Design and Performance Evaluation of an AI-Based W-Band Suspect Object Detection System for Moving Persons in the IoT Paradigm

KEPING YU, (Member, IEEE), XIN QI, (Member, IEEE), SAN HLAING MYINT, ZHENG WEN, (Member, IEEE), YUTAKA KATSUYAMA, KIYOHITO TOKUDA, (Member, IEEE), WATARU KAMEYAMA, (Member, IEEE), AND TAKURO SATO, (Fellow, IEEE)

1Global Information and Telecommunication Institute, Waseda University, Tokyo 169-8050, Japan
2School of Fundamental Science and Engineering, Waseda University, Tokyo 169-8050, Japan

Corresponding author: Xin Qi (samqixin@aoni.waseda.jp)

This work was supported by the Research Grant for Expanding Radio Wave Resources in FY2020 of the Ministry of Internal Affairs and Communications through a Contract for Research and Development of Radar Fundamental Technology for Advanced Recognition of Moving Objects for Security Enhancement under Grant JP1900254.

ABSTRACT The threat of terrorism has spread all over the world, and the situation has become grave. Suspicious object detection in the Internet of Things (IoT) is an effective way to respond to global terrorist attacks. The traditional solution requires performing security checks one by one at the entrance of each gate, resulting in bottlenecks and crowding. In the IoT paradigm, it is necessary to be able to perform suspicious object detection on moving people. Artificial intelligence (AI) and millimeter-wave imaging are advanced technologies in the global security field. However, suspicious object detection for moving persons in the IoT, which requires the integration of many different imaging technologies, is still a challenge in both academia and industry. Furthermore, increasing the recognition rate of suspicious objects and controlling network congestion are two main issues for such a suspicious object detection system. In this paper, an AI-based W-band suspicious object detection system for moving persons in the IoT paradigm is designed and implemented. In this system, we establish a suspicious object database to support AI technology for improving the probability of identifying suspicious objects. Moreover, we propose an efficient transmission mechanism to reduce system network congestion since a massive amount of data will be generated by 4K cameras during real-time monitoring. The evaluation results indicate that the advantages and efficiency of the proposed scheme are significant.

INDEX TERMS Suspicious object detection, moving persons, artificial intelligence, millimeter-wave imaging, network congestion control, Internet of Things.

I. INTRODUCTION

In the war of counterterrorism, the international community is still facing a series of serious new challenges and new issues that need to be resolved. In the face of the technological and ideological trends of international terrorism, the use of advanced technologies such as artificial intelligence (AI) [1], 5G [2], [3], information-centric networking [4]–[6], blockchain [7], Internet of Things (IoT) technology and other means to counteract terrorism is a top priority, and the focus should be on improving antiterrorism systems’ early warning mechanisms, rapid response mechanisms and consequence processing mechanisms. At present, a new generation of equipment for criminal investigation, technical investigation, anti-explosive measures, identification and security, continuous upgrading of computer hardware and software, and efficient wireless and satellite communications is enriching the front line of antiterrorism
operations. Among them, suspicious object detection systems are an effective way to respond to global terrorist attacks and should be given high priority.

Taking public transportation security checks as an example, there are several different methods [8], [9] of performing security checks of both people and carry-on items, such as metal detector gates, X-ray checkers for carry-on items, explosive detectors and human body imagers. Human body imagers can be divided into different categories depending on the different kinds of technology they use, including X-ray fluoroscopy, X-ray backscatter imaging, and millimeter-wave (MMW) imaging technologies. X-ray fluoroscopy is widely used in the medical field and is a very effective way to obtain real-time moving images of the interior of an object such as the human body. However, it requires a large amount of radiation and thus is not an ideal method for security checks. X-ray backscatter imaging is an advanced X-ray-based imaging technology that is much safer than fluoroscopy. This technology has been widely used in security checks in public areas for many years. MMW imaging also follows a similar imaging principle as that of backscatter machines. However, MMW scanners emit a special class of microwaves instead of X-rays; these microwaves pass through clothing and bounce off the person’s skin as well as any potentially threatening objects, allowing them to be discovered. These traditional solutions require people to line up and wait to be checked one by one. This process greatly increases the waiting time and creates a bottleneck in the security check area. Therefore, with the increasing worldwide population and the expansion of international communication, it is necessary to develop a suspicious object detection system that does not require people to stop at a particular point.

To address this need for a “no-stop” suspicious object detection system, many countries began to develop fast and enhanced suspicious object detection technologies for security checks several years ago [10], [11]. An early study showed the importance of avoiding the interruptions caused by X-ray screening operations at security checkpoints [12]. To this end, new technologies are required to improve the security detection accuracy and the system performance. IoT technology [13], sensor networks [14], [15] and machine learning technologies [16], [17] were all applied in the early work in this field. Moreover, current technologies such as MMW imaging and AI networks are expected to be key elements to improve suspicious object detection [18].

With the increasing numbers of nodes in IoT sensor networks [19] and the higher resolution of surveillance images [20], the networks for suspicious object detection systems are facing pressure in terms of network resources and processing power. The aim for future suspicious object detection systems is to achieve high performance and low congestion. Therefore, one of the important points in developing a new type of suspicious object detection system is to build it on the basis of a low-congestion principle to efficiently control network traffic [21].

The contributions of this paper are listed as follows.

1) We design and implement an AI-based W-band (75GHz-110GHz) suspicious object detection system for moving persons in the IoT paradigm. Compared with the traditional solutions, which require security checks to be performed one by one at the entrance of each gate and consequently result in bottlenecks and crowding, the proposed system enables automatic, no-stop inspections on the basis of W-band unidentified object detection for locations with large numbers of people, such as subway/airport lobbies, shopping malls, and concert venues.

2) A suspicious object database based on simulations and active/passive imagers has been established to support AI-based suspicious object recognition technology. A performance evaluation shows that it can significantly increase the probability of identifying suspicious objects.

3) We propose a low-congestion suspicious object/person network system to ensure efficient and safe data transmission for real-time monitoring data. This system can automatically track a suspicious person between different areas in real time. Moreover, an efficient transmission mechanism is proposed to reduce network congestion. Evaluation results show that the advantages and efficiency of the proposed scheme are significant.

The remainder of this paper is organized as follows. Section II presents previous related studies, including research on surveillance network systems, W-band unidentified object detection, artificial intelligence for image recognition, and low-congestion video transmission for the IoT paradigm. In section III, we present our proposal of a suspicious object detection system for moving persons, including the objective, the system architecture, the suspicious object database to support AI-based recognition technology, and the suspicious object/person network system. Then, in section IV, we evaluate this suspicious object detection system from two perspectives: AI-based suspicious object recognition and the suspicious object/person network system. Finally, we conclude our work.

