

Video Analytics Technology for Disaster Management and Security Solutions

The rising demand for safety and security in Japan and elsewhere has created a need for more advanced fire and security solutions based on the application of video analytics technology to surveillance cameras. In response, Hitachi is developing fire detection technologies that can recognize indistinct features in images indicative of smoke, as well as security technologies that include human behavior recognition and the detection and tracking of people. By automating incident detection through the use of video analysis to identify objects or people and assess situations, these technologies facilitate incident response, such as alerting a person to check what is happening. This article presents an overview of these technologies and examples of where they have been trialed, and also describes the outlook for the future.

Tomokazu Murakami, Ph.D.

Duan Qingzhu

Yuki Watanabe, Ph.D.

Atsushi Hiroike, Ph.D.

1. Introduction

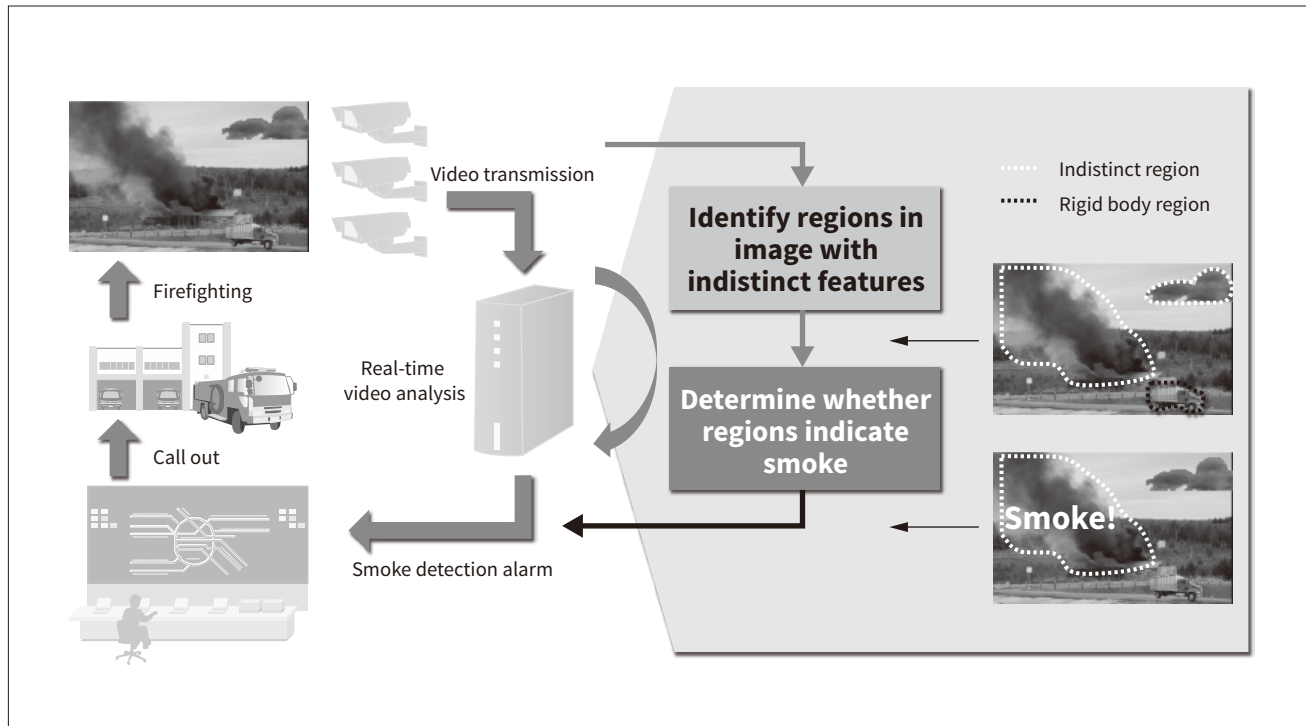
A wide variety of disaster prevention and security technologies are needed to prevent disasters, accidents, criminal acts, and other threats to people's way of life and the infrastructure of society, and to detect and respond quickly when they do occur. In the case of disaster prevention, there is a need to deal with a wide variety of threats, encompassing not only major earthquakes but also volcanic activity, forest fires, and rivers bursting their banks due to heavy rain. Likewise for crime prevention and security, there are growing calls for greater certainty in the maintenance of safety and security in response to the increasing incidence of

terrorism overseas, greater international mobility due to globalization, and the rising numbers of inbound tourists to Japan.

