

Article

Device-Free Indoor Location Estimation System Using Commodity Wireless LANs

Yuan Zhou, Minseok Kim , Hideaki Momose and Satoru Yasukawa

Graduate School of Science and Technology, Niigata University, Niigata 950-3198, Japan; f20c503g@mail.cc.niigata-u.ac.jp (Y.Z.); oloutionda@au.com (H.M.); strk956@gmail.com (S.Y.)

* Correspondence: mskim@eng.niigata-u.ac.jp

Abstract: In recent years, propagation channel characteristics have been effectively used in several applications such as motion sensing and position detection. Considerable attention has been paid to channel-sounding methods that are easy to utilize using low-cost devices. This paper presents a device-free indoor location estimation method using the spatio-temporal features of radio propagation channels using a 2.4 GHz-band three-by-three multiple-input-multiple-output (MIMO) channel sounder developed using commodity wireless local area network (WLAN). The measurement results demonstrated a reasonable performance of the proposed method with a small number of antennas.

Keywords: WiFi sounder; channel state information (CSI); MIMO; indoor location estimation; array signal processing; machine learning; support vector machine (SVM)



Citation: Zhou, Y.; Kim, M.; Momose, H.; Yasukawa, S. Device-Free Indoor Location Estimation System Using Commodity Wireless LANs. *Telecom* **2021**, *2*, 181–198. <https://doi.org/10.3390/telecom2020012>

Academic Editor: Sotirios K. Goudos

Received: 16 March 2021

Accepted: 15 April 2021

Published: 21 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The technical development of the Internet of Things (IoT) has enhanced people's demand for location services. In the IoT era, position information is crucial. Among these services, indoor positioning-as-a-service, closely related to people's lives, has attracted a great deal of attention. The indoor positioning system (IPS) can be widely used in home surveillance systems, nursing care for the elderly, and patient monitoring in hospitals [1]. Most existing surveillance systems use cameras for monitoring purposes [2]. However, these have insufficient coverage and blind areas. To improve the performance, it is necessary to increase the number of cameras. However, this also brings an excessive cost and complex configuration. Further, privacy is still a problem that cannot be ignored.

To address these problems, various methods utilizing radio waves have been considered. By capturing the fluctuation of wireless communication channel characteristics due to a person's movement and applying machine learning, the estimation of a person's position and behavior pattern can be achieved. However, positioning in an indoor environment still has many technical challenges owing to different propagation conditions from outdoor applications [3]. For example, angle-of-arrival (AOA)-based localization has received great attention in the past decades. However, the major difficulty of AOA-based positioning is the influence of multipath propagation. The complexity of estimating the unknown channel parameters (amplitudes, delays, and angles) grows as the number of paths increases [4]. The time-of-arrival (TOA)-based system also has many challenges in obtaining high localization accuracy due to the difficulty of the order of one nanosecond measuring of the TOA using low-cost hardware in multipath-rich environments. Such systems suffer from time and frequency offsets between the local clocks in different nodes [5]. Furthermore, the range-based indoor positioning system is easily degraded in non-line-of-sight (NLOS) propagation environments [6]. In addition, the commonly used localization technologies such as the Global Positioning System (GPS) do not work effectively (due to low accuracy, a bad signal, and other issues) in an indoor environment owing to signal fading and multipath effects caused by building structures. Therefore, we focused on wireless devices that have already been widely deployed, such as WiFi (WLAN using

IEEE802.11 standards). WiFi is widely used in homes, hotels, cafes, airports, shopping malls, and other various types of large or small buildings, which makes it very attractive for indoor positioning applications. The universality and affordability of WiFi devices will enable the popularization of indoor localization systems.

Recently, various studies researching indoor localization systems using WiFi devices have been reported. For instance, the IPS using the fingerprint method of received signal strength indication (RSSI) has been discussed [7,8]. However, owing to the temporal and spatial variation of the indoor environment [9], several access points (APs) are required to improve the accuracy; moreover, the configuration would become complex [10]. Channel state information (CSI)-based IPS has been developed [11–14]. However, owing to the limited bandwidth of WiFi devices, several APs are required to maintain the performance of the system [15]. Note that the latest WiFi standards allow a greater bandwidth, but no open-source tools for CSI acquisition have been provided. Therefore, we still need to use specific chipsets for CSI acquisition such as Qualcomm Atheros 9380/9580 and Intel 5300 [16], which support only IEEE 802.11n. Moreover, a microwave system using an antenna array has already been developed [17] where predetermined events are identified by monitoring the fluctuation of the signal subspace spanned by the eigenvectors. Because the eigenvector represents the spatial structure of the multipath propagation, spatial filtering by the first eigenvector can reduce the effects of noise, but a large number of antennas is needed to resolve the multipath components more precisely. In device-free localization (DFL) techniques, the performance is significantly influenced by the strong line-of-sight (LOS) path and is prone to bias depending on the person's location [18].

To cope with these problems, this study developed an indoor DFL system based on commodity WiFi devices, which achieves a reasonable performance by using machine learning for the spatio-temporal features of wireless channels. The technical contributions of this study are as follows. First, a 2.4 GHz-band three \times three MIMO channel sounder was developed where the channel bandwidth was extended to approximately 68 MHz by concatenating the CSI taken at six consecutive WiFi channels to achieve high resolution in the delay time domain. Second, we developed an indoor location estimation method using the spatio-temporal features of multipath propagation characteristics, which were obtained by separating the multipath components into three angle taps and two delay taps. This can reduce the dependency of the performance on the change in the LOS path. Third, a support vector machine (SVM) was applied to identify a person's location.

The remainder of this paper is organized as follows. In Section 2, a channel-sounding system based on IEEE 802.11n wireless LANs is presented. Next, we present an indoor location estimation method in Section 3. The identification results of the sub-area where a person is present via two measurement campaigns are presented in Section 4. This demonstrates the performance of the proposed system and its feasibility. Finally, Section 5 concludes the paper.

2. Channel Sounding Using IEEE 802.11n Wireless LANs

2.1. IEEE 802.11n

IEEE 802.11n has been standardized [19] as a MIMO-OFDM-based wireless LAN standard and is the successor to IEEE 802.11a/g. The backward compatibility of existing standards provides a mixed mode [20] to achieve MIMO transmission, extending the legacy mode. The frame format of the mixed mode is shown in Figure 1. After L-SIG (legacy signal field), which sends wireless LAN frame information such as the total length of the wireless LAN frame, it sends the HT-SIG (high-throughput signal field), which includes transmission parameters for spatial multiplexing transmission. Furthermore, after HT-STF (high-throughput short training field) and HT-LTF (high-throughput long training field), it estimates the CSI, which is necessary for spatial multiplexing transmission. In IEEE 802.11n, the number of subcarriers over the 20 MHz bandwidth is increased from 54 to 56. Furthermore, it provides a 40 MHz-mode using channel bonding as an option.

Consequently, by four-stream spatial multiplexing with a 40 MHz bandwidth, it can achieve transmission throughput up to 600 Mbps (64QAM, Code Rate 5/6).

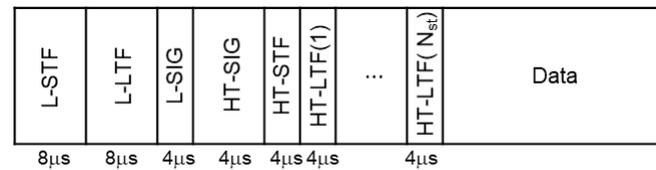


Figure 1. Frame format of IEEE 802.11n hybrid mode.

2.2. CSI Acquisition

In this study, we used an open-source software package called the CSI tool [21] to build our system. In IEEE 802.11n, the CSI was estimated in the receiver and then fed back to the transmitter, so the transmitter can use the CSI to complete calibration and beamforming. In other words, by setting the CSI estimate flag in the transmit packet, we can manage the receiver to report the CSI estimation to the transmitter. The CSI tool is a device driver for the Atheros network interface cards (NICs), ath9k (kernel module), to extract the CSI using IEEE 802.11n. In the CSI tool, the transmitter repetitively sends the CSI estimate flag packet for CSI acquisition, and the receiver reports the CSI acquisition result to the user space program whenever the CSI is acquired. Here, the CSIs acquired for every combination of transmitter and receiver antennas are estimated as complex numbers for 56 subcarriers, including transmitter/receiver characteristics, propagation channel characteristics, and antenna characteristics.

