A UWB-Based Method for Identification and Error Compensation of Non-Line-of-Sight Signals for Indoor Positioning of Ships

Jinping Chen, Xiaofei Yang, Zhen Su, Zujun He, Qinlan Xue and Feilong Huang

June 19, 2022

A UWB-based method for identification and error compensation of non-line-of-sight signals for indoor positioning of ships

1st Chen Jinping
School of electronic and information (of Affiliation)
Jiangsu University of Science and Technology
Zhenjiang, China
15298812761@163.com

2nd Yang Xiaofei
School of electronic and information (of Affiliation)
Jiangsu University of Science and Technology
Zhenjiang, China
yxfei_0809@just.edu.cn

3rd Su Zhen
Marine Equipment Research Institute
(of Affiliation)
Jiangsu University of Science and Technology
Zhenjiang, China
suzhen415@126.com

4th He Zujun
Marine Equipment Research Institute
(of Affiliation)
Jiangsu University of Science and Technology
Zhenjiang, China
hzj@just.edu.cn

5th Xue Qinlan
School of electronic and information (of Affiliation)
Jiangsu University of Science and Technology
Zhenjiang, China
xueqinlan88@163.com

6th Huang Feilong
School of electronic and information (of Affiliation)
Jiangsu University of Science and Technology
Zhenjiang, China
hfl190529@163.com

***Abstract*—With the advancement of ship informatization, precise positioning has become a trend for indoor positioning of ships. In the complex ship indoor environment, UWB signals are easily blocked, refracted and reflected during propagation, resulting in degraded positioning accuracy and affecting the positioning effect. In this paper, we consider the influence of non-line-of-sight signals (NLOS) on UWB signals in the UWB ranging process, focus on the main causes of non-line-of-sight errors in the ranging process, identify line-of-sight signals (LOS) and non-line-of-sight signals (NLOS) using Gaussian mixture model unsupervised learning clustering method, study the least squares-based ranging error compensation algorithm using existing data samples, and conduct an experimental validation of the studied The experimental validation of the method is carried out. The experimental results show that the studied method can significantly improve the ranging accuracy of UWB in practical applications, and the ranging error is reduced by 61.91% compared with the traditional ranging algorithm.**

Keywords—ultra wide band, indoor positioning of ships, NLOS error, Gaussian mixture model, least-squares

I. Introduction

With the rapid development of wireless sensors, positioning technologies such as WIFI, RFID, Bluetooth and ZigBee have emerged in indoor positioning [1]. Although these positioning technologies have been successfully applied in their respective fields, they are less applied in the field of ships, mainly because of the huge size of large ships, many compartments and complex structure, more signal shields, serious electromagnetic interference, and even high temperature and humidity in some cabins and cargo holds, and serious noise interference, which makes the positioning accuracy and effect unsatisfactory and difficult to meet the demand for high-precision positioning in ship cabins.

UWB technology has shown excellent performance in indoor positioning in recent years [2].UWB technology has the advantages of high transmission rate, low power consumption, strong anti-interference, strong multipath resolution, and strong penetration capability [3], which can meet the high accuracy positioning needs of ship cabin personnel. However, in the complex environment of ship cabins, the influence of factors such as numerous cabins and dense obstacles causes UWB signal propagation to switch randomly between two forms of Line of Sight propagation and Non Line of Sight propagation [4]. In the NLOS propagation method, the signal cannot arrive in a straight line propagation, which makes the propagation time of the signal prolonged leading to a large range value, which greatly affects the positioning accuracy. Therefore, it is important to study the NLOS environment identification problem to improve the accuracy of personnel positioning in ship cabins [5].

