



# Device-Free WLAN Based Indoor Localization Scheme With Spatially Concatenated CSI and Distributed Antennas

Osamu Muta , *Member, IEEE*, Keisuke Takata, *Member, IEEE*, Kazuki Noguchi, *Student Member, IEEE*, Tomoki Murakami , *Member, IEEE*, and Shinya Otsuki, *Member, IEEE*

**Abstract**—Various machine learning (ML) based localization schemes using channel state information (CSI) in wireless local area networks (WLANs) have been investigated recently. Adopting a proper feature selection technique is important to achieve further improvement in detection accuracy. As described herein, we propose a device-free indoor localization scheme using a lightweight ML model with compressed spatially concatenated CSI in WLAN systems with distributed antennas. In this scheme, feedback beam-forming weights (BFWs) are collected at a CSI capture terminal. Then, current and past BFWs are concatenated as accurate feature data to characterize the object behavior. Additionally, we propose the use of a frequency-domain sampling scheme for a low-complexity real-time target position detection with a small number of datasets. Using ray-trace based simulation analysis and experimentally obtained results from an indoor environment, we demonstrate that the proposed scheme using the concatenated CSI is effective not only for achieving more accurate real-time detection, but also for reducing the necessary complexity for both off-line training and on-line classification compared with other reference schemes.

**Index Terms**—Beam-forming weight, localization, spatially concatenated channel state information, wireless local area network.

## I. INTRODUCTION

WIRELESS sensing technologies integrated with wireless communications are key technologies for development of 6 G and beyond [1], [2], [3]. To be more specific, future wireless networks are expected to provide not only data transmission but also additional functions to support new application services such as sensing by radio signals. Recently, indoor object detection and localization using radio signals of existing communication infrastructure such as wireless local area

networks (WLANs) have received much attention [4], [5], [6]. The basic principle of object detection using radio signals is to characterize target object behaviors as fluctuations of wireless channels caused by target objects. To achieve that characterization, various wireless sensing schemes having different bases have been elucidated in the literature, such as radio-frequency identification (RFID) based [8], Bluetooth based [9], ZigBee based [10], and WLAN based [6], [7], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26] schemes. Among them, WLAN-based sensing is a cost-effective approach because existing access points (APs) can be exploited for sensing purposes.

Actually, WLAN-based sensing techniques using measured received signal strength (RSS) or channel state information (CSI) have been investigated widely [6], [7]. Particularly an effective way for characterizing target object behaviors more accurately is to utilize subcarrier-wise CSI in orthogonal frequency division multiplexing (OFDM), i.e., amplitude and phase fluctuations in a frequency-selective fading channel. Furthermore, combined with multi-input multi-output (MIMO) technologies, CSI is expected to enhance wireless sensing accuracy further by exploiting rich channel information obtained through multiple antennas [18].

To this end, an effective device-free CSI acquisition scheme for WLAN-based sensing is proposed wherein feedback frames conveying beam-forming weights (BFWs) are collected at an off-the-shelf WLAN device, after which they are used to train a machine learning (ML) model and to detect object positions and their behaviors [21], [22]. Although this method is effective for collecting CSI without explicit measurements, the available CSI and its achieved accuracy are limited when WLAN is equipped with only a single antenna. Although the localization accuracy can be improved by increasing the number of antennas in WLAN systems [24], [26], the achieved performance might deteriorate when a lightweight algorithm with a small dataset is used. In addition, because radio propagation characteristics are sensitive to antenna locations and their surrounding environments, investigating an effective feature selection technique for ML-based object detection is important. However, as described above, the effects of surrounding radio propagation environments on the achievable performance have not been analyzed sufficiently. To enable real-time sensing at a mobile WLAN device with limited

Manuscript received 24 February 2022; revised 23 June 2022 and 8 August 2022; accepted 1 September 2022. Date of publication 12 September 2022; date of current version 16 January 2023. This work was supported by collaborative research with the NTT Corporation. The review of this article was coordinated by Dr. M. Shahgir Ahmed. (*Corresponding author: Osamu Muta.*)

Osamu Muta is with the Center for Japan–Egypt Cooperation in Science and Technology, Kyushu University, Fukuoka 819-0395, Japan (e-mail: muta@ait.kyushu-u.ac.jp).

Keisuke Takata and Kazuki Noguchi are with the Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka 819-0395, Japan (e-mail: takata@mobcom.ait.kyushu-u.ac.jp; noguchi@mobcom.ait.kyushu-u.ac.jp).

Tomoki Murakami and Shinya Otsuki are with the NTT Corporation, Yokosuka 239-0847, Japan (e-mail: tomoki.murakami.nm@hco.ntt.co.jp; shinya.otsuki.ma@hco.ntt.co.jp).

Digital Object Identifier 10.1109/TVT.2022.3205767

hardware and energy resources, the required complexity and dataset must be minimized. Therefore, further investigation and analyses are necessary to develop a lightweight algorithm with a small dataset. The original contributions of this paper are three, as presented below.

- As described in this paper, we aim to design a real-time device-free indoor ML-based localization scheme with compressed spatially concatenated CSI in WLANs with distributed antennas. In this scheme, CSI feedback frames are collected continuously at an off-the-shelf WLAN device similarly to a process described by Murakami et al. [21]. Then current and past BFWs are concatenated as more accurate feature information. Subsequently, they are used as training data for an ML model. Unlike the conventional scheme [21], more accurate and low-complexity detection is possible by learning concatenated CSI that includes channel information in a spatial domain.
- Second, we propose the application of frequency-domain sampling-based simple compression to concatenated BFWs to reduce the inherent complexity. In this method, the collected BFWs are sampled at every several subcarriers in the frequency domain so that the inherent computational complexity can be reduced while improving the object detection accuracy to a considerable degree. Additionally, we theoretically discuss an optimum compression ratio that minimizes the data size (i.e., required complexity) under a given object detection performance, based on the fact that frequency-domain sampling of BFWs is equivalent to delay-time domain windowing of its inverse Fourier transform.
- After implementing the proposed design to an IEEE802.11ac-based WLAN, we conduct experimental evaluations and ray-trace based simulation to demonstrate the effectiveness of the proposed approach in an indoor environment under a multi-user (MU-)MIMO scenario. By collecting feedback frames and by building a database of the concatenated CSI, we clarify that the proposed approach achieves much faster execution while achieving object position detection more precisely than when using the conventional scheme.

*Notation:* Vectors and matrices are expressed respectively as lower case and upper case letters in bold typeface. Superscript  $H$  denotes the Hermitian transpose of a matrix.  $\mathbb{R}^{a \times b}$  and  $\mathbb{C}^{a \times b}$  respectively denote the real and complex matrix fields of dimension  $a \times b$ .  $\mathbb{Z}$  represents the set of integers.  $*$  stands for the convolution integral. The notation and variables used for this study are listed in Table I.

## II. RELATED WORK

Various WLAN-based sensing techniques have been proposed in the literature [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [24], [25], [26]. One typical approach is to use measured RSS data to characterize target objects. In one earlier study [13], an improved RSS fingerprinting algorithm for indoor positioning is presented, for which a clustering algorithm is adopted to delete noisy samples. The authors in [14] earlier proposed a data-rate

TABLE I  
NOTATION OF PARAMETERS AND VARIABLES

Symbol	Description
$\mathbf{A}^T$	Transpose of a matrix $\mathbf{A}$
$\mathbf{A}^H$	Hermitian transpose of a matrix $\mathbf{A}$
$\mathbb{R}^{a \times b}$	Real matrix fields of dimension $a \times b$
$\mathbb{C}^{a \times b}$	Complex matrix fields of dimension $a \times b$
$\mathbb{Z}$	Set of integers
$N$	Number of stations (STAs)
$M$	Number of antennas at APs
$S$	Number of streams
$K$	Number of subcarriers per symbol
$R$	Number of area divisions where the target exists
$U$	Spatially concatenated CSI length
$R_c (= 1/D_s)$	Compression ratio of CSI
$d$	Antenna spacing at AP
$D_s$	Sampling interval
$L$	Number of subcarriers after sampling ( $\lfloor K/D_s \rfloor$ )
$\mathbf{H}_k \in \mathbb{C}^{N \times M}$	MU-MIMO channel matrix at the $k$ -th subcarrier
$\mathbf{h}_{k,n} \in \mathbb{C}^{M \times 1}$	Channel vector at the $k$ -th subcarrier and the $n$ -th STA
$\mathbf{v}_{k,n} \in \mathbb{C}^{M \times 1}$	Beam-forming weights at the $k$ -th subcarrier and the $n$ -th STA
$\mathbf{0}_M^T \in \mathbb{C}^{M \times 1}$	The zero vector (length is $M$ ).
$R$	Number of regions

based fingerprint method that achieves comparable localization performance without measuring RSS, where the transmission power and resultant data rate are used in the fingerprint database. In one earlier reported method [15], RSS data of large amounts measured using the distributed massive MIMO systems are exploited for estimating the location of a user, where RSS data clustering is applied to reduce complexity. As investigated in the existing works described above, RSS data and other related information are available in most WLAN devices. They are readily applicable for sensing purposes. However, exploiting RSS data alone is insufficient for accurate sensing because those results indicate only the received power averaged over the signal bandwidth. An alternative approach to achieving additional performance improvement is to exploit CSI, which exhibits radio propagation characteristics. The authors of one earlier study [16] presented an overall system design for location-oriented activity identification, where CSI in OFDM systems obtained through existing WLAN devices is exploited for a walking user's recognition of various activities. One earlier report [17] presents a design of a smoking detection system in which a specific motion such as smoking activity is detected by extracting meaningful CSI variation information from WLAN signals. Articles in [27], [28], [29] proposed hybrid schemes that solve the localization problem through hybrid RSSI and time of arrival (TOA) measurements (or time difference of arrival). Unlike these approaches assuming that the target has wireless devices, the proposed method in this paper relies on a device-free concept that can detect an object with no wireless device. Another report [18] presents a through-the-wall human detection system design using commodity WLAN devices by which principal component analysis based filtering is applied to extract meaningful features from CSI, where correlated subcarriers are selected to extract features for robust human

detection. Nevertheless, the main emphasis of subcarrier selection is detection of the existence of human beings. Therefore, an effective subcarrier selection scheme is still necessary for achieving lightweight device-free localization. In [30], a subcarrier power allocation scheme for OFDM-based localization is developed in which subcarrier power allocation at the anchor nodes is optimized to reduce positioning errors. In [31], orthogonal multicarrier-based probing signal designs for base station (BS)-based or mobile station (MS)-based localization are presented where non-overlapping subcarrier allocation is necessary for MS based localization to differentiate signals from different BSs. In [32], a RSSI based positioning scheme is proposed in which RSSI of a selected reference subcarrier is exploited for trilateration-based positioning. Unlike the works described above, this paper presents consideration of simple but effective frequency-domain sampling (i.e., subcarrier selection) that is applicable to compressed CSI specified in IEEE802.11, i.e., beam-forming weights. Reportedly, more reliable results can be achieved by adopting a majority-vote-based detection scheme using MIMO technique. In another report [19], a fingerprint quality classification method is presented for improving CSI-based positioning accuracy, whereas a convolutional neural network based approach has been applied for CSI-based positioning in another earlier report [20]. One remaining challenge of CSI-based approaches is acquisition of a sufficient number of CSI samples. In some earlier reports [21], [22], [23], [24], [25], [26], effective IEEE802.11ac-based CSI acquisition schemes are presented in which feedback CSI frames from all nearby devices are collected and analyzed for sensing purposes. The experimentally obtained results are also discussed in an earlier report [26]. Consequently, the sensing area can be expanded easily without installing additional measuring stations (STAs). In one earlier study [23], a key feature extraction scheme using PCA is adopted for acquired CSI samples for WLAN-based human detection with a deep neural network and numerous CSI samples. Although this method is used to reduce the dimensions of huge CSI samples, it remains unclear whether this method is effective for extracting key features from a few CSI samples. Unlike deep-learning-based methods trained using massive datasets, the method described in this paper is a lightweight approach with a small data set. Spectrogram-based detection and estimation techniques have been investigated in the literature [33], [34], where multidimensional analysis of time-frequency signals is considered. In [34], spectrogram-based analysis is reported as effective for R-R Interval estimation in vital sign monitoring in the health care field. According to these existing works, the detection performance is expected to be enhanced by analyzing channel characteristics multidimensionally.