The CSI report in one packet transmitted with a bandwidth of 20 MHz is obtained in the form of a complex matrix of size $N_R \times N_T \times N_C$; if the number of antennas at the receiver and transmitter is N_R and N_T , respectively. Furthermore, information such as the timestamp, channel number, transmission rate, number of antennas (N_R , N_T), number of sub-carriers N_C , noise floor, PHY error, the received signal strength indicator (RSSI) of every antenna at the receiver, and the length of the payload is added. Because the real and imaginary parts of the CSI are each represented by 10 bits, their amplitude is automatically scaled according to the magnitude of the received power. Further, the phase of CSI fluctuates independently for each packet, but the amount of offset is common to all subcarriers (common phase offset).

2.3. MIMO Channel Sounding

Figure 2 shows the MIMO transceiver configuration of the IEEE 802.11n wireless LAN. The stream parser creates N_{st} spatial streams from the encoded bit stream, and the quadrature is modulated to the symbol stream by constellation mapping. The transmitter can send N_{st} ($\leq \text{rank}(H) \leq \min(N_T, N_R)$) parallel streams, where N_T and N_R denote the number of antennas in the transmitter and receiver, respectively. Then, the spatial streams are shifted cyclically by cyclic delay diversity (CDD), extending the communication area by transmitting the same signal at different carrier frequencies. This prevents beamforming that might be made in a specific direction when all the antennas send the signal with common header information. Spatial mapping (SM) is performed when the number of antennas is larger than the number of spatial streams ($N_T > N_{st}$). The frequency samples of each transmission branch are transformed into time-domain signals by the inverse Fourier transform and are transmitted simultaneously by all antennas.

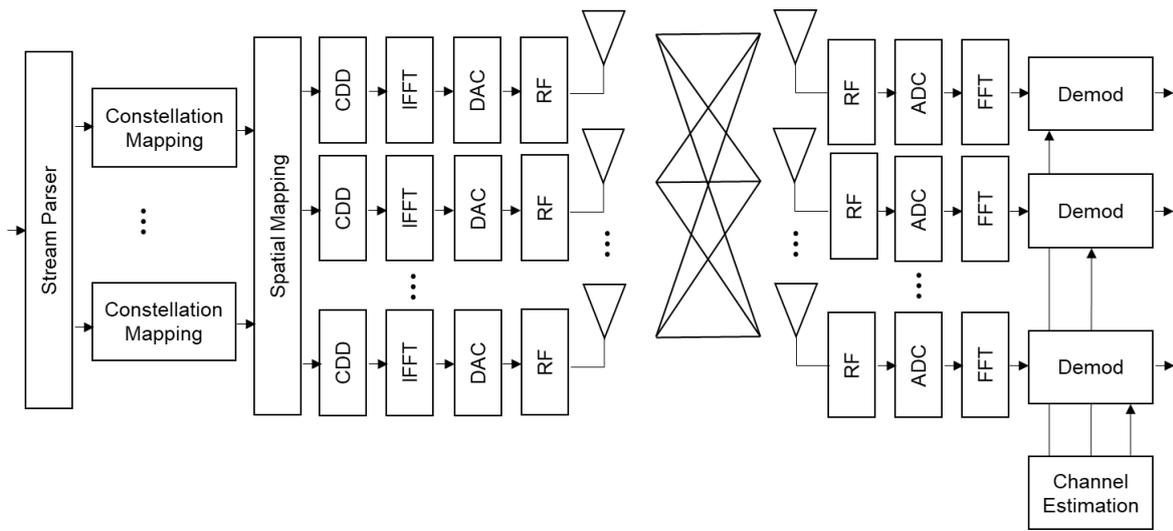


Figure 2. IEEE 802.11n transceiver configuration.

The MIMO propagation channel matrix H is estimated using the training signal (HT-LTF), which is contained in the frame format as shown in Figure 1. Figure 3 shows the block diagram of the three \times three MIMO channel-sounding system. Assuming that the number of spatial streams and the number of transmit antennas are the same ($N_T = N_{st} = 3$), the received signal vector at the k th ($k = 1, \dots, N_c = 56$) sub-carrier is expressed as:

$$y^{(k)} = G_{agc} H^{(k)} \Phi^{(k)} dx^{(k)} + n^{(k)} \in \mathbb{C}^{3 \times 1} \tag{1}$$

where G_{agc} , $H^{(k)}$, and $\Phi^{(k)}$ denote the gains of the automatic gain control (AGC), MIMO channel matrix, and CDD matrix, respectively. They are expressed as:

$$G_{agc} = \text{diag} \left(\left[\sqrt{G_{agc,1}} \quad \sqrt{G_{agc,2}} \quad \sqrt{G_{agc,3}} \right]^T \right), \tag{2}$$

$$H^{(k)} = \begin{bmatrix} H_{11}^{(k)} & H_{12}^{(k)} & H_{13}^{(k)} \\ H_{21}^{(k)} & H_{22}^{(k)} & H_{23}^{(k)} \\ H_{31}^{(k)} & H_{32}^{(k)} & H_{33}^{(k)} \end{bmatrix}, \tag{3}$$

$$\Phi^{(k)} = \text{diag} \left(\left[1 \quad e^{-j\frac{2\pi}{N_f} \delta_2 k} \quad e^{-j\frac{2\pi}{N_f} \delta_3 k} \right]^T \right), \tag{4}$$

where $x^{(k)}$ is the transmitted signal and $n^{(k)} = \left[n_1^{(k)} \quad n_2^{(k)} \quad n_3^{(k)} \right]^T$ is the noise vector. Here, δ_2 and δ_3 are specified as $\delta_2 = 8$ (400 ns) and $\delta_3 = 4$ (200 ns), and N_f is the number of FFT points and is set to 64.

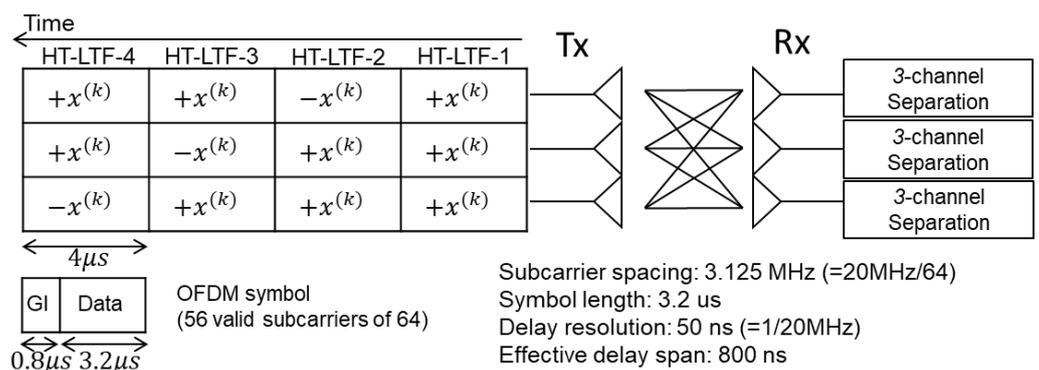


Figure 3. Three times three MIMO channel sounding.