II. PRELIMINARIES

A. SURVEILLANCE NETWORK SYSTEMS

As an essential technology that is currently changing our lives, IoT devices such as sensor networks have been widely researched in recent years. A video surveillance network is generally divided into three levels: front-end access, media exchange, and user access. More specifically, such a network is composed of a front-end coding unit, central business platform, network recording unit, client unit, and decoding unit. Historically, the security surveillance industry has experienced four stages of development: the analog surveillance era, the digital surveillance era, the network high-definition era, and the current intelligent surveillance era. Each industry update relies on upstream technological innovation and component cost reduction.

Compared with traditional video surveillance, modern surveillance network systems make it more convenient for
computers to perform automatic processing of video information, such as compression, storage, analysis and display, to achieve automated operation. This enables remote monitoring through such a network platform, even from thousands of kilometers away. The ability to use advanced digital software systems to complete a large amount of data analysis in a few minutes improves monitoring efficiency and yields more realistic and clearer digital images, facilitating more convenient and practical monitoring management and maintenance.

B. W-BAND UNIDENTIFIED OBJECT DETECTION

The MMW band is generally defined as the frequency band between 30 GHz and 300 GHz, which lies between the infrared and microwave regions in the frequency spectrum. Compared with visible light and infrared light, this spectral band offers a certain penetrability of most nonmetallic objects while also providing a resolution that microwaves cannot; therefore, technology operating in this band has recently become a focus of research in academia and industry. MMW body imaging technology is an advanced technology in the global security field. It has been used in passenger security screening at airports in the United States, the United Kingdom, the Netherlands, Australia, and Japan, but previously, only the United States and the European Union had access to related technology standards. Such a device can effectively detect suspicious objects hidden on various parts of the human body under the cover of clothing without requiring direct contact with the body, especially nonmetallic items, and the shapes, sizes, and positions of suspicious objects can be determined from the acquired images. In addition, MMW body imaging equipment is harmless to humans and has strong penetrating power. Its transmission power is less than one-thousandth of the electromagnetic wave radiation of mobile phones. It can accurately identify objects carried on the human body, effectively improving objectivity and accuracy, reducing the labor of security inspectors and improving security efficiency.

There are two main approaches to MMW imaging: one is passive, and the other is active. Passive imaging uses millimeter waves emitted by the human body itself to perform imaging. Such imaging is effective in an outdoor setting, but it is difficult to obtain a satisfactory image indoors because the temperature differences between the millimeter waves are limited. The principle of active imaging is similar to that of the flash of a camera. A MMW source emits radiation onto an object to be detected, and the field strength reflected by the object is detected for imaging. This approach can produce more satisfactory results for indoor imaging. Such MMW cameras can be widely used in airport security inspection, airborne imaging radar, industrial inspection, and other fields.

C. ARTIFICIAL INTELLIGENCE FOR IMAGE DETECTION

As AI devices become increasingly involved in our everyday lives, machine learning is enabling improvements in image tagging and object identification. Computer vision is the science of computer systems recognizing and analyzing different images and scenes. A key component of computer vision is object detection. Object detection is used to perform various AI tasks, such as facial recognition, vehicle detection, security scanning, and self-driving. Successful object detection has been demonstrated using several AI algorithms. The most well-documented algorithms related to object detection processes include Regions with Convolutional Neural Networks (R-CNN), Fast R-CNN, and Faster R-CNN. Given an image, R-CNN uses selective search to generate approximately 2000 region proposals from which to compute features by means of a convolutional neural network (CNN). Region proposals are regions that include potential objects. Each region proposal is wrapped as a $227 \times 227$ RGB image patch for input to a CNN. Feature extraction is then performed in the CNN layers, and the results are passed to multiple binary classifiers to determine the class of each particular region. A CNN is designed to focus on pixels within an image that are located next to each other. Each image is passed through the network as an input and is then sent back as an output in which each object is classified. A challenge in AI systems is to achieve accurate object detection in the case of fast-moving objects such as vehicles.

Based on future improvements in imaging technology and deep-learning-based object detection, we anticipate future suspicious object detection systems with the ability to alert about hidden threats and help security groups prevent terrorism [22].

D. LOW-CONGESTION VIDEO TRANSMISSION FOR THE IoT

Many kinds of data are collected and exchanged in a surveillance network, such as image data, video data, and processed information data. The largest-scale and least effective data are streaming video data. A video stream is a continuous stream of data that contains mostly useless information. Therefore, reducing the volume of video data is an easy way to save network bandwidth resources and improve overall network efficiency.

There are currently many ways to reduce the data volume involved in video transmission, such as different types of video encoding. Reducing the raw video data volume was the legacy approach used when people were still required to manually surveil all video and extract information. Now, with the increased usage of AI technologies [20] and IoT devices, it is easier and more efficient to process video data in a decentralized network [23] for application in a surveillance system. In addition, a previous study has shown that it is more efficient to allow IoT-based edge nodes to process the raw data [24].

III. PROPOSED SYSTEM DESIGN

A. OBJECTIVE

The objective of this paper is to design and implement an AI-based W-band suspicious object detection system for moving persons in the IoT paradigm. First, we attempt to use
current AI-based suspicious object recognition technology to improve the recognition rate of the proposed suspicious object detection system. To achieve this goal, we build a suspicious object database by means of simulations and experiments. Then, we design and implement a low-congestion suspicious object/person network system. It can effectively transmit the monitoring information generated by visible light cameras and hybrid imagers. Moreover, the proposed system can effectively reduce network congestion.

In recent years, the threat of terrorism has spread throughout the world, and the situation has become dire. To enhance the safety of public places, highly efficient suspicious object detection should be automatically performed. However, it is not wise to perform security checks one by one at the entrance of each gate since this will result in bottlenecks and crowding, especially in subway/airport lobbies, shopping malls, concert venues, etc. It is therefore necessary to perform automated suspicious object detection on moving people. As illustrated in Figure 1, the goal of this paper is to develop sensing/imaging technology that operates in the W-band to recognize suspicious objects and visualize humans or hidden suspicious objects from a distance. This system can be used in places with large numbers of people to automatically detect suspicious objects while greatly reducing the processing time.