Unlike the works described above, the main emphasis on this paper is the design of an effective spatially concatenated CSI-based localization scheme that works with a small dataset<sup>1</sup>. To this end, we propose a device-free WLAN based localization

<sup>1</sup>Although this paper mainly presents consideration of random forest as an ML model, we confirmed the effectiveness of the proposed approach with other machine learning models, as discussed in Fig. 22.

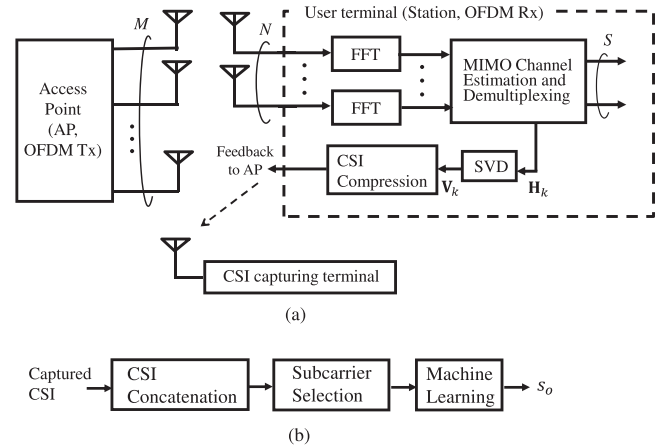


Fig. 1. IEEE802.11ac-based system block diagram, where  $M$  antenna elements are equipped with AP serving  $N$  user devices. Feedback BFWs are extracted at a CSI capturing terminal.

scheme with spatially concatenated CSI and frequency-domain sampling. We also discuss the achieved detection performance through simulation and results obtained through real-time experimentation in an indoor MU-MIMO scenario.

### III. SYSTEM DESCRIPTIONS

Fig. 1 depicts a WLAN-based block diagram with AP, a single antenna user device (STA: station), and a CSI capturing terminal to extract feedback BFWs from the captured feedback frames and analyze them, where IEEE802.11ac-based OFDM transceivers [36] are used. Here,  $M$ ,  $N$ , and  $S$  respectively denote the number of transmit antennas at AP, the number of users, and the total number of streams. Here,  $S \leq \min(M, N)$ . Also,  $\min(a, b)$  is a function that selects a smaller value (either  $a$  or  $b$ ). This section defines signal representations at AP and user devices. Details of the CSI capturing terminal are presented along with the proposed design in the next section.

The OFDM signal with  $K$  subcarriers is transmitted through  $M$  antennas at AP. Let  $\mathbf{H}_k = [\mathbf{h}_{k,1}, \dots, \mathbf{h}_{k,N}] \in \mathbb{C}^{N \times M}$  denote the MU-MIMO channel matrix at the  $k$ -th subcarrier, where  $\mathbf{h}_{k,n} = [h_{k,n,1}, \dots, h_{k,n,M}]^T \in \mathbb{C}^{M \times 1}$ ,  $k = 1, \dots, K$ , and where  $K$  represents the number of subcarriers per OFDM symbol. Here,  $h_{k,n,m}$  is the channel coefficient between the  $m$ -th transmit antenna and  $n$ -th user device at the  $k$ -th subcarrier. Consequently, the concatenated channel matrix over subcarriers in the frequency domain can be expressed as  $\mathbf{H} = [\mathbf{H}_1, \dots, \mathbf{H}_k, \dots, \mathbf{H}_K]$ .

On the STA side, after passing through the MIMO channel, the received OFDM signal is demodulated with FFT for channel estimation and data detection. The estimated channel matrix at the  $k$ -th subcarrier  $\hat{\mathbf{H}}_k \in \mathbb{C}^{N \times M}$  is decomposed by singular value decomposition (SVD) into  $\hat{\mathbf{H}}_k = \mathbf{U}_k \Sigma_k \mathbf{V}_k^H$ , where  $\mathbf{U}_k \in \mathbb{C}^{N \times N}$ ,  $\mathbf{V}_k \in \mathbb{C}^{M \times M}$ , and  $\Sigma_k \in \mathbb{C}^{N \times M}$  respectively represent the left-singular matrix, right-singular matrix, and diagonal matrix for which the diagonal element is a singular value of the channel. If  $M > N$ , then the right-singular matrix is reduced to  $\mathbf{V}_k = [\mathbf{v}_{k,1}, \dots, \mathbf{v}_{k,N}, \mathbf{0}_M^T, \dots, \mathbf{0}_M^T] \in \mathbb{C}^{M \times M}$ . The first  $N$

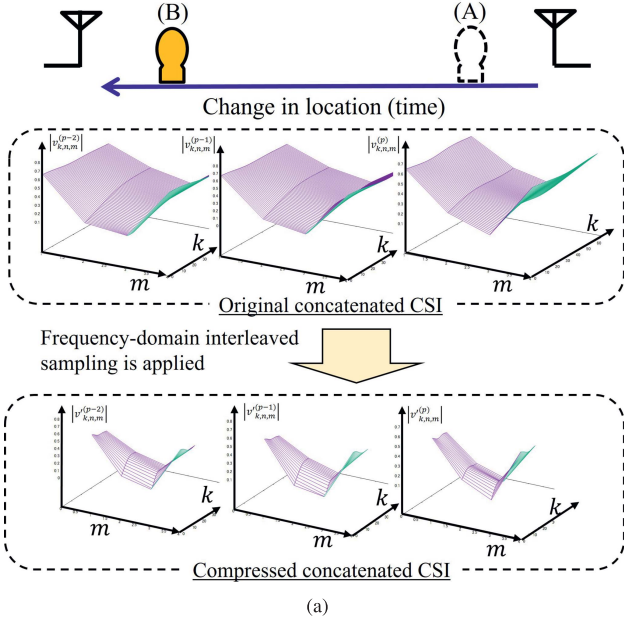


Fig. 2. Concept of the proposed approach for device-free WLAN-based object detection and localization, where subscripts  $k$  and  $m$  respectively denote the subcarrier index and transmit antenna index. (a) Illustrative example of concatenated CSI construction and frequency-domain compression. (b) Frequency-domain sampling scheme.

column vectors  $\mathbf{v}_{k,n}$  are used as beam-forming weights for  $N$  user devices. Here,  $\mathbf{v}_{k,n} = [v_{k,n,1}, \dots, v_{k,n,M}]^T \in \mathbb{C}^{M \times 1}$  and  $\mathbf{0}_M^T = [0, \dots, 0]^T \in \mathbb{C}^{M \times 1}$  denotes the zero vector with length of  $M$ . In IEEE802.11ac standards, a compressed version of right-singular matrix  $\mathbf{V}_k$  is fed back to AP side as BFWs. To this end, the right-singular matrix  $\mathbf{V}_k$  is converted to an angle information sequence  $(\phi_{j,i}^k$  and  $\psi_{j,i}^k$  defined in Appendix A) by application of Givens rotation to  $\mathbf{V}_k$  as a linear transformation to create a zero-entry in a matrix [36] as

$$\mathbf{V}_k = \left[ \prod_{i=1}^{\min(S,M-1)} \mathbf{D}_i^k \left[ \prod_{j=i+1}^M (\mathbf{G}_{ij}^k)^T \right] \right] \tilde{\mathbf{I}}_{M \times S}, \quad (1)$$

where  $i$  and  $j$  respectively represent the transmitting and receiving antenna indices. Furthermore,  $\mathbf{D}_i^k$  and  $\mathbf{G}_{ij}^k$  respectively denote the  $M \times M$  diagonal matrix and the Givens rotation matrix [36]. The quantized and compressed CSI  $\hat{\phi}_{j,i}^k$  and  $\hat{\psi}_{j,i}^k$  are fed back from STA to AP to perform downlink beam-forming. The analyses presented herein assume that  $\hat{\phi}_{j,i}^k$  and  $\hat{\psi}_{j,i}^k$  are quantized respectively by 6 and 4 bits. This paper examines a device-free localization problem, i.e., the target person has no

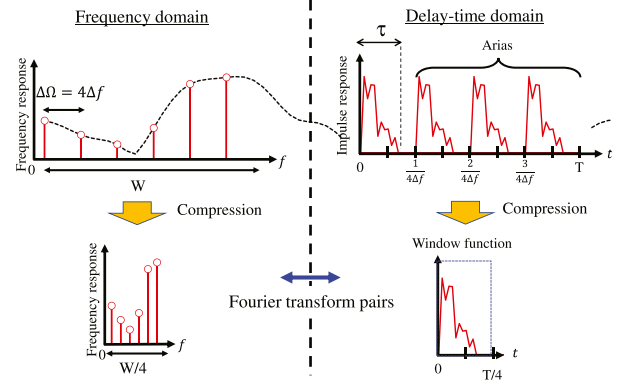


Fig. 3. Equivalence relation of sampling-based compression in the frequency-domain and windowing-based compression in a delayed-time domain.

wireless devices. We consider a multi-user MIMO-based system serving multiple STAs.

#### IV. PROPOSED APPROACH FOR EFFECTIVE LOCALIZATION

This section explains concepts of the proposed localization approach using the concatenated CSI. The CSI capturing terminal in Fig. 1 captures feedback frames from user devices and detects the BFW matrix  $\hat{\mathbf{V}}_k^{(p)}$  at the  $p$ -th time instance at the  $k$ -th subcarrier. Let  $\hat{\mathbf{V}}^{(p)} = [\hat{\mathbf{V}}_1^{(p)}, \dots, \hat{\mathbf{V}}_K^{(p)}]$  signify the  $p$ -th concatenated BFW matrix.

Fig. 2(a) presents the concept of the proposed approach using spatially concatenated CSI, where current and past BFW matrices are concatenated to obtain more accurate feature data. Here, “spatially concatenated CSI” denotes a group of multiple CSI samples obtained when the target person is located at different positions. As this figure shows, when the target object is moved from point (A) to point (B), the channel state between the transmitter and receiver will change in time. Because this paper presents consideration of a device-free scheme that requires no additional functions of user devices for sensing purposes, the CSI capturing terminal analyzes channel fluctuation in a spatial domain by capturing these BFWs sequentially. The concatenated CSI over space and frequency domains is expressed as

$$\hat{\mathcal{V}}^{(p)} = \left[ \hat{\mathbf{V}}^{(p-(U-1))}, \dots, \hat{\mathbf{V}}^{(p)} \right], \quad (2)$$

where  $U$  is defined as the concatenated CSI length, which shows the number of BFWs in each concatenated CSI, and where  $p > U$  is assumed.

To reduce the required complexity, we propose application of a simple frequency-domain sampling scheme to the concatenated BFWs by compressing the total data size. The concept of the frequency domain sampling based compression is presented in Fig. 2(b), where BFWs over frequency-domain  $\mathbf{V}_k$  are selected (sampled) at every  $D_s$  subcarrier before application to an ML block. Thereby, the data can be compressed to  $1/D_s$ . For purposes of explanation in this paper, we define the compression ratio by frequency-domain sampling as  $R_c = 1/D_s$ . Application of the frequency-domain sampling can reduce the total data size of the concatenated CSI while avoiding loss of the

essential channel information. After concatenating current and past  $U - 1$  BFWs as single feature data, they are used for both off-line training and on-line detection.