To separate the transmitted signals from multiple transmitting antennas, the signal $x^{(k)}$ is transmitted four times with each transmitting antenna by setting \mathbf{d} as $\mathbf{d}[1] = [1 \ 1 \ 1]^T$, $\mathbf{d}[2] = [-1 \ 1 \ 1]^T$, $\mathbf{d}[3] = [1 \ -1 \ 1]^T$, $\mathbf{d}[4] = [1 \ 1 \ -1]^T$ in Equation (1), then the matrix of the received signal can be obtained as:

$$\mathbf{Y}^{(k)} = \mathbf{G}_{\text{agc}} \mathbf{H}^{(k)} \mathbf{\Phi}^{(k)} \mathbf{D} x^{(k)} + \mathbf{N}^{(k)} \in \mathbb{C}^{3 \times 4}, \quad (5)$$

where:

$$\mathbf{Y}^{(k)} = \begin{bmatrix} \mathbf{y}^{(k)}[1] & \mathbf{y}^{(k)}[2] & \mathbf{y}^{(k)}[3] & \mathbf{y}^{(k)}[4] \end{bmatrix}, \quad (6)$$

$$\mathbf{N}^{(k)} = \begin{bmatrix} \mathbf{n}^{(k)}[1] & \mathbf{n}^{(k)}[2] & \mathbf{n}^{(k)}[3] & \mathbf{n}^{(k)}[4] \end{bmatrix}, \quad (7)$$

and:

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}. \quad (8)$$

Then, the CSI is obtained using Equation (5) as:

$$\hat{\mathbf{H}}_{\text{CSI}}^{(k)} = \mathbf{Y}^{(k)} \mathbf{D}^{-1} / x^{(k)}. \quad (9)$$

Note that unlike data transmission, the signal-to-noise power ratio (SNR) should sufficiently be large in channel sounding because the target to be discovered is not data, but the unknown radio channel. However, in a low-SNR regime, the SNR can be improved by taking the average of the channel snapshots under the static channel condition. It is necessary to remove the effects of the amplifier and CDD to obtain the MIMO channel matrix, which is expressed as:

$$\hat{\mathbf{H}}^{(k)} = \mathbf{G}_{\text{agc}}^{-1} \hat{\mathbf{H}}_{\text{CSI}}^{(k)} \bar{\mathbf{\Phi}}^{(k)}. \quad (10)$$

Owing to the nonlinear distortions of the transmitter and receiver circuits in the CSI, further calibration to Equation (10) is required.

2.4. Developed System

The bandwidth of IEEE 802.11n is approximately 20 MHz with an optional 40 MHz mode. Therefore, the delay time resolution is theoretically limited to 50 ns (=1/20 MHz), which is not sufficient to apply WiFi to sensing with subtle variation of the propagation path. Bandwidth expansion with channel bonding has been proposed [22], and the toolkit has been published as open-source [23]. Utilizing this toolkit, the channel bandwidth could be extended to approximately 67.8125 MHz (delay resolution: approximately 15 ns, distance resolution: approximately 4.5 m) by concatenating the CSIs taken at six consecutive WiFi channels, as shown in Figure 4.

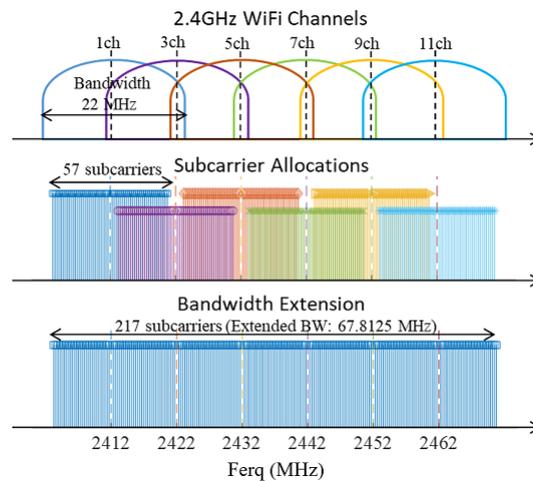


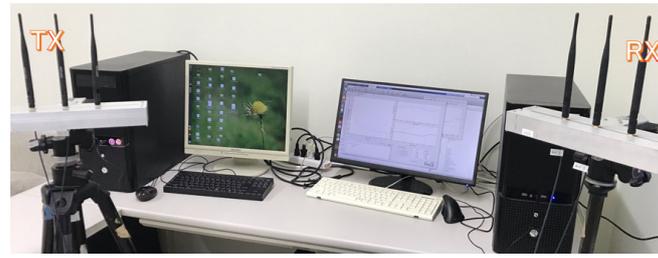
Figure 4. Bandwidth extension by six-channel bonding.

2.4.1. System Configuration

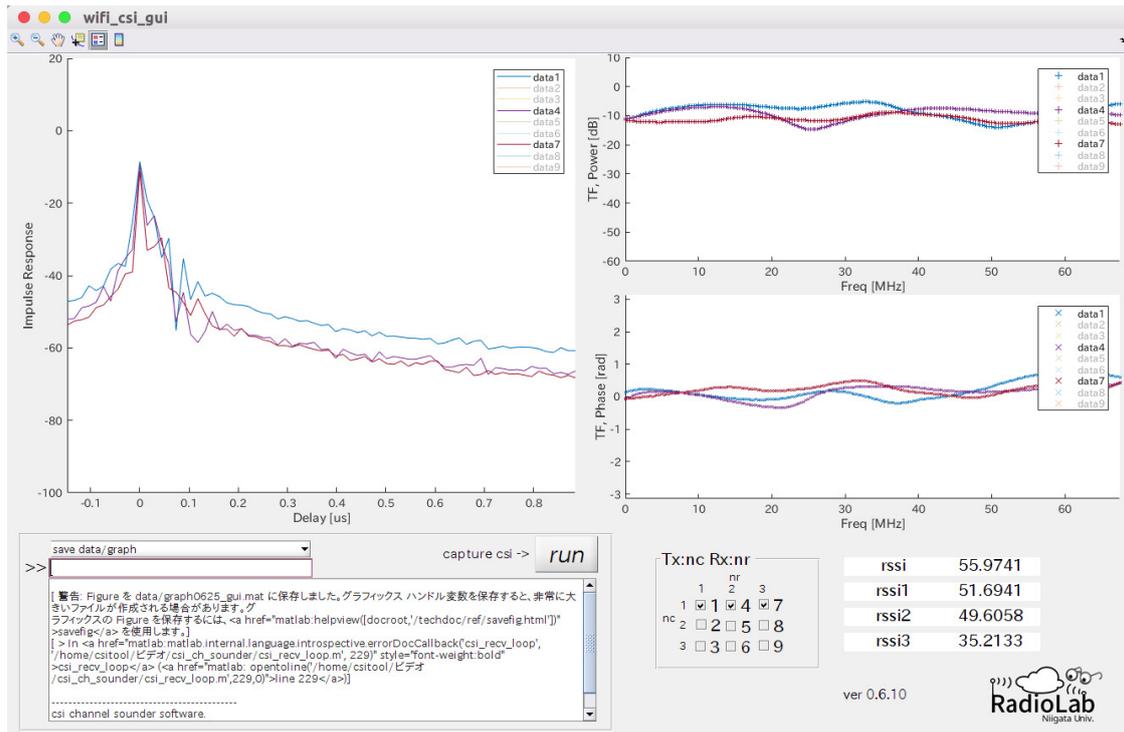
The system specifications are listed in Table 1. Two Linux PCs equipped with an Atheros AR9380 WLAN card were used for the transmitter and receiver, as shown in Figure 5. To extend the bandwidth, fifty packets are transmitted and received continuously at every channel while sweeping the channel from 1 to 11 at two-channel intervals. Separately averaging the amplitude and phase of the CSI value at each subcarrier filters out abrupt fluctuations and improves the SNR. Here, the beacon interval was set to a minimum value of 15 ms for fast channel switching.

Table 1. System setup.

Item	Values
Hardware	Linux (Kernel 4.1 Modified version) PC × 2 (equipped with Rabortw AR9380 WiFi card)
Software	AP (transmitter): hostapd, hostapd_cli STA (receiver): WPA_supplicant
WiFi channels	1 ch (2412 MHz), 3 ch (2422 MHz), 5 ch (2432 MHz), 7 ch (2442 MHz), 9 ch (2452 MHz), 11 ch (2462 MHz)
Sub-carriers	217
Extend bandwidth	67.8125 MHz (bonded by 6 channels)
Beacon interval	15 ms (usually, 100 ms)
Packet transmission method	Send 50 packets in a row, and wait for 50 ms Repeat while switching channels



(a)



(b)

Figure 5. Developed channel sounding system. (a) Hardware setup. (b) Software.

2.4.2. Back-To-Back Calibration

The phase variation of the CSI occurs because of hardware imperfections such as the nonlinear distortion of the amplifiers, power control uncertainty, and phase ambiguity [22,24], which should be removed to obtain the genuine propagation channel characteristics. However, because the six CSIs taken at consecutive WiFi channels are concatenated to extend the bandwidth, it is necessary to calibrate the inter-channel discontinuity in the amplitude and phase values.

