learning algorithm to reduce false detections. The device is also provided with a function for sending information on suspicious individuals and abnormal behavior to terminals on a real-time basis.

As future work, we are planning to add an abnormal behavior classification function to the surveillance device. We will continue to work on the development of further sophisticated surveillance devices that has increased reliability in detecting suspicious individuals and abnormal behaviors.

References

(1) Y. Ono, M. Higashikubo, M. Moriguchi, Y. Ogiuchi, and M. Umemura, "Video surveillance system combining wide-angle and PTZ cameras," Proc. SSII2012, DS2-10, 2012 (Japanese)
(2) A. Shimada, D. Arita, and R. Taniguchi, "Increment and Decrement of Gaussians in Adaptive Mixture-of-Gaussian Background Models," Proc. MIRU2006, pp. 746-751, 2006 (Japanese)
(3) N. Dalal, and B. Triggs, "Histograms of Oriented Gradients for Human Detectio," Proc. CVPR, San Diego, California, USA, pp. 886–893, 2005
(4) V. N. Vapnik, "The Nature of Statistical Learning Theory," Springer, pp. 138-146, 1995

Contributors (The lead author is indicated by an asterisk (*).)

K. WATANABE*
• Ph.D.
Information & Communications Laboratories

M. UMEMURA
• Information & Communications Laboratories

M. HIGASHIKUBO
• Group Manager, Information & Communications Laboratories