Fig. 3 presents basic principle of frequency-domain sampling and its equivalence relation to delay-time domain windowing. As shown in this figure, we assume that frequency response  $H(\omega)$  is sampled at every  $\Delta\Omega$  duration. Therefore, the discrete-frequency-domain response is given as  $H(\omega)\delta_D(\omega)$ , where  $\delta_D(\omega) = \sum_n \Delta\Omega\delta(\omega - n\Delta\Omega)$ . Assuming that the Fourier transform pair  $H(\omega) \leftrightarrow h(t)$  is given, the inverse Fourier transform of  $H(\omega)\delta_D(\omega)$  is given as the following.

$$\begin{aligned} h_s(t) &= \mathcal{F}^{-1} \left[ H(\omega) \sum_n \Delta\Omega\delta(\omega - n\Delta\Omega) \right] \\ &= \mathcal{F}^{-1} [H(\omega)] * \mathcal{F}^{-1} \left[ \sum_n \Delta\Omega\delta(\omega - n\Delta\Omega) \right] \\ &= \sum_n h(\omega - n\Delta T), \end{aligned} \quad (3)$$

Therein,  $\Delta T = \frac{2\pi}{\Delta\Omega} = \frac{1}{D_s\Delta f}$ . Also,  $D_s\Delta f$  and  $\mathcal{F}^{-1}[\cdot]$  respectively denote the sampling interval in the frequency-domain and inverse Fourier transform operation. As presented in (3) and Fig. 3, the inverse Fourier transform of  $H(\omega)\delta_D(\omega)$  is a periodic function with period  $\Delta T$ . Therefore, it is clear that the CSI size can be reduced to  $R_c = 1/D_s$  without loss of feature information if the length of impulse response  $h(t)$  is less than  $\Delta T$ , i.e., the frequency domain sampling rate should meet the following condition:

$$\text{frequency-domain sampling rate} = \frac{1}{D_s\Delta f} \geq \tau. \quad (4)$$

As discussed later in Sect. V, the required execution time is increased with the increase of the sampling rate. Therefore, the sampling rate can be optimized by selecting  $\tau$  as a proper value to minimize the required complexity while avoiding the loss of key feature information in frequency-selective fading channel environments. In addition, because the CSI compression by frequency-domain sampling is equivalent to delay-time domain windowing of impulse response  $h(t)$ , the achievable minimum compression ratio in the frequency-domain is the same as  $1/D_s$ . It is readily apparent that the same compression effect is obtainable in either a time domain or a delay-time domain.<sup>2</sup>

As an example of space-frequency channel matrix and its right-singular matrix (i.e., BFW), three-dimensional plots of  $|h_{k,n,m}|$  and  $|v_{k,n,m}|$  with respect to the subcarrier index  $k$  and the antenna index  $m$  in terms of different user index  $u$  are shown respectively in Fig. 4. These plots show that right-singular matrices (i.e., BFWs) fluctuate depending on the current channel condition ( $\mathbf{H}$ ). Therefore, they are useful as feature information to characterize object behaviors. The  $l_2$ -norm of BFWs over users is normalized (i.e.,  $\|\mathbf{V}\|_{k,m} = 1$ ), unlike the  $l_2$ -norm of the channel matrix over users  $\|\mathbf{H}\|_{k,m}$ , where

<sup>2</sup>The delay-time domain windowing has also been confirmed as effective for reducing the data size, similarly to the confirmation presented in this paper for frequency-domain sampling.

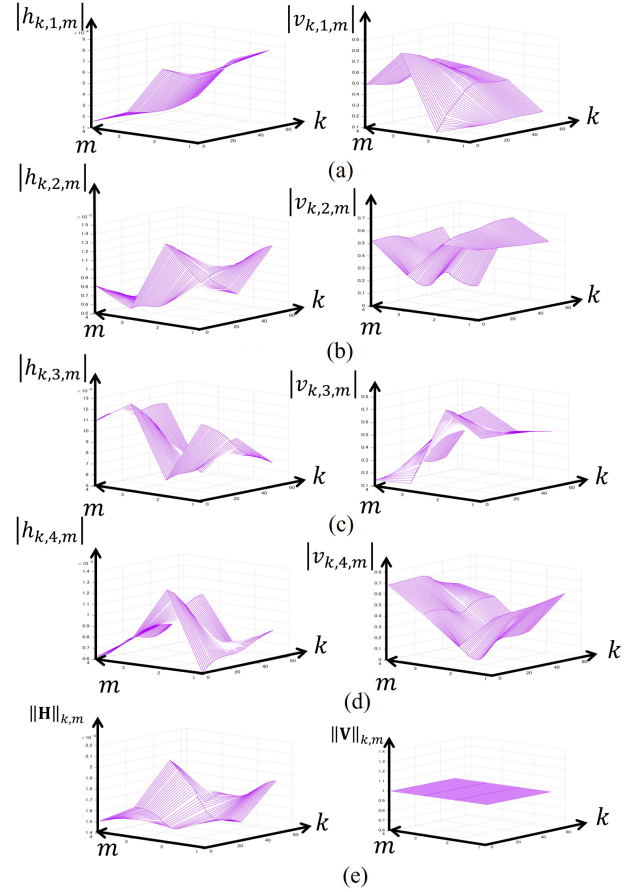


Fig. 4. Example of the space-frequency domain channel matrix  $\mathbf{H}$  and its right-singular matrix (beam-forming weights)  $\mathbf{V}$ , where subscripts  $k$  and  $m$  respectively denote the subcarrier index and transmit antenna index. (a)  $n = 1$  (user 1). (b)  $n = 2$  (user 2). (c)  $n = 3$  (user 3). (d)  $n = 4$  (user 4). (e)  $\|\mathbf{H}\|_{k,m} = \sqrt{\sum_{n=1}^N |h_{k,n,m}|^2}$ ,  $\|\mathbf{V}\|_{k,m} = \sqrt{\sum_{n=1}^N |v_{k,n,m}|^2}$ .

$\|\mathbf{V}\|_{k,m} = \sqrt{\sum_{n=1}^N |v_{k,n,m}|^2}$ . In other words, although antenna and subcarrier-wise channel strength information in  $\mathbf{H}$  is lost in  $\mathbf{V}$ , it is noteworthy that each beam-forming vector is a channel-dependent weighting factor as indirect space/frequency-domain channel state information.

## V. PERFORMANCE EVALUATION

### A. Simulation Setup and Scenario

We conducted a computer simulation to clarify the effectiveness of the proposed schemes. The simulation block diagram is the same as that presented in Fig. 1. Simulation parameters are presented in Table II. To calculate various channel impulse responses (CIRs) in indoor radio propagation environments, a three-dimensional ray launching algorithm is used [37]. The maximum numbers of reflections and the maximum number of diffraction are set respectively to three and one. The other environment setting and physical property values of objects are described in Fig. 5. For this simulation, we assume a device-free type localization, i.e., the target object has no wireless device for sensing. In this context, CSI between AP and the user device

TABLE II  
SIMULATION PARAMETERS

Number of antennas ( $M, N$ )	(4, 1), (4, 4)
Antenna heights at AP and STA [m]	1
Antenna spacing [m]	$d = 1$
Bandwidth [MHz]	20
Center frequency [GHz]	5.18
Number of subcarriers	52
Number of reflections	3
Number of diffractions	1
Machine learning model	Random forest

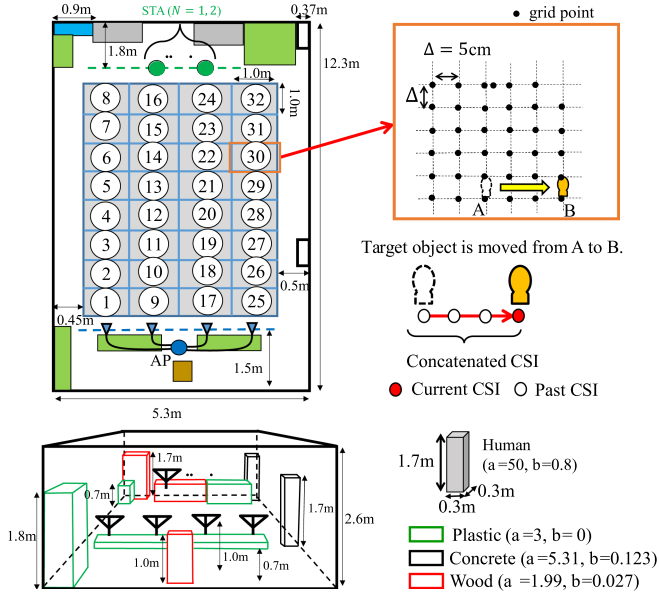


Fig. 5. Ray-trace simulation setting in an indoor environment (room) with AP and STA, where  $\alpha$  and  $\beta$  respectively denote the relative permittivity and electrical conductivity (S/m) of the materials.

is measured when a target object is placed at one of the grids in Fig. 5, where the grid size is set as  $\Delta = 0.05$  m, so that 400 different CIR measurements (OFDM symbols) per area are obtained. Stated more specifically, an indoor field is divided into  $R$  areas labeled as  $r = 0 \cdots R$ . A target object is located in one of the areas. Label numbers  $r = 1 \cdots R$  correspond to the area number in Fig. 5. The label number  $r = 0$  is defined to represent a case in which no target object exists in any area. It is noteworthy that  $r = 0$  corresponds to a case in which no target object exists in any area. We consider a multi-class classification problem to detect the label of the area in which a single target exists. Random forest with four-split cross-validation is used as a supervised machine learning model and its evaluation scheme. We consider an IEEE802.11ac-based WLAN system [36] in which feedback frames containing CSI ( $\phi_{j,i}^k$  and  $\psi_{j,i}^k$  defined in Appendix A) are sent from the user device to the AP<sup>3</sup>.

As a performance metric, we define the probability of position detection as a conditional probability that the detected label

<sup>3</sup>Feedback CSI includes 6 angle values per subcarrier when  $(M, N) = (4, 1)$  ( $\phi_{1,1}^k, \phi_{2,1}^k, \phi_{3,1}^k, \psi_{2,1}^k, \psi_{3,1}^k$ , and  $\psi_{4,1}^k$ ), and includes 12 angle values when  $(M, N) = (4, 4)$  ( $\phi_{1,1}^k, \phi_{2,1}^k, \phi_{3,1}^k, \psi_{2,1}^k, \psi_{3,1}^k, \psi_{4,1}^k, \phi_{2,2}^k, \phi_{3,2}^k, \psi_{3,2}^k, \psi_{4,2}^k, \phi_{3,3}^k$ , and  $\psi_{4,3}^k$ ).

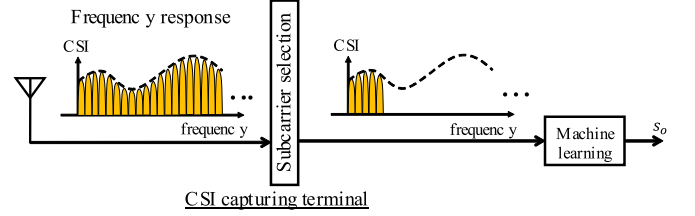


Fig. 6. Sampling in a localized manner (localized sampling).