**B. SYSTEM ARCHITECTURE**

To achieve the goal of “ensuring a sufficient security level without stopping the flow of people,” the proposed suspicious object detection system uses a staged screening method (primary screening/secondary screening) to recognize suspicious objects concealed by humans (Figure 3). In the primary screening stage, a W-band radar is used to measure suspicious objects approximately 15 meters away, and for the secondary screening stage, a W-band imager will be developed to measure suspicious objects concealed by humans within 5 meters. By integrating the sensing/imaging information with the existing visual surveillance camera images during these two screening processes, suspicious objects concealed by moving persons can be visually recognized.
allow security personnel to quickly obtain the location of a suspicious person and request a direct inspection. In brief, the entire safety inspection process using the proposed suspicious object detection system is as follows. First, visible light cameras should be able to detect each person, and active radars are used to conduct a primary screening of the people in the area to check whether they are carrying metal objects. Second, the system tracks each suspicious person, and the identified suspicious people are guided to secondary screening points. Then, it is necessary to associate the MMW images collected in the secondary screening stage with the corresponding visible light images. Finally, the suspicious objects are identified from the MMW images obtained during secondary screening; during this process, AI technologies are utilized to improve the identification probability. Moreover, this system supports the tracking of suspicious people in different areas.

The two main contributions of this paper are the suspicious object database and the suspicious object/person network system. Specific technical details will be described in later sections. Since a large number of 4K cameras will be used in this suspicious object detection system, massive amounts of data will be generated during real-time monitoring. To improve transmission efficiency, we propose an efficient transmission mechanism for use in the suspicious object/person network system to reduce network congestion.

C. SUSPICIOUS OBJECT DATABASE TO SUPPORT AI-BASED RECOGNITION TECHNOLOGIES

For this suspicious object detection system, it is crucial to establish a suspicious object database to assist AI-based recognition technologies to improve the identification probability. First, it is necessary to build a suspicious object database by means of simulation and the collection of images from active/passive imagers.

1) IMAGE GENERATION VIA SIMULATION

To improve public security, passive and active imaging technologies have become a key solution for overcoming the challenges faced by suspicious object detection systems, such as concealed objective detection. Imaging technologies can be classified into passive and active sensors. In this section, we present a passive imaging simulator. Since the amplitude of the detected radiation depends on the target object’s emissivity and temperature, these two parameters are considered the main parameters for the radiometric calculation. Then, we implement the preamp receiver, which includes a preamplifier for direct detection. Finally, we evaluate the simulation results by using raw data images from a real imager. The detailed simulation process is described in Figure 4.

First, we design the parameters for the input data [25], which consist of images that include the human body and various objects, to be used as training data for image recognition by a neural network and further research related to this project. Here, we consider the radiometric parameters as key parameters for passive imaging. The parameter values

are expressed in Table 1, where $T$ is the temperature of the human body or a metallic object, $\varepsilon$ is the emissivity, $\rho$ is the reflectivity, and $T_{room}$ is room temperature. In this research, we consider knives, scissors, forks and bottles as the concealed objects represented in the input data or training data. After designing the parameters, we create the dataset for the subsequent simulation process. The created dataset includes 245280 images, and the dataset size is 21.8 gigabytes. All input images are settled in frame size 192(H) × 512(V) pixel resolution by transforming in grayscale that simulates the actual image size of the passive imager.

![Figure 4. Flow chart of passive imaging simulation.](image-url)

![Figure 5. Basic physical method for passive imaging simulation.](image-url)
The adjusted value of the object, \( T_j \), is calculated as shown below:

\[
T_j = \frac{1}{N} \sum_{i} \sum_{j} T_k(i, j) \ast (x, y)
\]

(2)

where \( m(x, y) \) is a mask operator. Then, the output of the receiver with the amplifier, \( l(i, j) \), can be expressed as shown below:

\[
l(i, j) = F(T_o(i, j)) = a \ast T_o(i, j) \ast k(i) + off(i) + r
\]

(3)

where \( a \) is the coefficient for the amplitude of the amplifier, \( k(i) \) is the variation parameter of the amplitude for sensor \( i \), \( off(i) \) is the variation parameter of the offset for sensor \( i \) and \( r \) is Gaussian noise.

As the final step, the received signals are converted into the original image by using an image processing technique. We adopt a backprojection algorithm for image reconstruction based on the Radon transform, the inverse Radon transform and the projection-slice theorem. Here, the process of image reconstruction by calculating the average over \( N \) frames, \( J(i, j) \), can be expressed as shown below. Finally, the adjusted value \( L(i, j) \) is obtained to fit the digital data for the output image.

\[
J(i, j) = \frac{1}{N} \sum_{n=1}^{N} l_n(i, j)
\]

(4)

where \( N \) is the number of frames.

\[
L(i, j) = 255 \ast (J(i, j) - J_{low}) / (J_{high} - J_{low})
\]

(5)

where \( J_{low} \) is the minimum temperature value and \( J_{high} \) is the maximum temperature value.

**FIGURE 6. Results of the created dataset.**

After the simulation, we can compare the results obtained using a dataset collected from a real imager with the results obtained using our created dataset. We analyze the results obtained using the two different input datasets, i.e., the raw dataset generated by the real imager and our created dataset, as shown in Figure 6. According to the results, the proposed passive imaging simulator can produce good results that are very similar to the original images for each dataset. On the other hand, the created dataset yields significantly improved simulation results.

2) SUSPICIOUS OBJECT DATABASE

AI-based suspicious object recognition technology can be used to improve the identification probability for suspicious objects in the proposed suspicious object detection system. To achieve this purpose, a suspicious object database should be built for network training and evaluation through both simulation and image collection with active/passive imagers.

![Prototype of a suspicious object database generated via simulation.](image)

**FIGURE 7.** Prototype of a suspicious object database generated via simulation.

\( a: \)-generation of a suspicious object database via simulation

Passive sensor images can be generated via simulation to build a suspicious object database. There are two kinds of parameters that should be considered, and a prototype is illustrated in Figure 7. First, we should consider the whole simulation environment, including temperature, reflection, blur, variation, and noise. Second, the parameters of each suspicious item should be considered, including the suspicious object’s type and size as well as various rotations and transformations. Accordingly, as summarized in Table 2, we generated 44 kinds of bottles, 41 kinds of forks, 37 kinds of knives and 45 kinds of scissors via simulation, resulting in a total of 10516, 9799, 8843, and 10755 samples, respectively.

