Development of network communication system for locomotion robots

1.Abstract

The aim of the project is to develop a network system for communication between autonomous robots and a cloud server, and among these robots. The most critical issue is to build a high-accuracy positioning system with an error of up to centimeters, which means that the accuracy of the GPS solution cannot meet the requirements. Here we build a localization system for robots. After comparing the pros and cons of various technologies, we chose to use the UWB (Ultra Wide Band) scheme for the localization system. We analyzed the communication algorithm of base stations and tags in the UWB module. Then, we build a prototype localization system. The system has four base stations to show the 3-dimensional coordinates of the tags. Two tags are put in the system to verify the stability, and the number of tags can be increased to 20. After doing experiments in different base station locations and heights, we achieved the expected accuracy in a small range of about 2 by 2 by 2 meters. The system also runs stably in a larger measuring area of about 20 by 20 by 2 meters, but the accuracy under that condition needs to be improved.

2. Localization Principle behind UWB Scheme

2.1. Basic principle of UWB ranging

(1) TOF (Time-of-Flight ranging method): ranging method belongs to the two-way ranging technology, it mainly uses the signal in two asynchronous transceivers between the Flight Time to measure the distance between nodes. Because in the line-of-sight environment, the TOF ranging method shows a linear relationship with distance, so the result will be more accurate. We denote the time difference between the packets sent and received by the sender as T_{TOT}, and the time difference between the packets received and sent by the receiver as T_{TAT}. Consequently, the one-way flight time of packet in the air can be calculated as: $T_{TOF} = \frac{T_{TOT}-T_{TAT}}{2}$ (see Figure 1).

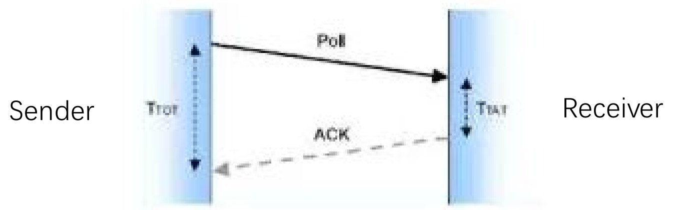


Figure 1. TOF ranging method.

But pure TOF algorithm has a relatively strict constraint: the sending and the receiving devices must be running clocks at identical frequencies.

(2) TW-TOF (Two-way Time-of-Flight Method):

(a) SS-TWR (Single-sided Two-way Ranging) is a simple measure of the time of a single round-trip message. Device A (the Sender) actively sends data to Device B (the Receiver), and Device B returns data in response to Device A (see Figure 2).

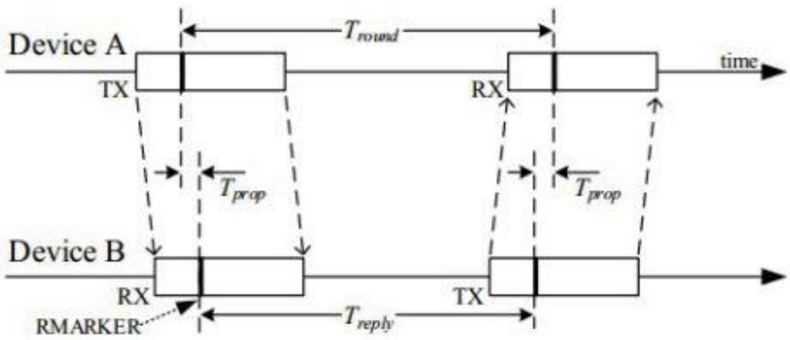


Figure 2. SS-TWR.

Ranging process: Device A actively sends (TX) data and records the sending time-stamp. After receiving the data, Device B records the receiving time- stamp. After the T_{reply} delay, Device B sends the data and records the sending time-stamp, while Device A receives the data and records the receiving time-stamp. Based on the two local clocks running at the same nominal frequency f_{clk} Therefore, two nominal time differences data can be obtained, respectively T_{tround} and T_{reply} of Device A and Device B. Finally, the nominal flight time T_{prop} of the wireless signal is obtained as follows:$T_{prop} = \frac{T_{tround-}T_{reply}}{2}$.