To remove the hardware imperfections contained in the CSI, we measure the system characteristics including the hardware imperfection in advance by back-to-back measurement. This measurement is conducted by connecting the transmit and receive antenna ports directly via an RF cable with known characteristics. We then obtain the channel transfer function by dividing the CSI by the system characteristics, which is called back-to-back calibration (B2B calibration). Here, CSI $\hat{H}_{\text{CSI,B2B},j,i}^{(k)}$ is obtained by connecting the i th transmit and j th receive antenna ports directly with a cable and is performed for every antenna combination. The MIMO channel matrix can be obtained as:

$$\hat{H}_{j,i}^{(k)} = \frac{\hat{H}_{\text{CSI},j,i}^{(k)} / \sqrt{G_{\text{agc},j}}}{\hat{H}_{\text{CSI,B2B},j,i}^{(k)} / \sqrt{G_{\text{agc},\text{B2B},j}}} e^{j \frac{2\pi}{N_f} \delta_{i,k}}, \quad (11)$$

where $G_{\text{agc},j}$ denotes the gain of the AGC amplifier of the j th receiving branch. It is assumed that the values of the denominator in Equation (11) should usually be measured in advance (upon initial setup). In commodity WiFi devices, the AGC amplifier operates automatically according to the received signal level to maintain the signal power at a certain level; thus, the CSI values are scaled as in Equation (11). The gain of the AGC amplifier can be deduced by referring to the amplitude of the CSI and RSSI as:

$$G_{\text{agc},j} = \frac{1}{P_j} 10^{\frac{\text{RSSI}_j}{10}}, \quad (12)$$

where $P_j = E[|y_j^{(k)}|^2]$.

The B2B calibration allows the CSIs acquired on multiple channels to be connected continuously because the phase and amplitude should become zero and unit after calibration. However, a phase jump occurs between adjacent channels owing to the independent random offset of the CSI phase characteristics in every packet. This can be removed by calculating the inter-channel phase difference using overlapped sub-carriers in adjacent channels. The amplitude of each channel also fluctuates owing to the AGC gain uncertainty. Using the averaged values of the overlapped subcarriers, the amplitudes were concatenated smoothly. The calibration procedure is shown in Figure 6a,b.

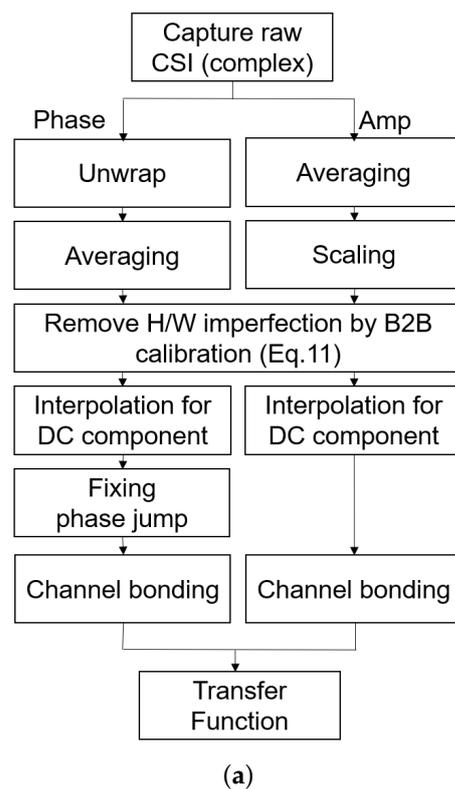


Figure 6. Cont.

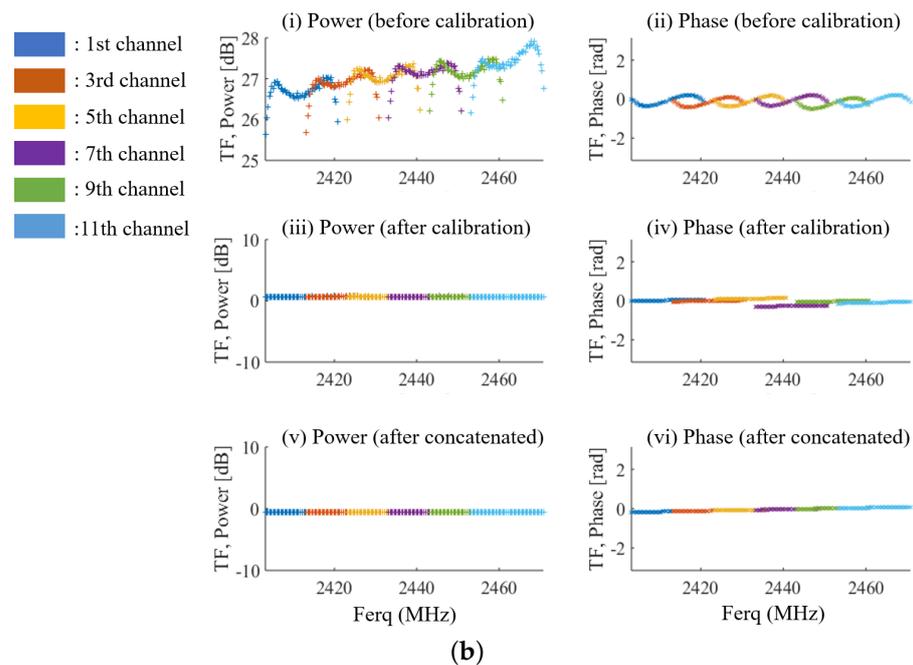


Figure 6. B2B calibration. (a) Calibration procedure. (b) Calibration result.

2.5. Evaluation

The channel impulse response test was conducted for system validation purposes. An unmatched T-junction (BL41-6203-00, Orient Microwave Corp.) [25] was used for the measurements. Figure 7a shows the measurement configuration. Transmitter antenna port 1 (Tx1) was connected to a power splitter. Port 1 of the splitter was connected to receiver antenna port 1 (Rx1), which was used for the reference measurement. Port 2 of the splitter was connected to port A of the T-junction, and the signal from port B was input to receive antenna port 2 (Rx2). Here, port C was connected to a 6 m cable that left the end opened, which generated reflected waves at the cable end. The attenuation per 1 m of the cable used (Enviroflex_316, Huber Shuner, Corp.) [26] was approximately 1.62 dB/m. The measurement results are shown in Figure 7b. Only direct waves were observed in Tx1–Rx1, and the first reflected wave from the cable end could also be observed in Tx1–DUT–Rx2, which was in good agreement with the calculated value. Here, the calculation was done considering the propagation speed of electromagnetic waves in coaxial cables to be about 77 % of that in a vacuum and the path length and attenuation rate of the cable (delay time: $\frac{12}{3 \times 10^8 \times 0.77} = 52$ ns, power attenuation: $12 \times 1.65 = 20$ dB).

To evaluate the inter-channel phase stability, the Tx1 signal was distributed to Rx1, Rx2, and Rx3 using a splitter. After B2B calibration, two-hundred ten channel transfer functions were obtained. The variation of $\angle \hat{H}_{2,1}$ and $\angle \hat{H}_{3,1}$ based on the reference value of $\hat{H}_{1,1}$ is shown in Figure 8. The left figure shows the relative phase difference of 217 sub-carriers for 210 time samples, and the right figure shows the phase difference of all sub-carriers acquired for 210 time samples. It can be seen that the phase variation was less than approximately 20 degrees.

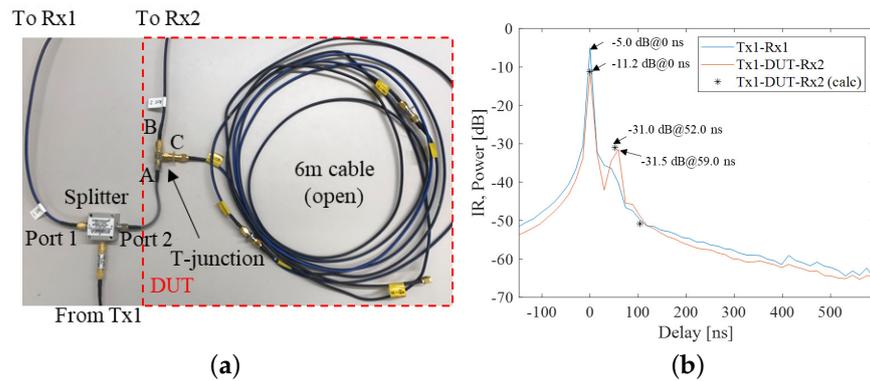


Figure 7. DUT test. (a) DUT test setup. (b) DUT test result.

