Research on Indoor Visible Light Positioning Algorithm Based on K-means Clustering

ZHANG Rui¹ᵃ, ZHANG Yerong¹ᵇ
¹College of Electronic and Optics Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China
ᵃ1018020820@njupt.edu.cn
ᵇzhangyr@NJUPT.edu.cn

Abstract—In order to improve the indoor positioning accuracy and reduce the impact of a reflected light on the positioning performance, based on the analysis of the existing RSS positioning algorithm, an improved positioning algorithm based on K-means clustering is proposed. According to the indoor visible light communication system model and optical power distribution, 10*10 receiving points are uniformly selected on the receiving surface of 5m*5m. Through the collection of the received optical power of each receiving point, the K-means clustering algorithm is used to classify 100 receiving points: the first cluster of points is less affected by a reflected light, and the second cluster of points is greatly affected by a reflected light. The second cluster of receiving points is weighted after classifying, and the simulation results show that the algorithm significantly reduces the room edge positioning error, and the positioning performance of the entire room is improved by 28%.

1. Introduction
In recent years, people's demand for navigation and indoor positioning services has increased. In order to meet this ubiquitous development demand, countries all over the world are intensively investing in research. At present, the most commonly used indoor positioning technologies are infrared radiation (IR), radio frequency identification (RFID), wireless local area network (WLAN), Bluetooth, and ultra-wideband (UWB)[1]. The above-mentioned traditional positioning technology has disadvantages such as high cost, insufficient positioning accuracy, and susceptibility to electromagnetic interference. VLC indoor positioning has become an emerging positioning technology because of its immunity to electromagnetic interference, good confidentiality, high transmission rate, and no need for spectrum authentication, which has been studied extensively[2].

VLC indoor positioning encodes and modulates the light signal emitted by the LED, and uses pulse modulation to add the modulated signal to the driving current of the LED lamp. When the receiver is located in the illuminated area, a photodetector on the receiver decodes and demodulates the received signal to restore the original signal, and then the location of the receiver is obtained through the positioning algorithm. Commonly used positioning algorithms are roughly divided into four categories: geometric measurement method, scene analysis method, approximate perception method and image sensor imaging method[3]. The geometric measurement method mainly includes time of arrival (TOA)[4], time difference of arrival (TDOA)[5], angle of arrival (AOA)[6], and received signal strength (RSS)[7]. TOA can achieve higher positioning accuracy, but requires transmitters and receivers to be completely synchronized, while TDOA has lower requirements for synchronization. AOA is
complicated and requires additional hardware and high cost to support. RSS calculates the transmission distance by measuring the received signal strength, combining with the wireless optical communication channel model, and use the trilateration positioning method to obtain the mobile target position. However, RSS positioning is susceptible to environmental changes and reflection interference. In order to solve the problem that the reflected light has a great influence on the RSS positioning algorithm, this paper proposes an indoor visible light positioning method based on K-means clustering.

2. SYSTEM MODEL

2.1. VLC system model

The VLC system used in this article is shown in Figure 1. The indoor location is a 5m × 5m × 3m room area. Four LED lights are evenly distributed on the ceiling, and a photodetector (PD) receives the signal emitted from the LED on the plane. The illumination radiation areas of the four LEDs overlap each other to ensure that each positioning node in the radiation area can simultaneously receive the signals emitted by the four LEDs. Suppose that four LED emitters' \( (i=1, 2, 3, 4) \) coordinates are \( A(x_a, y_a, z_a) \), \( B(x_b, y_b, z_b) \), \( C(x_c, y_c, z_c) \), \( D(x_d, y_d, z_d) \) respectively, and the coordinates of the positioning node PD are \( (x_p, y_p, z_p) \). In order to ensure that the positioning node can distinguish the signals from different LEDs, the signals can be coded with on-off keying (OOK) \(^8\), pulse position modulation (PPM) and color shift keying (CSK) \(^9\).

![Figure1 Indoor VLC System Model](image1)

2.2. VLC channel model

Generally, in the VLC indoor positioning system, the signal propagation path of LED is mainly divided into line-of-sight path (LOS) and non-line-of-sight path (NLOS) \(^{10}\). Considering the actual situation, this article uses LOS plus NLOS channel model. Assuming that the LED lamp is Lambertian radiation, its Lambertian transmission model is shown in Figure 2.