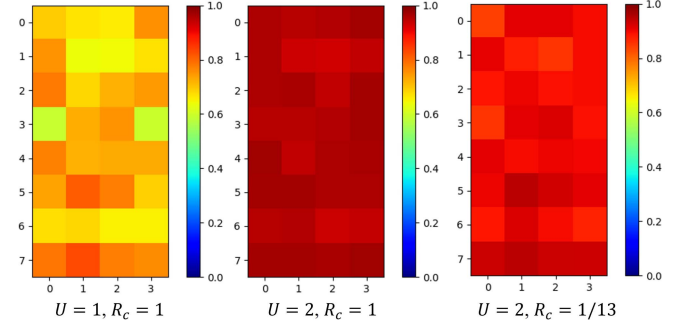


Fig. 7. Heat map of the object detection probability for area classification problems in terms of different  $U$  and  $R_c$ , where  $(M, N)$  is set to  $(4, 1)$ .

number  $s_o$  is matched to the actual one  $s_a$ , which is

$$\text{Prob}(s_o = r | s_a = r), \quad (5)$$

where  $r \in \{0, 1, \dots, R\}$ .

For comparison, we consider another subcarrier selection scheme, designated as ‘‘localized sampling (LS)’’. The sampling in a localized manner is presented in Fig. 6, where the signal after sampling includes only a portion of the overall frequency response (i.e.,  $L$  adjacent subcarriers), unlike the proposed approach. Here,  $L = \lfloor \frac{K}{D_s} \rfloor$ , where  $\lfloor \times \rfloor$  is a floor function that outputs a maximum integer value that is less than or equal to  $x$ .

## B. Simulation Results

Fig. 7 presents heatmaps of the average object detection probability in Eq. (5) for different conditions ( $U = 1, 2$  and  $R_c = 1, 1/13$ ), where  $(M, N) = (4, 1)$  is assumed. Here,  $U = 1$  corresponds to the conventional scheme in [21]. These results indicate that the proposed scheme ( $U = 2$ ) achieves higher average detection performance at most areas than the conventional scheme ( $U = 1$ ), even when frequency-domain sampling ( $R_c = 1/13$ ) is applied. This finding implies that frequency-domain sampling is effective at reducing the required data size (i.e., computational complexity) while achieving detection performance comparable to that of the case of  $R_c = 1$ .

Fig. 8 shows the average detection probability of the proposed scheme as a function of compression ratio  $R_c$  in terms of the concatenated CSI length  $U$ , where the number of antennas at AP is  $M = 4$ . For comparison, the average detection probability of a case with localized sampling is also shown. The label ‘‘Interleaved’’ denotes a case with frequency-domain sampling used in the proposed scheme. It is apparent from this figure that higher detection performance is obtained as  $U$  increases. Additionally,

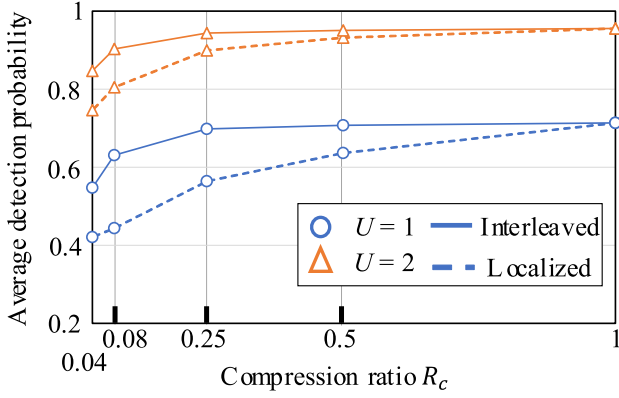


Fig. 8. Average detection probability of the proposed scheme as a function of compression ratio  $R_c$  in case with  $(M, N) = (4, 1)$ , where  $U$  denotes the spatially concatenated CSI length and where  $U = 1$  corresponds to a conventional scheme [21].

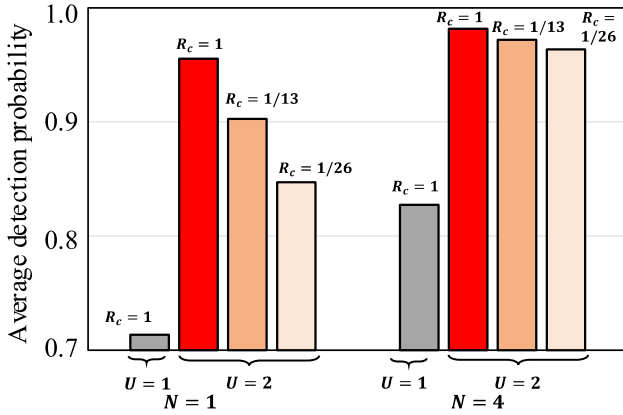


Fig. 9. Average detection probability of the proposed scheme in terms of  $R_c$  and  $N$  in MU-MIMO scenario, where  $(M, N)$  is set to  $(4, 1)$  and  $(4, 4)$ .  $U = 1, 2$  is assumed.

we can confirm that “interleaved” sampling achieves higher detection performance in a lower compression ratio region because the overall spectrum information is not contained in CSI after localized sampling, unlike the interleaved one in the proposed scheme.

Fig. 9 shows the average detection probability of the proposed scheme in terms of the number of users  $N$  and the concatenated CSI length  $U$  in the MU-MIMO scenario, where the compression ratio is set as  $R_c = 1, 1/13$ , and  $1/26$ , respectively, where  $N = 1, 4$  and  $U = 1, 2$  are assumed. It is apparent from this figure that the proposed scheme with  $U = 2$  achieves better detection performance than the conventional scheme ( $U = 1$ ), even when the compression ratio is  $R_c = 1/26$ . Similar results are also obtained in cases for which  $U = 4$ . Particularly, one can find that the proposed scheme achieves good detection performance comparable to that obtained in the case without sampling ( $R = 1$ ), even when the compression ratio of  $R = 1/26$  is adopted.

### C. Experiment Setup and Scenario

For an experimentation-based demonstration of the performance achieved using the proposed scheme, we implemented

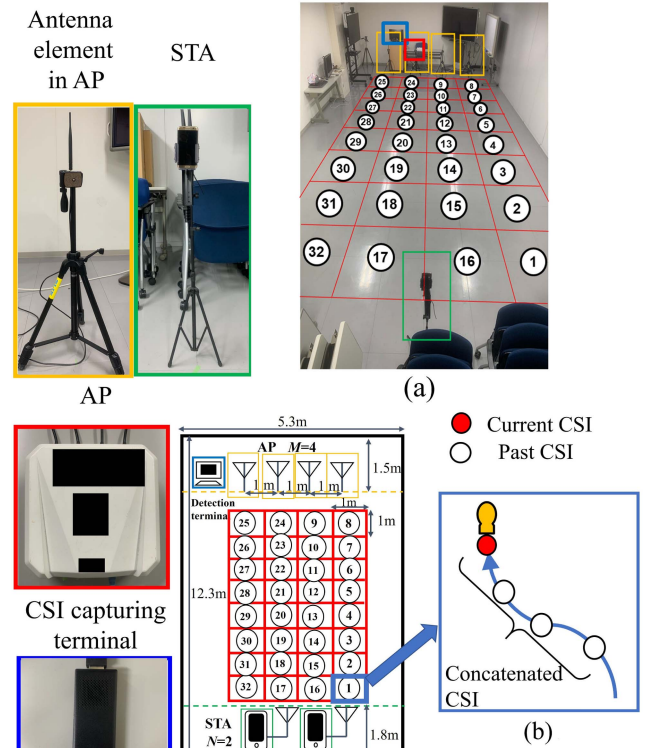


Fig. 10. Experiment setting of an indoor environment (a room) with an AP, a CSI capturing terminal, and two stations (STA). (a) Experimental setup. (b) Acquisition of concatenated CSI in experiment.

TABLE III  
EXPERIMENT SETUP

Number of antennas at AP	$M = 4$
Number of antennas at STA	$N = 2$
Antenna height [m]	1
Antenna spacing [m]	$d = 1$
Bandwidth [MHz]	20
Center frequency [GHz]	5.18
Number of subcarriers	52
Machine learning model	Random forest

our designed algorithm to an IEEE802.11ac-based system as shown in Fig. 1 and conducted experiments in an indoor environment. Fig. 10 shows the experiment setup and an indoor environment with an AP, a CSI capturing terminal, and user devices (STA), where IEEE802.11ac-based AP and STA are placed on both sides of the room. As the figure shows, four antennas and a single antenna are equipped respectively with AP and STA. Some system parameters used for experimentation are presented in Table III. The AP equips a linearly placed antenna array, where the distance between adjacent antenna elements can be extended using coaxial extension cables. The user device (Galaxy S7 edge; Samsung Electronics) is fixed with the tripods as shown in Fig. 10. It regularly sends feedback frames to the AP. At the CSI capturing terminal, the required functions are implemented on a stick-type PC (Compute Stick STK2M364CC; Intel Corp.) to capture the feedback frames and to build the measured CSI database. Random forest is used as a supervised

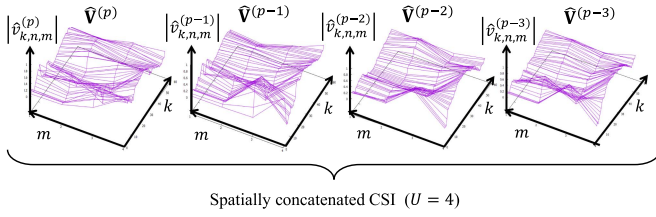


Fig. 11. Example of measured beam forming weights included in a concatenated CSI.

machine learning algorithm. For experiments, off-line training is conducted to construct the model beforehand, whereas on-line object detection is conducted in real time. In offline training processing, the person is moving within area- $i$  while the object detection terminal captures (overhears) feedback frames from STAs to BS and extracts the CSI. After applying preprocessing and frequency-domain sampling to the extracted CSI, they are labeled as “ $i$ ”. The person moves from area 1 to 32 so that CSI dataset for labels 1–32 are constructed. The machine learning model is then trained by the constructed dataset. This device-free detection scheme obviates the need for user devices to have any dedicated function for sensing purposes. In addition, no need exists for a target object to have any wireless device.

Similarly to the simulation scenario, we consider a multi-classification problem for detecting a single object position<sup>4</sup> (i.e., the corresponding label number), where the indoor area is divided into  $R = 32$  regions labeled as  $r = 0, \dots, R$ . A human is placed at one of the areas as a target object.<sup>5</sup> It is noteworthy that  $R = 0$  corresponds to a case in which the object does not exist. To evaluate the detection accuracy that we achieved, we define the average position detection probability of the target object as in (5). For comparison, we also evaluate the detection performance in a case with the localized sampling scheme. As illustrated in Sect. IV, frequency-domain sampling rate is set to  $\frac{1}{D_s \Delta f}$ , where  $\Delta f = 3.8 \times 10^5$  Hz.

#### D. Experiment Results

Fig. 11 presents an example of measured space-frequency transmitting antenna weights for concatenated BFWs, where the number of transmitting antennas  $M = 4$ , the number of streams per user  $S = 1$ , and the number of subcarriers  $K = 52$  are used. The weight coefficients (BFWs) in this figure are calculated from measured CSI samples. The results imply that the concatenated BFWs exploit spatial channel fluctuation information for more accurate characterization of object behaviors.

Fig. 12 presents the average object detection probabilities of the proposed scheme for different compression ratios  $R_c$  in terms of the number of concatenated BFWs  $U$  and the number of users

<sup>4</sup>This paper presents consideration of a multi-area classification problem for single target detection. Based on that principle, it can be extended to a multi-area classification problem for multiple target detection at the cost of required complexity if the corresponding dataset for multiple targets is available. As for this problem, a theoretical framework for a multiple target detection has been presented recently in [38]. Extension to multiple target detection is left as a subject for our future study.

<sup>5</sup>We confirmed that similar experimentally obtained results are obtained even when a different person is the target.

