Accuracy improvement of positioning data in greenhouse for agricultural machinery via optimisation algorithm

Jing Geng¹, Xiang Wang¹, Saihua Zhu¹, Xiangze Lin¹

¹Department of Electrical Engineering, College of Engineering, Nanjing Agricultural University, Nanjing, People’s Republic of China
E-mail: xzlin@njau.edu.cn

Abstract: Although several indoor positioning systems for greenhouse have been developed in automated agricultural machinery monitoring, the data measured by positioning systems still has large errors. In order to solve the problem, in this study, improved weighted-centroid algorithm and maximum likelihood estimation (MLE) algorithm were developed to optimise positioning data of agricultural machinery in the greenhouse, which is measured by ultra-wide band-based indoor positioning system according to different agricultural machinery working states. Compared with real coordinate data, when agricultural machinery stays still, the average static error of the data was 66.4 mm with the improved weighted-centroid algorithm; when the machinery is in dynamic state, the average dynamic error of the data was 61 mm and the probability that error <60 mm was 0.848 according to the MLE. Hence, the optimisation algorithms presented in this study improve the positioning accuracy significantly and the position of agricultural machinery is able to be estimated better compared with the raw data.

1 Introduction

Indoor positioning technique has played an important role in mechanisation and intelligentisation of agricultural machinery operations in the greenhouse [1–3]. Currently, GPS is widely applied in indoor positioning; however, it was found that GPS signal is severely attenuated after passing through indoor obstacles and not suitable for indoor positioning [4, 5]. In this case, many researchers began to focus on new wireless positioning technology to solve indoor positioning problems [6, 7].

Since the E-911 targeting criteria were presented by Federal Communications Commission in 1996, indoor positioning technology has developed rapidly. Various indoor wireless positioning technologies were put forward based on Bluetooth, radio frequency identification, WLAN, ultra-wide band (UWB) and many other wireless technologies [8–11]. Unlike those wireless positioning technologies, UWB positioning technology has the advantages of strong penetration, low power consumption, simple structure, work reliability and operation stability [12–14], so lots of scholars have paid attention to applying UWB positioning technology to indoor positioning system [15–17]. Meanwhile, in the choice of ranging algorithm, time difference of arrival (TDOA) was widely used due to its low complexity and high accuracy [18]. However, some experiments have shown that the measurement results of UWB positioning system are affected by many factors such as manufacturing errors, installation errors, sensor biases, sensitivity errors, noise errors and errors caused by ranging algorithms [19]. In order to reduce measurement error, positioning ranging error model and signal transmission error model and other related models were established. Thus, the aim of this study is to develop efficient optimisation algorithms to optimise positioning data measured by UWB-based indoor positioning system of agricultural machinery in the green house according to the different movement state of agricultural machinery [20, 21].

2 Data acquisition and housing

In order to obtain practical data, the position experiment was conducted in a tomato greenhouse, located in Jiangsu Academy of Agricultural Sciences, from June 7 to July 3, 2018. In our experiment, the indoor positioning system is built by Ubisense Series 7000 IP Sensors, Ubisense Series 7000 Industrial tag and positioning target, the experimental scene is shown in Fig. 1. The Ubisense tag is attached to the self-propelled spray robot, developed by Professor Wang Xiaoan's team of Nanjing Agricultural University, in greenhouse. Thus, this robot is monitored by a Ubisense tag, which transmits UWB pulses in extremely short duration, and remote sensors which enable location to be mapped by using TDOA techniques. The positioning system was set to cover an area of 10 m × 25 m in the green house with three Ubisense sensors.

The schematic diagram of positioning equipment installation in the greenhouse is shown in Fig. 2, where BS₁(x₁, y₁), BS₂(x₂, y₂), BS₃(x₃, y₃) are Ubisense sensors, and the Ubisense tag MS is attached to the spray robot, which can reflect real-time coordinate position of spray robot. The specific steps for collecting data are as follows. First, a master sensor is arbitrarily selected among three sensors. Then, the time signals of the master sensor and the other slave sensors are synchronised, and time information is transmitted to the computer through switchboard. Finally, the coordinates are obtained according to the TDOA ranging algorithm.

3 Data analysis

3.1 Static optimisation algorithm

When the robot stays still, the measured coordinate data will be distributed around the real point due to the presence of measurement error, so the center of all measured points can be found to improve the positioning accuracy. However, the
where \( y \) is the real coordinate. Hence, weighted-centroid algorithm is used to optimise positioning data for static operations of agricultural machinery in this paper.

![Fig. 2 Simulation of UWB equipment installation](image)

![Fig. 3 Agricultural machinery operation simulation route](image)

3.1.1 Time complexity: Time complexity of an algorithm is the time cost of running the algorithm and its function is to solve the scale of the problem. The purpose of applying time complexity is to choose the optimal algorithm with the shorter running time, and the time complexity of different clustering algorithms are shown in Table 1. From Table 1, it can be seen that different from those above clustering algorithms, weighted-centroid algorithm is more suitable for processing large data due to its low time complexity. Hence, weighted-centroid algorithm is used to optimise positioning data for static operations of agricultural machinery.

