

# Device-Free Multi-Person Indoor Localization Using the Change of ToF

Atsushi Nomura  
Tokyo Institute of Technology  
Tokyo, Japan  
nomura@miubiq.cs.titech.ac.jp

Masato Sugasaki  
Tokyo Institute of Technology  
Tokyo, Japan  
sugasaki@miubiq.cs.titech.ac.jp

Kota Tsubouchi  
Yahoo Japan Corporation  
Tokyo, Japan  
ktsubouc@yahoo-corp.jp

Nobuhiko Nishio  
Ritsumeikan University  
Shiga, Japan  
nishio@cs.ritsumei.ac.jp

Masamichi Shimosaka  
Tokyo Institute of Technology  
Tokyo, Japan  
simosaka@miubiq.cs.titech.ac.jp

**Abstract**—Indoor position information, which is difficult to obtain by GPS, can be used for various services and applications, thus indoor localization methods have been widely researched. Among them, device-free indoor localization does not require the target persons to possess a localization device, such as a smartphone, which can support the localization of all persons in the environment. Many methods using RSSI and Wi-Fi CSI have been proposed as device-free indoor localization using radio waves. However, RSSI is easily affected by environmental factors, such as multipath propagation. Wi-Fi CSI also has the disadvantage that there is no standard, so it relies on special hardware and software. Therefore, we propose device-free multi-person indoor localization using ToF information, which is less susceptible to noise. ToF information in indoor localization has mainly been used for highly accurate estimation of the receiver position, while this paper proposes a new application. In addition, distance measurements using ToF of Wi-Fi and UWB have been standardized and can be implemented with commercial equipment, which is highly practical. In this research, we considered the change in the distance measurement due to radio wave occlusion by a human body and built a device-free indoor localization system that can estimate multi-person position even if the model is trained on the data of only one person.

**Index Terms**—Wi-Fi, UWB, Round Trip Time, Time of Flight, device-free indoor localization, indoor positioning, density estimation

## I. INTRODUCTION

Indoor position information, which is difficult to obtain by GPS, can be used for various services and applications, so indoor localization methods have been widely researched. Among indoor localization methods, many methods have been proposed that use radio wave information received from access points by the devices that the targets possess [1]–[5]. However, these methods cannot localize persons who do not have such devices or have not installed the corresponding application. Therefore, device-free indoor localization, which does not require the targets to possess a localization device, is a very important method given a lack of localization.

Currently, cameras are often used for device-free indoor localization in the real world. However, there are places not suitable for installation of a camera from the viewpoint of

privacy. Therefore, many methods using radio waves to solve the privacy issue have been proposed.

For example, methods that use changes of RSSI (Received Signal Strength Indicator) due to radio wave interference caused by humans have been proposed [6], [7]. However, they are very susceptible to environmental disturbance and have low accuracy.

In recent years, many methods using Wi-Fi CSI (Channel State Information) have been proposed [8], [9]. CSI responds sensitively to changes in radio waves in the environment, so it is possible to estimate persons' positions with high accuracy. However, there is no standardization for CSI, so we have to rely on special equipment and software to acquire it in a practical setting.

In this paper, we propose a device-free multi-person indoor localization method using distance calculated by ToF (Time of Flight) information, which is less susceptible to noise. The RTT (Round Trip Time) protocol can calculate the distance from the communication time between the transceiver and receiver with low noise. IEEE 802.11mc standardizes distance measurement using Wi-Fi as Wi-Fi RTT. In addition to Wi-Fi, distance measurement is also possible with UWB (Ultra Wide Band), which is standardized by IEEE 802.15.4a. In this manner, the measurement using the RTT protocol is standardized and does not require special equipment or software, so it is more practical than the CSI-based method.

Using this protocol, many researchers have achieved accurate indoor localization with devices that the target persons possess [10]–[13]. However, in this manner, the use of ToF is mainly focused on high-precision estimation of receiver position, and device-free indoor localization with it has not been discussed. This paper proposes a new application of ToF information to device-free indoor localization.

However, it is not clear how the distance measurement results change due to human radio wave interference or how they can be used for localization. In this research, we constructed a localization system after considering the change.

The contributions of this study can be summarized as follows:

- We investigated the changes due to radio wave occlusion caused by a human body in the measured distance and signal strength of Wi-Fi and UWB.
- We propose a Gaussian Median-Max feature as a function to appropriately express the degree of human presence and absence between antennas.
- We built a localization model that can estimate multi-person positions by learning only one person's data in the environment.
- We performed device-free indoor localization with the proposed model and evaluated its accuracy.

#### *Related work*

*Device-based indoor localization with radio wave:* As indoor localization methods using devices owned by target persons, there are methods using the strength of radio waves called RSSI (Received Signal Strength Indicator). Wi-Fi RSSI-based indoor localization has been researched for the last 20 years [1]–[5]. To localize the target devices in this framework, a model matches an input fingerprint, which is a vector of RSSIs, to the training data. This framework has been actively explored owing to its practicality and the availability of Wi-Fi signals in many environments.

ToF distance measurement technologies are being considered as new protocols for indoor localization in recent years [10], [12]–[15]. Both Wi-Fi RTT and UWB are standardized by IEEE 802.11mc and IEEE 802.15.4a, so they can be obtained easily with commercial equipment. ToF measurement is also known to have less noise than RSSI.

However, device-based indoor localization methods cannot localize persons who do not have devices such as smartphones or have not installed the corresponding application.

*Device-free localization with Wi-Fi RSSI and CSI:* As device-free indoor localization methods using the changes of Wi-Fi radio waves, methods using changes in RSSI due to human interference [6], [7] have been researched. Additionally, there are methods using Wi-Fi CSI (Channel State Information) [9], [16]–[18], which is detailed information on radio waves, such as the amplitude and phase difference of radio waves acquired by multiple antennas of Wi-Fi transceivers and receivers. In particular, CSI can respond to changes in radio waves in the environment sensitively, so its use has been researched not only for high-precision localization of persons but also for various fields, such as human activity recognition [19]. However, RSSI is very susceptible to environmental factors, such as multipath propagation, and handling the noise becomes a problem. In addition, there is no standardization for CSI, so we have to rely on special equipment and software to acquire it in practical use.