Assuming the actual frequencies of the clocks are denoted respectively by $(1 + e_A)f_{clk}$ and $(1 + e_A)f_{clk}$, the difference time recorded by the local clocks are $T_{tround} = (1 + e_A)T_{tround}$ and $T_{reply} = (1 + e_A)T_{reply}$. Consequently, the recorded flight time T_{prop} is given by

$$\begin{aligned}
T_{prop} &= \frac{T_{tround} - T_{reply}}{2} \\
&= \frac{\left(\hat{T}_{tround} - \hat{T}_{reply}\right) + \left(e_A\hat{T}_{tround} - e_B\hat{T}_{reply}\right)}{2} \\
&= \frac{1}{2}(1 + e_A)\left(\hat{T}_{tround} - \hat{T}_{reply}\right) - \frac{1}{2}(e_B - e_A)\hat{T}_{reply} \\
&= (1 + e_A)\hat{T}_{prop} - \frac{1}{2}(e_B - e_A)\hat{T}_{reply}
\end{aligned}$$

Assuming $e_A \ll 1$ such that $1 + e_A \approx 1$ one obtains

$$T_{prop} \approx \hat{T}_{prop} - \frac{1}{2}(e_B - e_A)\hat{T}_{reply}$$

Consequently, the ranging error ε is given by

$$\varepsilon = \hat{T}_{prop} - T_{prop} \approx \frac{1}{2}(e_B - e_A)\hat{T}_{reply}$$

The smaller the $\hat{T}_{reply}$, the more accurate the ranging. In addition, $\hat{T}_{reply}$ covers not only the receiving and sending time of Device B, but also the loading and sending time of data. (In addition to supporting location, UWB can also transmit data: 128 bytes in the Standard Mode, and 1024 bytes in the Extended Mode.) The typical error relationship is as follows: 1 ns is approximately equal to the 750 px ranging error (see Figure 3).

clock error / T_{reply}	2 ppm	5 ppm	10 ppm	20 ppm	40 ppm
100 μs	0.1 ns	0.25 ns	0.5 ns	1 ns	2 ns
200 μs	0.2 ns	0.5 ns	1 ns	2 ns	4 ns
500 μs	0.5 ns	1.25 ns	2.5 ns	5 ns	10 ns
1 ms	1 ns	2.5 ns	5 ns	10 ns	20 ns
2 ms	2 ns	5 ns	10 ns	20 ns	40 ns
5 ms	5 ns	12.5 ns	25 ns	50 ns	100 ns

Figure 3.

Let the image resolution is a, the number of pixels is b, then the distance of ranging error c is given by

$$c = \frac{b}{a}$$

It can be seen that ε increases with the increase of $\hat{T}_{reply}$, e_A and e_B, resulting in inaccurate ranging. Therefore, unilateral two-way ranging (SS-TWR) is not commonly used. It can be used for certain applications when the demand for accuracy is low but that on a shorter ranging time is high.

(b) DS-TWR(Double-sided Two-way Ranging) Double-sided two-way Ranging is an extended ranging technique that records Two round-trip time-stamps to get the flight time. This increases the response time but reduces the ranging error. Double-sided two-way ranging can be divided into two methods according to the number of messages sent: 4 Messages and 3 Messages.

i. 4 Messages: Device A initiates the first ranging message and Device B responses, generating 2 time-stamps T_{round1} and T_{reply1}; then after some time, Device B initiates the ranging, and Device A responds, generating two more time stamps T_{round2} and T_{reply2} (see Figure 4).

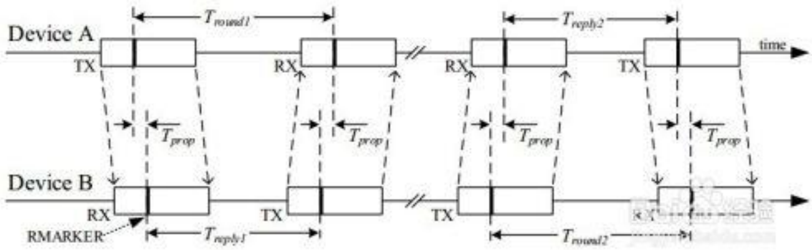


Figure 4: DS-TWR 4 message method.

ii. 3 Messages: Compared with the 4 message mode, the initiation of the second ranging is avoided. When Device A receives the data, it will immediately return the data, and finally, the following four time differences can be obtained: T_{round1}, T_{reply1}, T_{round2}, T_{reply2} (see Figure 5).