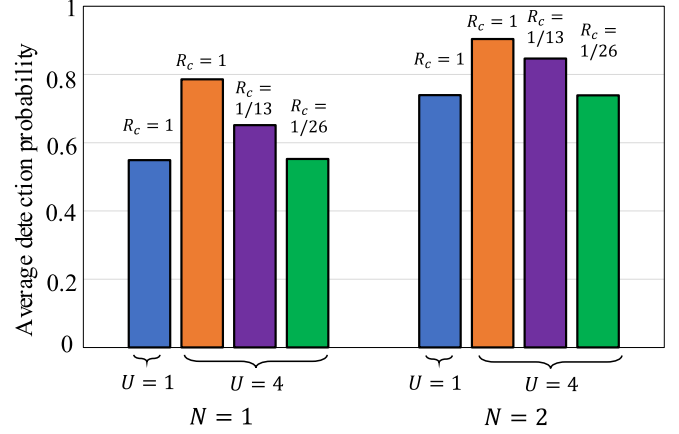


Fig. 12. Object detection probability of MU-MIMO system using the proposed scheme  $((M, N) = (4, 1)$  and  $(4, 2)$  in terms of  $U$  and  $R_c$ , where  $U = 1, 4$  and  $R_c = 1, 1/13, 1/26$ .

$N$ , where  $U = 1, 4$  and  $(M, N) = (4, 1)$  and  $(4, 2)$ . One can find that the detection performance is improved by analyzing the CSI samples from multiple user devices. Similarly to simulation results presented in Fig. 9, one can observe that the proposed scheme with the concatenated CSI and frequency-domain sampling achieves better detection performance than in the case without concatenating CSI ( $U = 1$ ), even when  $R_c$  takes a small value.

Fig. 13 shows the area-wise detection probability of the proposed scheme in the cases of  $U = 1$  and 4 for three experiments (experiment 1, experiment 2, and experiment 3), where  $R = 32$  and  $(M, N) = (4, 1)$  and  $(4, 2)$ . The two experimentally obtained data from experiment 1 and experiment 2 are measured in the same room on different days. Consequently, experiment setups in both experiments are the same, but the CSI dataset differs. All experiments were conducted for the same scenario. Although the overall performance of experiment 1 was found to be better than that of experiment 2, the results clarify that the proposed scheme with  $U = 4$  achieves much better performance than the conventional one ( $U = 1$ ) [21]. The results also indicate that the overall detection performance can be improved by increasing the number of user devices  $N$  in MU-MIMO systems because the object behavior is characterized more accurately by collecting various CSI samples from different user devices. Hereinafter, results in experiment 2 are used for additional discussions.

Fig. 14 presents confusion matrices for classification results in cases with the proposed scheme ( $U = 4$ ) and the conventional scheme ( $U = 1$ ). In Figs. 14(b) and 14(c), a compression ratio of  $R_c = 1/13$  is adopted for the proposed ( $U = 4$ ) and the conventional case ( $U = 1$ ). For these figures, the probability of detection is drawn as a heatmap, where the horizontal and vertical axes show  $s_o$  and  $s_a$  in (5). The results indicate that more accurate detection is possible using the concatenated CSI ( $U = 4$ ), even when the CSI data are compressed as  $R_c = 1/13$ .

Fig. 15 shows the average detection probabilities of the proposed scheme as a function of compression ratio  $R_c = \frac{1}{D_s}$  in terms of  $U$ , where  $U = 1, 2, 3$ , and 4 are used. For simplicity



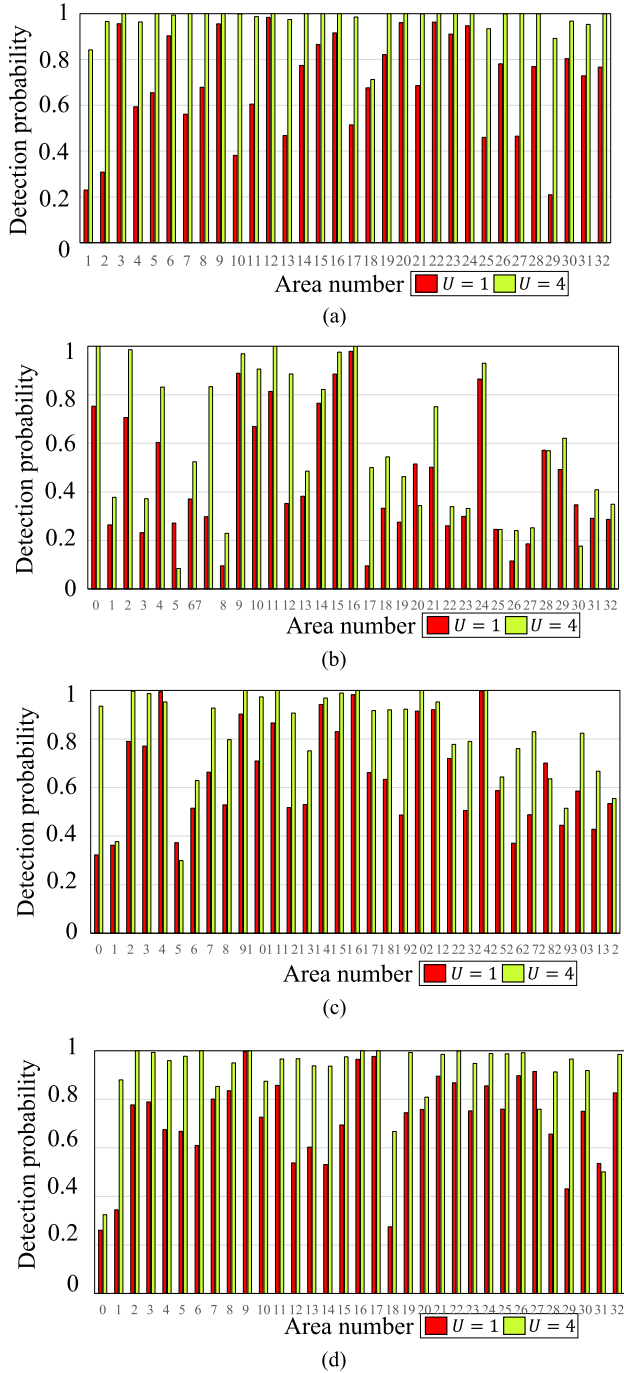


Fig. 13. Area-wise detection probability for the case of  $U = 1$  and 4 for two experiment trials (experiment 1 and experiment 2), where the number of antennas at AP is  $M = 4$ . (a)  $N = 1$  (Experiment 1). (b)  $N = 1$  (Experiment 2). (c)  $N = 2$  (Experiment 2). (d)  $N = 2$  (Experiment 3).

of discussions,  $D_s$  is selected among the divisors of 52 carriers (i.e., 1, 2, 4, 13, 26, and 52). Solid and dotted lines respectively represent cases of interleaved sampling and localized sampling. The results indicate that the proposed scheme with the concatenated CSI and frequency-domain sampling achieves better detection performance than in the case with localized sampling. They also indicate that the data size can be compressed by choosing a proper compression ratio for a given  $U$  and  $N$ . The

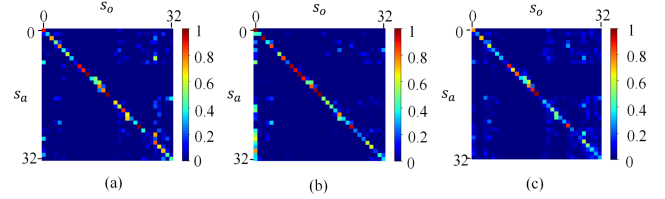


Fig. 14. Confusion matrix of the classification problem in the proposed scheme ( $U = 4$ ) and the conventional scheme ( $U = 1$ ), where the number of antennas at AP is  $M = 4$  and the number of stations is  $N = 1$ .

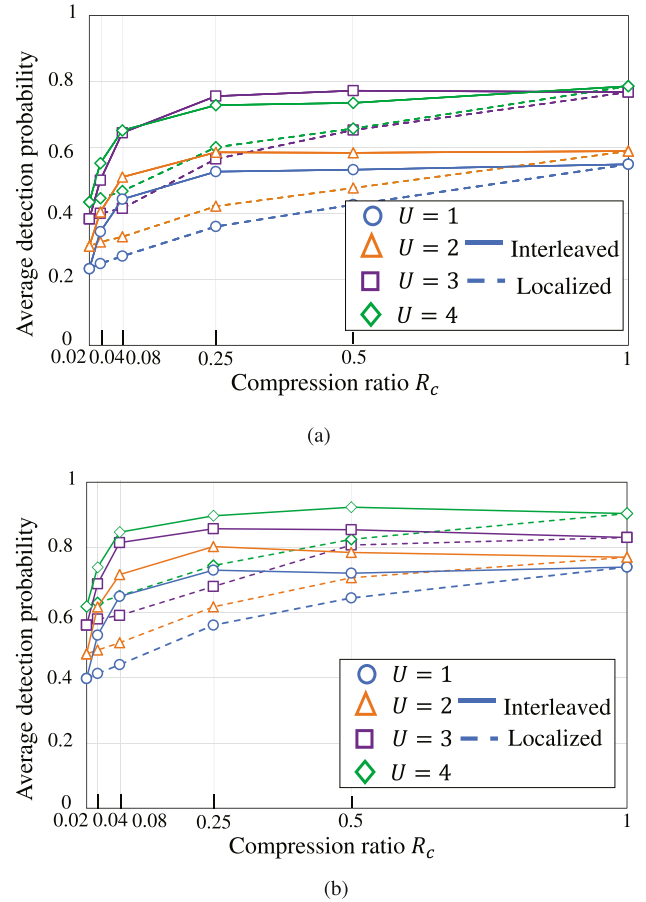


Fig. 15. Average detection probabilities as a function of compression ratio  $R_c$  for  $N = 1$  and 2 in an indoor experiment scenario, where the number of antennas at AP is  $M = 4$ . (a)  $N = 1$ . (b)  $N = 2$ .

results also clarify that almost the same detection probability is achieved when  $R_c$  is higher than 0.25 (i.e.,  $D_s \leq 4$ ). This result indicates that  $R_c = 0.25$  ( $D_s = 4$ ) is the near optimum value that approximately meets the condition in (4)<sup>6</sup>.

Fig. 16 shows the average detection probabilities of the proposed schemes in terms of different antenna spacing at AP, where  $N = 1$  and 2. The orange and blue bars respectively correspond to the cases of  $U = 1$  and 4. Results indicate that the achieved detection performance tends to be improved as the antenna spacing at the AP side increases. It is also apparent that the achieved performance is improved by collecting CSI from

<sup>6</sup>We confirmed that frequency domain sampling rate approximately satisfies the condition in (4), i.e.,  $\frac{1}{D_s \Delta f} \geq \tau$  when  $R_c$  is higher than 0.25.

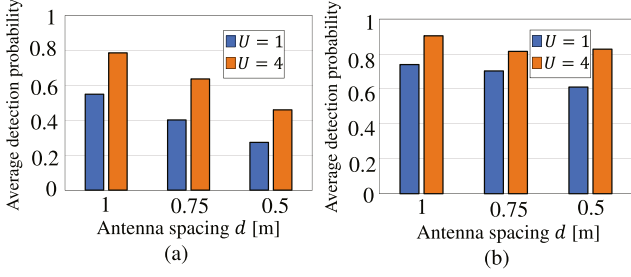


Fig. 16. Average detection probabilities for different antenna spacing  $d$  in an indoor experiment scenario, where the number of antennas at AP is  $M = 4$ . (a)  $N = 1$  (b)  $N = 2$ .