*Applied technology for human detection using ToF information:* There have been several studies that have applied ToF information to human detection [20], [21]. Choi et al. proposed a system that detects persons passing through an entrance by measuring distances with a UWB transceiver and

receiver attached to the ceiling [20]. They also constructed a system to detect persons in a small area [21]. Although these methods have proposed new application methods of detecting the presence of persons, they are limited to the detection of one person in a small area. Therefore, they cannot be applied to multi-person localization or the estimation of congestion.

In this paper, we propose device-free multi-person indoor localization using ToF information.

## II. DISTANCE MEASUREMENT WITH TOF INFORMATION

### A. Overview of Wi-Fi RTT and UWB

Wi-Fi RTT is a protocol standardized by IEEE802.11mc that measures the distance between a Wi-Fi access point and Wi-Fi device based on ToF of Wi-Fi radio waves. Because it uses time information, it is known to be less susceptible to noise than Wi-Fi RSSI. Standardization has been done, so there are merits that Wi-Fi CSI does not have, such as being able to collect on general Android devices.

UWB (Ultra Wide Band) is a wireless technology and is standardized by IEEE 802.15.4a. It uses signals distributed over a wide band and can measure distance from ToF in the same way as Wi-Fi RTT. It is often used for localization and radar. In recent years, UWB antennas have also been installed on the iPhone. The output of the transmitted UWB radio wave is weak, and it is only noise for other radio waves, so there is little problem of radio wave interference.

These ranging techniques are often used for accurate receiver position estimation in the field of indoor localization [10], [12]–[15].

### B. How to apply ToF information to human detection and its concerns

Both Wi-Fi RTT and UWB use the communication time between antennas for distance measurement. Due to radio wave diffraction and reflection, ToF is longer in the NLoS (None Line of Sight) state where there is an obstacle between the antennas than in the LoS (Line of Sight) state where there is nothing between them. As a result, the measured distance also becomes longer. By focusing on the change in the measured distance, it becomes possible to determine whether there is a person between the antennas or not. In our research, we realize device-free indoor localization by applying this.

However, there are concerns about the use of this ranging result for device-free indoor localization. Figure 1 summarizes the results of 50 distance measurements of 7 m with Wi-Fi RTT (left) and UWB (right) in a boxplot. Measured distance by Wi-Fi RTT has a variation of approximately 1 m even when there is no obstacle, as shown in the left of Figure 1. This is because Wi-Fi RTT is susceptible to multipath. Multipath is a phenomenon in which transmitted radio waves pass through multiple propagation paths and are received multiple times. If it is susceptible to this, there is a problem that the radio waves that have passed through multiple paths are recognized as being slightly shifted and overlapped due to the difference in arrival time, and variation occurs in measured values. There is

concern that this measurement variation may lead to a decrease in localization accuracy.

Meanwhile, UWB radio wave is strong against multipath, so the error is very small, as shown on the right of Figure 1. However, UWB is generally known to be able to measure accurate distance even if there are obstacles, and it is not clear whether it can be used for human detection between antennas.

Therefore, in this research, after verifying how much the measured distances of Wi-Fi RTT and UWB change due to radio wave occlusion by human (Section IV-A), we constructed a localization system.

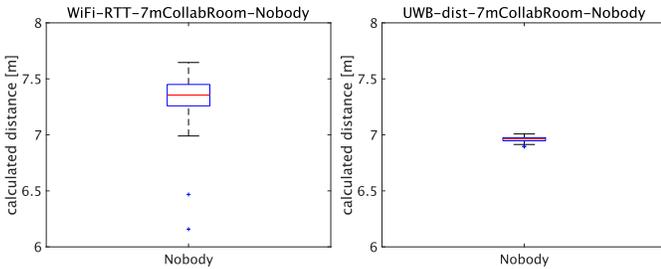


Fig. 1: 50 distance measurements of 7 m with Wi-Fi RTT (left) and UWB (right)

### III. DEVICE-FREE MULTI-PERSON INDOOR LOCALIZATION

#### A. Flow of proposed localization system

The localization model proposed in this research consists of three important points: environment sensing, data featurization, and density-based device-free localization. Figure 2 shows the flow of our ToF-based device-free indoor localization. First, we monitor the target environment with installed transceivers and receivers by determining whether radio waves are occluded for each transceiver-receiver pair. The ToF measured distance is affected if a direct path (line of sight (LoS)) changes to an indirect path (non line of sight (NLoS)) because of occlusion by a human body. To detect the location of an individual, we combine the NLoS information of the transceiver-receiver pairs to observe the locations of target persons.

Then, we extract their possible locations from the combination. However, we cannot explicitly determine whether the signal state is LoS or NLoS because we do not know the LoS and NLoS ranging value explicitly. The ranging value contains errors due to offsets depending on the devices, measurement noise, environment structure effect, and different deterioration values caused by difference of person. To handle ranging noise, we encode how likely it is for the ranging value to be LoS or NLoS by using median or maximum ranging results in a training dataset for each transceiver-receiver pair. Details are described in Section III-D.

Finally, for multi-person localization, we estimate the target location with a density estimation model. Here, although regression directly models the location and provides an accurate estimation, it cannot model locations of multiple persons. In this research, to match the localization of multiple persons,

we define the localization model as a density estimation. We estimate the density for each small mesh, which divides a target environment. The estimated multi-person location is obtained by thresholding the density estimation results. This localization procedure does not place an upper bound on the number of estimated persons. Details of the model are described in Section III-C.

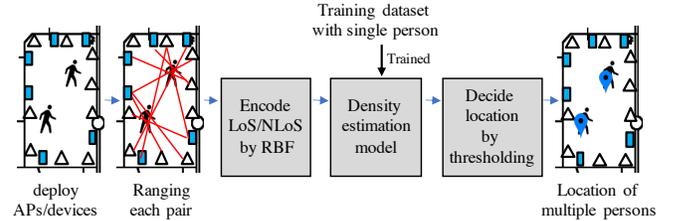


Fig. 2: Flow of ToF-based indoor localization

#### B. Scheme of device-free multi-person indoor localization

The cross point of multiple NLoS signals is the location where a person is most likely to be. Figure 3 illustrates an example of the person detection scheme; Figure 3 (a) shows an environment. Figure 3 (b) shows the signal properties when one person is in the environment. The two ranging results become NLoS because LoS is occluded by the person. From the ranging results, the location where the person is most likely to be is the lower-left one.