Along with this, a large number of surveillance camera systems are being used to monitor rivers, forests, and volcanos, and to prevent crime in public places such as railway stations, airports, and shopping complexes, for example. Although some of these are infra-red or other special-purpose cameras for monitoring at night or in other situations with poor lighting, most are visible-light camera. While many cameras have been installed in a variety of locations and for diverse purposes to enable the continuous and detailed monitoring and recording of what is happening over large areas, this still largely relies on people to view the images. However, as future increases in the

Figure 1 — Overview of Fire Detection System Based on Video Analysis

The system detects fire by using real-time analysis of video from surveillance cameras to identify regions in an image with indistinct features and determine whether they indicate smoke. This assists surveillance staff in the early detection and fighting of fires.



number of installed cameras to provide greater anti-crime and security capabilities will only place an even greater burden on the people doing the monitoring, there is a need to find ways of using video analytics technology to automate this work.

Along with supplying networked video monitoring systems like these, Hitachi has also responded by developing functions that use video analytics for event detection and situation assessment. Among the things able to be detected are water levels and the state of the flow in rivers, and the flames or smoke associated with forest fires or volcanic activity. In security, demand is especially strong for ways of assessing people's behavior, as well as situation assessment applications such as monitoring the movements of people and vehicles and identifying when objects have been abandoned.

This article describes video analytics technologies developed by Hitachi, its plans for future activities, and the outlook for the technology. These include a technology for detecting fire based on the recognition of indistinct features such as smoke, a technology for recognizing people's behavior, and a technology that uses machine learning for the detection and tracking of individuals.

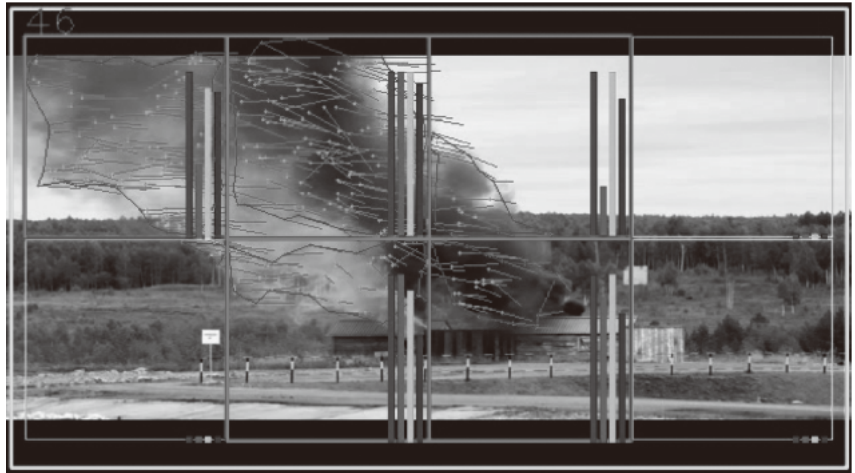
2. Detection of Fire by Recognition of Indistinct Features

Forest fires becoming larger and more frequent due to climate change and warming has become a concern in recent years. While detecting fires early, raising the alarm, and commencing firefighting are important for preventing fires from spreading, in developing nations in particular, the lack of adequate communications infrastructure means that, in many cases, there is little prospect of residents or others being able to raise the alarm. Accordingly, the practice in countries with frequent forest fires is to install surveillance cameras able to oversee a wide area and have fire department staff visually monitor them. Unfortunately, this visual monitoring is labor-intensive and even then, fires are sometimes overlooked. The new technology helps reduce the workload of surveillance staff by using video analysis for the automatic detection of the smoke from fires in surveillance video that covers a wide area, usually of forest (see **Figure 1**).