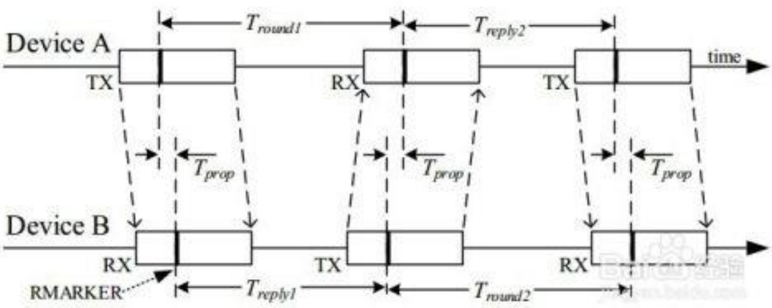


Figure 5: DS-TWR 3 message method.

As the nominal flight time for both 4 and 3 message protocols is given by

$$\hat{T}_{prop} = \frac{1}{2}(\hat{T}_{tround1} - \hat{T}_{reply1}) = \frac{1}{2}(\hat{T}_{tround2} - \hat{T}_{reply2})$$

Hence $\hat{T}_{tround1,2} = 2\hat{T}_{prop} + \hat{T}_{reply1,2}$.Consequently, one obtains

$$\hat{T}_{tround1}\hat{T}_{tround2} = \left(2\hat{T}_{prop} + \hat{T}_{reply1}\right)\left(2\hat{T}_{prop} + \hat{T}_{reply2}\right)$$

$$\hat{T}_{tround1}\hat{T}_{tround2} - \hat{T}_{reply1}\hat{T}_{reply2} = \hat{T}_{prop}\left(2\hat{T}_{reply2} + 2\hat{T}_{prop} + 2\hat{T}_{reply1} + 2\hat{T}_{prop}\right)$$

$$= \hat{T}_{prop}\left[\left(\hat{T}_{reply2} + 2\hat{T}_{prop}\right) + \hat{T}_{reply2} + \left(\hat{T}_{reply1} + 2\hat{T}_{prop}\right) + \hat{T}_{reply1}\right]$$

$$= \hat{T}_{prop}\left(\hat{T}_{tround2} + \hat{T}_{reply2} + \hat{T}_{tround1} + \hat{T}_{reply1}\right)$$

$$\hat{T}_{prop} = \frac{\hat{T}_{tround1}\hat{T}_{tround2} - \hat{T}_{reply1}\hat{T}_{reply2}}{\hat{T}_{tround2} + \hat{T}_{reply2} + \hat{T}_{tround1} + \hat{T}_{reply1}}$$

Similarly, the recorded flight time is given by

$$T_{prop} = \frac{T_{tround1}T_{tround2} - T_{reply1}T_{reply2}}{T_{tround2} + T_{reply2} + T_{tround1} + T_{reply1}}$$

$$= \frac{(1+e_A)\hat{T}_{tround1}(1+e_B)\hat{T}_{tround2} - (1+e_B)\hat{T}_{reply1}(1+e_A)\hat{T}_{reply2}}{(1+e_B)\hat{T}_{tround2} + (1+e_A)\hat{T}_{reply2} + (1+e_A)\hat{T}_{tround1} + (1+e_B)\hat{T}_{reply1}}$$

For $e_A, e_B \ll 1$, one ignores the second-order term in e_A and e_B when expanding the numerator and obtains

$$T_{prop} \approx \frac{(1+e_A+e_B)(\hat{T}_{tround1}\hat{T}_{tround2} - \hat{T}_{reply1}\hat{T}_{reply2})}{(1+e_B)\hat{T}_{tround2} + (1+e_A)\hat{T}_{reply2} + (1+e_A)\hat{T}_{tround1} + (1+e_B)\hat{T}_{reply1}}$$

$$= \frac{(1+e_A+e_B)(\hat{T}_{tround1}\hat{T}_{tround2} - \hat{T}_{reply1}\hat{T}_{reply2})}{(\hat{T}_{tround2}+\hat{T}_{reply2}+\hat{T}_{tround1}+\hat{T}_{reply1})[1+\frac{e_A\left(T_{tround1}+T_{reply2}\right)+e_B\left(T_{tround2}+T_{reply1}\right)}{T_{tround2}+T_{reply2}+T_{tround1}+T_{reply1}}]}$$

$$= \hat{T}_{prop}\frac{(1+e_A+e_B)}{[1+\frac{e_A\left(T_{tround1}+T_{reply2}\right)+e_B\left(T_{tround2}+T_{reply1}\right)}{T_{tround2}+T_{reply2}+T_{tround1}+T_{reply1}}]}$$