We perform multi-person sensing in the same way as the single person case. In Figure 3 (c), the NLoS signals cross at the upper-right location in addition to the lower-left location of the Figure 3 (b) setting. From the ranging result, the locations where the persons are most likely to be are the upper-right and lower-left. We can differentiate the one-person setting from the multi-person setting with NLoS combinations in this manner.

Here, we also discuss the position of the receivers and transceivers placed in the environment. When two or more signals are occluded by the target person, the position of the person will be most likely estimated to be the intersection of the paths. Therefore, it is essential that there are at least two intersections of direct paths in the target mesh. Figure 11 shows an example of the localization environment. In Figure 11, the intersections of the direct paths are evenly distributed in the environment, and the meshes for localization also have them. Therefore, the environment is suitable for our indoor localization.

#### C. Device-free indoor localization model

This section describes the details of the localization model. We use changes in ToF measurements to estimate the positions of persons. However, it involves noise due to the number of persons in the environment and the passage of time. In addition, as it is not clear how much the ranging results change depending on occlusion by a human body, it is not possible to explicitly acquire the state between antennas. In other words, it is necessary to deal with the changes due to

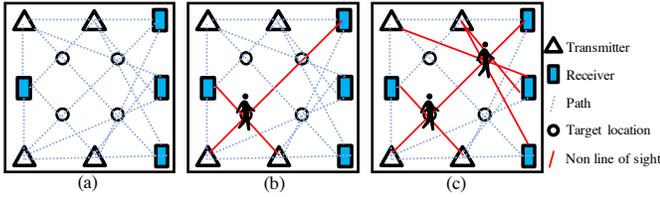


Fig. 3: Scheme of device-free multi-person indoor localization: (a) environment with four target locations, (b) NLoS when one person is at lower-left location, (c) NLoS when one person is at lower-left location and one person is at upper-right location

the occlusion as the probability of human presence, rather than directly detecting human presence at the intersection of direct paths. To construct a localization model that is robust against the noise, we propose a density estimation model. It is also possible to deal with multi-person localization by using a density estimation model.

As shown in Figure 11, we place  $N^{(T)}$  transceivers and  $N^{(R)}$  receivers in the environment and use the ranging results obtained from the pairs as input to construct a density estimation model. We define the measurement results such as ToF measured distance and RSSI as  $x_{(i,j)}$  acquired by the pair of transceiver  $i$  and receiver  $j$ . The localization model is expressed by Equation 1. In the localization model, the vector  $\mathbf{y} = [y_1, \dots, y_{|\mathcal{L}|}]^T$ , which indicates the density of persons in each localization target mesh, is the objective variable. We use  $\phi(\mathbf{x})$  as an explanatory variable.  $\phi$  is a function, which featurizes a vector of measurement results  $\mathbf{x} = [x_{(1,1)}, x_{(1,2)}, \dots, x_{(N^{(T)}, N^{(R)})}]^T$  acquired by each pair.

$$\mathbf{y} = \mathbf{W}\phi(\mathbf{x}). \quad (1)$$

$\mathcal{L}$  is a set of localization target meshes,  $y_i$  [number of persons /  $m^2$ ] is a density of persons in the mesh with a size of  $1 \times 1$  m, and  $\mathbf{W}$  is a weight matrix. The design of the feature function  $\phi$  is discussed in Section III-D, the optimization of the weight  $\mathbf{W}$  is discussed in Section III-E, and the inputs suitable for localization are discussed in Section IV-A.

After estimating density of persons in each target mesh, using the threshold  $a$  determined by Equation 2, the positions of the target persons are estimated by the threshold function  $h$  of Equation 3.

$$a = \theta(\max(\mathbf{y}) - \min(\mathbf{y})) + \min(\mathbf{y}), \quad (2)$$

$$h(y_i) = \begin{cases} 1 & (y_i > a) \\ 0 & (\text{otherwise}), \end{cases} \quad (3)$$

where  $\theta$  is a real number such that  $0 < \theta < 1$ , and  $\min(\mathbf{y})$  and  $\max(\mathbf{y})$  are the minimum and maximum values of the elements of vector  $\mathbf{y}$ , respectively. As a result of the threshold processing, it is estimated that the target person exists in mesh  $l_i$  where  $h(y_i) = 1$ .

In this model, multi-person localization can be realized only by learning the training data in which one person exists in the environment. This is achieved by defining the model for each target mesh and learning the density of persons independently in each mesh for an input  $\mathbf{x}$ . In other words, if the training data on all mesh contains human existing data, device-free indoor multi-person localization can be achieved by learning only one person's data in the environment.

#### D. Formulation of features considering human occlusion

To estimate the locations of the target persons with ToF measured distance, we need to know whether the signal state is LoS or NLoS. However, we cannot determine LoS or NLoS directly from the ranging result for each transceiver-receiver pair. This is because the ranging value has measurement noise, variations due to differences in each person, environmental structures, etc. To handle this problem, we propose a novel feature representation using the median and max values of the training dataset, which are likely to correspond to LoS and NLoS conditions.

If we look at the amount of LoS and NLoS data in the training dataset for each transceiver-receiver pair, most of the data is LoS, and a small amount is NLoS. This stems from the fact that only a few signals are occluded in the dataset acquired under the single-person condition. From this fact, the median value of the dataset for each pair reflects the LoS condition. In comparison, under the NLoS condition, as discussed in Section IV-A, the ranging result becomes longer than the expected ranging result. From this fact, the max value of the dataset for each pair is considered to be acquired under the NLoS condition.