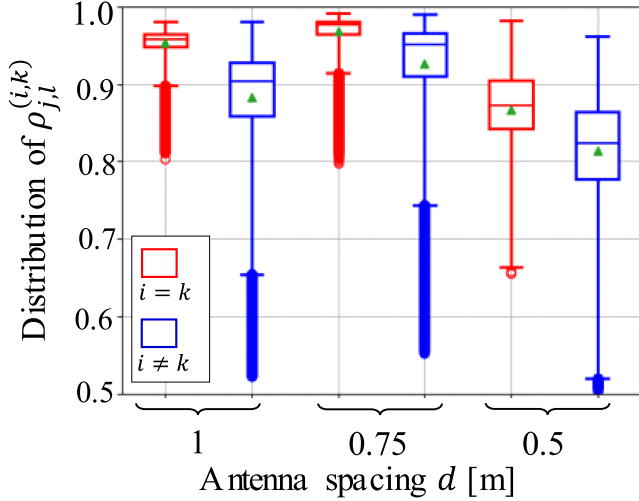


Fig. 17. Correlation between the CSI samples for offline training and those for online detection in terms of antenna spacing  $d$ , where  $U = 4$  and  $N = 1$ .

multiple user devices irrespective of the antenna spacing  $d$ . To analyze the results further, we define the correlation between CSI samples for off-line training and those for online detection as

$$\rho_{j,l}^{(i,k)} = \frac{\mathbf{x}_j^{(i)} (\mathbf{y}_l^{(k)})^T}{\|\mathbf{x}_j^{(i)}\| \|\mathbf{y}_l^{(k)}\|} \quad (6)$$

where  $\mathbf{x}_j^{(i)} = [x_{j1}^{(i)}, \dots, x_{jQ}^{(i)}] \in \mathbb{R}^{1 \times Q}$  denotes the  $j$ -th training CSI labeled to the  $i$ -th area. Here,  $Q$  denotes the number of features per CSI sample.  $\mathbf{y}_l^{(k)} = [y_{l1}^{(k)}, \dots, y_{lQ}^{(k)}] \in \mathbb{R}^{1 \times Q}$  denotes the  $l$ -th CSI for online detection when the target exists in the  $k$ -th area. Here,  $\rho_{j,l}^{(i,i)}$  represents the correlation between CSI samples collected when the target is located in the  $i$ -th area for offline training and online detection. Fig. 17 shows correlation between the CSI samples for offline training and those for online detection in antenna spacing  $d$  where  $U = 4$  is used. Here, “ $i = k$ ” (red box) and “ $i \neq k$ ” (blue box) denote the correlation between the same label numbers (i.e., correct label) and that between different label numbers (i.e., incorrect label). Results presented in Figs. 16 and 17 indicate that, as the antenna spacing becomes wider,  $\rho_{j,l}^{(i,i)}$  ( $i = k$ ) takes higher values than

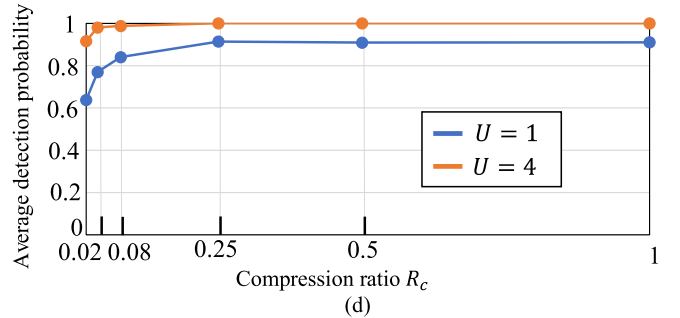
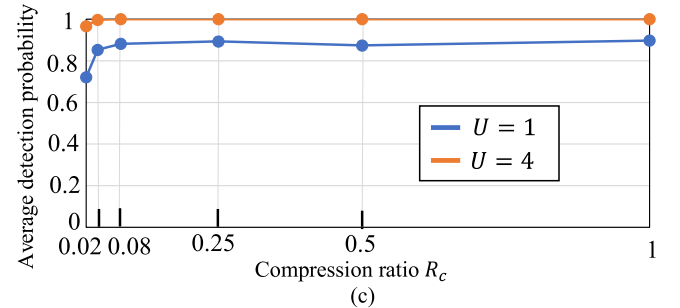
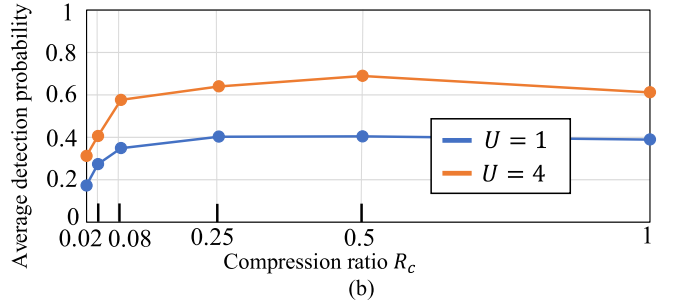
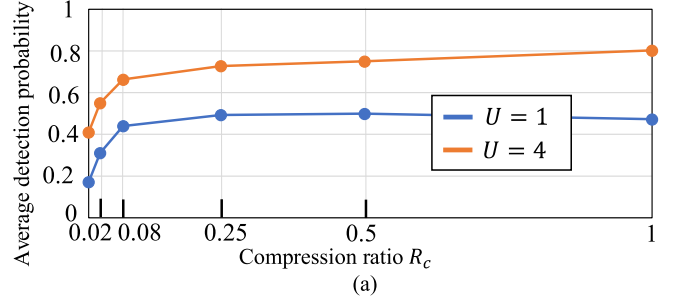


Fig. 18. Average detection probabilities for different targets (either person A or person B) in the cases of  $U = 1$  and 4, where machine learning model is trained for person A. (a)  $N = 1$ , Detection target: Person A. (b)  $N = 1$ , Detection target: Person B. (c)  $N = 4$ , Detection target: Person A. (d)  $N = 4$ , Detection target: Person B.

$\rho_{j,l}^{(i,k)}$ ,  $i \neq k$ . Consequently, the detection performance improves as the antenna spacing widens.<sup>7</sup>

In the evaluations described above, CSI samples for both off-line training and online detection are collected from the same person. To clarify the detection performance when the trained model is applied for different persons, i.e., when CSI for online detection is collected from different persons, the average

<sup>7</sup>We evaluated the average detection performance for different antenna heights and confirmed that the detection performance is not always improved when all antenna heights are uniformly changed because the relative relation among antennas is not changed.

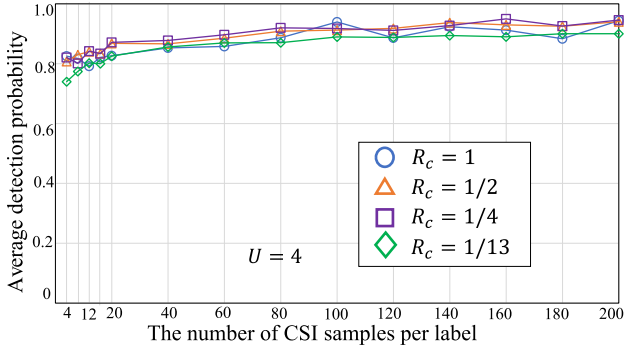


Fig. 19. Average detection probability as a function of the number of CSI samplers per label ( $N = 1$ ) in terms of compression ratio  $R_c$ , where  $U = 4$ .

TABLE IV

EXECUTION TIME FOR OFF-LINE TRAINING [s] ( $M, N$ ) = (4, 1)

model	$U = 1$		$U = 4$	
	$R_c = 1$	$R_c = 1/4$	$R_c = 1$	$R_c = 1/4$
RF	3.38	1.62	1.88	0.781
LR	25.3	6.69	15.9	4.06
SVM	14.4	4.08	4.96	1.24
LSVM	7.11	2.64	9.24	2.05
DT	1.07	0.280	1.53	0.335
KN	0.323	0.101	0.281	0.0833
GNB	0.193	0.0790	0.185	0.0676

TABLE V

EXECUTION TIMES FOR OFF-LINE TRAINING [s] ( $M, N$ ) = (4, 2)

model	$U = 1$		$U = 4$	
	$R_c = 1$	$R_c = 1/13$	$R_c = 1$	$R_c = 1/13$
RF	4.72	1.22	3.16	0.603
LR	42.2	3.79	31.4	2.35
SVM	25.5	2.69	10.8	0.852
LSVM	12.0	2.00	20.4	1.31
DT	2.28	0.164	3.25	0.208
KN	0.642	0.0694	0.559	0.0555
GNB	0.376	0.0546	0.360	0.0458

detection probabilities of the proposed schemes for different targets (persons A and B) are shown in Fig. 18. Here, the model is trained using CSI collected for person A (170 cm height, 50 kg weight) and applied to on-line detection of person B (174 cm height, 58 kg weight). From these results, it is readily apparent that the proposed method achieves similar detection performance for both persons A and B. This result implies that online detection performs well if CSI is collected for similar targets, even when the model is trained using CSI for different persons.

Fig. 19 presents the average detection performance as a function of the number of training CSI samples in cases of  $U = 4$ . Here, each CSI sample includes feature information corresponding to beam-forming matrices for all subcarriers. Results show that the proposed method achieves similar average detection performance, even when the number of CSI samples per label is reduced to around 20. The results reveal that the proposed method works well with a small dataset.

To clarify the implications of channel aging, we evaluate the average detection performance as a function of elapsed time after

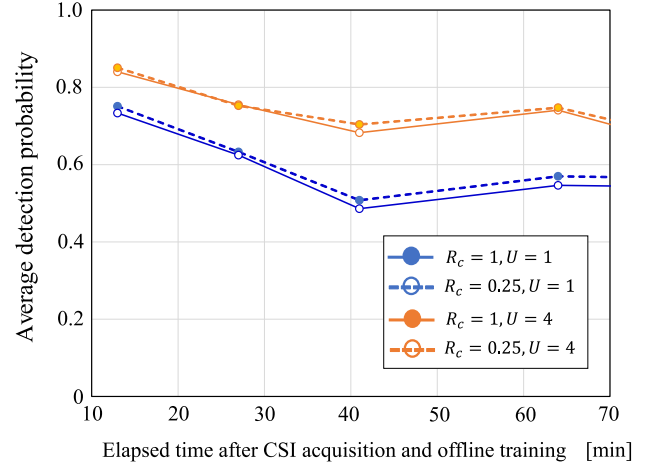


Fig. 20. Effects of channel aging on average detection performance in terms of concatenation size  $U$  and compression ratio  $R_c$  where  $N = 1$ .

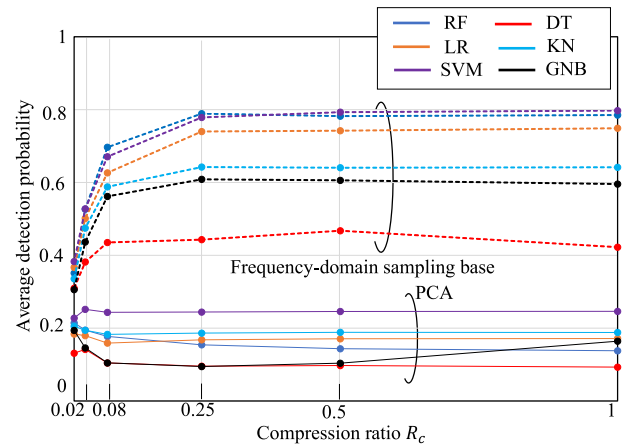


Fig. 21. Average detection probability of the frequency-domain sampling-based method and the PCA-based method as a function of the compression ratio in cases with various machine learning models where  $N = 1$ .

CSI acquisition and training, as depicted in Fig. 20. Results show that the detection performance is degraded gradually because of channel aging. The environmental conditions might change over time even if the state inside the room remains unchanged. However, it is also apparent that the proposed method using the concatenated CSI ( $U = 4$ ) achieves higher detection performance than for the case of  $U = 1$ .