Again for $e_A, e_B \ll 1$ and ignoring the second-order terms in e_A and e_B, one obtains

$$T_{prop} \approx \hat{T}_{prop}(1+e_A+e_B)\left[1 - \frac{e_A\left(\hat{T}_{tround1} + \hat{T}_{reply2}\right) + e_B\left(\hat{T}_{tround2} + \hat{T}_{reply1}\right)}{\hat{T}_{tround2} + \hat{T}_{reply2} + \hat{T}_{tround1} + \hat{T}_{reply1}}\right]$$

$$\approx \hat{T}_{prop}\left[1 + e_A + e_B - \frac{e_A\left(\hat{T}_{tround1} + \hat{T}_{reply2}\right) + e_B\left(\hat{T}_{tround2} + \hat{T}_{reply1}\right)}{\hat{T}_{tround2} + \hat{T}_{reply2} + \hat{T}_{tround1} + \hat{T}_{reply1}}\right]$$

Consequently

$$\varepsilon = \hat{T}_{prop} - T_{prop}$$

$$\approx -\hat{T}_{prop}\left[e_A + e_B - \frac{e_A\left(\hat{T}_{tround1} + \hat{T}_{reply2}\right) + e_B\left(\hat{T}_{tround2} + \hat{T}_{reply1}\right)}{\hat{T}_{tround2} + \hat{T}_{reply2} + \hat{T}_{tround1} + \hat{T}_{reply1}}\right]$$

The above-ranging mechanisms are asymmetric ranging methods because they do not have to be the same for response time. Even with a 20 ppm crystal, the clock error is in the picosecond level. The error formula is as follows:

$$error = \hat{T}_{prop}(1 - \frac{k_a + k_b}{2})$$

Device A runs at the required frequency k_a f_{clk}, and Device B runs at the required frequency k_b f_{clk}. Both k_a and k_b are close to the value 1.

To figure out the value of the error, if Devices A and B are running with poor crystal oscillation (20 ppm error), for example, Device A is 20-ppm slower, Device B is 20 ppm faster, or put over, this will result in a total error of 40 ppm, so k_a and k_b may be 0.99998 or 1.00002.

Even if the UWB operates over a large range, say 100 m, the air flight time of the wireless signal is about 333ns, because the error is: $20 \times 10^{-6} \times 333 \times 10^{-9} = 6.7 \times 10^{-12} = 6.7\text{ps}$, which is only 2.2 mm after the range is converted.

Note that the response times do not have to be equal and that the T_{reply1} does not have to be equal to the T_{reply2}, which provides a lot of convenience for MCU system processing.

The main source of error must be whether the timestamp of the received data is correct. Not the PPM of the crystal.

2.2.UWB positioning principle

The more mature positioning algorithms are TOA (time of arrival), TDOA (time difference of arrival), AOA (Angle of arrival or called DOA estimation) positioning technology, and the mixture of these three technologies. Here we adopt the TDOA method.

TDOA positioning does not need to synchronize between base stations and mobile terminals, but only between base stations. Because base stations are fixed, it is much easier to synchronize between base stations than between base stations and mobile terminals. This makes TDOA positioning much easier to implement than TOA positioning, so TDOA positioning is widely used.

It can locate by measuring the transmission delay difference between two different base stations and mobile terminals. Assuming that the distance difference between the position of the mobile terminal and Base Stations 1 and 2 is $R_{21} = R_2 - R_1$, then the position of the mobile terminal must be on the hyperbola with the two base stations as the focus and the distance difference between the two focal points is constant R_{21}. That is if the position of the mobile terminal is (X_0, Y_0), the position of Base Station 1 is (X_1, Y_1), and the position of Base Station 2 is (X_2, Y_2), then they satisfy the relation:

$$R_{21} = R_2 - R_1 = \sqrt{(X_0 - X_2)^2 + (Y_0 - Y_2)^2} - \sqrt{(X_0 - X_1)^2 + (Y_0 - Y_1)^2}$$

Then another set of hyperbolas can be obtained through the TDOA of another pair of mobile terminals: R_{13} for Base Stations 1 and 3, or R_{32} for Base Stations 2 and 3. The two sets of hyperbolas will produce at most two intersection points, and then the location of the mobile terminals can be determined according to prior knowledge, such as the distance between the mobile terminal and base stations R_2 and R_1 (see Figure 6).