![Fig. 4 Intercept time route simulation](image)

3.1.2 Improved weighted centroid: If the measurement errors are fixed, the weighted-centroid algorithm can be used to predict the real coordinate with those measured data [22]. Weighted centroid can be calculated according to the following equation:

\[
x_c = \frac{\sum o_i x_i}{\sum o_i}, \quad y_c = \frac{\sum o_i y_i}{\sum o_i},
\]

(1)

where \( o_i \) is weight value, \( x_i \) and \( y_i \) are horizontal and vertical coordinates of the equipment data, respectively; \( x_c \) and \( y_c \) are weighted coordinates.

However, the distribution of data are not ideal in practice and in order to get the optimal results which are as close as possible to the real coordinates, an improved weighted-centroid algorithm is proposed, which is shown as

\[
x = \frac{x_1 + x_2 + \cdots + x_n}{n}, \quad y = \frac{y_1 + y_2 + \cdots + y_n}{n},
\]

(2)

where \( x_n \) and \( y_n \) are the weighted coordinates calculated by (1); \( x \) and \( y \) are mean weighted coordinates, which can be used to avoid uneven distribution of data. Hence, the measurement errors can be easily calculated by

\[
E = \sqrt{(x - x_0)^2 + (y - y_0)^2},
\]

(3)

where \((x_0, y_0)\) is the real coordinate.

### Table 1: Time complexity and characteristics of the center algorithm

<table>
<thead>
<tr>
<th>Center algorithms</th>
<th>Time complexity</th>
<th>Handle large data</th>
<th>The effect of receiving the order of data input</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURE</td>
<td>( O(n^2) )</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>DBSCAN</td>
<td>( O(n \log n) )</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>wave-cluster</td>
<td>( O(n) )</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>hyper-graphic</td>
<td>( O(n^2) )</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>CLARANS</td>
<td>( O(n^2) )</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>K-means</td>
<td>( O(n^2) )</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>SNN</td>
<td>( O(n^2) )</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>GA</td>
<td>according to the fitness</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>weighted centroid</td>
<td>( O(1) )</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

3.2 Dynamic optimisation algorithm

When the robot is in motion, the coordinates of tag are always changing and only one pair coordinate can be measured at the time. The maximum likelihood estimation (MLE), a regression algorithm, can be used to predict the positioning information of the dynamic operation due to its consistency, validity and invariance [23, 24]. However, since MLE algorithm is able to predict only one value at a time and horizontal and vertical coordinates of the measured data change simultaneously [25], the coordinates need to be represented by only one parameter, i.e. \((x, f(x))\) where \( f(\cdot) \) is a certain function. Therefore, a pre-processing algorithm is now developed to solve the problem of representing coordinates with a single parameter.

#### 3.2.1 Data pre-treatment: Before optimising the measured data, these data must be pre-processed to obtain more accurate positioning data. In Fig. 3, it shows that a very short time interval \([t_1, t_2]\) is chosen from the period when the robot moves from point \(T_1\), the moment \(t_1\), to point \(T_2\), the moment \(t_2\).

As the chosen time interval is short enough, the curve between the point \(T_1\) and \(T_2\) can be seen as a straight line. When the robot moves to the point \(T_1\), the point \(M\) is measured by UWB-based indoor positioning system. As shown in Fig. 4, the distance \(d\) from point \(M\) to the line \(y = kx + b\), can be easily gained and point \(P\), the pre-processed point, can be obtained according to the positional relationship between point \(M\) and the line \(y = kx + b\). According to the geometric theorem, it is not difficult to see that point \(P\) is closer to \(M(x, y)\) than \(M\). On the other hand, the prediction of abscissa and ordinate is converted to the positional relation between point \(T_2\) and the straight line \(y = kx + b\), i.e. the point \(M(x, y)\) is converted to the point \(P(x, kx + b)\), and the specific implementation can refer to our previous work Location information collection and optimisation for agricultural vehicle based on UWB, which has accepted by the Journal of Agricultural Machinery. This algorithm is applicable to simple dynamic routes, such as straight line and simple curve movement. Due to the greenhouse environment, agricultural machinery will move in a straight line, thus, this algorithm can further optimise these measured coordinate data.
3.2.2 Maximum likelihood estimation algorithm [24]: Assume that random samples \( X_1, X_2, \ldots, X_n \) have a normal distribution, i.e. \( X_i \sim N(\mu, \sigma^2) \). The likelihood \( L(\theta) \) can be calculated according to the equation:

\[
L(\theta) = \prod_{i=1}^{n} f(X_i | \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(X_i - \mu)^2}{2\sigma^2}},
\]

so under this condition, the log likelihood \( LL(\theta) \) can be calculated according to (5):

\[
LL(\theta) = \sum_{i=1}^{n} \log \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(X_i - \mu)^2}{2\sigma^2}} = \sum_{i=1}^{n} \left[ -\frac{1}{2} \log(2\pi\sigma) - \frac{(X_i - \mu)^2}{2\sigma^2} \right],
\]

where \( \bar{\mu} = (1/n) \sum_{i=1}^{n} X_i \) and \( \bar{\sigma}^2 = (1/n) \sum_{i=1}^{n} (X_i - \mu)^2 \).