The effect of non-line-of-sight has been explored by researchers, and reference [6] investigated the establishment of a Gaussian model to identify NLOS signals . Reference [7] proposed a low-cost non-line-of-sight identification and suppression technique using Fresnel zone applicable to ultra-wideband ranging in static and dynamic environments. Reference [8], an improved particle filtering algorithm based on residual analysis is proposed to reduce the impact of NLOS errors on positioning accuracy. Reference [9], an improved statistical fingerprint analysis method is proposed for energy-assisted TOA estimation and reduction of NLOS ranging errors by fourth-order cumulative volume technique using artificial neural networks. Reference [10] proposes to use maximum likelihood estimation for localization, consider the probability of occurrence of LOS and NLOS propagation, derive the position error bound of the localization system using Cramer-Rao Lower Bound, and evaluate the effect of NLOS propagation on the position error bound.

In this paper, a recognition method based on Gaussian mixing model is proposed to construct a LOS/NLOS recognition model based on the position information of each node, which can accurately identify the NLOS signal, and then perform least-squares fitting on the ranging value and ranging error to compensate the ranging error and improve the positioning accuracy of the system.

II. System and Ranging Model Description

A. UWB ranging model

In this paper, a Double-Sided Two-Way Ranging based on the time-of-flight method is used to calculate the distance between the base station and the tag. This ranging method requires both the base station and the tag to have signal transceiver functions, and the base station receives the reply tag signal in real time and measures the distance between them, as shown in Figure 1.

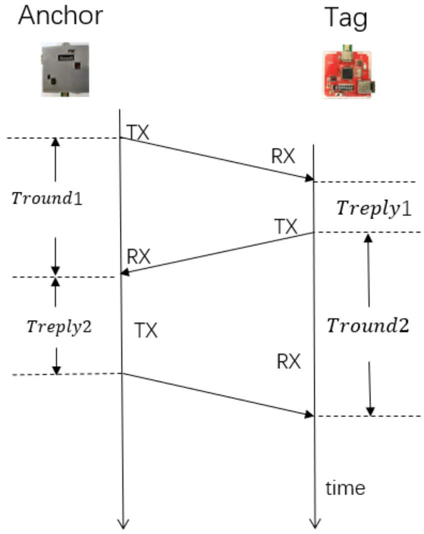

Fig. 1. A Double-Sided Two-Way Ranging based on the time-of-flight method

In order to achieve continuous communication between the base station and the tag, initialization is required prior to ranging. After initialization, the base station starts sending polling messages. When the tag receives the polling message, it records the reception T_{round1} time at that time and initiates a response at T_{reply1} time to return the data to the base station. After receiving the response message from the tag, the base station records the reception time T_{round2} and returns the data to the tag at T_{reply2} time. The signal round-trip times T_{prop2} and T_{prop1} are calculated from the recorded times.

$$\mathrm{T}_{prop1} = \frac{1}{2}(\mathrm{T}_{round1} - \mathrm{T}_{reply1}) \tag{1}$$

$$\mathrm{T}_{prop2} = \frac{1}{2}(\mathrm{T}_{round2} - \mathrm{T}_{reply2}) \tag{2}$$

Then the true value of the round-trip TOF is obtained from formula (3).

$$T_{prop} = \frac{T_{round1} \times T_{round2} - T_{replay1} \times T_{replay2}}{T_{round1} + T_{round2} + T_{replay1} + T_{replay2}} \tag{3}$$

Since the response time is not required to be the same, Double-Sided Two-Way Ranging is an asymmetric ranging method. The errors are as follows.

$$\begin{aligned}\overline{T}_{prop} &= \frac{T_{round1}(1+\mathrm{E}_A) \times T_{round2}(1+E_B) - T_{replay1}(1+E_B) \times T_{replay2}(1+E_A)}{T_{round1}(1+E_A) + T_{round2}(1+E_B) + T_{replay1}(1+E_B) + T_{replay2}(1+E_A)} \\ &= \frac{(4T_{prop}^2 + 2T_{prop}(T_{replay1} + T_{replay2}))(1+E_A)(1+E_B)}{4T_{prop} + 2(T_{replay1} + T_{replay2}) + (2T_{prop} + T_{replay1} + T_{replay2})(E_A + E_B)} \\ &= \frac{2(1+E_A)(1+E_B)}{(1+E_A)+(1+E_B)} T_{prop}\end{aligned} \tag{4}$$