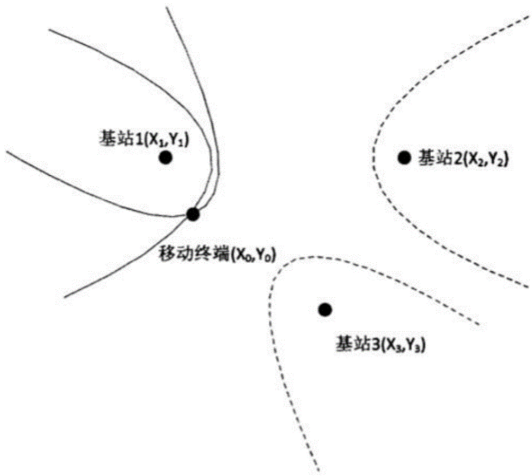
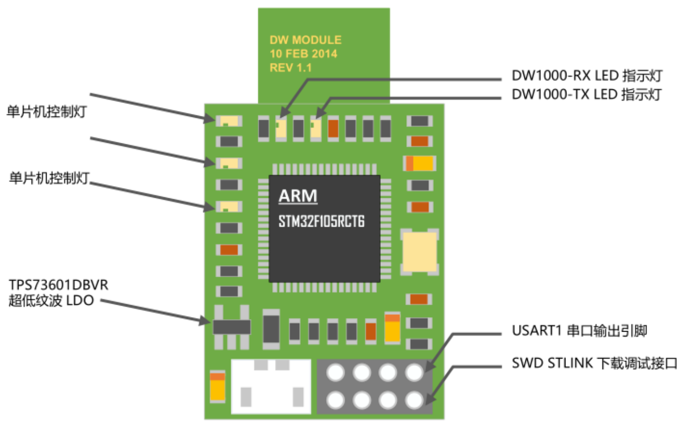


Figure 6. The basic principle of the TDOA method.

3.UWB positioning prototype system design

3.1. Materials used

The list shows the overview of the module used in this project.

YCHIOT UWB Mini 4 module series	
Module picture	

Main control chip	STM32F103RCT6 microcontroller
Peripheral circuits	DWM1000 module, power module, LED indicator module, reset circuit, etc.

Figure 7 shows the hardware parameters of the module.

基本参数		无线参数	
PCB 工艺	4 层板-环氧树脂	通讯速率	110 kbit/s, 850 kbit/s, 6.8 Mbit/s
供电接口	micro-USB(5.0V) / 接线柱	工作频率	3.5 GHz ~ 6.5 GHz
通讯接口	micro-USB(5.0V) / 串口 (3.3V TTL)	工作频道	6
下载接口	SWD (VCC SDIO SCK GND)	发射功率	-35dbm/MHZ ~ -62dbm/MHZ 可程控
主控制器	STM32F103RCT6(64pin)	最大包长	1023 字节
外部晶振	8Mhz	通讯距离	Mini4-CA 为 30m，Mini4-SMA 为 80 米
PCB 尺寸	35mm * 24mm	数据抖动	典型±10cm，一般遮挡±30cm

Figure 7. Hardware parameters of the module.

3.2. Operating procedure

(1) Hardware platform networking. All tags are powered by a charger; Base stations A1/A2/A3 are powered by a power bank.

(2) Install the virtual serial port driver VCP1.4.0 on the computer.

(3) A0 base station is directly connected to the computer with USB;

(4) Open the upper computer software DecaRangeRTLS.exe, input the base station coordinates, record the tag location data, that is, the measuring coordinates.

(5) Measure the actual location of tags, calculate the actual coordinate.

3.3. Experiment process and results

(1) Test the distance measuring system of UWB hardware. Using one base station and one tag.

(2) Using the UWB hardware, I designed a simple positioning system, using four base stations and two tags.

(3) In my first time in the base station setting, because of the small site conditions, I built the system in a space about 1 by 1 by 1 meter. The coordinates of base stations are: A0(0.00,0.00,0.00), A1(0.00,1.00,0.00), A2(1.00,0.00,0.00), A3(1.00,1.61,0.50). The system successfully operated. The errors of the x-axis and y-axis

are around 0.1 to 0.2 m. and the errors on the z-axis are even higher (see Figure 8).