\( b: \) generation of a suspicious object database by means of active/passive imagers

In addition to simulation, we can also generate real experimental images through active or passive MMW imaging for suspicious object databases. As shown in Figure 8, we built a simple anechoic chamber with active/passive imagers (left panel) to generate image data. The right panel shows an example of generated passive imager data. Using this chamber, we performed an experiment with subjects carrying different...
numbers and types of suspicious objects (knives, bombs, guns, liquids, phones, etc.). The relative positions of the subjects and the orientations of the suspicious objects were also varied. Usually, suspicious objects are hidden in clothes or bags; therefore, we also used different kinds of items (cloth, cotton, etc.) to wrap the suspicious objects during this experiment. Accordingly, as summarized in Table 3, we generated 52 samples using the active imager and 1009 samples using the passive imager. In this MMW imaging system, passive imager, based on the prototype of [26], can shoot 1 image per 2 seconds. In contrast, we use a very primitive prototype of the active imager that consists of a single antenna, a receiver and mechanical scanning structure. It will take 6 hours to generate one sample from the active imager. Therefore, the number of samples between active imager and passive imager is significantly different.

3) AI-BASED SUSPICIOUS OBJECT RECOGNITION ALGORITHM

To improve the probability of identifying suspicious objects in the proposed suspicious object detection system, we use AI technology based on our developed suspicious object database. Among the available AI technologies, CNNs represent one of the main categories of deep learning for performing image recognition and image classification. Two critical features of CNNs that distinguish them from other neural networks are their reduced computational complexity and translational invariance. A CNN comprises two main sections. The first section is used to extract features and includes convolutional layers, pooling layers, and batch normalization layers. The second section works in the same way as a traditional neural network and is used to perform classification; it includes one or more flatten layers and one or more fully connected layers. In this paper, we directly use CNN technology, meaning that the performance of the AI component depends on the CNN. In addition, we consider the application of AI with and without a noisy environment in this work. Because CNNs are a mature technology for image recognition and this paper focuses on the construction of the suspicious object database rather than the CNN, the technical details of the CNN are not described here.

D. SUSPICIOUS OBJECT/PERSON NETWORK SYSTEM

1) SYSTEM STRUCTURE

In recent years, the threat of terrorism has spread all over the world; hence, strengthening suspicious object detection technologies has become an urgent issue. To this end, we are developing a low-congestion suspicious object detection system for moving persons that integrates various sensors and imagers. Figure 3 has already described the whole system architecture, which includes AI-based suspicious object recognition technology and a suspicious object/person network system. In this section, we will introduce the suspicious object/person network system in detail. As shown in Figure 9, the operation process of this system mainly includes the following five steps: a) person detection, b) tracking, c) association, d) suspicious object recognition, and e) tracking through different areas.

a: PERSON DETECTION

This step is performed only in the primary screening stage. The purpose of this step is mainly to detect people by means of visible light cameras (4K cameras) and to initially recognize suspicious persons who are carrying metal objects by means of active radar detection. At the same time, each suspicious person’s location is recorded. In this step, corresponding information such as the “area ID,” “person ID,” “image,” “location (x,y),” “time,” “suspicious object type” and “suspicious person confidence” is generated.

b: TRACKING

A large amount of traffic will be generated in the network system if we transmit the original 4K video. Therefore, it is necessary to develop a content-based extraction and production process for the video data in the suspicious object/person network system to reduce network congestion. In this step, person-only regions are cropped from different RGB images in multiple frames and integrated to reduce network traffic. Meanwhile, corresponding information such as “groups of
images,” “groups of locations” and “groups of times” are added to the suspicious object database.

c: ASSOCIATION

Suspicious persons detected to be carrying metal in the primary screening stage will be guided by security personnel to a secondary screening point. The MMW images generated during secondary screening are then associated with the corresponding camera RGB images. Meanwhile, “MMW images” are recorded as new information for each suspicious person.

d: SUSPICIOUS OBJECT RECOGNITION

By means of the suspicious object database in combination with AI-based suspicious object recognition technology, each suspicious object is accurately identified, and the information of the “dangerous object type” and “suspicious person confidence” is recorded.

e: TRACKING THROUGH DIFFERENT AREAS

Since a suspicious person’s movements may pass through different monitoring areas, this system supports monitoring a given suspicious person in different areas.

The overall scope of this system is to design and implement a suspicious object/person detection and tracking system. This section describes the recognition, tracking and communication platform, which is a central part of our project. As shown in Figure 10, there are four specific components of the overall system: a) Person Detection & Tracking, b) Tracking Management, c) Suspicious Person/Object Binding, and d) Data Management. The network system should

FIGURE 9. System structure of the suspicious object/person network system.

FIGURE 10. Function design for the suspicious object/person network system.
be designed with flexible structures to handle a maximum of 100 RGB cameras, 100 infrared cameras, 100 external suspicious person detection systems, and 100 hybrid imagers. For instance, it should be possible to add additional PCs for the Person Detection & Tracking process, as shown in Figure 10. Figure 11 presents the detailed function modules of the suspicious object/person network system. Moreover, we also provide an example of information flow in the proposed suspicious object/person network system (Figure 12).

To track a suspicious person through different areas, a local person ID will be generated when the person is first detected (for example, at the airport entrance). At the same time, the end point of tracking (for example, the boarding gate) is predetermined. A local person ID exists independently in each area, but the global person ID must be unified; for this purpose, it is necessary to associate local person IDs in different areas. Therefore, each pair of adjacent areas is defined with some overlap for the association of local person IDs. Facial recognition or some other image-feature-based identification method is needed to handle complicated situations. However, these approaches are not in the scope of the first system. Additionally, location data should be defined using both a global location and a local location. For the example of an airport, the global location indicates the absolute position in the whole airport, and the local location indicates the position within the surveillance range of one camera. Given the area ID and local location, the global location can be calculated using the predetermined allocation parameters. In the table in Figure 11, only two (local) areas are described. However, overall location data should be recorded that can be acquired from every frame of camera data.

2) CONTENT-BASED ASSOCIATION OF VARIOUS SENSORS IN THE SUSPICIOUS OBJECT/PERSON NETWORK SYSTEM

In the suspicious object detection system, a large number of sensors will be used. It is vital to associate different types of sensor data acquired from the same object. To address this issue, we study a method of estimating a person’s position from each type of sensor data. Figure 13 illustrates a scenario similar to our system. It includes an RGB camera and a MMW imager, which can detect a person from distances of 6 m to 15 m and 1 m to 5 m, respectively.
detected from the same person by the RGB camera and the MMW imager.