Therefore, we propose a Gaussian Median-Max feature, which uses median and maximum ranging values for each transceiver  $j$  and receiver  $k$  pair:

$$\phi^{(M)}(x_{(j,k)}) = \exp\left(\frac{-(x_{(j,k)} - \tilde{\mu}_{(j,k)})^2}{2\sigma^2}\right), \quad (4)$$

$$\phi^{(U)}(x_{(j,k)}) = \exp\left(\frac{-(x_{(j,k)} - \mu_{(j,k)}^{(U)})^2}{2\sigma^2}\right), \quad (5)$$

where  $\tilde{\mu}_{(j,k)}$  is the median of ranging results between transceiver  $j$  and receiver  $k$ , and  $\mu_{(j,k)}^{(U)}$  is the maximum ranging result between transceiver  $j$  and receiver  $k$ .  $\sigma$  is defined as 1.0 when Wi-Fi RTT ranging results are used and 0.1 when UWB ranging results are used. Figure 4 shows an example of featurization on the Gaussian Median-Max feature.

The featurization function for the LoS signal (Equation 4) returns a value of nearly 1.0 when the input is nearby the median value and nearly 0.0 when the input is far from the median value. The featurization function for the NLoS signal (Equation 5), in a similar way as the LoS case, returns a value of nearly 1.0 when the input is nearby the maximum value and nearly 0.0 when the input is far from the maximum value. This featurization function can sufficiently express whether the radio wave is occluded or not.

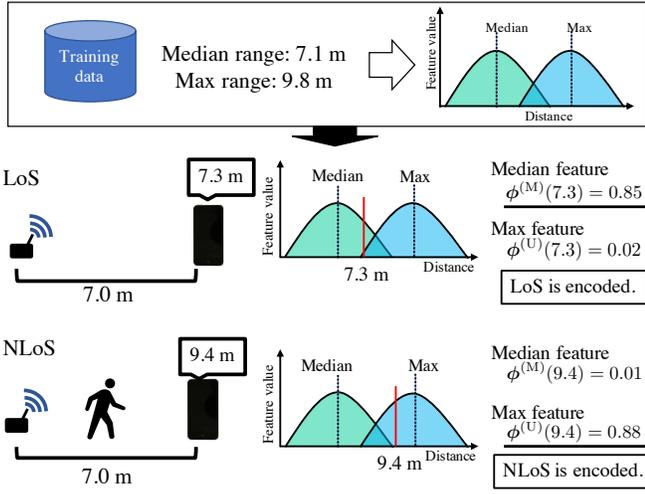


Fig. 4: Example of featurization on Gaussian Median-Max feature

#### E. Parameter optimization of the proposed system

In this research, we used the dataset  $\mathcal{D} = (\mathbf{x}, \mathbf{y})$  composed by a vector of ranging results  $\mathbf{x} = [x_{(1,1)}, x_{(1,2)}, \dots, x_{(N^{(T)}, N^{(R)})}]^T$ , which is acquired by the transceiver-receiver pair, and a vector  $\mathbf{y} = [y_1, \dots, y_{|\mathcal{L}|}]^T$  indicating the density of persons in each target mesh to design the density estimation model. To optimize this model using the dataset  $\mathcal{D}$ , we define the L2 regularized least square regression loss function for each location as

$$\arg \min_{\mathbf{W}} \frac{1}{2} \|\mathbf{W}\phi(\mathbf{x}) - \mathbf{y}\|^2 + \lambda \|\mathbf{W}\|_2^2, \quad (6)$$

where  $\mathbf{W}$  is the weight of each featurized data, and  $\lambda$  is the coefficient of the regularization term.  $\phi$  uses both median and max functions discussed in Section III-D, and it is defined as

$$\phi(\mathbf{x}) = [\phi^{(M)}(x_{(1,1)}), \phi^{(M)}(x_{(1,2)}), \dots, \phi^{(M)}(x_{(N^{(T)}, N^{(R)})}), \phi^{(U)}(x_{(1,1)}), \phi^{(U)}(x_{(1,2)}), \dots, \phi^{(U)}(x_{(N^{(T)}, N^{(R)})})]^T. \quad (7)$$

We construct the model by solving this optimization problem.

### IV. PERFORMANCE EVALUATION

This chapter summarizes the verifications of Wi-Fi and UWB measurements to realize robust device-free indoor localization. In addition, we evaluate the localization accuracy of the proposed method.

#### A. Verification of measurements for robust localization

In this section, we summarize the measurement results acquired by Wi-Fi and UWB transceiver-receiver pair in the situation when there is a person or not between the antennas.

We collected data in two places: a corridor about 2 m wide and a room of about  $5 \times 7$  m. We used Google Wi-Fi as

a Wi-Fi transceiver (access point) and the Pixel3 is Android device as a Wi-Fi receiver. For UWB, we used EVK1000, which is a two-way ranging evaluation kit made by Decawave, as transceiver and receiver. The equipment was installed on a tripod, the Wi-Fi devices were set at 1.3 m from the floor, and the UWB devices were 1.1 m from the floor so that a standing person could block the direct waves sufficiently. The equipment was set on the tripod 7 m apart, with a person stood upright on the straight line connecting the antennas, as shown in the Figure 5, and we collected the data.



Fig. 5: An example of data acquisition in a corridor (left) and a sufficiently large room (right)

1) *Change in ToF measured distance:* First, we summarize the distance measurement results calculated by ToF. The results of each 50 measurements are plotted in a boxplot with the distance from the receiver to the obstacle ( $x$  in Figure 6). The horizontal axis is changed by 0.5 m except for "Nobody", which shows the results when there is no person between the antennas.