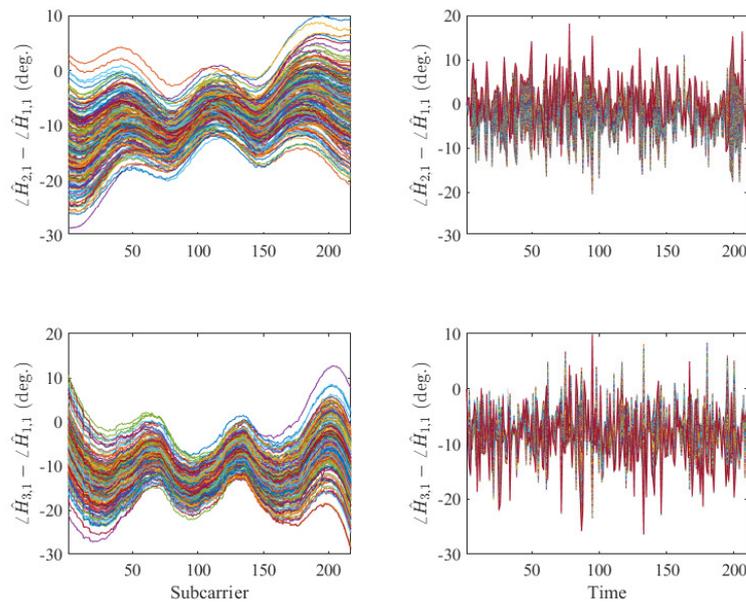


Figure 8. Phase stability.

3. DF Indoor Location Estimation

The DF indoor location estimation system was developed using the WiFi sounder described above. The receiving antennas were arranged linearly at approximately 6.3 cm intervals (half-wavelength of the minimum frequency 2.403 GHz). The bandwidth-broadened channel impulse responses were obtained by concatenating the CSIs taken at six consecutive WiFi channels to extend the bandwidth to approximately 68 MHz (delay resolution: approximately 15 ns, distance resolution: approximately 4.5 m).

3.1. Experiment Scenario

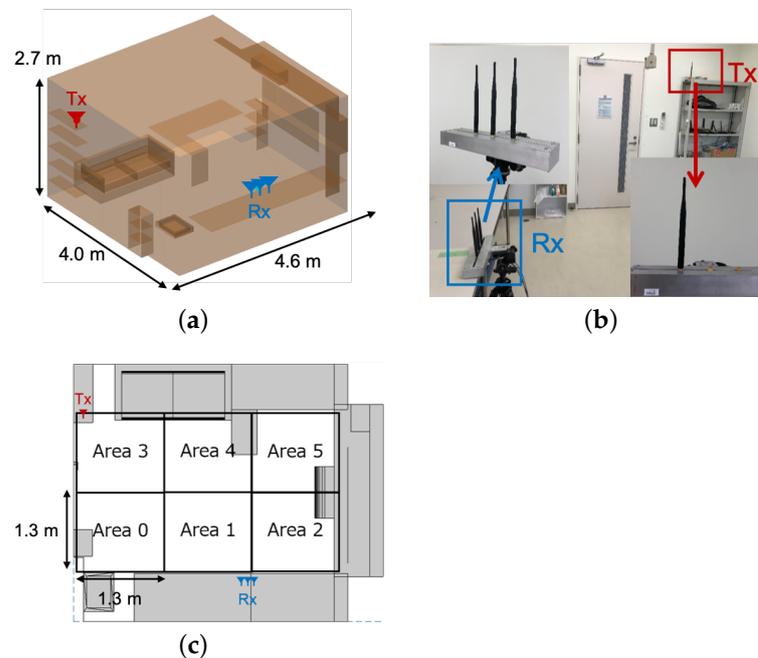
The experiment was conducted in a small office and a medium-sized conference room. The measurement specifications are presented in Table 2. Note that WiFi communication was performed between the access point (AP) and station (STA) fixed in the environment, and a target person moving in the environment was not equipped with any device. Here, the CSI was acquired 10 times for each target position and averaged to reduce the effects of noise.

Table 2. Measurement specifications.

Room Type	Small Office	Conference Room
Size	4.0×4.6 (m ²)	7.8×8.8 (m ²)
Bandwidth	67.813 MHz (6 channel bonding)	
Transmitting antennas	1	2
Receiving antennas	3-element linear array	
Element pattern	Omni-directional	

3.1.1. Small Office

The room model of the small office and the position of the antenna are as shown in Figure 9a,b. The receiving and transmitting antennas were placed at a height of 0.9 m and 1.9 m, respectively. For this environment, thirty CSIs were captured while a person moved randomly in a sub-area; thus, one-hundred eighty samples were captured in total.

**Figure 9.** Small office condition. (a) Room model. (b) Antenna position. (c) Defined area.

3.1.2. Conference Room

The room model of the conference room and the position of each antenna are shown in Figure 10a,b. The receiving antenna was a 3-element linear array placed at a height of 2.0 m. One transmitting antenna (Tx1) was placed at a height of 0.9 m inside the room, and one antenna (Tx2) was placed at a height of 1.6 m outside the room. In this environment, the CSI was captured at each position when a person moved along a predetermined route. Here, the person moved at 0.5 m intervals with a random orientation. One-hundred fifty-one CSIs were captured in total.

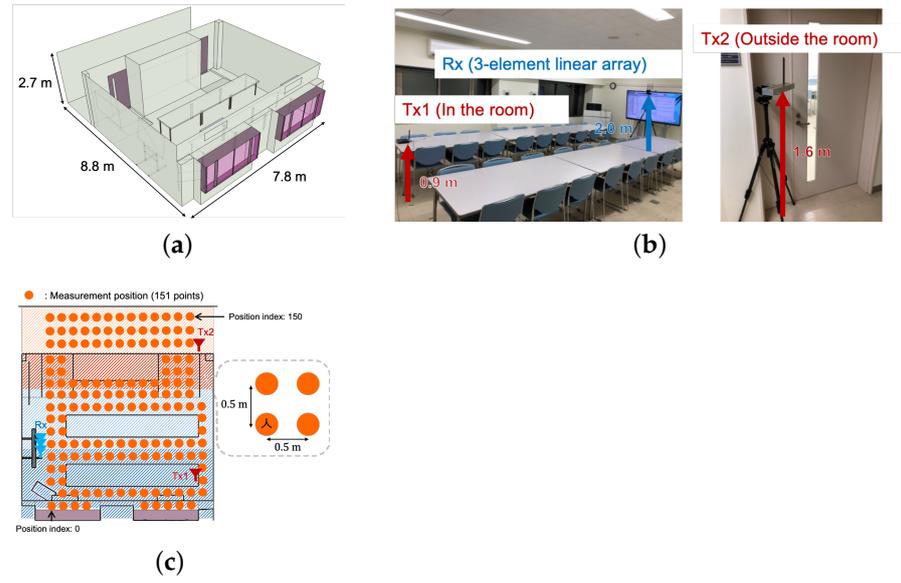


Figure 10. Conference room condition. (a) Room model. (b) Antenna position. (c) Human position.