Face detection [27] is used to estimate the person’s position from an image. As shown in Figure 14, based on the initial vertical position $V_{\text{face}}$ of the rectangle obtained via face detection and the height $L$ of the face rectangle, the foot position $V_{\text{foot}}$ of the detected person in the image can be calculated as shown in equation 6 below.

$$V_{\text{foot}} = V_{\text{face}} + k \times L$$  \hspace{1cm} (6)

where $k$ is the proportion coefficient. Based on a previously published human dimension database [28], the average height of adolescents is 1699.1 mm, and the average morphological face height is 121.1 mm. Therefore, $k$ can be simply considered to be 14.

Here, the morphological face height is the distance from the nose to the lowest point of the lower jaw. In this study, this distance is estimated from the face detection area. It is considered equivalent to the distance from the eyebrows to the mouth. The horizontal position of the foot in the image, $h_{\text{foot}}$, is taken to be the midpoint of the face rectangle. Then, a projection transformation matrix is obtained for the image captured by the visible light camera, and the position ($h_{\text{foot}}, V_{\text{foot}}$) in the image can be converted into the position ($x, y$) from the camera.

3) REDUCTION OF NETWORK TRAFFIC VOLUME FOR THE SUSPICIOUS OBJECT/PERSON NETWORK SYSTEM

To reduce risks for soft targets in public areas, a suspicious object detection system that can handle a large number of cameras and various sensors, including 4K cameras and MMW imagers, is needed. In this section, we study an efficient transmission mechanism to reduce the network congestion in the system since a massive amount of data will be generated by the 4K cameras during real-time monitoring.

As illustrated in Figure 3, the operation process of the suspicious object detection system mainly includes five steps: person detection, tracking, association, dangerous object recognition and tracking through different areas. Network traffic reduction is implemented in the person detection and tracking steps. Many related kinds of research have already been developed and proven to be useful for the current network system [29]. We have reviewed and selected several effective methods to reduce network traffic during data transmission.

Figure 15 illustrates the flow chart for the reduction of network congestion. This system uses 4K RGB cameras in combination with a large number of sensors to detect suspicious objects and track suspicious persons. The 4K RGB cameras generate $3840 \times 2160$ resolution images at 5 fps. Then, these images are passed to the “Person Detection Function” (Figure 16(a)) and the “Tracking Function” (Figure 16(b)). We set several options for each process to
seek an effective method of network congestion reduction. For the Person Detection Function, we provide three methods of reducing network traffic: 1) crop a person-only region from the whole image, 2) compress the image obtained from 1), or 3) resize the cropped image from 1) to dimensions smaller than $64 \times 256$ pixels and then compress the resulting image. The reason why we consider resizing the cropped image is that the size of a cropped person-only image varies with the distance between the camera and person. Resizing is an effective way to reduce the image size, especially when the person is close to the camera. For the Tracking Function, we propose two methods of alleviating network congestion: a) select consecutive frames but remove images with noise and detection errors or b) in addition to a), select only the latest frame as a representative image if a group of images look sufficiently alike.

For the Person Detection Function, we provide three methods of reducing network traffic: 1) crop a person-only region from the whole image, 2) compress the image obtained from 1), or 3) resize the cropped image from 1) to dimensions smaller than $64 \times 256$ pixels and then compress the resulting image. The reason why we consider resizing the cropped image is that the size of a cropped person-only image varies with the distance between the camera and person. Resizing is an effective way to reduce the image size, especially when the person is close to the camera. For the Tracking Function, we propose two methods of alleviating network congestion: a) select consecutive frames but remove images with noise and detection errors or b) in addition to a), select only the latest frame as a representative image if a group of images look sufficiently alike.

IV. EVALUATION

A. EVALUATION OF AI-BASED RECOGNITION TECHNOLOGY

This section describes the implementation of an image recognition system using a CNN. Figure 17 shows the configuration of the CNN used to evaluate the proposed suspicious object detection system. In addition to the input and output layers, it includes five main types of layers: convolution, pooling, batch normalization, flatten, and fully connected. The network parameters for each layer are given in Table 4. In this experiment, we used four types of suspicious objects for the training and evaluation of the CNN: bottles, forks, knives and scissors. All images had dimensions of $32 \times 32$ pixels resolution to fit the image size of the actual passive imager with three types of deterioration: normal, blurred and noisy (examples are shown in figure 18). Table 5 specifies the numbers of images used for the training and evaluation of the CNN. We used a total of 39913 images for training, including 10516 bottle images, 9799 fork images, 8843 knife images and 10755 scissors images. For evaluation, we used 44 bottle images, 41 fork images, 37 knife images and 45 scissors images, for a total of 167 images.

The recall is best (92.2%) when normal images are used for both training and evaluation. Generally, the recall is higher when the same deterioration type is used for both training and evaluation, but the noisy case is different. When we use noisy images for training, the recall is 75.4%, 74.9% and 73.1% when evaluation is performed using images with different deterioration types (normal, blurred, and noisy), and the image recognition performance was evaluated using the evaluation images for each type of deterioration. Table 6 shows the evaluation results for different deterioration types. According to these result, the recall is best (92.2%) when normal images are used for both training and evaluation. Generally, the recall is higher when the same deterioration type is used for both training and evaluation, but the noisy case is different. When we use noisy images for training, the recall is 75.4%, 74.9% and 73.1% when evaluation is performed using images with different deterioration types (normal, blurred, and noisy), and the image recognition performance was evaluated using the evaluation images for each type of deterioration. Table 6 shows the evaluation results for different deterioration types.
TABLE 7. Object identification results with blur and noise.

|          | Evaluation results |          |          |          |          |          |          |
|----------|--------------------|----------|----------|----------|----------|----------|----------|
|          | Knives  | Bottles | Forks | Scissors | Total | Recall | Average recall |
| Input    |         |         |        |          |        |        |               |
| Knives   | 61      | 3       | 5      | 5        | 74     | 82.4%  | 80.2%    |
| Bottles  | 0       | 84      | 0      | 4        | 88     | 95.5%  |          |
| Forks    | 26      | 8       | 46     | 2        | 82     | 56.1%  |          |
| Scissors | 8       | 3       | 1      | 78       | 90     | 86.7%  |          |
| Total    | 95      | 98      | 52     | 89       | 334    |         |          |
| Precision| 64.2%   | 85.7%   | 88.5%  | 87.6%    |         |         |          |
| Average precision | 81.5% |         |         |          |         |         |          |

Performed using normal, blurred, and noisy images, respectively. In actual suspicious object detection systems, most of the obtained images will be blurred or noisy. Table 7 lists the object identification results obtained with blur and noise. In this case, the CNN was trained on blurred and noisy images, and the evaluation was also performed using blurred and noisy images. The recall and precision were determined for each object category separately. As seen from this table, the average recall is 80.2%, and the average precision for the four categories is 81.5%. Therefore, the target accuracy of at least 50% was achieved.

