Featured Articles

Image Sensing Techniques for Analysis and Interpretation of Surveillance Video

Satoshi Nakata
Tomokazu Murakami, Ph.D.
Wataru Ebihara
Ichiro Ote
Hiroshi Koya

OVERVIEW: Video surveillance systems, which have been heading for digital formats and higher image quality along with a switch from analog cameras to network ones for the purposes of crime prevention, disaster response, anti-terrorist measures, and tighter internal control are expanding in terms of their introduction and use at enterprises and public facilities[1), 2)]. They are also spreading into a wide range of other business areas through the utilization of video content analysis and various databases. Moreover, activity involving the solving of management issues, such as operational efficiency or quality of product improvement, has been initiated recently in the industrial field with the progress of IoT technology, accordingly. In addition to security purposes, Hitachi is also proposing plant surveillance video for things such as product quality improvement and plant safety by making use of image sensing for analyzing and recognizing humans, vehicles, and other objects from surveillance video.

INTRODUCTION

THERE have been two changes in the video surveillance systems field in the context of a transition from analog cameras to network ones in recent years: higher resolutions and larger-scale camera networks that connect a central data center with hundreds of cameras. Under such circumstances, video content analysis technology for identifying humans, vehicles, and other objects in video images acquired by surveillance cameras has become practical to use. In addition, the demand for automatic monitoring systems with multiple cameras to supersede human monitoring is increasing.

Meanwhile, in the field of industry, the German Industrie 4.0 project, representing the fourth industrial revolution, and the Internet of things (IoT) are becoming more popular and more familiar. Consequently, there is a growing need for video content analysis functions such as surveillance cameras with sensors that are able to recognize and collect needed pieces of information when humans are present on video, for example. Such information can be utilized to increase productivity or improve quality, efficient allocation of manufacturing equipment, or optimal posting of workers in a factory, by analyzing them together with big data coming from the production management systems.

This article describes the technological direction of image sensing that can recognize and analyze humans, vehicles, and other objects on surveillance video and Hitachi's activity in this field, as well as the problem posed by analysis processing load for the implementation of surveillance video systems and its solution.

FROM IMAGE PROCESSING TO VIDEO CONTENT ANALYSIS

Hitachi started researching image processing in 1968, and has been turning its research results into commercial and industrial applications on dedicated hardware for positioning objects and checking the shape of components mounted on circuit boards, or wafer inspection by large-scale integration (LSI) production equipment. These technologies have developed into video content analysis techniques that are used for realtime detection of road conditions through in-vehicle cameras through the evolution of information technology (IT) and embedded technology.

Meanwhile, in the video surveillance sector, the speed of networks and central processing units (CPUs) has been increasing in recent years since analog cameras have begun to be replaced with network ones. As a result, the way image processing algorithms are

implemented has changed from dedicated hardware to software, high-speed searching of scenes from large amounts of video recorded on hard disks, and realtime video content analysis of live video have become possible on personal computers (PC) and servers.

IMAGE SENSING TECHNIQUES

Trends in Technology for Video Content Analysis

The technology for measuring and identifying humans, vehicles, and objects from cameras or sensors is defined as "image sensing" technology in this article. Video content analysis techniques using cameras will be mainly explained in the following two sections.

Video content analysis has been subjected to extensive research and using video surveillance systems for security purposes is the most widespread use in this research field. The typical applications of the system are intrusion detection for finding humans that enter into a specific area, facial recognition, measuring traffic volumes of vehicles, recognizing vehicle license plates, and so on. The flow of technical development in these applications can be broadly divided into the three steps described below (see Fig. 1).

The first step is "presence sensing." At the beginning, it is necessary to recognize the presence of human bodies or faces, and vehicles, determine their locations, or count the number of them in video images in order to perform analysis. Applications such as detection for finding humans intruders and measurement of traffic volumes can be handled in this step.

The second step is "status and attribution recognition." Information gathered by status, type, or attribution is classified or checked against pre-registered information. Applications such as facial recognition and recognizing vehicle license plates can be handled in this step.

The last step, which, unlike the first and second steps, has not established sufficient technologies, and is expected to be a subject of further study, is "behavior understanding and prediction." This step features the recognition of interactions or contexts, including temporal and circumstantial changes between detected objects and the prediction of their future statuses by means of analysis using a large amount of data.

Details of Hitachi's techniques in the three steps will be introduced in the following section.

Hitachi's Video Content Analysis Techniques

(1) Presence sensing

Hitachi has developed a technique that enables the identification of congested situations by analyzing existing surveillance camera video (see Fig. 2). Additional sensors or other equipment are not needed for the technique because it only requires rough information.

If the technique was introduced into surveillance cameras at a manufacturing factory, it could be utilized for being notified of the high possibility that an accident has occurred, for example, in places that are not usually crowded with people.