3.2. Signal Processing

The array output $y(t)$ of the K -element linear array is expressed as:

$$y(t) = \mathbf{w}^H \mathbf{x}(t), \quad (13)$$

$$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_K(t)]^T, \quad (14)$$

$$\mathbf{w} = [w_1, w_2, \dots, w_K]^T. \quad (15)$$

where $[\cdot]^H$ denotes the complex conjugate transpose and $\mathbf{x}(t)$ and \mathbf{w} denote the input signal and the weight of each antenna, respectively. Then, the output power can be expressed as:

$$P_{out} = \frac{1}{2} E[|y(t)|^2] = \frac{1}{2} \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w}. \quad (16)$$

where \mathbf{R}_{xx} denotes the correlation matrix of the received signal and is expressed as:

$$\mathbf{R}_{xx} = E[\mathbf{x}(t) \mathbf{x}^H(t)]. \quad (17)$$

To direct the main beam of the array antenna to an arbitrary angle ϕ , each weight expressed in Equation (15) can be set as:

$$w_k = \exp\left(-j \frac{2\pi}{\lambda} d_k \sin \phi\right), \quad (18)$$

then the weight vector can be expressed as:

$$\mathbf{w} = \left[\exp\left(-j \frac{2\pi}{\lambda} d_1 \sin \phi\right), \dots, \exp\left(-j \frac{2\pi}{\lambda} d_K \sin \phi\right) \right]^T \equiv \mathbf{a}(\phi). \quad (19)$$

Here, $\mathbf{a}(\phi)$ denotes the mode vector, and angle ϕ is a variable. Then, the output power is expressed as:

$$P_{out} = \frac{1}{2} \mathbf{a}^H(\phi) \mathbf{R}_{xx} \mathbf{a}(\phi). \quad (20)$$

3.2.1. Existing Method

This subsection describes an existing method that uses spatial features [17]. We assumed that L waves were arriving from angles $\theta_1, \dots, \theta_L$. According to Equation (14), the array signal can be expressed as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (21)$$

$$\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_L)], \quad (22)$$

where \mathbf{A} denotes the direction matrix, $\mathbf{a}(\phi)$ denotes the mode vector defined by Equation (19), and $s(t)$ and $n(t)$ denote the signal vector and internal noise, respectively. We assumed that the internal noise component was independent in each antenna with *mean* = 0 and *variance* = σ^2 . Then, the correlation matrix of the input signal is expressed as:

$$\mathbf{R}_{xx} = E[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H + \sigma^2\mathbf{I}. \quad (23)$$

The eigenvalue λ_k and eigenvector \mathbf{e}_k ($k = 1, 2, \dots, K$) can be obtained via the eigenvalue decomposition (EVD) of the correlation matrix \mathbf{R}_{xx} , which satisfies the following equation:

$$\mathbf{R}_{xx}\mathbf{e}_k = \lambda_k\mathbf{e}_k \quad (k = 1, 2, \dots, K). \quad (24)$$

Then, Equation (23) can be expressed as:

$$\mathbf{R}_{xx} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H = \sum_{k=1}^K \lambda_k \mathbf{e}_k \mathbf{e}_k^H, \quad (25)$$

where:

$$\mathbf{E} := [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_K], \quad (26)$$

$$\mathbf{\Lambda} := \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_K\}. \quad (27)$$

Since \mathbf{R}_{xx} is a Hermitian matrix, the eigenvalues can be sorted as:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq \lambda_{L+1} = \dots = \lambda_K = \sigma^2. \quad (28)$$

Because $\text{rank}[\mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H] = 1$, the eigenvalue distribution is:

$$\lambda_1 > \lambda_2 = \dots = \lambda_L > \lambda_{L+1} = \dots = \lambda_K = \sigma^2. \quad (29)$$

Thus, the space spanning by the eigenvector matrix \mathbf{E} can be separated into a signal subspace and a noise subspace.

The first eigenvalue separated by the process above and the corresponding vector (first eigenvector) were used in machine learning. The feature value for machine learning [17] is expressed as:

$$P = |\mathbf{v}_0^H \cdot \mathbf{v}|, \quad (30)$$

$$Q = 1 - \frac{\lambda' - \lambda'_0}{\lambda'_0}, \quad (31)$$

$$\mathbf{v} := \mathbf{e}_1 \propto \mathbf{a}(\phi), \quad (32)$$

$$\lambda' := \lambda_1, \quad (33)$$

where \mathbf{v}_0 and λ'_0 denote the first eigenvector and first eigenvalue obtained in advance, respectively. The first eigenvector is considered the optimal weight vector for the maximum ratio combination to maximize the output SNR. Therefore, it is strongly influenced by the direct wave. However, there is a problem in that the features hardly fluctuate when blocking the multipath component with low power.

3.2.2. Proposed Method

To address the problem mentioned above, we proposed a method that treats the multipath components separately in the delay time domain and beam space. Here, beamforming with three orthogonal beams was used to separate the multipaths in the three beamspace. The output signal obtained at multiple delay taps and multiple beamspace is expressed as:

$$y_{m,n}(t) = \mathbf{w}_m^H \mathbf{x}_n(t) = \gamma_{m,n}(t) e^{j\theta_{m,n}}, \quad (34)$$

where $\mathbf{w}_m \in \mathbb{C}^{3 \times 1}$ denotes the weight vector for the m th orthogonal beam ($m = 1, 2, 3$) and $\mathbf{x}_n(t) \in \mathbb{C}^{3 \times 1}$ denotes the n th delay tap of the channel impulse responses ($n = 1, 2$). The weight vector is obtained by (18) where $\phi \in \{-42^\circ, 0^\circ, 42^\circ\}$, and the beam patterns are as shown in Figure 11a. The target space was divided into three parts to receive multiple waves. γ and θ denote the resulting amplitude and phase, respectively. Figure 11b shows the condition when the single-bounce reflection wave is blocked by a human body. Here, because the multipath components were treated separately and received, its fluctuation can be detected when blocked, as seen in Figure 11c. Finally, the amplitude and phase of the received signal obtained at multiple delay taps and multiple beamspace were used as spatio-temporal feature values for machine learning.

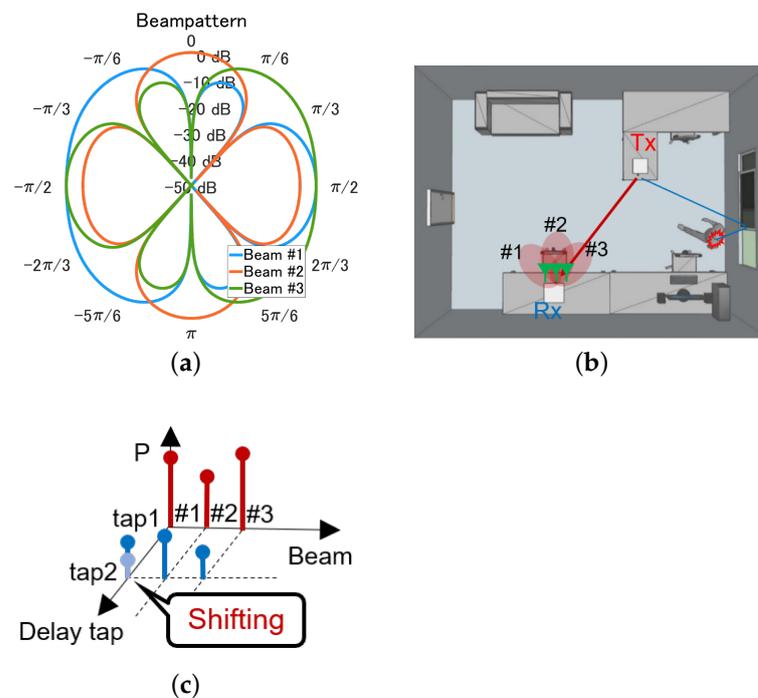


Figure 11. Concept of the proposed method. (a) Beam pattern. (b) Event detection. (c) Multipath separation.

4. Result and Discussion

The machine learning used to create the classification model was the support vector machine (SVM) [27,28] and random forest [29]. The implementation used scikit-learn [30] Python [31] library and the hyperparameters for the model were determined by a framework called Optuna [32]. Figure 12 shows the impulse response captured by the WiFi channel sounder under different conditions (with and without a person). It was observed that the channel bonding extended the delay resolution to approximately 15 ns. Therefore, the first delay tap was set to 0 ns, and the second delay tap was set to approximately 15 ns considering the dimension of the room.