Moreover, the ROC curve calculated based on Table 7 was used to evaluate the system. The calculation method is as follows. When an image is input for recognition, the system will output a similarity for each of the four different kinds of suspicious objects (knives, bottles, forks, and scissors). The similarity for the category that is the same as the input category is set to OK, and the similarities for the other categories are set to NG. From the recognition results for multiple input images, the probability density functions of the sets of OK and NG are calculated, and the threshold value of the probability is varied to set the false positive rate and true positive rate (Figure 19).

Figure 20 shows the ROC curve. In this figure, the horizontal axis is the false positive rate, and the vertical axis is the true positive rate; the values from Table 7 are plotted in this coordinate system to obtain the ROC curve. Three cases are considered. (a) The conditions of the evaluation images are the same as those of the training images, as is the case in Table 7. (b) The images used for evaluation have a background gray value of 130, an object gray value of 100, and a noise level of 20 dB or 10 dB stronger than that in (a) and are subjected to a $21 \times 21$ blur filter to enhance the blur. (c) The images used for evaluation have a background gray value of 255, an object gray value of 0, a noise level of 10 dB, and enhanced blur due to a $21 \times 21$ blur filter. From Figure 20, we can obtain the following conclusions. In case (a), if the true positive rate (accuracy) of the recognition results is approximately 90%, the false positive rate (rate of incorrect answers that are regarded as correct) is approximately 15%.

Similarly, if the true positive rate of the recognition results is approximately 50%, the false positive rate is reduced to approximately 2%. With different thresholds, this system can provide corresponding true positive rates and false positive rates. The threshold should be chosen based on the specific needs of different situations. With higher image degradation, as represented in (b) and (c), the accuracy of recognition will decrease. When the true positive rate is 50%, the false positive rate is 2% in (a), 20% in (b) and 24% in (c). According to these results, to achieve high accuracy, the suspicious object detection system should be used to evaluate images acquired under conditions similar as possible to those of the images used for training and with as little deterioration as possible.

B. EVALUATION OF THE SUSPICIOUS OBJECT/PERSON NETWORK SYSTEM

1) SYSTEM EVALUATION

In the suspicious object/person network system, three types of processing time ($T_a, T_b, T_d$) were measured to evaluate the system’s performance, as shown in figure 21.
The processing time $T_a$ is the time required for person information detection and the integrated and aggregated packet generation for a detected person for (a) tracking processing. The processing time $T_b$ is the time required to aggregate the hybrid imager data into an integrated packet associated with the same person for (b) association processing. The processing time $T_d$ is the time required for the suspicious object recognition result data to be processed by the AI-based recognition server, received, saved and distributed for (c) data management. $T_c$ is the processing time by suspicious object detection server, which is out scope of network system and assumed maximum is 1 second. Since the network system starts operating when a person reaches a sufficiently close distance of 5 m, i.e., the working distance of the hybrid imagers, under the assumption of an average walking speed of 1 m/s, the target total time from $T_a$ to $T_d$ should be within 5 seconds. We used a 4K-resolution clip of 6 people walking for this evaluation. The whole process generated 8 tracking packets, and $T_a$, $T_b$ and $T_d$ are shown in Figure 22. Although there is one $T_a$ sample exceeding 1 second, the average time remains within 1 second (0.826 s). The results for $T_b$ and $T_d$ are all within 1 second; thus, the overall evaluation time is acceptable for the suspicious object/person network system.

2) CONTENT-BASED ASSOCIATION OF VARIOUS SENSORS IN THE SUSPICIOUS OBJECT/PERSON NETWORK SYSTEM

Regarding the integration of data from various sensors, we evaluate the accuracy of position estimation for visible light camera images. As shown in figure 23 (a), six walking persons are detected in a 4K video ($3840 \times 2160$ pixels) at 5 frames/sec. The red arrow in figure 23 (b) indicates the correct distance value obtained from the perspective projection transformation of the evaluated video. It is compared with the estimated distance value calculated using the proposed method to evaluate the distance estimation error. Figure 24 illustrates the distance estimation errors for persons at different distances from the camera. The six persons are represented by differently colored dots.

| Person | 1 | 2 | 3 | 4 | 5 | 6 |
|--------|---|---|---|---|---|---|
| RMSE (m) | 1.18 | 0.67 | 0.59 | 1.2 | 0.60 | 0.85 |

Table 8 shows the root mean square errors (RMSEs) of the estimated positions for the six persons. These results confirm that the position of a walking person can be estimated via face detection with an error of approximately 1 m. In this study, we investigated how to estimate the position of a person from the coordinates obtained through face detection based on an average height and face size. We believe that it is possible to correlate the outputs of different sensors based on a person’s position.
3) REDUCTION OF NETWORK TRAFFIC VOLUME FOR THE SUSPICIOUS OBJECT/PERSO NETWORK SYSTEM

We used a 20-second, 34.6 megabyte (MB) mp4-encoded 4K (3840 × 2160 resolution) video depicting six persons walking (samples from this video are shown in figure 16(a)) to evaluate the proposed transmission mechanism for network traffic reduction. In a total of 96 frames, the "person detection" process yielded 207 extracted regions, and in the "tracking" process, these regions were integrated into 14 groups based on different persons. As a result, 90 regions identified as noise during the tracking process were eliminated.

**TABLE 9. Evaluation results for network traffic reduction.**

| Method               | Data size (MB) |
|----------------------|----------------|
| Person detection     | Tracking       |
| (1) Crop             | 200.4          |
| (2) Crop, Zip        | 91.1           |
| (3) Crop, Resize, Zip| 6.5            |
| (3) Crop, Resize, Zip| 6.5            |

| Method               | Data size (MB) |
|----------------------|----------------|
| Tracking             | Person detection |
| (a) Select consecutive frames but remove images with noise | 122.8 |
| (a) Select consecutive frames but remove images with noise | 55.1 |
| (b) Select the latest frame of each group as a representative image | 3.8 |
| (b) Select the latest frame of each group as a representative image | 0.5 |

The evaluation results for network traffic reduction are shown in Table 9. Instead of sending the full-length 4K video (34.6 MB), the packet size can be reduced to 0.5 MB if we resize and compress the cropped images in addition to selecting only the latest frame of each group as a representative image. These results show that the network traffic can be efficiently reduced for the suspicious object/person network system.