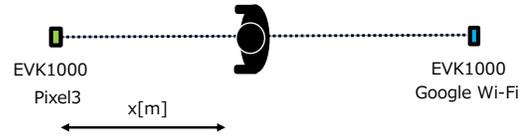


Fig. 6: View from top of the data acquisition environment

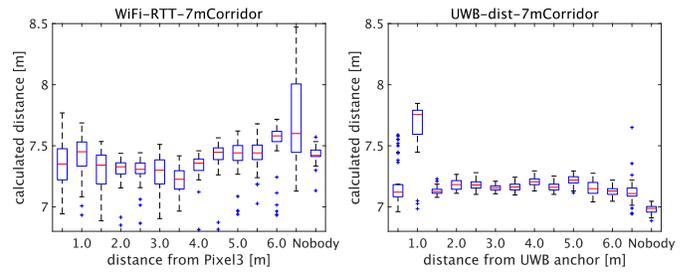


Fig. 7: Wi-Fi RTT (left) and UWB (right) ToF distance measurements in the corridor

Figure 7 shows the results of ToF distance measurements acquired in the corridor. As shown in the left of Figure 7, the ranging results by Wi-Fi RTT have a large variation, and it is observed that the phenomenon that the measured distance is extended due to human occlusion is not so much. Meanwhile, as shown in the right of Figure 7, the ranging results by UWB have less data variation than the results of Wi-Fi RTT, and the

phenomenon that the measured distance is extended due to human occlusion is observed clearly.

Next, we summarize the ranging results acquired in the room with a sufficient distance to the wall. In the same manner, the results of each 50 measurements are plotted on a boxplot. Figure 8 shows the results of ToF distance measurements acquired in the room.

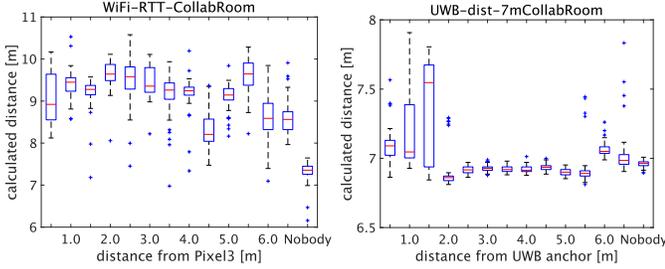


Fig. 8: Wi-Fi RTT (left) and UWB (right) ToF distance measurements in the sufficiently large room

As shown in the left of Figure 8, although the ranging results of Wi-Fi RTT have a large variation, the results are greatly increased by the human body compared to the results in the corridor. Meanwhile, as shown in the right of Figure 8, although the ranging results of UWB change at the point near the antennas of UWB, there is almost no change at the points far away.

As described above, the results vary greatly depending on the environment in which the data are acquired. The reason for this is considered as follows.

First, the occlusion of a human increases the number of radio waves that pass through paths other than the direct path, regardless of radio waves and location. Among them, in UWB, an increase in the measured distance was observed in the corridor, and there was no change in the large room. This is because the arrival time of UWB radio waves varies only slightly in the corridor, where the length of the path of direct wave and the path of the reflected wave do not differ much. Therefore, it is difficult to distinguish between direct and reflected waves, and misrecognition occurs, so the measured distance increased. On the contrary, in the room where the reflected path is sufficiently longer than the direct path, it is easy to distinguish the direct wave, so the increase in the measured distance was not observed.

In contrast, Wi-Fi is inherently difficult to distinguish between direct and reflected waves, so the measured distance increased significantly in the room where the distance to the wall is long and the paths are often long.

From these results, it can be seen that the distance measurement by ToF increases due to radio wave occlusion by the human body. However, the increases are not remarkable depending on the environment.

2) *Change in direct wave attenuation:* Next, we consider the attenuation of radio waves by the human body. As can be seen from the previous results of distance measurement, distance measurement using UWB is robust against multipath

because it is easy to distinguish direct waves. This property also means that we can easily obtain data on the radio waves that have passed through the direct path.

Here, we focus on FSL, which is a UWB measurement value that indicates the estimated strength of the first received radio wave. The radio wave that arrives at a receiver first passes through the shortest path. If there is a person on the direct path, the first radio wave should be received after being attenuated by the human body. Therefore, it would be possible to determine the presence or absence of a person by measuring how much FSL weakened in comparison with RSL, which is a UWB measurement value of the maximum radio wave intensity.

In the same way as the distance measurements, we plot the difference between RSL and FSL measurements (RSL-FSL [dBm]) for 50 times on a boxplot. Figure 9 shows the results.

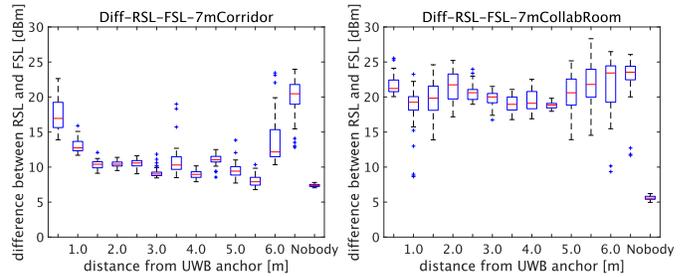


Fig. 9: Difference between UWB RSL and FSL in the corridor (left) and large room (right)

As shown in the corridor result on the left of Figure 9, it is observed that the difference when there is a person is larger than when there is no person. This is because the occlusion of the human body weakens the direct wave. The same tendency is observed in the result of the room shown in the right of Figure 9. The results show that the difference between RSL and FSL does not depend on the environment and can be used as a feature value to determine the presence of a person.

Next, we also consider Wi-Fi RSSI. Figure 10 shows the result of plotting 50 RSSI measurements on a boxplot in the same way.

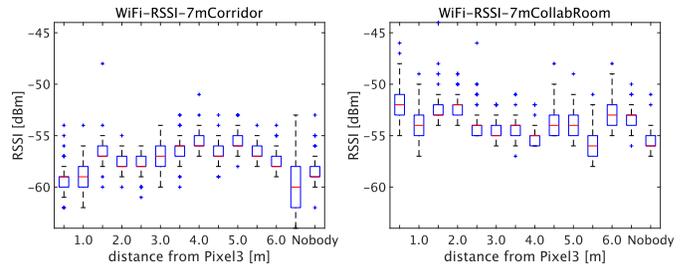


Fig. 10: Wi-Fi RSSI in the corridor (left) and large room (right)

Here, unlike the plots so far, RSSI is attenuated by the presence of persons compared to "Nobody". In other words, it should be noted that the measured value is expected to be small when there is a person. As shown in Figure 10,

in both environments, although the RSSI value is attenuated when there is a person near the antenna, it is observed that there is little change when there is a person at the central part between the antennas. Therefore, Wi-Fi RSSI is observed to be unsuitable for device-free indoor localization.