Ways of detecting fires in surveillance video that have been proposed in the past include using the

Figure 2 — Example Detection of Smoke by Recognition of Indistinct Features in Images

The overall image is split into blocks and the block feature descriptors used to check for the presence of indistinct features. Those blocks that contain indistinct features are then subject to further detailed analysis to detect whether they show smoke caused by a fire.



difference from the background to identify changes such as in the color or area of moving regions of the image (feature descriptors). The problem with this approach, however, is that it is difficult to tell the difference between clouds and smoke or to deal with changes in lighting, with the result that misidentifications are frequent.

In this case, the method used to improve the accuracy of fire detection was to first identify blocks in the image that contain indistinct features such as smoke or clouds, and then to focus the detailed analysis on these parts. Whether or not a region of image contains an indistinct feature is determined on the basis of feature descriptors representing the change in area of the moving region, variation in the direction of movement, linearity of movement, and spatial frequency. A model was also created to identify the presence of smoke from a fire that works by obtaining a “dense trajectory” feature descriptor for blocks identified as containing an indistinct feature and using previously collected images to perform learning. When tested on about 50 test videos of actual fires, this achieved a detection accuracy of more than 80%, with no misidentifications. **Figure 2** shows an example of detection results. Smoke detection is performed by checking each block of the divided image for regions containing indistinct features and then evaluating the feature descriptors for each of these.

Through use in forest or other monitoring systems, Hitachi believes that this technology can help improve the efficiency of surveillance work for preventing fires from spreading.

3. Human Behavior Recognition and Detection and Tracking of People in Security Applications

Security cameras are already widely used for security and progress is being made on the use of video analytics to automate detection. Hitachi has past experience of commercializing a variety of such practices, including a way of determining the level of crowding at railway stations and disseminating this information. The following section focuses in particular on the use in anti-crime applications of video analytics for human behavior recognition and the detection and tracking of people.

3.1

Human Behavior Recognition

Among the requirements for maintaining passenger and operational safety at locations such as railway stations is the ability to detect people carrying suspicious objects or moving erratically on platforms so that action can be taken before an incident or accident occurs. The first step in detecting specific types of behavior like these is to identify which regions of video images represent people. The movements of these people are then tracked over a number of frames and their images analyzed to determine things like what they are carrying and characteristics of their appearance.

A number of technologies for recognizing people’s behaviors have been developed to work within the limitations of the systems available at such sites

and to deliver the required accuracy. The technology described here has lightweight execution requirements and performs high-speed detection with minimal constraints due to system resources. It works by identifying “tracklet” feature descriptors (localized paths) from sequential camera frames and then applies cluster analysis to these using kernels that allow for occlusion (people’s images overlapping). This enables the regions of an image that contain people to be identified with high accuracy while still only requiring a low volume of computation. It also means that the movements of each person in the camera image can be tracked, and therefore that it is possible to detect people who are moving erratically by applying a geometric transform to these movements to calculate the degree of horizontal variation in their forward progress. Similarly, rule-based image evaluation, with color and shape as criteria, can be used to check for the presence of dangerous goods such as gasoline cans (see **Figure 3**).

Key features of this technology are that it can process camera video in real time and provide comparatively robust detection even under crowded conditions.