![Figure2 VLC Channel Model](image2)

The LOS channel gain of the optical signal from the transmitter to the receiver can be expressed as:
The NLOS channel gain of the optical signal from the transmitter to the receiver can be expressed as:

\[
H(0) = \begin{cases} 
\frac{A_r (m + 1)}{2 \pi d^2} \cos^m(\varphi) T_r(\psi) g(\psi) \cos(\psi), & 0 < \psi < F_{\text{FOV}} \\
0 & \text{otherwise}
\end{cases}
\]  
(1)

Where \( A_r \) is effective area of the optical receiver, \( m \) is the number of modes of the radiation lobe, \( d \) is the linear distance from the LED transmitting end to the receiving end, \( d_i \) is the linear distance from the LED to the reflection point on the wall, \( d_{2i} \) is the linear distance from the reflection point to the receiving end optical receiver, \( \theta_i \) is the emission angle, \( \psi_i \) is the reception angle, \( \theta_i' \) and \( \psi_i' \) are the emission angle and the incidence angle of the reflection point from the LED to the wall, respectively, \( \theta_2 \) and \( \psi_2 \) are the emission angle and the incidence angle of the reflection point on the wall to the optical receiver, respectively. \( T_r(\psi) \) is the gain of the optical concentrator, and \( \varphi_{1,2} \) is the half power angle \([11]\). Therefore, the received power \( P_r \) can be expressed as:

\[
P_r = P_t \times (H(0) + H_{\text{off}}(0))
\]  
(3)

3. Improved RSS positioning algorithm based on K-means clustering

3.1. Principle of RSS positioning algorithm

The RSS positioning algorithm calculates the distance by detecting the received light intensity to achieve indoor positioning. From the indoor LED lighting model shown in Figure 1, the receiver RX receives the signals emitted by the three LEDs. Let the coordinates of the three LEDs be \( A(x_a, y_a, z_a) \), \( B(x_b, y_b, z_b) \), and \( C(x_c, y_c, z_c) \). The coordinates of RX is \( (x, y, z) \). The distance from \( TX_i \) \((i=1,2,3)\) to RX are \( d_i \):

\[
\begin{align*}
\left|d_1\right|^2 &= (x_a - x)^2 + (y_a - y)^2 + (z_a - z)^2 \\
\left|d_2\right|^2 &= (x_b - x)^2 + (y_b - y)^2 + (z_b - z)^2 \\
\left|d_3\right|^2 &= (x_c - x)^2 + (y_c - y)^2 + (z_c - z)^2
\end{align*}
\]  
(4)

It can be seen from Figure 1 that the TXs are coplanar and the RX is on the ground, so \( z_a = z_b = z_c = h \) \((h \text{ is the height of the room), } z = 0\). Therefore, the coordinates of RX on the \( XY \) plane can be obtained by solving equations (4) \([12]\). Simplify formula (4) as:

\[
\begin{align*}
2x(x_a - x_a) + x_a^2 - x_a^2 + 2y(y_a - y_a) + y_a^2 - y_a^2 &= d_1^2 - d_1^2 \\
2x(x_b - x_b) + x_b^2 - x_b^2 + 2y(y_b - y_b) + y_b^2 - y_b^2 &= d_2^2 - d_2^2
\end{align*}
\]  
(5)

As long as the three LEDs are not collinear, formula (5) has an only solution, that is the positioning coordinates of RX.