From these results, accurate localization can be achieved by using the difference between RSL and FSL as a feature quantity in addition to the ToF measured distance.

### B. Experiment to evaluate localization accuracy

In this experiment, we constructed an indoor localization system using the ToF measured distance by Wi-Fi RTT and UWB as feature values, and performed it to compare the localization accuracy. In addition, we compare with Wi-Fi RSSI device-free localization using the featurization function 8 proposed by Youssef et al. [6]. The featurization function is defined as

$$\phi^{(A)}(x_{(j,k)}) = \left| \frac{x_{(j,k)} - a_{(j,k)}}{a_{(j,k)}} \right|, \quad (8)$$

where  $x_{(j,k)}$  is RSSI acquired by receiver  $j$  and transceiver  $k$  pair, and  $a_{(j,k)}$  is the average of RSSI acquired by receiver  $j$  and transceiver  $k$  pair in the training data.

We also evaluated the performance of localization using both ToF measured distance and signal strength. We define  $\sigma$  in Equations 4 and 5 as 1.5 for using Wi-Fi RSSI and 1.0 for the difference between UWB FSL and RSL.

1) *Experimental settings*: The localization experiment was performed in a room about  $5 \times 7$  m. As shown in Figure 11, when using Wi-Fi, we set up 9 Google Wi-Fis as access points and 6 Pixel 3s as receiving devices by the wall. When using UWB, we set up 9 EVK1000s as transceivers and 6 EVK1000s as receivers at the same position. In addition, a total of 35  $1 \times 1$  m meshes were set as the localization target meshes. As discussed in Section III-B, these radio transceivers, receivers, and the meshes are arranged so that at least two direct paths (blue lines in Figure 11) pass through each mesh. Additionally, in this experiment, we set  $\theta$  in Equation 2 to 0.8.

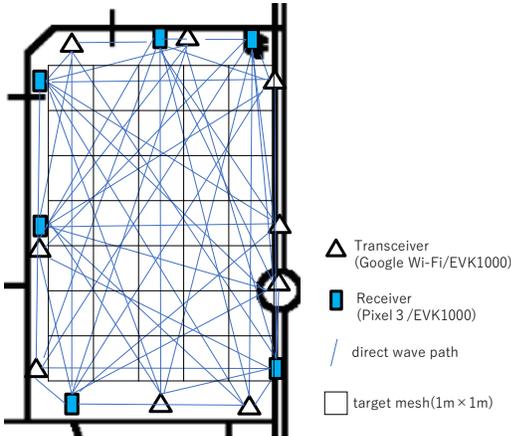


Fig. 11: Plan view of the indoor localization environment

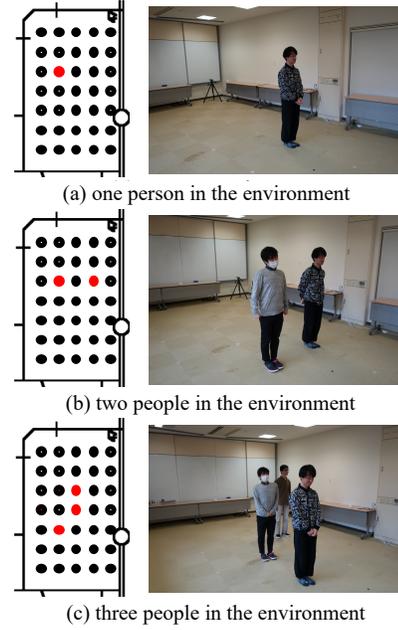


Fig. 12: Data acquisition patterns (left: positions of the target persons in the environment; right: photo of corresponding setting)

2) *Acquired data*: We acquired data for Wi-Fi and UWB on separate days. We acquired each data with one to three target persons in the environment. The target persons stood upright in the center of the target mesh. Figure 12 shows examples of data acquisition in the environment. We acquired a total of 35 patterns for each target mesh as data of one person. The data of two persons were acquired in a total of 32 patterns, 1 m and 2 m apart, at the center, left and right columns. We also acquired 52 data of three-person case at randomly selected target meshes. We set zero or one person per mesh and acquired 40 samples for each pattern.

3) *Evaluation metrics*: We used three metrics to evaluate multi-person localization as follows.

*Mean Estimated Number Error (MNE)*: We evaluated the difference between the number of estimated persons and that of actual persons using the mean estimated number error. We define MNE as

$$\frac{1}{N} \sum_i^N |n_i - \sum h(\hat{y})|, \quad (9)$$

where  $N$  is the number of data,  $n_i$  is the number of persons in data  $i$ ,  $h$  is a threshold function defined as Equation 3, and  $\hat{y}$  is estimated human density vector.

*Mean Average Estimated Location Error (MALE)*: We evaluated the localization accuracy with error between the estimated and actual locations, which is represented as the mean of the average estimated location error for the location

of each actual person. We defined MALE as

$$\frac{1}{N} \sum_i \frac{1}{|\mathcal{L}_i|} \sum_j \min ||l_{(i,j)} - r(h(\hat{\mathbf{y}}))_k||, \quad (10)$$

where  $N$  is the number of data,  $\mathcal{L}_i$  is a set of meshes of the target persons' position in data  $i$ ,  $h$  is a threshold function defined as Equation 3, and  $\hat{\mathbf{y}}$  is an estimated human density vector. Additionally,  $r$  is a mapping function from the vector indicating the density of persons in each mesh after threshold processing to a mesh set of the positions where targets are.

We define the estimated location error as the distance from the actual position of each target to the closest estimated position.

*Mean Estimated Number on Each Location Error (MNELE):* We computed the average estimation error of the number density of persons for each target mesh. We defined MNELE as

$$\frac{1}{N|\mathcal{L}|} \sum_i |y_i - \hat{y}_i|, \quad (11)$$

where  $N$  is the number of data,  $\mathcal{L}$  is a set of all localization target meshes,  $y_i$  is a vector representing the density of persons in each mesh, and  $\hat{y}_i$  is estimated human density vector.