V. CONCLUSION

In this paper, we have designed and implemented an AI-based W-band suspicious object detection system for moving persons. By means of W-band unidentified object detection, this system can perform no-stop suspicious object detection automatically, making it suitable for densely populated places such as subway/airport lobbies, shopping malls, and concert venues. With the goal of using a CNN to improve the probability of identifying suspicious objects, a suspicious object database has been established by means of simulations and the collection of images from active/passive imagers. Moreover, we have proposed an efficient transmission mechanism to reduce network congestion in the system since a massive amount of data will be generated by the 4K cameras during real-time monitoring. The evaluation results show that the designed suspicious object detection system can achieve good performance in suspicious object detection for moving persons. In addition, the established suspicious object database can support CNN analysis to significantly increase the probability of identifying suspicious objects, and the proposed low-congestion transmission mechanism can improve the network transmission efficiency.

ACKNOWLEDGMENT

This article was presented in part at the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, May 2020 [30].

REFERENCES

[1] Z. Zhou, H. Liao, B. Gu, K. M. S. Huq, S. Mumtaz, and J. Rodriguez, “Robust mobile crowd sensing: When deep learning meets edge computing,” *IEEE Netw.*, vol. 32, no. 4, pp. 54–60, Jul. 2018.
[2] L. Zhen, T. Sun, G. Lu, K. Yu, and R. Ding, “Preamble design and detection for 5G enabled satellite random access,” *IEEE Access*, vol. 8, pp. 49873–49884, 2020.
[3] S. H. Myint, K. Yu, and T. Sato, “Modeling and analysis of error process in 5G wireless communication using two-state Markov chain,” *IEEE Access*, vol. 7, pp. 26391–26401, 2019.
[4] K. Yu, S. Eum, T. Kurita, Q. Hua, T. Sato, H. Nakazato, T. Asami, and V. P. Kafle, “Information-centric networking: Research and standardization status,” *IEEE Access*, vol. 7, pp. 126164–126176, 2019.
[5] X. Qi, Z. Wen, K. Yu, K. Murata, K. Shibata, and T. Sato, “Content-oriented disaster network utilizing named node routing and field experiment evaluation,” *IEICE Trans. Inf. Syst.*, vol. E102.D, no. 5, pp. 988–997, May 2019.
[6] K. Yu, M. Arifuzzaman, Z. Wen, D. Zhang, and T. Sato, “A key management scheme for secure communications of information centric advanced metering infrastructure in smart grid,” *IEEE Trans. Instrum. Meas.*, vol. 64, no. 8, pp. 2072–2085, Aug. 2015.
[7] Z. Zhou, B. Wang, M. Dong, and K. Ota, “Secure and efficient vehicle-to-grid energy trading in cyber physical systems: Integration of blockchain and edge computing,” *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 1, pp. 43–57, Jan. 2020.
[8] Q. Liu, B. Kang, K. Yu, Q. Ji, L. Si, Wang, H.-A. Li, “Contour-Maintaining-Based image adaption for an efficient ambulance service in intelligent transportation systems,” *IEEE Access*, vol. 8, pp. 12644–12654, 2020.
[9] H.-A. Li, M. Zhang, K. Yu, J. Zhang, Q. Hua, B. Wu, and Z. Yu, “Combined forecasting model of cloud computing resource load for energy-efficient IoT system,” *IEEE Access*, vol. 7, pp. 149542–149553, 2019.
[10] Y. O. Yildiz, D. Q. Abraham, K. Panetta, and S. Agaian, “A new concept of airport security screening,” in *Proc. IEEE Conf. Technol. Homeland Secur.*, May 2008, pp. 444–448.
[11] N. Hattenschwiler, S. Merks, and A. Schwanger, “Airport security X-ray screening of hold baggage: 2D versus 3D imaging and evaluation of an on-screen alarm resolution protocol,” in *Proc. Int. Carnahan Conf. Secur. Technol. (ICCST)*, Oct. 2018, pp. 1–5.
[12] M. L. Janer and M. D. Rossetti, “Simulation modeling of alternatives to avoid interruptions of the X-ray screening operation at security checkpoints,” in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2016, pp. 3005–3016.
[13] X. Qi, Y. Su, K. Yu, J. Li, Q. Hua, Z. Wen, J. Lopez, and T. Sato, “Design and performance evaluation of content-oriented communication system for IoT network: A case study of named node networking for real-time video streaming system,” *IEEE Access*, vol. 7, pp. 88138–88149, 2019.
[14] S. Chen, L. Zhang, C. Shen, K. Yu, S. H. Myint, and Z. Wen, “On scheduling policies with heavy-tailed dynamics in wireless queuing systems,” *IEEE Access*, vol. 8, pp. 32137–32149, 2020.
[15] X. Qi, W. Sun, J. Wang, Q. Li, G. Luo, and K. Yu, “RVFL-LQP: RVFL-based link quality prediction of wireless sensor networks in smart grid,” *IEEE Access*, vol. 8, pp. 7829–7841, 2020.
[16] L. Martin, A. Tuyluzoglu, W. C. Karl, and P. Ishwar, “Learning-based object identification and segmentation using dual-energy CT images for security,” *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4069–4081, Nov. 2015.
[17] H. Liao, Z. Zhou, X. Zhao, L. Zhang, S. Mumtaz, A. Jolfaei, S. H. Ahmed, and A. K. Bashir, “Learning-based context-aware resource allocation for edge computing-empowered industrial IoT,” *IEEE Internet Things J.*, early access, Dec. 31, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8946558, doi: 10.1109/JIOT.2019.2963371.
K. Yu et al.: Design and Performance Evaluation of an AI-Based W-Band Suspicious Object Detection System

KEPing Yu (Member, IEEE) received the M.E. and Ph.D. degrees from the Graduate School of Global Information and Telecommunication Studies, Waseda University, Tokyo, Japan, in 2012 and 2016, respectively. He was a Research Associate and a Junior Researcher with the Global Information and Telecommunication Institute (GITI), Waseda University, from 2015 to 2019 and from 2019 to 2020, respectively, where he is currently a Researcher with GITI. He has hosted and participated in more than ten projects, is involved in many standardization activities organized by ITU-T and ICNRG of IRTF, and has contributed to ITU-T Standards Y.3071 and Supplement 35. His research interests include smart grids, information-centric networking, the Internet of Things, blockchain, and information security. He was the General Co-Chair and the Publicity Co-Chair of the IEEE VTC2020-Spring EBTSM workshop, the TPC Co-Chair of the SCML2020, the Local Chair of the MONAMI 2020, the Session Chair of the CCs2020, and the Session Chair of the ITU Kaleidoscope 2016. He has served as a TPC Member for more than ten international conferences, including ITU Kaleidoscope, the IEEE Vehicular Technology Conference (VTC), the IEEE Consumer Communications and Networking Conference (CCNC), and the IEEE Wireless Communications and Networking Conference (WCNC). He has been a Lead Guest Editor for Sensors, Peer-to-Peer Networking and Applications, and Energies. He is an Editorial Board Member of the IEEE Open Journal of Vehicular Technology (OJVT).