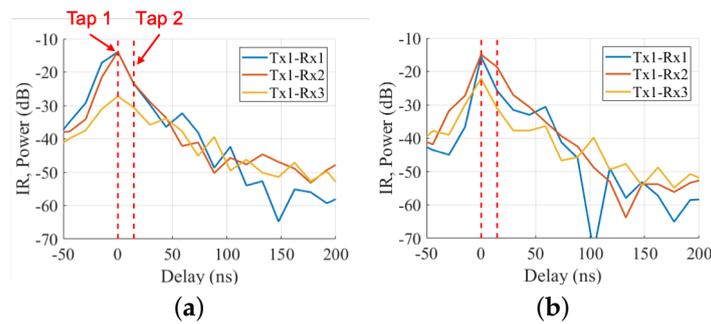


Figure 12. Fluctuation of the impulse response. (a) Without a human. (b) With a human.

4.1. Small Office

Figure 13a shows the features obtained by the existing method in the same office. Here, the first eigenvector and the first eigenvalue obtained in the environment without a person were used as reference data to calculate the features. The features, P , which are the correlation of the eigenvector obtained in environments with and without a person, fluctuated significantly in three locations, in Area 1 due to the shielding effect of the direct wave. However, Areas 3 and 0 were relatively stationary compared to Area 1, even though the direct wave also passed these areas due to the height of the transmitting antenna, which was higher than the subject. Therefore, the direct wave might not be shielded in Area 3 and not completely shielded depending on the orientation of the person's body in Area 0.

Figure 13b shows the features obtained by the proposed method in a small office. The solid and dotted lines denote Delay Taps 1 and 2, respectively. The colors blue, orange, and green represent Beams 1, 2, and 3, respectively. Relatively large phase fluctuations can be seen on the first delay tap of Beam 1, as the direct wave was shielded. Furthermore, many fluctuations can also be confirmed at the second delay tap, which was considered to be the result of shielding the single-bounce reflection wave showing that the events can be detected even when multiple waves with a small power were blocked.

The classification results of the existing and proposed methods are shown in Figure 14. The classification was performed using the support vector machine. For the existing method, the success rate was 35.8%. The average F-score was 0.31, which was not correctly classified due to insufficient features, feature P being relatively stationary. For the proposed method, the success rate and average F-score improved to 64.5% and 0.63, respectively. However, misclassification occurred and was considered to be due to the effect of similar propagation channel conditions, which could improve if the specific area of the room, such as the area covered by the sofa, table, and corridor, were well-defined rather than considering the room to be equally divided.

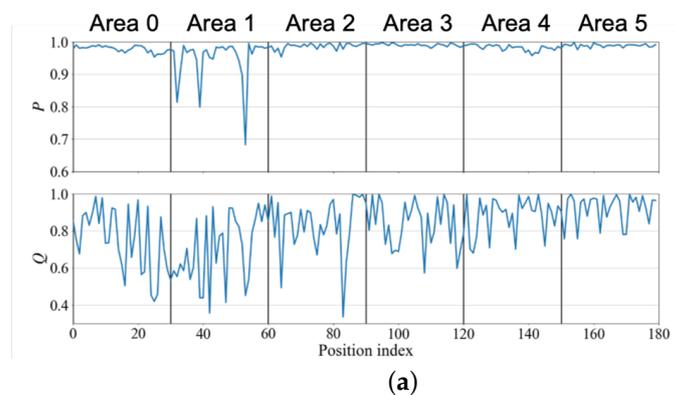


Figure 13. Cont.

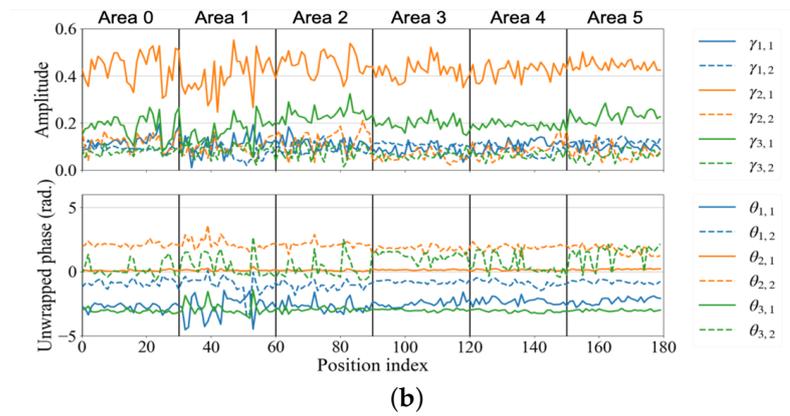


Figure 13. Features of the small office. (a) Existing method. (b) Proposed method.

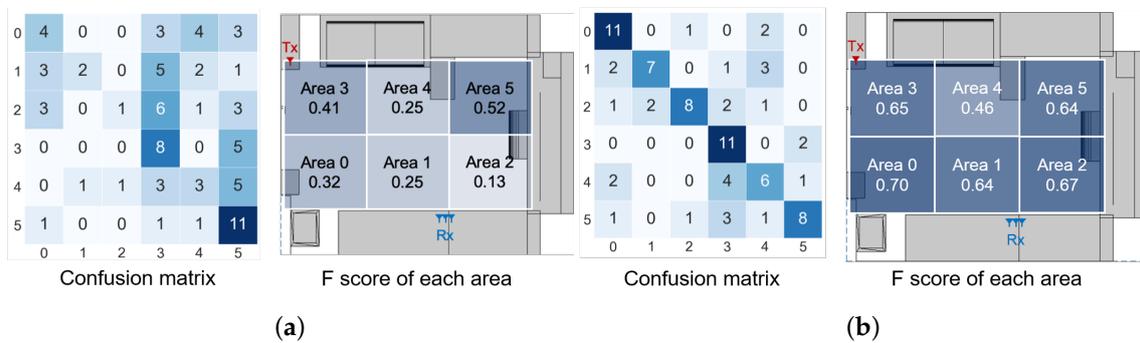


Figure 14. Classification result. (a) Existing method. (b) Proposed method.

4.2. Conference Room

Figure 15 shows the features obtained by the proposed method in the conference room. As shown in Figure 15b, the amplitude of Tx2 fluctuated as it was placed outside the room and was greatly affected by noise. The classification results are shown in Figure 16a. The success rate was 88.2, and the average F-score for each area was 0.84. Here, the features of 76 positions were used to build the machine learning model, and the rest were used as test data. Notably, the positions of each area were randomly selected. However, the ratio of positions used for training to those used for prediction should be the same. As shown in Figure 16b, most of the misclassified positions were near the subarea boundary.

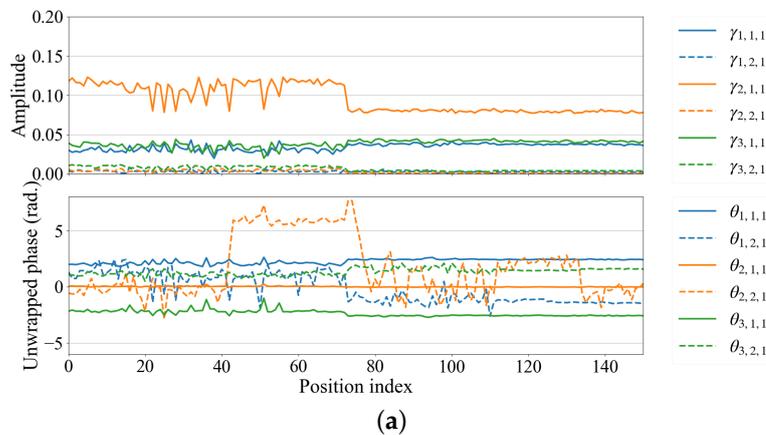


Figure 15. Cont.

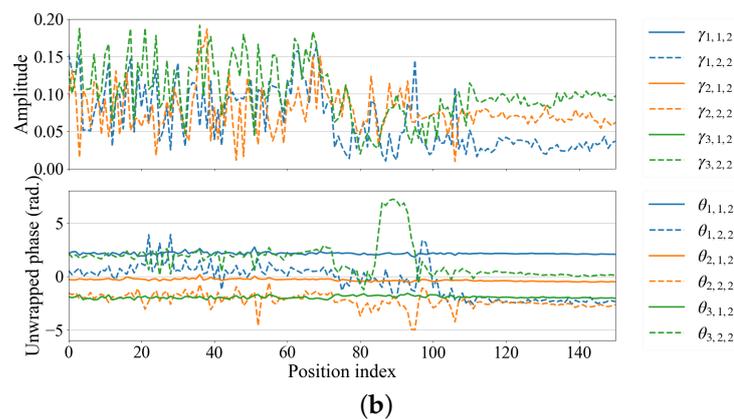


Figure 15. Features of the conference room. (a) Features of Tx1. (b) Features of Tx2.