3.2

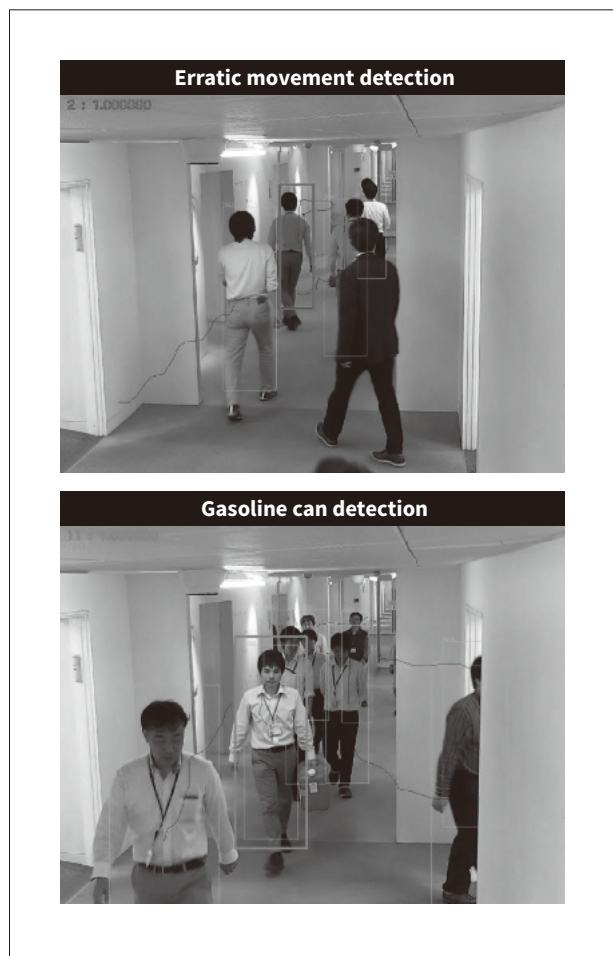
Detection and Tracking of People

Along with assessing events captured by individual cameras, large public spaces such as railway stations, airports, and stadiums where many security cameras are installed also need to be able to identify particular people and trace their movements, such as when tracking down the location of a suspicious individual or lost child within a large area. Whereas past practice when an incident occurred was to utilize a large number of people to view video footage based on witness accounts in order to locate and track the individual, Hitachi has developed a machine learning technology for person detection and tracking that can facilitate this process.

The technology uses a two-step process to quickly trace the movements of a specific individual (see **Figure 4**). The first step is able to rapidly narrow down a large amount of video to those scenes that show people who match the characteristics reported in witness accounts. It uses a deep neural network for attribute recognition that has already been trained

Figure 3 — Example Use of Video Analysis for Detection of Erratic Movement and Gasoline Cans

After identifying regions of the video image that show people, erratic behavior is identified by analyzing their path through the space, and gasoline cans by color and shape.



to perform real-time recognition of more than 100 features covering 12 categories (including gender, age group, hairstyle, type and color of clothing, and items carried) from the appearance of the people who appear in the surveillance video. The results of this are saved in a database. It can, for example, find all the video showing people who match the description of a “male aged about 30 to 45 wearing blue jeans and a green top.”

The next step involves choosing who in the selected candidate scenes to track. The behavior of the chosen individual can then be tracked across the areas covered by multiple cameras. This is done by first obtaining feature descriptors specific to the detected person from the intermediate layer output of the deep neural network based on images that show their entire body, and then comparing these feature descriptors against

Figure 4 — Overview of Person Detection and Tracking

In the past, responding to an incident required that people view a large amount of video footage based on witness accounts. In contrast, this technology can rapidly identify the movements of a person of interest by using numerous attributes to narrow down the candidates and by utilizing full body images to track them across multiple cameras.