3.2. Principle of K-means clustering algorithm

K-means clustering algorithm belongs to the category of unsupervised learning in machine learning.
The basic idea of the algorithm is: first select K sample points as clustering centers, and then gather the sample points to each center according to a certain clustering criterion to obtain the initial cluster; then judge whether the initial classification is reasonable, if not, modify the classification; repeat the iterative algorithm of modifying the clustering until it is reasonable \[^{[13]}\]. The specific implementation algorithm steps are as follows:

**STEP1:** Randomly select K initial cluster centers \(z_1(1), z_2(1), \ldots, z_k(1)\);

**STEP2:** Assign each sample set to K cluster centers according to the principle of minimum distance. That is, at the m-th iteration, if \(\|x - z_j(m)\| < \|x - z_i(m)\|, i, j = 1, 2, \ldots, k\) \(x\) belongs to the clustering domain represented by the j-th cluster center after the m-th iteration. The new cluster center is \(z_i(m) = \frac{1}{N_i} \sum_{x \in f_i(m)} x, i = 1, 2, \ldots k\), Where \(N_i\) is the number of samples in the i-th clustering domain, and \(f_i(m)\) is the clustering domain represented by the j-th cluster center at the m-th iteration.

**STEP3:** The mean vector in STEP2 is used as the new cluster center, let \(J = \sum_{x \in f_i(m)} \|x - z_j(m+1)\|^2, i = 1, 2, \ldots, k\) reach the minimum value, J is sum-of-squared-error criterion function;

**STEP4:** Judge the sum-of-squared-error criterion function in STEP3. At that time \(z_i(m+1) = z_i(m)\), the algorithm converges and the iteration is completed; otherwise, return to STEP2 to continue the iteration.

### 3.3. Improved RSS positioning algorithm

Only consider LOS links, RSS algorithms for positioning can achieve ideal positioning accuracy. However, when considering the NLOS link, due to the presence of reflected light, the received signal strength at the receiver is the sum of the LOS plus the NLOS component, and the positioning error of the receiving point in the area affected by the reflection is significantly increased. Therefore, an optimization algorithm based on K-means clustering is proposed on the basis of RSS positioning algorithm.

According to the VLC system model introduced in the first section, the RSS of each receiving point can be obtained. According to the principle of K-means algorithm, the centroid distance between RSS of each receiving point and centers of K clusters is calculated \[^{[14]}\]. In order to divide the receiving plane into two areas, one is the area that is less affected by the reflected light, and the other is the area that is greatly affected by the reflected light. So K is 2 in this paper. According to the divided areas, only the positioning results in the second type of area are weighted. The flowchart of specific algorithm shown in Figure3.
4. Simulation results and analysis

4.1. Simulation and analysis of RSS positioning algorithm

Set up four LED emitters A, B, C, D on the ceiling of the 5*5*3m room, each LED emitter has a unique ID so that the receiver can distinguish them. Suppose the four LED arrays are 2.6m away from the ground, and the coordinates are (-1.25, 1.25, H), (1.25, 1.25, H), (1.25, -1.25, H), (-1.25, -1.25, H) respectively. Each LED emitter contains 3600 lamp beads, and the light power of each lamp bead is 20mW.

According to formula (2), the power distribution reflected light on the receiving surface can be obtained. As shown in Figure 4, it can be seen that most of the reflected power is concentrated in the corner of the room or near the wall, and the maximum value is about -1.54dBm. However, the central area of the room is less affected by reflected light, and the minimum value of the received once reflection light power is about -10.25dBm.

According to formula (3), the total received optical power distribution under the LOS and NLOS channels is shown in Figure 5. The optical power of the entire receiving surface presents a distribution with high center and low wall edges and corners. The highest is 1.91dBm below the four LEDs, and the lowest is about 0.04dBm near the corner of the wall. The average received optical power of the entire receiving surface is 1.29dBm.
The simulation obtains the total received optical power including one reflected light, and the estimated distance from all the points to the four LEDs can be measured. According to the RSS positioning algorithm, then 10*10 receiving points on the receiving surface are located. The positioning results and error distribution are shown in Figure 6 and Figure 7, respectively. It can be seen that the positioning error on the edges and corners of the room is relatively large. The maximum error is 1.58m, located in the corner of the room, while the minimum error is only 0.007m, located in the center. The average positioning error of the receiving point on the entire receiving surface is 0.39 m. The main reason is that the receiving points located at the edges and corners are greatly affected by the reflected light, which makes the estimated distance error from the LED to the receiving point larger and ultimately leads to the difference in positioning accuracy between the center area and the edge area.
4.2 Simulation and analysis of improved positioning algorithm