Compared to the previous purpose of surveillance cameras, to record what happened in the past, this

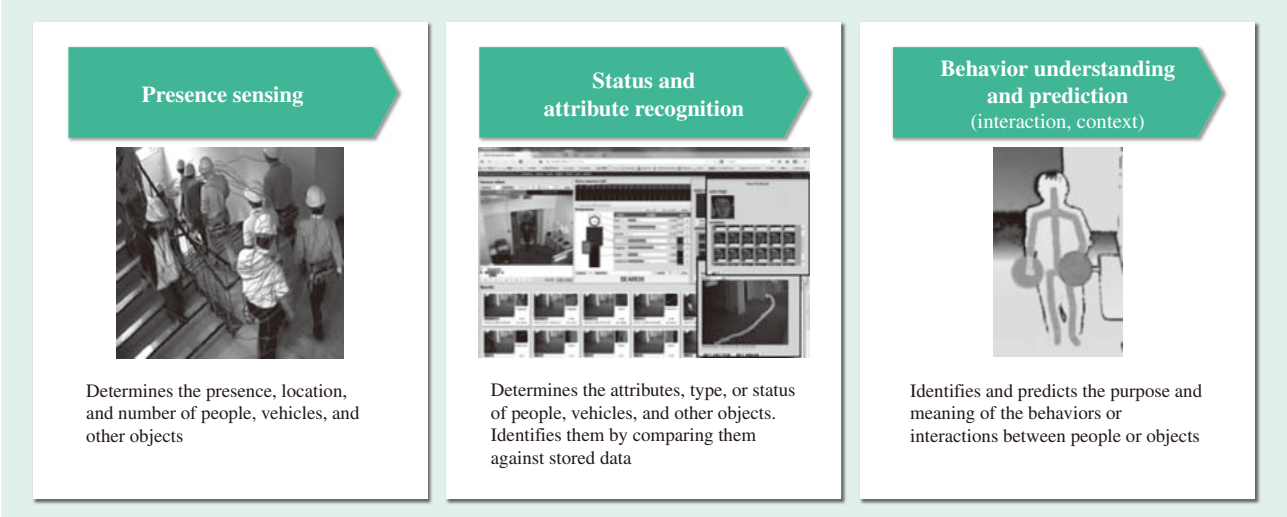

Fig. 1—Classification of Video Content Analysis Techniques.
The sequence of video content analysis can be explained in three steps: "presence sensing," "status and attribution recognition," and "behavior understanding and prediction."

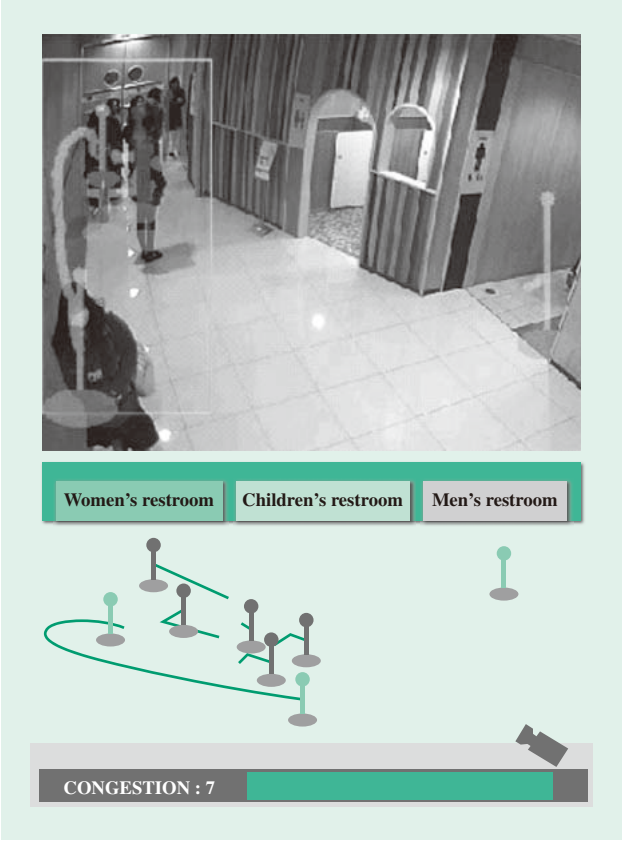

Fig. 2—Example Analysis of Congested Situations in Restrooms.
The system can identify the congested situation in the white rectangle by detecting the presence of humans on existing surveillance camera video.

Fig. 3—Multi-perspective Search Technique.
The technique can search for a specific individual on video from among a large number of surveillance cameras based on the person's clothes and objects the person is carrying.