$$T_{prop} = \frac{(1+\mathrm{E}_A)+(1+E_B)}{2(1+E_A)(1+E_B)}\overline{T}_{prop} \tag{5}$$

$$\begin{aligned}Error &= \overline{T}_{prop} - T_{prop} \\ &= \left(1 - \frac{(1+\mathrm{E}_A)+(1+E_B)}{2(1+E_A)(1+E_B)}\right)\overline{T}_{prop} \\ &= \frac{E_A + E_B + 2E_A E_B}{2(1+E_A)(1+E_B)}\overline{T}_{prop}\end{aligned} \tag{6}$$

Since $\mathrm{E}_A \ll 1, E_B \ll 1$, formula (6) can become formula (7).

$$Error \approx \frac{\mathrm{E}_A + E_B}{2}\overline{T}_{prop} \tag{7}$$

From this, it can be found that the error is only related to the clock drift and the flight time.

B. *Cause of NLOS error*

From the analysis of the ranging model, the ranging accuracy of TOA estimation is affected by clock drift and flight time. The Double-Sided Two-Way Ranging can eliminate the error caused by clock drift to a certain extent, but the main cause of error is still related to flight time. In the environment of LOS, the medium propagating between the base station and the tag is free, and the TOA estimation is accurate. However, in the NLOS environment, there are interference factors such as noise, walls and obstacles between the base station and the tag, and the propagation path is NLOS, which makes the communication time between the two increase and the ranging error increases, resulting in inaccurate positioning accuracy. For most indoor positioning environments of ships, LOS and NLOS coexist, which has a great impact on achieving high-precision positioning. Therefore, it is necessary to distinguish between LOS and NLOS signals, compensate for NLOS errors, and improve positioning accuracy.

III. PROPOSED NLOS IDENTIFICATION AND ERROR COMPENSATION METHOD

A. *Gaussian mixing model*

In the complex ship indoor environment, the signal in NLOS environment undergoes various phenomena such as refraction, reflection, and energy attenuation, resulting in the attenuation of the localization path, so that the solved localization value deviates significantly from the actual value. Since the offset due to the fading path is in accordance with the normal model , a Gaussian mixture model [11] can be used to distinguish between LOS and NLOS signals.

Gaussian mixture model is a widely used clustering algorithm and belongs to unsupervised clustering model. It is a commonly used distribution model of variables, which is to quantify things precisely with Gaussian probability density function, and to decompose a will into several models formed based on Gaussian probability density function.

Gaussian mixture model is trained to classify the target by samples, with sample data $\mathrm{X} = (x_1, x_2, \cdots, x_\mathrm{i})$, i is the number of features, superimposed by K Gaussian probability densities, and the GMM expression is formula (8).

$$P(x)=\sum_{k=1}^{\infty} p(k)p(x\mid k)=\sum_{k=1}^{\infty}\pi_k\cdot N(x\mid \mu_k,\Sigma_\kappa) \quad (8)$$

where $N(x\mid u_k,\Sigma_k)$ is the probability density function of the kth Gaussian model, μ_k is the mean of each class; Σ_k is the covariance matrix; π_{k} is the mixing factor and is the weight of each Gaussian distributed sample, and sums to 1. The expression of the Gaussian density function for each sample data is formula (9).

$$N(x\mid \mu_k,\Sigma_\kappa)=\frac{1}{(2\pi)^{\frac{D}{2}}\left|\Sigma_\kappa\right|^{\frac{1}{2}}}e^{-\frac{1}{2}(x-\mu_k)\Sigma_\kappa^{-1}(x-\mu_k)^{\mathrm{T}}} \quad (9)$$