4) *Evaluation results:* We conducted the localization experiment by learning only one person's data in the environment. However, the training data was separated from the test data. The experimental results are shown in Tables I, II, III. Figure 13 shows an example of density estimation and localization estimation results plotted on a plan view that simulates the localization environment of a  $5 \times 7$  grid. Additionally, "combined" in the tables mean the result of using both ToF measured distance and signal strength.

First, we compare the accuracy of the method using Wi-Fi RSSI [6] with the proposed method using ToF measured distance. As shown in Tables I, II, III, focusing on MNE and MALE, which are evaluation indices after threshold processing, the proposed methods are much more accurate than Wi-Fi RSSI. Here, focusing on MNELE, which evaluates the density estimation result before threshold processing, it is observed that it takes a relatively small value in the result of Wi-Fi RSSI. As shown in the results of density estimation in Figure 13 (a), the reason is that the average value of density estimation for all meshes is small when featurization is performed by the Equation 8. Because there is nobody in most of the 35 meshes in this experiment, MNELE becomes small in localization with Equation 8, which estimated the absence of human close to 0 in density. In addition, when comparing the results of the method using Wi-Fi RTT ToF with UWB ToF, the accuracy of localization using UWB is better in all evaluation metrics. From this, it can be concluded that UWB is more suitable than Wi-Fi for device-free indoor localization using ToF.

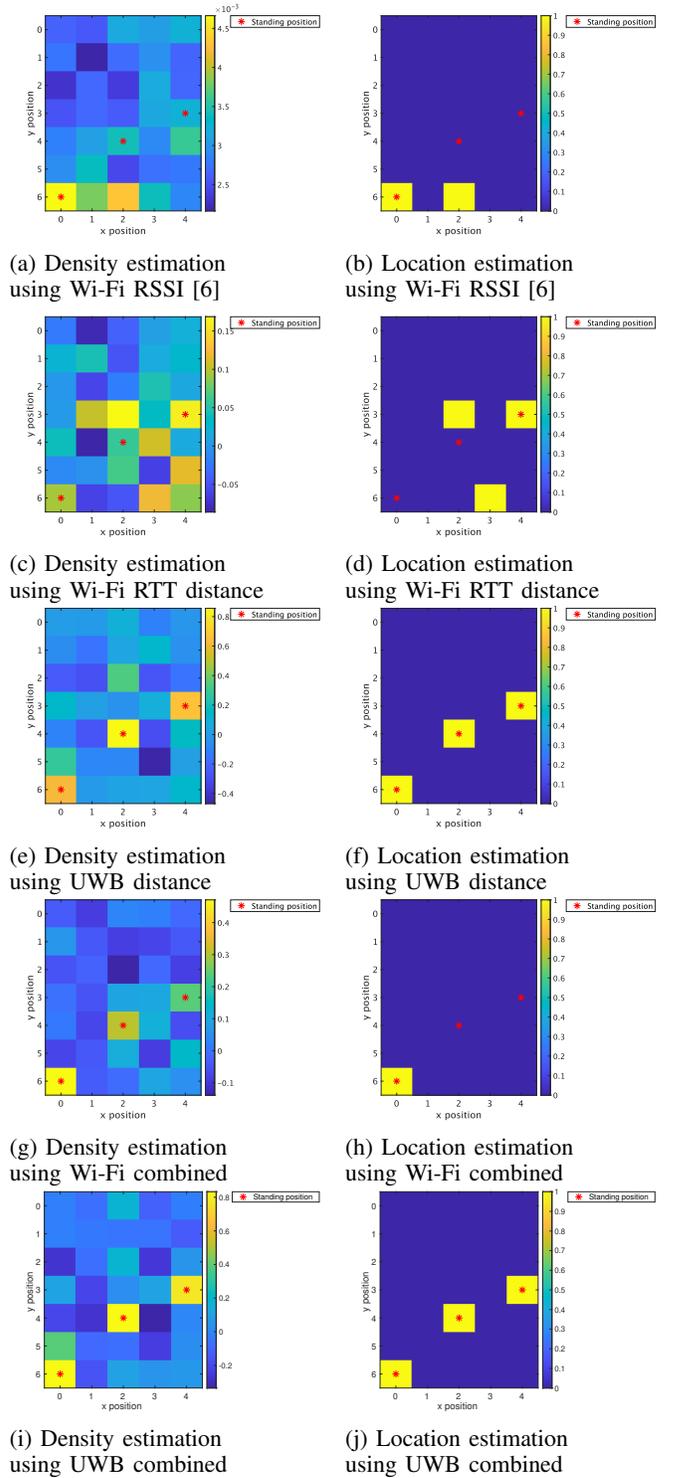


Fig. 13: Examples of three-person density estimation (left) and location estimation (right) for each method (red marks are actual persons' locations, yellow squares are estimated locations)

Next, we focus on the results when combining ToF measured distance and signal strength. Comparing the results using Wi-Fi RTT measured distance with Wi-Fi combined use in Tables I, II, III, although there is little change in the localization

for two or three people, a significant improvement in accuracy was observed in the localization for one person. Meanwhile, comparing the results using UWB measured distance with UWB combined use in Tables I, II, III, in general, the accuracy improved, and in the localization of one person, MNE and MALE were 0, that is, the accuracy was high enough to measure correctly in all test cases. From this result, we can conclude that the combined use of signal strength is very useful for improving the accuracy of ToF-based device-free localization.