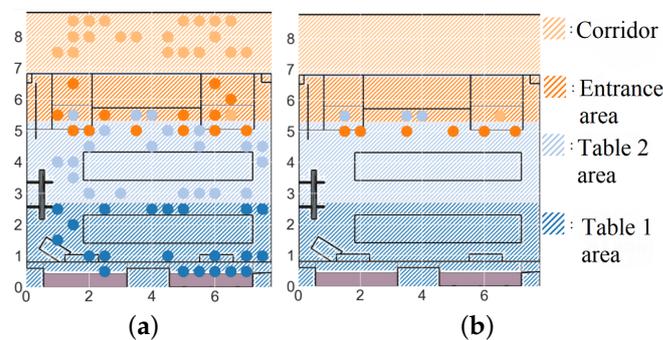


Figure 16. Classification result. (a) Classified position. (b) Misclassified position.

5. Conclusions

In this study, a three-by-three MIMO channel sounder was built, and the CSI taken at six consecutive WiFi channels was concatenated to extend the bandwidth and improve the delay time resolution. Based on this, a practical indoor location estimation system was proposed to cope with the problem that the features are significantly influenced by the strong LOS path. Then, we experimented with a small office and a conference room, and the results of person position estimation by machine learning were presented. The classification result showed that a rough classification of the subarea in a room where a target person is present is possible using a commodity WiFi device.

Author Contributions: Formal analysis, S.Y.; funding acquisition, M.K.; methodology, M.K. and S.Y.; project administration, M.K.; software, H.M.; supervision, M.K.; validation, Y.Z.; visualization, H.M.; writing—original draft, Y.Z. All authors read and agreed to the published version of the manuscript.

Funding: This research was partially funded by KDDI foundation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rainer, M. *Indoor Positioning Technologies*. ETH Zurich Research Collection: Zurich, Switzerland, 2012; pp. 11–14.
2. Nash, J. Global sales of video surveillance equipment projected to surpass 20 billion this year. *Biometric Update*, 10 January 2020.
3. Kaplan, E.; Hegarty, C. *Understanding GPS Principles and Applications*, 2nd ed.; Artech House: Norwood, MA, USA, 2005.
4. Fascista, A.; Coluccia, A.; Ricci, G. A Pseudo Maximum likelihood approach to position estimation in dynamic multipath environments. *Signal Process.* **2021**, *181*, 107907. [[CrossRef](#)]

5. Sathyan, T.; Humphrey, D.; Hedle, M. WASP: A System and Algorithms for Accurate Radio Localization Using Low-Cost Hardware. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2011**, *41*, 211–222. [CrossRef]
6. Li, S.; Hedley, M.; Collings, I.B.D. Humphrey. Indoor positioning based on ranging offset model and learning. In Proceedings of the 2017 IEEE International Conference on Communications Workshops (ICC Workshops), Paris, France, 21–25 May 2017; pp. 1265–1270.
7. Woo, S.; Jeong, S.; Mok, E.; Xia, L.; Choi, C.; Pyeon, M. Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR. *Autom. Constr.* **2011**, *20*, 3–13. [CrossRef]
8. Chan, S.; Shon, G. Indoor Localization Using Wi-Fi Based Fingerprinting and Trilateration Techniques for Lbs Applications. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, XXXVIII-4/C26, 1–5. [CrossRef]
9. Kaemarungsi, K.; Krishnamurthy, P. Analysis of WLAN's received signal strength indication for indoor location fingerprinting. *Pervasive Mob. Comput.* **2012**, *8*, 292–316. [CrossRef]
10. Brena, R.F.; García-Vázquez, J.P.; Galván-Tejada, C.E.; Muñoz-Rodríguez, D.; Vargas-Rosales, C.; Fangmeyer, J. Evolution of Indoor Positioning Technologies: A Survey. *J. Sens.* **2017**, *2017*, 359. [CrossRef]
11. Samadh, S.A.; Liu, Q.; Liu, X.; Ghourchian, N.; Allegue, M. Indoor Localization Based on Channel State Information. In Proceedings of the 2019 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNet), Orlando, FL, USA, 20–23 January 2019; pp. 1–4.
12. Wu, K.; Xiao, J.; Yi, Y.; Chen, D.; Luo, X.; Ni, L. CSI-Based Indoor Localization. *IEEE Trans. Parallel Distrib. Syst.* **2013**, *24*, 1300–1309. [CrossRef]
13. Wang, X.; Gao, L.; Mao, S.; Pandey, S. CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach. *IEEE Trans. Veh. Technol.* **2017**, *66*, 763–776. [CrossRef]
14. Altintas, B.; Serif, T. Improving RSS-Based Indoor Positioning Algorithm via K-Means Clustering. In Proceedings of the 17th European Wireless 2011—Sustainable Wireless Technologies, Vienna, Austria, 27–29 April 2011; pp. 1–5.
15. Zhang, Y.; Cui, Q.; Zhang, P.; Tao, X. Channel State Information Feedback with Zero-Overhead in Closed-Loop MIMO System. *IEICE Trans. Commun.* **2010**, *93*, 9–15. [CrossRef]
16. Linux 802.11n CSI Tool. Available online: <https://dhalperi.github.io/linux-80211n-csitool/> (accessed on 20 October 2019).
17. Ohtsuki, T.; Hong, J. Activity Recognition Based on Array Sensor. In Proceedings of the APSIPA ASC 2011, Xi'an, China, 18–21 October 2011.
18. Yasukawa, S.; Kim, M. Intruder Detection Using Radio Wave Propagation Characteristics. In Proceedings of the ICCE-Asia 2018, Jeju, Korea, 24–26 June 2018.
19. *IEEE Standard for Information Technology—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput*; IEEE Std 802.11n-2009 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, and IEEE Std 802.11w-2009); IEEE Std: Piscataway, NJ, USA, 29 October 2009; pp.1–565.
20. Yasushi, T.; Yusuke, A.; Yasuhiko, I. MIMO Technologies in Wireless LAN Systems. *J. Inst. Image Inf. Telev. Eng.* **2016**, *51*, 35–40.
21. Xie, Y. Atheros CSI Tool. Available online: <https://wands.sg/research/wifi/AtherosCSI/> (accessed on 10 October 2019).
22. Zhu, H.; Zhuo, Y.; Liu, Q.; Chang, S. π -Splicer: Perceiving Accurate CSI Phases with Commodity WiFi Devices. *IEEE Trans. Mob. Comput.* **2018**, *17*, 2155–2165. [CrossRef]
23. Lab of Innovation on Networking. π -Splicer. Available online: <http://lion.sjtu.edu.cn/project/projectDetail?id=5> (accessed on 10 October 2019).
24. Tadayon, N.; Rahman, M.; Han, S.; Valaee, S.; Yu, W. Decimeter Ranging with Channel State Information. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 3453–3468. [CrossRef]
25. Junction, T. BL41-6203-00, Orient Microwave Corp. Available online: <http://www.orient-microwave.com/index.html> (accessed on 10 October 2019).
26. Coaxial Cable Enviroflex_316. Huber+Shuner AG Corp. Available online: <http://hubersuhner.com/en/> (accessed on 20 September 2019).
27. Vapnik, V.N.; Chervonenkis, A.Y. Pattern recognition using generalized portrait method. *Autom. Remote Control* **1963**, *24*, 774–780.
28. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; pp. 144–152.
29. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. Available online: <https://doi.org/10.1023/A:1010933404324> (accessed on 25 October 2019). [CrossRef]
30. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
31. Python Software Foundation. Welcome to Python.org. Available online: <https://www.python.org/> (accessed on 30 October 2019).
32. Preferred Networks, Inc. Optuna—A Hyper Parameter Optimization Framework. Available online: <https://optuna.org> (accessed on 30 October 2019).