4 Experimental results

4.1 Raw data

Raw data obtained from UWB positioning system is presented in Table 2. From Table 2, it can be seen that most of the measurement errors are over 150 mm, which need to be improved.

4.2 Effect of static optimisation algorithm

When the robot stays still, three groups and 500 data for each group are collected in each static position and the three groups of data in a static position are shown in Fig. 5.

From Fig. 5, it can be seen that the three sets of data were roughly distributed in one area, but there are some deviations due to measurement errors and pulse characteristics of the positioning equipment. Hence, measurement errors will be distributed at a certain distance around the real point. Also, Fig. 6 shows the effect of the static optimisation algorithm.

In Fig. 6, the first and second groups of coordinate set after weighted-centroid algorithm were close to the real points, but the third group of coordinate set after weighted-centroid algorithm was deviated from the real coordinates. Hence, from Fig. 6, it is not difficult to see that the improved weighted algorithm can reduce the errors, caused by accidental situations, between measured data and real data.

The effect of static optimisation algorithm is shown in Table 3. From Table 3, it can be seen that the effect of optimisation is closely related to the initial data. If the measured data were close to the real coordinates, the effect of optimisation would be better. On the contrary, if the measured data were far away from the real coordinates, the effect of optimisation would be rather poor.

4.3 Effect of dynamic optimisation algorithm

4.3.1 Effect of pre-treatment: In this experiment, the robot moved along the planned route three times in total, during which the coordinate data were measured and recorded. Fig. 7 shows the cumulative probability, which is utilised to better observe and compare with measured data, raw data and the data after pre-treatment. The errors between the measured coordinate and real coordinate are represented by the horizontal axis and the cumulative probability (the error less than or equal to the error value) is indicated by longitudinal axis. As shown in Fig. 7, it can be seen that the effect of the pre-processing algorithm is very significant and the errors in pre-processing data is always better than the errors in initial data. Also Fig. 8 shows the optimisation path after pre-processing, which is similar to the planned route. Therefore, the pre-processing algorithm can be applied for data processing before MLE.

Remark 1: The current indoor positioning system cannot be applied to the measurement of complex motion curves. In this paper, one pair coordinate can be represented by a single parameter.
by pre-processing because of the particularity of the environment. Generally speaking, if the path of agricultural machinery in greenhouse is straight or simple curve, these measured data can be pretreated.

4.3.2 Effect of MLE: When the robot was in dynamic state, UWB positioning devices were used to measure data. In this paper, 10,000 measured coordinates and 10,000 real coordinates were used for training models; then, 4500 data were divided into nine groups to verify the model, and the performance of MLE is shown in Table 4. From Table 4, it is not difficult to find that coordinate data optimised by MLE is far better than initial data. What is more, the average error of optimisation by MLE is 61 mm. In the initial state, the probability of the error <60 mm is 0.098. Under the condition of pre-treatment, the probability of the error <60 mm is 0.371. When the MLE was applied, the probability of the error <60 mm is 0.848. By contrast, the pre-treatment algorithm can be further optimised and the probability of data accuracy can be maximised by MLE algorithm.

Remark 2: By comparing the static and dynamic positioning accuracy, it is found that the accuracy of static positioning is slightly lower than dynamic positioning, and the processing time of the static optimisation is much longer. However, there is a very obvious and irreplaceable advantage for static positioning optimisation, which is that there is no probability value with static optimisation and the optimisation effect is more stable than dynamic optimisation. In practical operation process, the agricultural machinery is in a static state for a large part of the time, and only a small part of the time is in motion. Hence, the stability requirement of static positioning accuracy is much higher than that of dynamic positioning.

5 Conclusion

In this paper, the UWB-based indoor positioning system was used to estimates agricultural machinery position in the green house. However, there were many measurement errors when the agricultural machinery was working in practical process. In order to reduce and control the errors, improved weighted-centroid algorithm and MLE algorithm have been developed and tested. When agricultural machinery was in static operation, the average error was 66.4 mm with improved weighted-centroid algorithm and MLE algorithm have been developed and tested. When agricultural machinery was in static operation, the average error was 66.4 mm with improved weighted-centroid algorithm. When agricultural machinery was in dynamic operation, the average error was 61 mm and the probability of the error <60 mm was 0.848 with MLE algorithm, which meets the practical demands in greenhouse operations. In a future study, other optimisation algorithms will be proposed to further reduce the errors of measured data.
6 Acknowledgments
This work was supported by Natural Science Foundation of China (Grant No. 61773216) and the Natural Science Foundation of Jiangsu Province of China (Grant No. BK20171386).

7 References