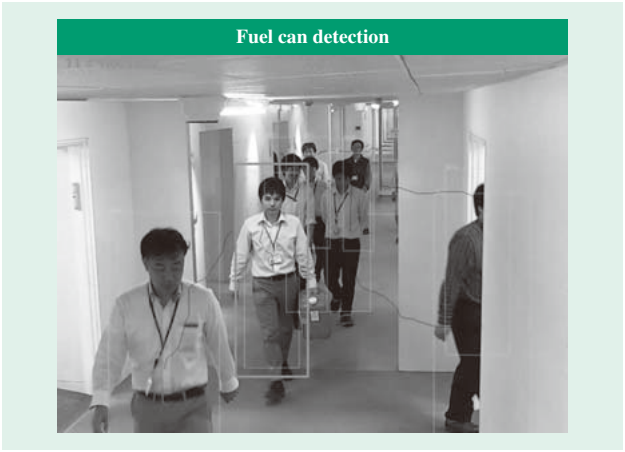

Fig. 4—Realtime processing.
The technique can detect people who are walking erratically or carrying prohibited hazardous objects.

technique adds new value by providing timely and prompt notification of accidents or other incidents.

(2) Status and attribution recognition

Hitachi has developed a multi-perspective search technique that enables searching for a specific person from among a large number of integrated multi-surveillance video images based on certain clues, including facial appearance, clothes, personal belongings, and route of movement (see Fig. 3).

Hitachi conducted a demonstration experiment and tested its effectiveness at tracking the route of movement of a suspicious person in a building using a group of ten or more surveillance cameras.

(3) Behavior understanding and prediction

Hitachi has a sophisticated and enhanced status recognition technique for recognizing a person's behavior in surveillance camera video images, so that realtime analysis of human behavior and features of the objects they are carrying can be achieved in addition to the existing techniques to obtain human location information through existing surveillance cameras (see Fig. 4).

This technique brings new applications for use in detecting people walking erratically or carrying prohibited hazardous objects such as fuel cans.

Moreover, Hitachi has been further studying human behavior analysis by means of a multi-model processing technique that combines surveillance camera video images and rangefinders in order to accurately analyze physical posture and physical positional relationships, because this cannot be done by using video images only (see Fig. 5).

Hitachi intends to utilize this technique to monitor work or to assist inexperienced workers at factories or site shops in the future.

Outlook for Image Sensing Technique

It is anticipated that image sensing techniques that determine information including the status of humans, vehicles, or objects using cameras, will be expand more into the workplace and continue to develop to a higher level in the future.

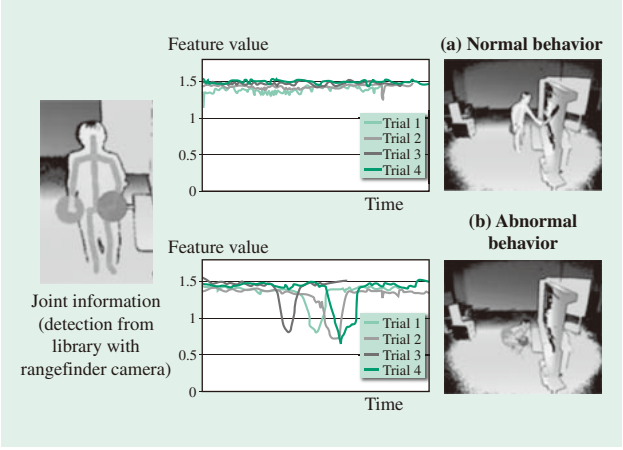

Fig. 5—Detection of Human Behavior Using Rangefinder.
The technique can detect specific human actions such as crouching down based on a measurement of the angles of the person's joints using rangefinder information.

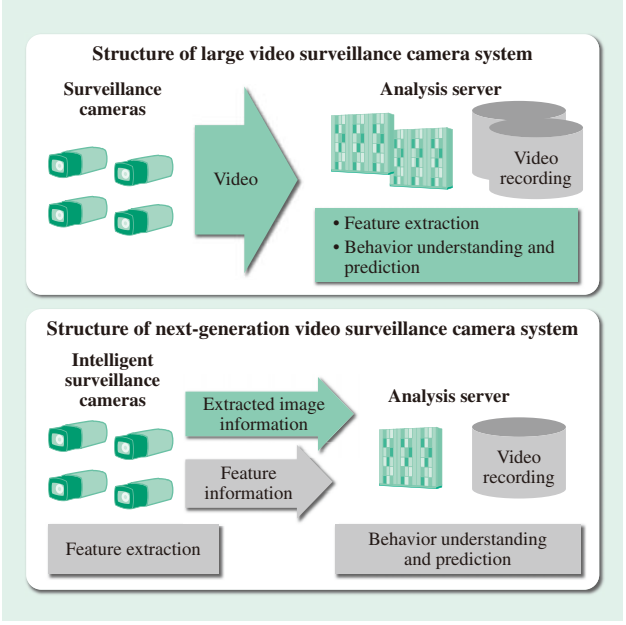

Fig. 6—Structure of Next-generation Video Surveillance Camera System.
The system can reduce the total amount of data with a feature extraction processing function added on to surveillance cameras.