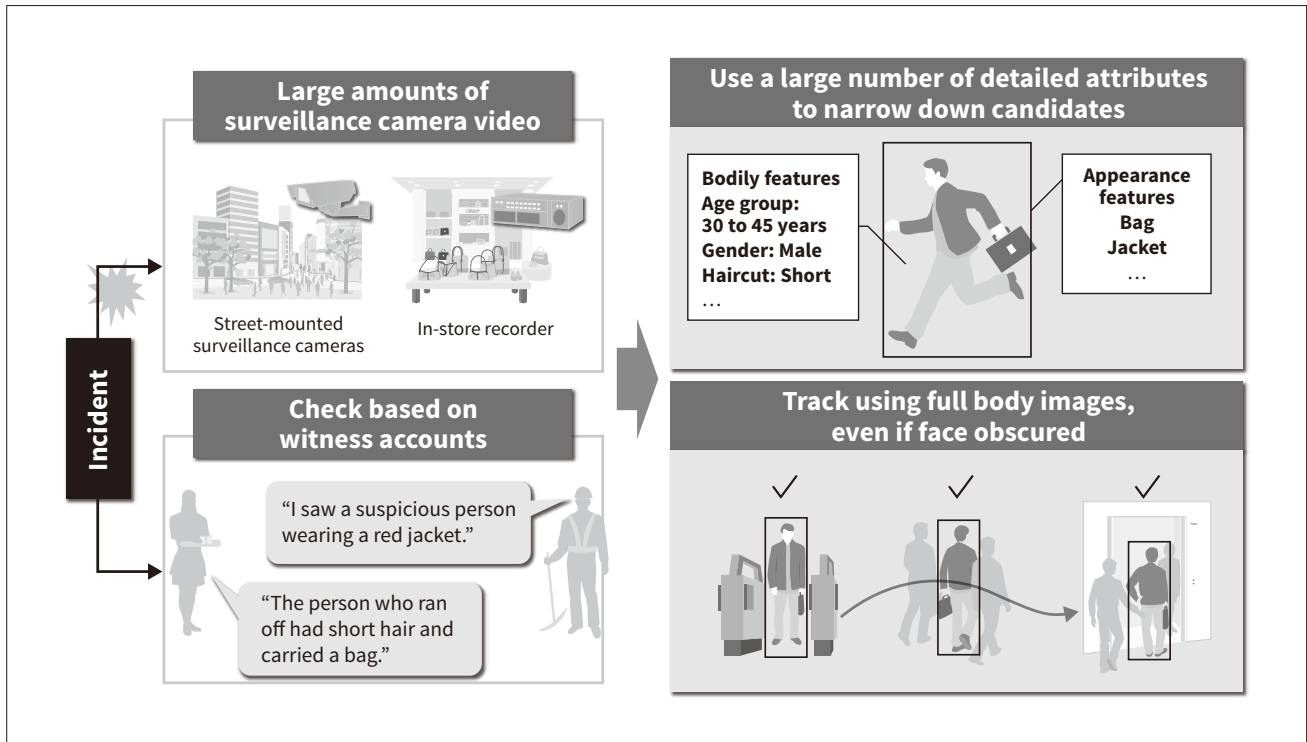


image data stored in a high-speed vector search database. Whereas past technologies that used facial recognition to track people frequently missed detection opportunities because they were unable to use images that did not clearly show the person's face, this new technology is capable of comprehensive tracking because it uses full-body images. The results of testing on data held by Hitachi found that the detection rate was three times higher than that when using facial recognition on its own, demonstrating an ability to perform searches of more than 10 hours of video in less than 1 second.

Figure 5 shows an example implementation of a system using this technology in which access was gained to a railway station and extras were hired to recreate the circumstances at the station. The trial demonstrated the tracking of a male acting in the role of a suspicious individual. As the technology can run in real time on a server connected to the surveillance cameras, it has the potential to be of major benefit as an aid for the tracking of suspicious people or searching for missing children at large public places.