TABLE I: Localization accuracy when there is one person in the environment

Feature	MNE	MALE	MNELE
Wi-Fi RSSI [6]	1.13	2.10	0.0013
Wi-Fi RTT distance	1.06	1.02	0.0019
UWB distance	0.07	0.08	<b>0.0005</b>
Wi-Fi combined	0.46	0.46	0.0019
UWB combined	<b>0.00</b>	<b>0.00</b>	<b>0.0005</b>

TABLE II: Localization accuracy when there are two persons in the environment

Feature	MNE	MALE	MNELE
Wi-Fi RSSI [6]	1.17	2.09	0.0006
Wi-Fi RTT distance	0.96	1.53	0.0020
UWB distance	0.76	0.56	0.0008
Wi-Fi combined	1.07	1.37	0.0017
UWB combined	<b>0.57</b>	<b>0.37</b>	<b>0.0004</b>

TABLE III: Localization accuracy when there are three persons in the environment

Feature	MNE	MALE	MNELE
Wi-Fi RSSI [6]	1.38	2.35	0.0024
Wi-Fi RTT distance	1.35	1.79	0.0035
UWB distance	1.29	1.39	<b>0.0015</b>
Wi-Fi combined	1.37	1.64	0.0036
UWB combined	<b>1.23</b>	<b>1.27</b>	0.0016

## V. CONCLUSION

In this paper, we proposed device-free multi-person indoor localization using the change of ToF. To realize device-free multi-person indoor localization, we used ToF ranging information to detect persons by installing transceivers and receivers in the environment. Moreover, we developed a mesh-wise person density estimation-based localization model and LoS/NLoS encoded feature representation for constructing a multi-person localization model with a single-person dataset.

In addition, we conducted an experiment to evaluate the accuracy of device-free multi-person indoor localization by learning only one person data. This experiment showed that the proposed method using ToF achieves higher accuracy than the existing method using Wi-Fi RSSI, and multi-person localization is possible with training data from only one person. At the same time, we also evaluated the localization accuracy using ToF ranging with two radio waves, Wi-Fi RTT

and UWB, and showed that the localization using UWB has higher accuracy than Wi-Fi RTT.

We are planning to realize high-precision localization with a large number of people (more than three) and localization in an environment with obstacles, such as furniture, in the future.

## REFERENCES

- [1] P. Bahl and V. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *In Proceedings of International Conference on Computer Communications (INFOCOM)*, vol. 2, 2000, pp. 775–784.
- [2] J. Yin, Q. Yang, and L. Ni, "Adaptive temporal radio maps for indoor location estimation," in *In Proceedings of International conference on pervasive computing and communications (PerCom)*, 2005, pp. 85–94.
- [3] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *Trans. on SMC(C)*, vol. 37, no. 6, 2007.
- [4] S. He and S.-H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, 2016.
- [5] J. Xiao, Z. Zhou, Y. Yi, and L. M. Ni, "A survey on wireless indoor localization from the device perspective," *ACM Computing Surveys*, vol. 49, no. 2, 2016.
- [6] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*, ser. MobiCom '07. New York, NY, USA: Association for Computing Machinery, 2007, pp. 222–229. [Online]. Available: <https://doi.org/10.1145/1287853.1287880>
- [7] M. Moussa and M. Youssef, "Smart ceives for smart environments: Device-free passive detection in real environments," in *2009 IEEE International Conference on Pervasive Computing and Communications*, 2009, pp. 1–6.
- [8] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using wifi," in *In Proceedings of the ACM Conference on Special Interest Group on Data Communication (SIGCOMM)*, 2015, pp. 269–282.
- [9] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie, "Lifs: Low human-effort, device-free localization with fine-grained subcarrier information," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 243–256. [Online]. Available: <https://doi.org/10.1145/2973750.2973776>
- [10] O. Hashem, M. Youssef, and K. A. Harras, "Winar: Rtt-based sub-meter indoor localization using commercial devices," in *In Proceedings of International Conference on Pervasive Computing and Communications (PerCom)*, 2020, pp. 1–10.
- [11] S. Yan, H. Luo, F. Zhao, W. Shao, Z. Li, and A. Crivello, "Wi-fi rtt based indoor positioning with dynamic weighted multidimensional scaling," in *In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2019, pp. 1–8.
- [12] L. Schauer, F. Dorfmeister, and M. Maier, "Potentials and limitations of wifi-positioning using time-of-flight," in *In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2013, pp. 1–9.
- [13] J. Choi, Y.-S. Choi, and S. Talwar, "Unsupervised learning technique to obtain the coordinates of wi-fi access points," in *In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2019, pp. 1–6.
- [14] M. Rea, A. Fakhreddine, D. Giustiniano, and V. Lenders, "Filtering noisy 802.11 time-of-flight ranging measurements from commoditized wifi radios," *In Transactions on Networking*, vol. 25, no. 4, pp. 2514–2527, 2017.
- [15] K. Jiokeng, G. Jakllari, A. Tchana, and A.-L. Beylot, "When ftm discovered music: Accurate wifi-based ranging in the presence of multipath," in *In Proceedings of International Conference on Computer Communications (InfoCom)*. IEEE, 2020, pp. 1857–1866.
- [16] W. Jiang, H. Xue, C. Miao, S. Wang, S. Lin, C. Tian, S. Murali, H. Hu, Z. Sun, and L. Su, "Towards 3d human pose construction using wifi," in *In Proceedings of the 26th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2020, pp. 1–14.

- [17] R. H. Venkatnarayan, M. Shahzad, S. Yun, C. Vlachou, and K.-H. Kim, "Leveraging polarization of wifi signals to simultaneously track multiple people," *In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 4, no. 2, pp. 1–24, 2020.
- [18] C. R. Karanam, B. Korany, and Y. Mostofi, "Tracking from one side: multi-person passive tracking with wifi magnitude measurements," in *In Proceedings of the 18th International Conference on Information Processing in Sensor Networks (IPSN)*, 2019, pp. 181–192.
- [19] Z. Chen, L. Zhang, C. Jiang, Z. Cao, and W. Cui, "Wifi csi based passive human activity recognition using attention based blstm," *IEEE Transactions on Mobile Computing*, vol. 18, no. 11, pp. 2714–2724, 2019.
- [20] J. W. Choi, X. Quan, and S. H. Cho, "Bi-directional passing people counting system based on ir-uwb radar sensors," *In Journal of IEEE Internet of Things*, vol. 5, no. 2, pp. 512–522, 2017.
- [21] J. W. Choi, D. H. Yim, and S. H. Cho, "People counting based on an ir-uwb radar sensor," *In Journal of IEEE Sensors*, vol. 17, no. 17, pp. 5717–5727, 2017.