Since the distinction is between LOS and NLOS, the Gaussian mixture model can be written in the following form.

$$\begin{gathered}P(x)=\pi_1\cdot N(x\mid \mu_k,\Sigma_\kappa)+\pi_2\cdot N(x\mid \mu_k,\Sigma_\kappa)\\ (1\le\pi_1\le 1;\quad 1\le\pi_2\le 1;\quad \pi_1+\pi_2=1)\end{gathered} \quad (10)$$

Assuming that the data set is D_N , each sample data is an independent event that conforms to the Gaussian distribution and can be expressed in terms of conditional probabilities, so the Gaussian mixture model maximum likelihood function is formula (11).

$$\mathrm{L}(\mu_k,\Sigma_\kappa)=\sum_{j=1}^{N}\log\Big(\sum_{k=1}^{2}\pi_k\cdot N(x\mid \mu_k,\Sigma_\kappa)\Big) \quad (11)$$

The maximum likelihood function is solved using the maximum expectation algorithm (EM). EM algorithm is to find the maximum likelihood estimate or the maximum a posteriori estimate of the parameters in a probabilistic model. The main procedure is as follows.

(1) Define the component array K and set the initial values of π_{k} , μ_k , Σ_k for each component k.

(2) E Step: Based on the set initial values, the posterior probability of the kth Gaussian model for each sample is calculated as formula (12).

$$\gamma(i,k)=\frac{\pi_k\cdot N(x_i\mid \mu_k,\Sigma_\kappa)}{\sum_{k=1}^{K}\pi_k\cdot N(x\mid \mu_k,\Sigma_\kappa)} \quad (12)$$

(3) M Step: Based on the calculated , and then calculate the new π_{k} , μ_k , Σ_k .

$$\pi_k=\frac{N_k}{N} \quad (13)$$

$$u_k=\frac{1}{N_k}\sum_{n=1}^{N}\gamma(i,k)x_i \quad (14)$$

$$\Sigma_k=\frac{1}{N_k}\sum_{n=1}^{N}\gamma(i,k)(x_i-\mu_k)(x_i-\mu_k)^T \quad (15)$$

(4) Repeat E step and M step until the parameters converge.

B. Distance measurement error compensation model based on least squares algorithms

Based on the above analysis, it can be assumed that there is a nonlinear relationship between the ranging error caused by the NLOS environment and the ranging results, based on the least squares nonlinear polynomial equation.

$$\mathrm{e}_i=a_0+a_1d_{\mathrm{i}}+a_2d_i^{\,2}+\cdots+a_md_i^{\,m} \quad (16)$$

Where, e_i is the ranging error, d_i is the desired ranging value, and a_m is polynomial coefficient.

If there are n ranging errors, this equation can be expressed as formula (17).

$$\begin{cases}\mathrm{e}_i=a_0+a_1d_1+a_2d_1^{\,2}+\cdots+a_md_1^{\,m}\\ \mathrm{e}_2=a_0+a_1d_2+a_2d_2^{\,2}+\cdots+a_md_2^{\,m}\\ \qquad\vdots\\ \mathrm{e}_n=a_0+a_1d_n+a_2d_n^{\,2}+\cdots+a_md_n^{\,m}\end{cases} \quad (17)$$

Write the equation in matrix form.

$$\mathrm{E}=D\cdot C \quad (18)$$

$$\mathrm{E}=\left[e_1,e_2,\cdots,e_n\right]^{\mathrm{T}} \quad (19)$$

$$\mathrm{D}=\begin{Bmatrix}d_1^0 & \cdots & d_1^m\\ \vdots & d_i^j & \vdots\\ d_n^0 & \cdots & d_n^m\end{Bmatrix} \quad (20)$$

$$\mathrm{A}=\left[a_0,a_1,\cdots,a_1\right]^T \quad (21)$$

When the condition n>(m+1) is satisfied, the polynomial coefficients are obtained using the least squares method [12] as formula (22).