In recent years, cameras, sensors, servers, and other equipment have been spreading throughout the market because of their falling prices. In terms of technological advances, great progress has been made on machine learning techniques, particularly with deep learning, which contributes to improving the accuracy of image recognition. In addition, data analysis and visualization have become easier to performing because of the progress made in big data processing technology and the availability of data processing platforms such as Pentaho*.

Hitachi has been pursuing research and development in this field in order to provide integrated solutions that enable the identification of anomalies or changes in large facilities by combining data collected from many cameras and sensors with cloud-based data processing platforms. These solutions also improve the accuracy of detection through video content analysis techniques using deep learning.

FUTURE ISSUES AND HOW TO OVERCOME THEM

Hitachi has been considering the practical realization of video content analysis not only for purposes of security, but also for all kinds of purposes.

As the scale of facilities becomes larger and the variation of video content analysis becomes greater, so too will the required number of surveillance cameras and volume of video data also increase. Through this, some issues are expected to be come up, including cost increases due to the larger storage capacity of servers that receive video data, cost increases due to increasing analysis processing loads, and increases in network loads due to large volume, high-resolution video (see Fig. 6).

In order to overcome these issues, surveillance cameras should be equipped with "presence sensing" and "status and attribution recognition" functions and steps for extracting feature information from video images alone, because distributed processing and data compression are effective ways of doing that. As a result of this, it will be possible to carry out the final step, "behavior understanding and prediction," with low cost and high speed, based on transferring the minimum amount of feature extraction video data and feature data to the analysis server.

For the practical realization of this, the interoperation of a large number of surveillance cameras for large facilities, realtime processing that enables prompt response, and combination processing techniques using multiple devices for detailed behavior assessment will be required.

With regard to next-generation video surveillance camera systems, a promising approach is to build a system with business intelligence (BI) tools such as Pentaho based on image sensing techniques that combine intelligent surveillance cameras/sensors with analysis servers.

* An open source business intelligence (BI) tool for professional use.

CONCLUSIONS

This article has described the utilization of image sensing techniques in surveillance cameras and sensors, and image sensing techniques for analyzing or recognizing humans or objects in video images for purposes beyond just security purposes.

Hitachi will continue contributing to the realization of a "safe and secure" society and applying image sensing techniques for public facilities and all kinds of fields behind the scenes at an international sports event in Tokyo, in 2020. Hitachi will also continue moving forward to resolve customers' operational efficiency improvement challenges and quality assurance improvement challenges utilizing these technologies.

REFERENCES

(1) S. Kanai et al., "Case Studies of Food Defense in Relation to Plant Security Management," Hitachi Review **64**, pp. 209–214 (Mar. 2015).
(2) T. Sugiura et al., "Industrial Solution Technologies for Overcoming Customer challenges," Hitachi Review **64**, pp. 182–187 (Mar. 2015).
(3) Y. Watanabe et al., "Person Retrieval System for Surveillance Video Analysis Using Multi-Perspective Features," Technical report of IEICE, Vol. 113, No. 493, PRMU2013-172, pp. 25–30 (Mar. 2014) in Japanese.

ABOUT THE AUTHORS

Satoshi Nakata
Physical Security Solution Design Department, Physical Security Solution Division, Security & Image System Solution Group, Hitachi Industry & Control Solutions, Ltd. He is currently engaged in the integrated security solutions business.

Tomokazu Murakami, Ph.D.
Media Systems Research Department, Center for Technology Innovation – Systems Engineering, Research & Development Group, Hitachi, Ltd. He is currently engaged in the development of image recognition and processing techniques. Dr. Murakami is a member of The Institute of Electronic, Information, and Communication Engineers (IEICE), The Institute of Image Information and Television Engineers (ITE), and The Virtual Reality Society of Japan (VRSJ).

Wataru Ebihara
Solution Planning Department, Security Business Planning Division, Security & Image System Solution Group, Hitachi Industry & Control Solutions, Ltd. He is currently engaged in product planning for security solutions.

Ichiro Ote
Business Management Department, Industrial Solutions Division, Industry & Distribution Business Unit, Hitachi, Ltd. He is currently engaged in product planning for security solutions.

Hiroshi Koya
Solution Planning Department, Security Business Planning Division, Security & Image System Solution Group, Hitachi Industry & Control Solutions, Ltd. He is currently engaged in business planning for security solutions.