In order to filter out the points at the edge of the room that are greatly affected by the reflected light, all the receiving points are clustered through the K-means clustering algorithm. The classification result is shown in Figure 8. Cluster 1 contains points in the central area of the room that are less affected by once reflected light. Cluster 2 is a collection of points at the edge area of the room, susceptible to once reflected light. It can be seen from Figure 7 that the positioning performance of points in the edge area of the room is poor. We can know from Figure 6 that due to the influence of the estimated distance error, the positioning results of points from cluster 2 are concentrated in the middle area, which required to weighting to decrease the positioning error. The weighting coefficient $n$ takes a value from 1 to 2 every 0.1. When $n$ is 1, the positioning result is the RSS positioning result. It can be seen from Figure 9 that as the weighting factor $n$ increases, the average positioning error first decreases and then increases. When $n$ is 1.5, the average positioning error is the smallest, 0.28m, and the positioning accuracy of the entire receiving surface is improved by about 28%. The positioning result is shown in Figure 10.
5. Conclusion

RSS algorithm is the most commonly used positioning algorithm. Although its implementation is simple, its accuracy is easily affected by reflection interference. This paper proposes an improved positioning algorithm based on the K-means clustering algorithm on the basis of the RSS positioning algorithm. By weighting the positioning results that are greatly affected by once reflected light, it effectively reduces the positioning error of the points at the edge of the room. The simulation results show that the positioning performance of the proposed positioning algorithm in the edge area of the room has been significantly improved, and the complexity of the algorithm is low. However, the weighting factor of this algorithm is to weight the positioning results by controlling variables. In the follow-up work, we will continue to study to find an adaptive weighting factor according to the relationship between the weighted value and the error.

REFERENCES

[1] Hu Qingqing. Research on Indoor Localization Algorithms Using Visible Light Communication Based on LEDs [D].

[2] Komine T, Nakagawa M. Fundamental analysis for visible-light communication system using LED lights[J]. Consumer Electronics IEEE Transactions on, 2004, 50(1):100-107.

[3] Zhao Jiaqi, Chi Nan. Comparative Research on Several Key Indoor Positioning Technologies Based on LED Visible Light Communication[J].Light&Lighting,2015, 000(001):34-41.
[4] Zhang Yanlong. Research on the Key Techniques of Indoor Location [D]. University of Science and Technology of China, 2014.

[5] Wann C D, Yeh Y J, Hsueh C S. Hybrid TDOA/AOA Indoor Positioning and Tracking Using Extended Kalman Filters[C]// Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd. IEEE, 2006.

[6] Yin J, Wan Q, Yang S, et al. A Simple and Accurate TDOA-AOA Localization Method Using Two Stations[J]. IEEE Signal Processing Letters, 2015, 23(1):144-148.

[7] Yu X, Wang J, Lu H. Single LED-Based Indoor Positioning System Using Multiple Photodetectors[J]. Photonics Journal, IEEE, 2018, 10(6):1-8.

[8] Jelena Vučić, Christof Kottke, Stefan Nerreter. 230 Mbit/s via a Wireless Visible-Light Link Based on OOK Modulation of Phosphorescent White LEDs[J]. IEEE, 2010, 45(2):1-3.

[9] Chen Beibei. Modulation Technology and Positioning Algorithm Based on Visible Light Communication[D].

[10] Zhang W, Chowdhury M I S, Kavehrad M. Asynchronous indoor positioning system based on visible light communications[J]. Optical Engineering, 2014, 53(4):045105.

[11] Zeng L, O’Brien D, Le-Minh H, et al. Improvement of Date Rate by using Equalization in an Indoor Visible Light Communication System[C]// IEEE International Conference on Circuits & Systems for Communications. IEEE, 2008.

[12] Cheng Shunli, Yan Kun, Li Zhao. Visible-Light Indoor Positioning Technology Based on the Parameters Optimization of Lambertian Model[J]. Study on Optical Communications, 2018, 209(05):73-77.

[13] Muhammad Saadi*, Touqeer Ahmad†, Yan Zhao‡, et al. An LED Based Indoor Localization System Using k-Means Clustering[C]// 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 2017.

[14] Wang Hui. Research on Positioning Method of Indoor Visible Light Based on Position Fingerprint[D].