$$\mathrm{A}=(D^TD)^{-1}D^TE \quad (22)$$

The compensated range value is obtained from formula 23.

$$\mathrm{d}_\lambda=\mathrm{d}_t-(a_0+a_1d_t+a_2d_t^{\,2}+\cdots+a_md_t^{\,m}) \quad (23)$$

IV. Simulation and Experimentation

A. NLOS recognition model testing

The experimental data in this part are 1000 real data of UWB positioning, including 500 LOS signals and 500 NLOS signals each. This experiment uses DWM1000 module developed by DecaWave as the positioning information acquisition equipment, and the experimenter places the UWB positioning module around a large empty ship cabin of 6.5m*8m*3m, and first acquires the positioning information in LOS environment with a sampling time of 5 min. when collecting the positioning information in NLOS, the positioning tag is placed in the iron box without the lid, and while collecting, the experimenter is in the positioner rotated around to create NLOS conditions. The collected localization coordinates are combined and used as features.

The obtained experimental data are clustered by K-means algorithm and GMM algorithm, and the clustering effect of each method is compared. Figure 3 and Figure 4 show the true distribution of the localization coordinates and K-means clustering of the LOS and NLOS signals, with the NLOS signal in red and the LOS signal in green; Figure 5 shows the clustering of the Gaussian mixture model, with the NLOS signal in red and the LOS signal in green. The effect of clustering can be seen from the graph that Gaussian mixture model can describe the distribution of data more clearly than K-means algorithm. k-means algorithm is measured by Euclidean distance between points, so the classification boundary in the graph tends to be circular. The Gaussian mixture model, on the other hand, clusters by expectation and variance, so the classification boundaries in the figure are elliptical in shape and more flexible.

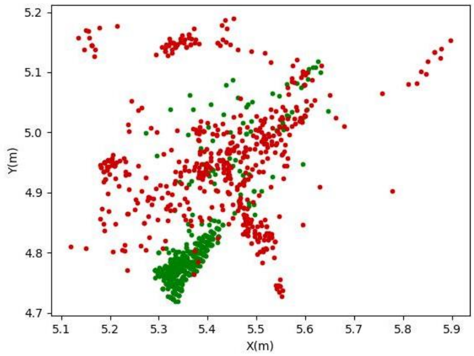

Fig. 2. True distribution

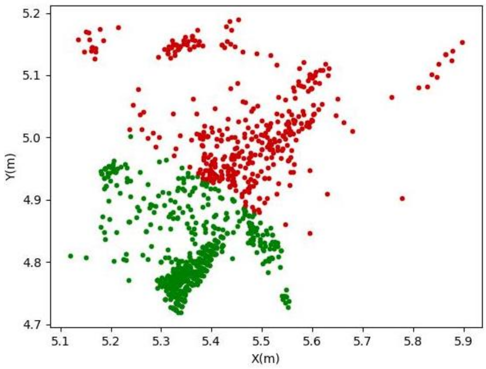

Fig. 3. K-means clustering

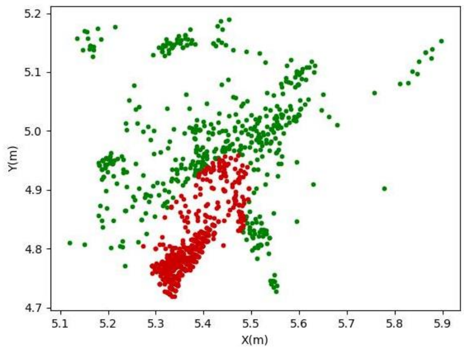

Fig. 4. Gaussian mixture model clustering

NLOS recognition is a binary classification problem that utilizes precision [13] and recall [14] as evaluation metrics. They are calculated by Equation 24 and Equation 25, respectively, where TP is the correctly recognized NLOS signal, FP is the incorrectly recognized NLOS signal, TN is the correctly recognized LOS signal, and FN is the incorrectly recognized LOS signal.

$$\mathrm{Pr} = \frac{TP}{TP + FP} \tag{24}$$

$$\mathrm{Re} = \frac{TP}{TP + FN} \tag{25}$$

Table I compares the clustering performance between the K-means and GMM algorithms. From Table I, it can be seen that the GMM algorithm classifies with higher accuracy than the K-means algorithm, and thus can better improve the accuracy of localization.