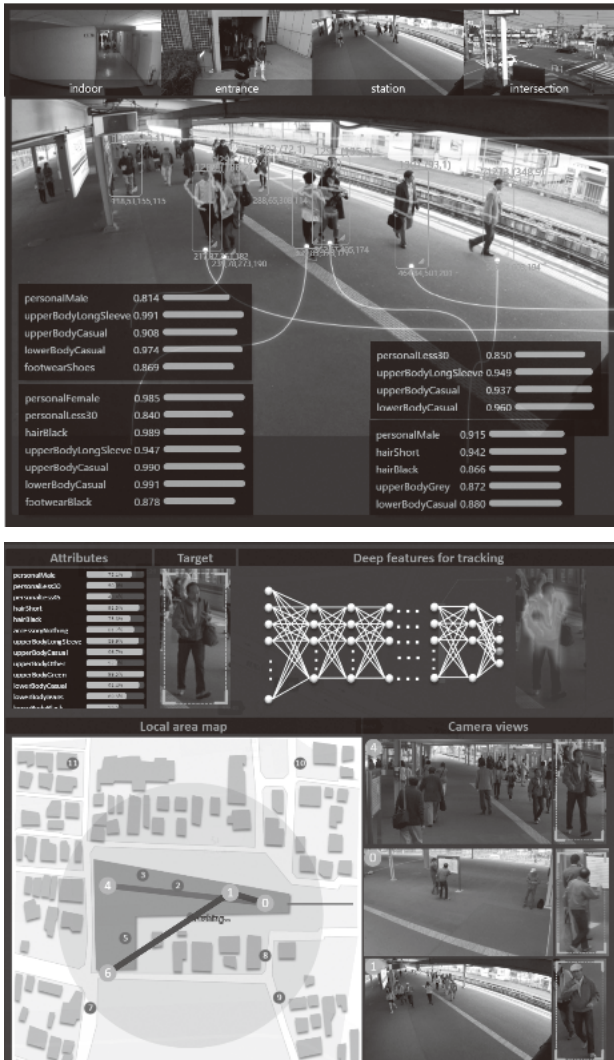
4. Conclusions

This article has described how video analytics technology underpins disaster management and security solutions in a technology for detecting fire based on the recognition of indistinct features such as smoke, and security applications for human behavior recognition and the detection and tracking of people.

Advances in machine learning such as convolutional neural networks, especially through international competitions, have led to significant improvements in accuracy for the recognition of objects and actions in recent years, and it is expected that these technologies will be put to practical use in the near future. While discussion of machine learning in this article has looked at technologies for recognizing people, it is anticipated that applications for the technology will spread to a wide range of other uses, including vehicles and general objects. Along with overcoming challenges such as the constraints on video analytics technology imposed by computing resources and

Figure 5 — Trial of Person Detection and Tracking

The tracking of a specific individual was trialed by installing a number of security cameras at a railway station and using extras to act as station users. The trial demonstrated that a function for tracking persons of interest can be incorporated into a security camera system by checking the person's attributes and extracting feature descriptors from full body images in real time.



making it more robust with respect to the environments in which it is used, Hitachi intends to continue developing the technology with the aim of supplying solutions that connect to security camera networks and provide high security even when staffing levels are low. The objective is to contribute to making society safe and secure by utilizing these technologies in applications such as disaster management over wide areas and the guarding and prevention of crime in public places.

References

- 1) Q. Duan et al., "Smoke Detection Method for Wide-area Surveillance by Indistinct Region Extraction Methods," the 2016 IEICE Society Conference (Sep. 2016) in Japanese.
- 2) S. Nakata et al., "Image Sensing Technologies for Analysis and Interpretation of Surveillance Video," Hitachi Review, 65, pp. 53–57 (Jun. 2016).
- 3) T. Sagawa et al., "Integrated Physical Security Platform Concept Meeting More Diverse Customer Needs," Hitachi Review, 65, pp. 353–358 (Sep. 2016).
- 4) Hitachi News Release, "Development of Image-analysis Technology with AI for Real-time People-detection and Tracking," (Mar. 2017), <http://www.hitachi.com/New/cnews/month/2017/03/170327.html>

Authors



Tomokazu Murakami, Ph.D.

Media Systems Research Department, Center for Technology Innovation – Systems Engineering, Research & Development Group, Hitachi, Ltd. *Current work and research:* Development of image recognition and processing technologies. *Society memberships:* The Institute of Electronics, Information and Communication Engineers (IEICE), The Institute of Image Information and Television Engineers (ITE), and The Virtual Reality Society of Japan (VRSJ).



Duan Qingzhu

Media Systems Research Department, Center for Technology Innovation – Systems Engineering, Research & Development Group, Hitachi, Ltd. *Current work and research:* Development of image recognition and processing technologies.



Yuki Watanabe, Ph.D.

Media Systems Research Department, Center for Technology Innovation – Systems Engineering, Research & Development Group, Hitachi, Ltd. *Current work and research:* Development of image retrieval and machine learning technologies.



Atsushi Hiroike, Ph.D.

Media Systems Research Department, Center for Technology Innovation – Systems Engineering, Research & Development Group, Hitachi, Ltd. *Current work and research:* Development of image retrieval technologies. *Society memberships:* The Japanese Psychological Association (JPA).