In a report of an earlier study [23], a CSI-based device-free human detection method using principal component analysis (PCA) was examined. For that method, the CSI samples are compressed with PCA and are used for training a deep-learning model. To clarify its effectiveness further, we applied this method to the same machine learning models with a small dataset and compared them with the proposed method. Fig. 21 shows the average detection performance of the frequency-domain sampling-based method and the PCA-based method as a function of the compression ratio in cases with various machine learning models. Here, we consider six typical ML models: Random Forest (RF), decision tree (DT) [39], logistic regression (LR) [40], support vector machine (SVM), K-nearest neighbors

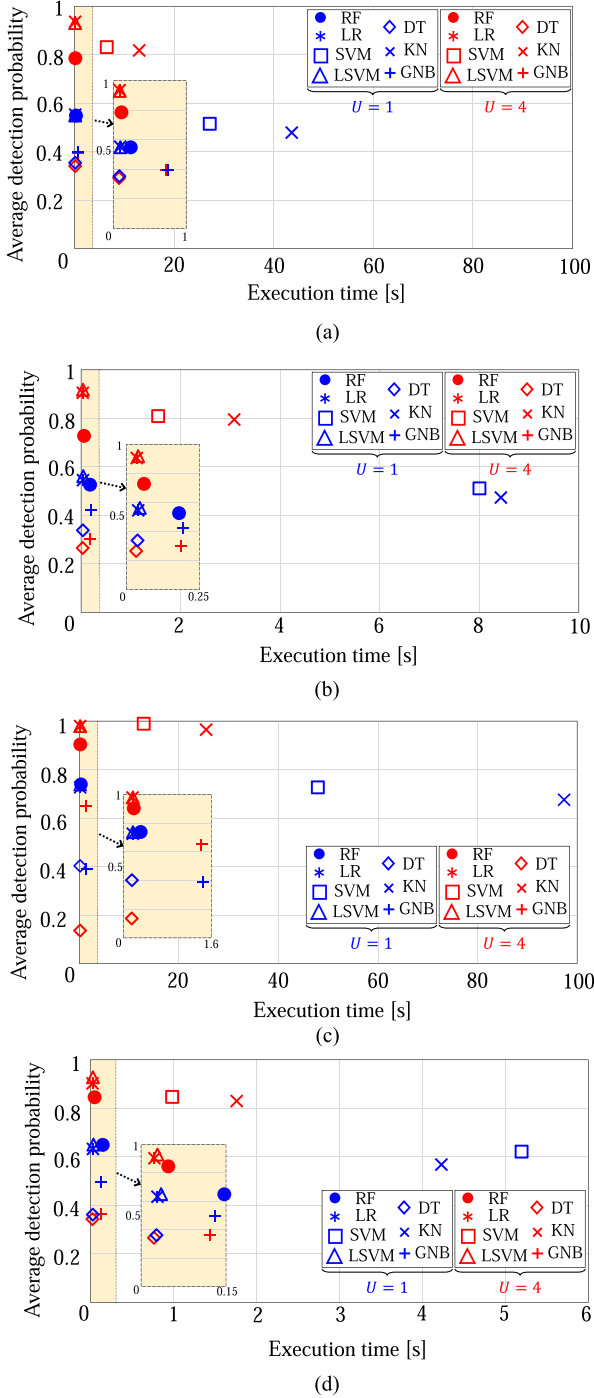


Fig. 22. Experimentally obtained results for a time-complexity and detection probability tradeoff, where the number of antennas at AP is  $M = 4$ . (a)  $N = 1$ ,  $R_c = 1$ . (b)  $N = 1$ ,  $R_c = 1/4$ . (c)  $N = 2$ ,  $R_c = 1$ . (d)  $N = 2$ ,  $R_c = 1/13$ .

(KN) [42], and Gaussian naive Bayes (GNB) [43]). Results show that the proposed method with frequency-domain sampling achieves higher detection performance than the PCA base. This is true because the frequency-domain sampling method reduces the CSI size better without losing key feature information in the case of a small dataset.

To assess the relation between the detection accuracy and the necessary complexity, we evaluate the respective execution

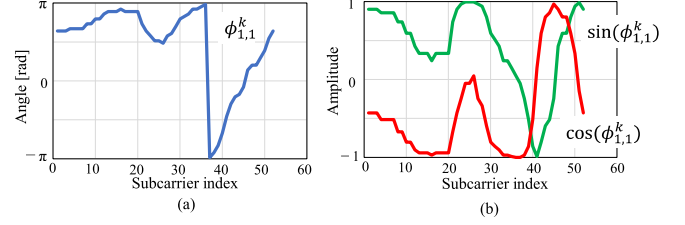


Fig. 23. Illustration of pre-processing of  $\phi_{1,1}^k$ . (a) An example of angle information  $\phi_{1,1}^k$ . (b) Trigonometric function expressions of  $\phi_{1,1}^k$ .

times found for off-line training and on-line detection. Tables IV and V respectively present the required execution times for off-line training of various ML models when  $N = 1$  and 2. Here, similar to Fig. 21, we consider seven typical ML models: RF, DT, LR, SVM, linear SVM (LSVM) [41], KN, and GNB. Execution times were measured using the same workstation (64 GB memory, Core i7-10750H CPU; Intel Corp.). The results presented in these tables clarify that a shorter execution time can be achieved by adopting frequency-domain sampling ( $R < 1$ ). It is also noteworthy that the proposed scheme ( $U = 2$ ) achieves a shorter execution time with higher detection performance than the conventional scheme with  $U = 1$ . This finding implies that concatenating multiple CSI is effective not only for precise detection, but also for accelerating the ML model training when the total amount of BFWs is the same.

The relation between the detection performance and execution time for on-line target detection is presented in Fig. 22, where  $(M, N)$  is set to  $(4, 1)$  and  $(4, 2)$ . Here, the execution time is defined as the total calculation time to obtain 7128 ML results. The yellow separated area in the figure represents an enlarged view of the yellow part of the figure. Similarly to the off-line training case, the results show that shorter required execution time can be achieved by adopting frequency-domain sampling with a compression ratio of  $R_c$ . It is noteworthy that the proposed scheme using concatenated BFWs ( $U = 4$ ) achieves not only better detection performance, but also lower execution time than the conventional scheme ( $U = 1$ ) in all cases. This is true because the available number of features (i.e., concatenated BFWs) in  $U = 4$  is one-fourth of the case with  $U = 1$ , which leads to a shorter execution time. However, as the number of user devices (STAs)  $N$  increases, the execution time increases because the total amount of CSI samples is increased by a factor of  $N$ .

## VI. CONCLUSION

As described herein, after developing a lightweight localization scheme with a small dataset, we used an indoor experiment and ray-tracing based simulation to demonstrate that it works in real time. The proposed approach uses concatenated BFWs as spatial feature information and for both training an ML model and for detecting the existence of a target object and its positions. Through both simulation and experimentally obtained results, we have obtained findings indicating the proposed approach as effective at enhancing the object position detection probability while reducing its inherent complexity where a supervised ML

model is adopted. Practical experiments also indicate that localization performance can be improved by applying the proposed schemes to an MU-MIMO system with a distributed antenna array and number of users. The concept is applicable to any wireless communication system, including single carrier systems, if the acquisition of CSI between transceivers is possible. Application of the developed algorithms with a more powerful ML model under more various scenarios such as outdoor environments is left as a subject for our future work.

#### APPENDIX A

##### PRE-PROCESSING OF ANGLE INFORMATION $\phi_{j,i}^k$

$\mathbf{D}_i^k$  and  $\mathbf{G}_{ij}^k$  in (1) are given respectively as

$$\mathbf{D}_i^k = \begin{bmatrix} \mathbf{I}_{i-1} & 0 & \cdots & \cdots & 0 \\ 0 & e^{j\phi_{j,i}^k} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & 0 & e^{j\phi_{M-1,i}^k} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{G}_{ij}^k = \begin{bmatrix} \mathbf{I}_{i-1} & 0 & 0 & 0 & 0 \\ 0 & \cos(\psi_{j,i}^k) & 0 & \sin(\psi_{j,i}^k) & 0 \\ 0 & 0 & \mathbf{I}_{j-i-1} & 0 & 0 \\ 0 & -\sin(\psi_{j,i}^k) & 0 & \cos(\psi_{j,i}^k) & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I}_{M-1} \end{bmatrix}, \quad (7)$$

where  $\mathbf{I}_x$  represents an  $x \times x$  identity matrix.  $\tilde{\mathbf{I}}_{M \times S} = [\mathbf{I}_S \mathbf{0}_{N, M-S}]^T$  is an expanded identity matrix in which extra elements take a value of zero if  $M \neq S$ . Here,  $\mathbf{0}_{a,b}$  is an  $a \times b$  zero matrix. The angle information  $\phi_{j,i}^k$  and  $\psi_{j,i}^k$ <sup>8</sup> are fed back to AP, where  $-\pi \leq \phi_{j,i}^k \leq \pi$  and  $0 \leq \psi_{j,i}^k \leq \pi/2$ . Fig. 23(a) presents an example of angle information  $\phi_{j,i}^k$ . It takes discontinuous values (significant change) at the boundary of  $\pi$  irrespective of the object status, which might degrade its object detection accuracy. To mitigate this negative effect, this paper adopts pre-processing that transforms the angle information to trigonometric functions  $\sin(\phi_{j,i}^k)$  and  $\cos(\phi_{j,i}^k)$ , as shown in Fig. 23(b). This difficulty can also be resolved by decompressing  $\phi_{j,i}^k$  and  $\psi_{j,i}^k$  to  $\mathbf{V}_k$ . Because transformation to trigonometric functions is simpler than decompression to  $\mathbf{V}_k$ , this paper adopts the pre-processing explained above<sup>9</sup>.