TABLE I. ALGORITHM PERFORMANCE COMPARISON

	TP	FP	TN	FN	Pre	rec
K-means	382	85	427	106	81.80	78.28
GMM	478	105	356	61	81.99	88.68

B. *Distance measurement error compensation model experiments*

In order to obtain accurate observation data, field tests was set up in the ship meeting room, and the experimental scene was laid out as shown in Fig. 5. Four locations with fixed coordinates were selected for base station deployment in the room, and a three-dimensional spatial coordinate system was established with the base station coordinate position in the lower right corner of the figure as the reference. Nine fixed points are set between base station 0 and base station 1, and the distance from these nine points to the four base stations is collected. Moreover, the existence of obstacles, people walking and wall interference factors are maintained during the experimental test to increase the complexity of the indoor environment, so that the test environment can meet the basic characteristics of the complex indoor environment as much as possible.

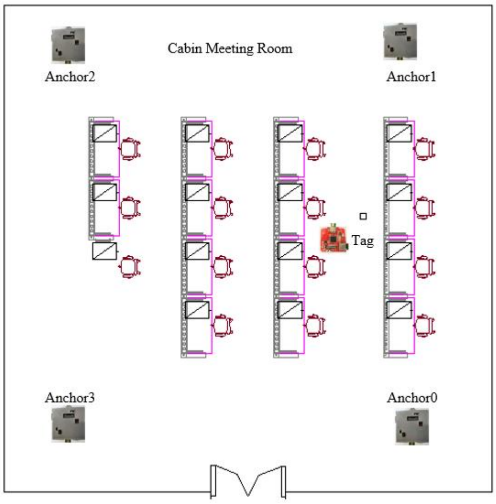

Fig. 5. The test scene

After processing the collected positioning data, the ranging error and the ranging value were fitted by least squares. The results of the fitting experiments are shown in Figure 6. By observing the magnitude of the error data in the figure, it can be found that the distance values obtained from the tag positioning show a relatively high aggregation, and the dispersion of the distance values of each group is low and remains within 0.2m. From the error compensation model curve, it can be found that the ranging error decreases with the increase of distance in the ranging of 1-5m.

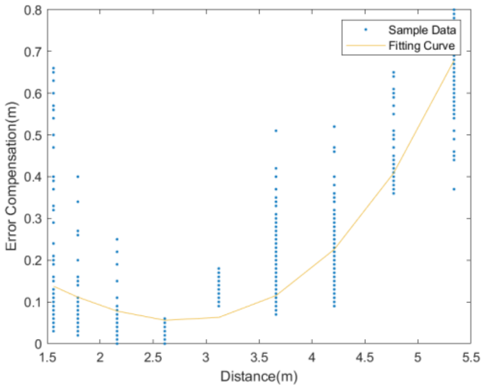

Fig. 6. Compensation model fitting results

To test the model compensation, the accuracy of the compensated and uncompensated data was compared. Table I shows that the average ranging errors of the two sets of data were 0.08 m and 0.21 m. The compensated and uncompensated ranging data reduced the ranging error by 0.13 m compared to each other. After compensation, the ranging error value was reduced by 61.91%, which is still a considerable effect.