Xin Qi received the M.E. and Ph.D. degrees from Waseda University, Tokyo, Japan, in 2016 and 2019, respectively. He is currently a Research Associate with the Global Information and Telecommunication Institute, Waseda University. His research interests include ICN and 3N for next-generation communication systems.

Toshosto Sato (Member, IEEE) received the B.S. degree in fine measurement engineering and the M.S. degree in systems engineering from the Nagoya Institute of Technology, Nagoya, Japan, in 1985 and 1987, respectively, and the Ph.D. degree in systems design engineering from the University of Fukui, Japan, in 2003. He joined Toshiba Corporation, Japan, in 1987, where he developed video processing systems for tasks, including facial recognition and vehicle detection.

San Hlaing Myint received the B.C.Tech., B.C.Tech. (Hons.), and M.C.Tech. degrees from the University of Computer Studies at Mandalay, Myanmar, in 2004, 2005, and 2008, respectively, and the Ph.D. degree from Waseda University, Tokyo, Japan, in 2019. He was a Tutor with the Ministry of Science and Technology, Myanmar, from 2007 to 2015. From 2009 to 2014, he was with the University of Computer Studies at Yangon, Myanmar, where he performed research focused on sustainable cloud computing. He is currently a Research Associate with the Global Information and Telecommunication Institute, Waseda University. His current research focuses on fifth-generation wireless networks and the analysis of channel coding methods.

Zheng Wen (Member, IEEE) received the B.E. degree in computer science and technology from Wuhan University, China, 2009, and the M.Sc. and Ph.D. degrees from Waseda University, Japan, in 2015 and 2019, respectively. In 2018, he became a Research Associate with Waseda University, Tokyo, Japan, where he is currently an Assistant Professor (a Lecturer) with the Department of Communication and Computer Engineering. His research interests include ICN/CCN for next-generation communication systems, AI, and the IoT.
YUTAKA KATSUYAMA received the B.E. and M.E. degrees in information engineering and the Ph.D. degree in information science from Tohoku University, Japan, in 1984, 1986, and 2014, respectively. From 1986 to 1992, he worked for Fujitsu Ltd., Japan. From 1992 to 2019, he worked for Fujitsu Laboratories Ltd., Japan, where he developed OCR systems. He is currently a Researcher with the Global Information and Telecommunication Institute, Waseda University, Japan. His main research interest includes pattern recognition. He is a member of the IEICE. He received the FIT2002 Award (Paper Award), in 2002.

Kiyohito Tokuda received the B.E., M.E., and D.E. degrees in electrical engineering from Hosei University, in 1981, 1983 and 1996, respectively. He joined the Communication Systems Laboratory, Research and Development Group, Oki Electric Industry Company, Ltd., in 1983. Since then, he has involved in research on ITS communication systems based on mobile data transmission technology and signal processing technology. He was the General Manager of the Wireless Technology Research and Development department of Oki and retired, in 2010. He has been appointed as a Guest Professor with the Global Information and Telecommunication Institute (GITI), Waseda University, a Project Researcher with RCAST, University of Tokyo, and a Project Researcher with ROEC, Tohoku University. He has also been established as the CEO of both Kurouto Networks Corporation, in 2010, and Wuxi Kurouto Networks Corporation, in 2012. He is currently a Senior Researcher with GITI.

Wataru Kameyama (Member, IEEE) received the bachelor’s, master’s, and D.Eng. degrees from the School of Science and Engineering, Waseda University, in 1985, 1987, and 1990, respectively. He joined ASCII Corporation, in 1992, and was later transferred to France Telecom CCETT, from 1994 to 1996, for a secondment. After joining Waseda University as an Associate Professor, in 1999, where he has been a Professor with the Department of Communications and Computer Engineering, School of Fundamental Science and Engineering, since 2014. He has also been involved in MPEG, MHEG, DAVIC, and TV-Anytime Forum activities. He was the Chairman of ISO/IEC JTC1/SC29/WG12 and has been the Secretariat and the Vice Chairman of the TV-Anytime Forum. He is a member of IEICE, IPSJ, ITE, IIEEJ, and ACM. He received the Niwa-Takayanagi Best Paper Award, in 2006, and the Niwa-Takayanagi Best Author Award, in 2009, from the Institute of Image Information and Television Engineers, and the International Cooperation Award from the ITU Association of Japan, in 2012.

Takuro Sato (Fellow, IEEE) was born in Niigata, Japan, in January 1950. He received the B.E. and Ph.D. degrees in electronics engineering from Niigata University. He has been a member of the Research and Development Laboratories, Oki Electric Industry Company, Ltd., Tokyo, Japan, where he worked on PCM transmission equipment, mobile telephone technology and the standardization of mobile data transmission, and CDMA systems for the international standardization committee. From 1977 to 1978, he developed AT&T AMPS (EIA/TIA-553) cellular phone equipment, Oki Electric Industry Company, Ltd. He developed a high-speed cellular modem for the AMPS cellular system, USA, in 1983. This technology was proposed to be standardized in the CCITT (now ITU) SG17 standard. In 1990, he developed a data transmission system based on digital cellular technology. He developed a W-CDMA system named IS-665 under the auspices of TIA for next-generation cellular systems. In 1990, the T1P1/TIA Joint Technical Committee (JTC) was organized to evaluate the proposed 2nd-generation, 1.9 GHz, Personal Communications Systems. He has proposed W-CDMA, which passed the evaluation tests and became TIA Standard IS-665 and T1P1 Standard J-STD-015, in 1996. He became a Professor with the Department of Information and Electronics Engineering, Niigata Institute of Technology, in 1995. He has contributed to the standardization process for the IEEE 802.11a. He has established the venture company Key Stream to provide LSI integrated circuits for 802.11 wireless LAN systems. In 2004, he became a Professor with the Faculty of Science Engineering, Waseda University. Recently, he has been interested in smart grid technologies in cooperation with ICT systems, including wireless communication. He is also interested in mobile edge computing technologies based on information-centric networking (ICN) and their applications in 5G mobile communication networks. He is a Fellow of IEICE and JSST, and a member of the Engineering Academy of Japan.