#### REFERENCES

- [1] NTT DOCOMO 6G White Paper Ver. 4.0, "5G evolution and 6G," Jan. 2022. [Online]. Available: [https://www.docomo.ne.jp/english/binary/pdf/corporate/technology/whitepaper\\_6g/DOCOMO\\_6G\\_White\\_PaperEN\\_v4.0.pdf](https://www.docomo.ne.jp/english/binary/pdf/corporate/technology/whitepaper_6g/DOCOMO_6G_White_PaperEN_v4.0.pdf)
- [2] I.F. Akyildiz, A. Kak, and S. Nie, "6G and beyond: The future of wireless communications systems," *IEEE Access*, vol. 8, pp. 133995–134030, 2020.
- [3] L. Leyva, D. Castanheira, A. Silva, A. Gameiro, and L. Hanzo, "Cooperative multiterminal radar and communication," *IEEE Veh. Technol. Mag.*, vol. 16, no. 4, pp. 38–47, Dec. 2021.
- [4] Z. Chaloupka, "Technology and standardization gaps for high accuracy positioning in 5G," *IEEE Commun. Standards Mag.*, vol. 1, no. 1, pp. 59–65, Mar. 2017.
- [5] K. J. Ray and B. Wang, *Wireless AI*. Cambridge, U.K., Cambridge Univ. Press, 2019.
- [6] A. Khalajmehrabadi, N. Gatsis, and D. Akopian, "Modern WLAN fingerprinting indoor positioning methods and deployment challenges," *IEEE Commun. Surv. Tut.*, vol. 19, no. 3, pp. 1974–2002, Jul.–Sep. 2017.
- [7] Y. Ma, G. Zhou, and S. N. Wang, "WiFi sensing with channel state information: A survey," *ACM Comput. Surv.*, vol. 52, no. 3, pp. 1–36, Jun. 2019.
- [8] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: Indoor location sensing using active RFID," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2003, pp. 407–415.
- [9] U. Bandara, M. Hasegawa, M. Inoue, H. Morikawa, and T. Aoyama, "Design and implementation of a bluetooth signal strength based location sensing system," in *Proc. IEEE Radio Wireless Conf.*, 2004, pp. 319–322.
- [10] B. Mrazovac, M. Z. Bjelica, D. Kukulj, B. M. Todorovic, and D. Samardžija, "A human detection method for residential smart energy systems based on ZigBee RSSI changes," *IEEE Trans. Consum. Elect.*, vol. 58, no. 3, pp. 819–824, Aug. 2012.
- [11] Q. Gao, J. Tong, J. Wang, Z. Ran, and M. Pan, "Device-free multi-person respiration monitoring using WiFi," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14083–14087, Nov. 2020.
- [12] M. Yamamoto and T. Otsuki, "Power-absorption-based model for power compensation in WLAN positioning using smartphone," *IEICE Trans. Commun.* vol. 98, no. 6, pp. 1125–1132, Jun. 2015.
- [13] K. Wang et al., "Learning to improve WLAN indoor positioning accuracy based on DBSCAN-KRF algorithm from RSS fingerprint data," *IEEE Access*, vol. 7, pp. 72308–72315, 2019.
- [14] Y. Duan et al., "Data rate fingerprinting: A WLAN-Based indoor positioning technique for passive localization," *IEEE Sensors J.*, vol. 19, no. 15, pp. 6517–6529, Aug. 2019.
- [15] S. S. Moosavi and P. Fortier, "Fingerprinting localization method based on clustering and gaussian process regression in distributed massive MIMO systems," in *Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun.*, 2020, pp. 1–7.
- [16] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, 2014, pp. 617–628.
- [17] X. Zheng, J. Wang, L. Shangguan, Z. Zhou, and Y. Liu, "Smokey: Ubiquitous smoking detection with commercial WiFi infrastructures," in *Proc. IEEE 35th Annu. Int. Conf. Comput. Commun.*, 2016, pp. 1–9.
- [18] H. Zhu, F. Xiao, L. Sun, R. Wang, and P. Yang, "R-TTWD: Robust device-free through-the-wall detection of moving human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1090–1103, May 2017.
- [19] J. M. B. Rocamora, I. W.-H. Ho, and M. Mak, "Fingerprint quality classification for CSI-based indoor positioning systems," in *Proc. ACM MobiHoc Workshop Pervasive Syst. IoT Era*, 2019, pp. 31–36.
- [20] S. D. Bast, A. P. Guevara, and S. Pollin, "CSI-based positioning in massive MIMO systems using convolutional neural networks," in *Proc. IEEE Veh. Technol. Conf. Spring*, 2020, pp. 1–5.
- [21] T. Murakami, M. Miyazaki, M. Ishida, and A. Fukuda, "Wireless LAN-Based CSI Monitoring System for Object Detection," *MDPI Electron.*, vol. 7, pp. 1–11, Nov. 2018.
- [22] M. Miyazaki, S. Ishida, A. Fukuda, T. Murakami, and S. Otsuki, "Initial attempt on outdoor human detection using IEEE 802.11ac WLAN signal," in *Proc. IEEE Sensors Appl. Symp.*, 2019, pp. 1–6.
- [23] R. Takahashi, S. Ishida, A. Fukuda, T. Murakami, and S. Otsuki, "DNN-based outdoor NLOS human detection using IEEE 802.11ac WLAN signal," in *Proc. IEEE Sensors*, 2019, pp. 1–4.
- [24] K. Takata, O. Muta, T. Murakami, and S. Otsuki, "Performance evaluation of object detection using channel state information in wireless LAN systems with distributed antennas," in *Proc. Int. Symp. Wireless Pers. Multimedia Commun.*, 2020, pp. 1–6.
- [25] H. Senji, O. Muta, T. Murakami, and S. Otsuki, "An interleaved channel state information clustering scheme for wireless LAN-based object detection systems," *IEICE Commun. Exp.*, vol. 10, no. 1, pp. 42–99, 2021.

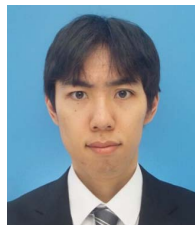
<sup>8</sup>Although the IEEE802.11ac-based feedback frame includes quantized CSI [36] and although it is adopted in simulation, to simplify the notation, the quantization is not considered in the Appendix.

<sup>9</sup>We have confirmed that similar detection accuracy is obtained even when  $\mathbf{V}_k$  is used instead of transformation to trigonometric functions.

- [26] K. Noguchi, O. Muta, T. Murakami, and S. Otsuki, "A CSI-based object detection scheme using interleaved subcarrier selection in wireless LAN systems with distributed antennas," in *Proc. IEEE Veh. Technol. Conf.*, 2021, pp. 1–5.
- [27] A. Catovic and Z. Sahinoglu, "The Cramer–Rao Bounds of Hybrid TOA/RSS and TDOA/RSS Location Estimation Schemes," *IEEE Commun. Lett.*, vol. 8, no. 10, pp. 626–628, Oct. 2004.
- [28] M. Laaraiedh, L. Yu, S. Avrillon, and B. Uguen, "Comparison of hybrid localization schemes using RSSI, TOA, and TDOA," in *Proc. 17th Eur. Wireless Conf.*, 2011, pp. 1–5.
- [29] A. Coluccia and A. Fascista, "Hybrid TOA/RSS range-based localization with self-calibration in asynchronous wireless networks," *MDPI J. Sensor Actuator Netw.*, vol. 8, pp. 1–21, May 2019.
- [30] A. Shahmansoori, G. Seco-Granados, and H. Wymeersch, "Power allocation for OFDM wireless network localization under expectation and robustness constraints," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 2027–2038, Mar. 2017.
- [31] D. Wang, M. Fattouche, F. M. Ghannouchi, and X. Zhan, "Quasi-optimal subcarrier selection dedicated for localization with multicarrier-based signals," *IEEE Syst. J.*, vol. 13, no. 2, pp. 1157–1168, Jun. 2019.
- [32] Y. Huang, X. Zhang, E. H. Aglzim, and L. Shi, "Target 5G visible light positioning signal subcarrier extraction method using particle swarm optimization algorithm," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast.*, 2021, pp. 1–6.
- [33] I. Reinhold, J. Starkhammar, and M. Sandsten, "The scaled reassigned spectrogram adapted for detection and localisation of transient signals," in *Proc. 25th Eur. Signal Process. Conf.*, 2017, pp. 907–911.
- [34] K. Yamamoto, K. Toyoda, and T. Ohtsuki, "Spectrogram-based non-contact RRI estimation by accurate peak detection algorithm," *IEEE Access*, vol. 6, pp. 60369–60379, 2018.
- [35] C. Liu et al., "RSS distribution-based passive localization and its application in sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2883–2895, Apr. 2016.
- [36] IEEE Computer Society, *IEEE Standard for Information Technology Telecommunications and Information Exchange Between Systems Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE 802.11-2020 (Revision of IEEE Std 802.11-2016), LAN/MAN Standards Committee, Dec. 2020.
- [37] EEM-RTM, [Online]. Available: <http://www.e-em.co.jp/rtm/index.html>
- [38] S. Bartoletti, Z. Liu, M. Z. Win, and A. Conti, "Device-free localization of multiple targets in cluttered environments," *IEEE Trans. Aerosp. Electron. Syst.*, early access, Nov. 18, 2021, doi: [10.1109/TAES.2021.3128964](https://doi.org/10.1109/TAES.2021.3128964).
- [39] Y. Amit and D. Geman, "Shape quantization and recognition with randomized trees," *Neural Comput.*, vol. 9, no. 7, pp. 1545–1588, 1997.
- [40] D. Lei, H. Zhang, H. Liu, Z. Li, and Y. Wu, "Maximal uncorrelated multinomial logistic regression," *IEEE Access*, vol. 7, pp. 89924–89935, 2019.
- [41] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, 1995.
- [42] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967.
- [43] A. H. Jahromi and M. Taheri, "A non-parametric mixture of Gaussian naive bayes classifiers based on local independent features," in *Proc. Artif. Intell. Signal Process. Conf.*, 2017, pp. 209–212.



**Osamu Muta** (Member, IEEE) received the Associate B.E. degree from the Sasebo Institute of Technology in 1994, the B.E. degree from Ehime University, Matsuyama, Japan, in 1996, the M.E. degree from the Kyushu Institute of Technology, Iizuka, Japan, in 1998, and the Ph.D. degree from Kyushu University in 2001. In 2001, he joined the Graduate School of Information Science and Electrical Engineering, Kyushu University as an Assistant Professor. Since 2010, he has been an Associate Professor with the Center for Japan-Egypt Cooperation in Science and Technology, Kyushu University. His research interests include signal processing techniques for wireless communications and powerline communications, MIMO techniques, interference coordination techniques, and nonlinear distortion compensation techniques for high-power amplifiers. He is a Senior Member of the Institute of Electronics, Information, and Communication Engineering (IEICE). He was the recipient of the 2005 Active Research Award in IEICE technical committee of radio communication systems, Chairperson's award for the best paper in IEICE technical committee of communication systems (2014, 2015, 2017), and the 2020 IEICE Communication Society Best Paper Award, respectively.



**Keisuke Takata** (Member, IEEE) received the B.E. degree from the Department of Electrical Engineering and Computer Science, School of Engineering, Kyushu University, Fukuoka, Japan, in 2020, and the M.E. degree from the Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan, in 2022. He was involved in research on object detection systems using channel state information of wireless LAN. He is currently enrolled in NEC Solution Innovators, Ltd.



**Kazuki Noguchi** (Student Member, IEEE) received the B.E. degree from the Department of Electrical Engineering and Computer Science, School of Engineering, Kyushu University, Fukuoka, Japan, in 2021. He is currently working toward the M.E. degree with the Graduate School of Information Science and Electrical Engineering, Kyushu University. His research interests include device-free WLAN-based object detection using channel state information. He is a Student Member of IEEE and the Institute of Electronics, Information and Communication Engineers.



**Tomoki Murakami** (Member, IEEE) received the B.E., M.E., and Dr. Eng. degrees from Waseda University, Tokyo, Japan, in 2006, 2008, and 2015, respectively. In 2008, he joined NTT Network Innovation Laboratories, Nippon Telegraph and Telephone Corporation, Yokosuka, Japan. He is currently a Distinguished Researcher with the Wireless Access Systems Project. His research interests are high efficiency technologies for future wireless systems. He was the recipient of the Young Engineer Award from the IEICE in 2010, the Active Research Award from IEICE

AP (Antenna and Propagation) in 2010, the Best Tutorial Paper Award from IEICE Communications society in 2014, Best Paper Award and the KIYASU-Zen'iti Award from the IEICE in 2015, and the Best Paper Award from IEICE on AP in 2016. He is a Member of IEICE.



**Shinya Otsuki** (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in communication engineering from Osaka University, Suita, Japan, in 1993, 1995, and 1997, respectively. He joined NTT in 1997. From 1997 to 2008, he studied wireless access systems, wireless LAN systems, and wireless systems for Internet services in trains. From 2008 to 2011, he was involved in international standardization efforts in evolved packet core and services using Internet Protocol multimedia subsystems with NTT Service Integration Laboratories. He has been with NTT Access

Network Service Systems Laboratories since 2011 and has been contributing to the activities of Working Parties 5A and 5C in Study Group 5 of ITU Radio communication Sector. His research interests include sensing technologies using wireless communication systems. He was the recipient the Young Engineer Award from the Institute of Electronics, Information and Communication Engineers in 2004 and the ITU-AJ International Activity Encouragement Award and the ITU-AJ Accomplishment Award from the ITU Association of Japan in 2014 and 2022, respectively. He is a Member of the Institute of Electrical and Electronics Engineers (IEEE) and IEICE.