TABLE II. DISTANCE MEASUREMENT ERROR COMPENSATION

Error	Min(m)	Max(m)	Mean(m)
Compensated	0	0.18	0.08
Uncompensated	0	0.8	0.21

V. CONCLUSION

In this paper, the causes of errors in the UWB localization process are analyzed by ranging models. In order to identify NLOS signals, a Gaussian mixture model is mainly used to distinguish NLOS signals, and the K-means algorithm is compared to compare the accuracy of the two algorithms in identifying NLOS signals. Finally, the least squares method is used to compensate the ranging error, and the error value is reduced by 61.91%, which improves the positioning accuracy. The disadvantage is that the influence of different ship cabin environments on UWB signals can vary greatly, and this paper does not analyze and study them by conducting multi-scene positioning data acquisition, and further research is needed.

REFERENCES

[1] Yang X , Wang J , Song D , et al. A Novel NLOS Error Compensation Method Based IMU for UWB Indoor Positioning System[J]. IEEE Sensors Journal, 2021, PP(99):1-1.

[2] Dong Mengyao. A Low-Cost NLOS Identification and Mitigation Method for UWB Ranging in Static and Dynamic Environments[J]. IEEE COMMUNICATIONS LETTERS,2021,25(7).

[3] Meiyu Liu,Xizhong Lou,Xiaoping Jin,Ruwen Jiang,Kaifeng Ye,Shubin Wang. Correction to: NLOS Identification for Localization Based on the Application of UWB[J]. Wireless Personal Communications,2021(prepublish).

[4] Chen M, Liu K, Jie M, et al. Spatio-Temporal Fingerprint Localization for Shipboard Wireless Sensor Networks[J]. IEEE Sensors Journal, 2018, PP(24):1-1.

[5] Carlo Galiotto,Nuno K. Pratas,Linda Doyle,Nicola Marchetti. Effect of LOS/NLOS propagation on 5G ultra-dense networks[J]. Computer Networks,2017,120.

[6] Minghao Si,Yunjia Wang,Shenglei Xu,Meng Sun,Hongji Cao. A Wi-Fi FTM-Based Indoor Positioning Method with LOS/NLOS Identification[J]. Applied Sciences,2020,10(3).

[7] Dong Mengyao. A Low-Cost NLOS Identification and Mitigation Method for UWB Ranging in Static and Dynamic Environments[J]. IEEE COMMUNICATIONS LETTERS,2021,25(7).

[8] Long Cheng,Liang Feng,Yan Wang. A Residual Analysis-Based Improved Particle Filter in Mobile Localization for Wireless Sensor Networks[J]. Sensors,2018,18(9).

[9] Liang X , Zhu W , Deng J . An Improved Statistics Fingerprint Analysis Method for Time Delay Estimation in Multipath NLOS Environment[J]. Wireless Personal Communications, 2021, 123(2):1855-1869.

[10] Xiufang Shi,Guoqiang Mao,Zaiyue Yang,Jiming Chen. MLE-based localization and performance analysis in probabilistic LOS/NLOS environment[J]. Neurocomputing,2017,270.

[11] Chuanqi CHENG,Xiangyang HAO,Jiansheng LI,Peng HU,Xu ZHANG.A Robust Gaussian Mixture Model for Mobile Robots' Vision-based Pose Estimation[J].Journal of Geodesy and Geoinformation Science,2019,2(03):79-90.

[12] Jingyu HUA,Yejia YIN,Anding WANG,Yu ZHANG,Weidang LU.Geometry-based non-line-of-sight error mitigation and localization in wireless communications[J].Science China(Information Sciences),2019,62(10):149-163.

[13] Cung Lian Sang,Bastian Steinhagen,Jonas Dominik Homburg,Michael Adams,Marc Hesse,Ulrich Rückert. Identification of NLOS and Multi-Path Conditions in UWB Localization Using Machine Learning Methods[J]. Applied Sciences,2020,10(11).

[14] Xiaofeng Yang,Feng Zhao,Tiejun Chen. NLOS identification for UWB localization based on import vector machine[J]. AEUE - International Journal of Electronics and Communications,2018,87.