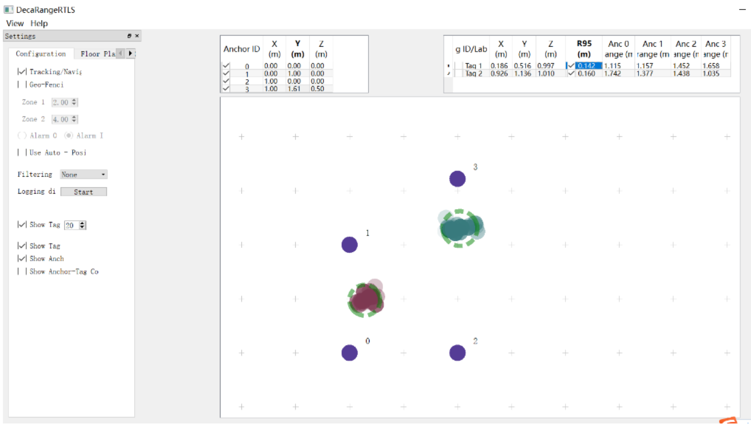


Figure 8.

It is not a satisfying result. My guess on the errors is, the space is too small for the system to operate. Next time I will try to build the system in a bigger place, about 10 by 10 by 3 meters. A stable indoor place, with few disturbances and enough space, is needed.

(4) To test the system in a suitable space, I went to the local government affairs hall. Building a system with a bigger range, I collected the data. The distribution of base stations and the data is as follows:

(a) The first test, set the base stations to be regular shape in a space about 8 meters by 8 meters by 3 meters. The coordinates of base stations are: A0(0.00,0.00,0.00), A1(0.00,8.00,0.00), A2(8.00,0.00,0.00), A3(8.00,8.00,1.35). (see Figure 9).

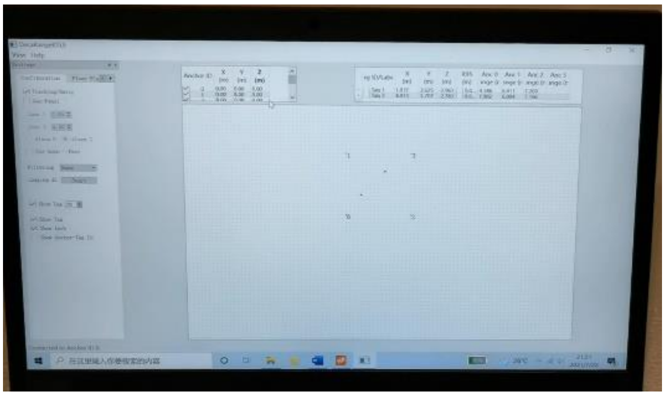


Figure 9.

The test result is as follows (2-dimensional).

	Measuring Coordinates	Actual Coordinates
Tag 1	(1.8,2.6)	(1.8,2.9)

	(3.0,-0.1) (4.1,4.0)	(3.0,0.2) (4.0,4.0)
Tag 2	(4.7,5.7) (8.1,4.4) (1.7,13.1)	(4.4,5,4) (7.9,4.3) (1.5,12.7)

(b) The second test set the base stations to be an unregular shape. The coordinates of base stations are: A0(0.00,0.00,1.35), A1(1.50,10.00,1.35), A2(20.00,2.00,1.35), A3(9.70,-2.35,0.8) (see Figure 10).

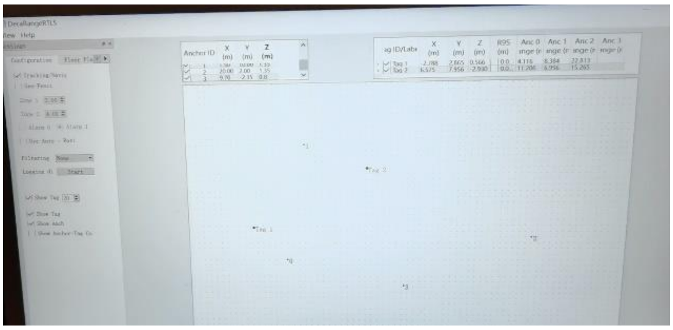


Figure 10.

The test result is as follows (2-dimensional).

	Measuring Coordinates	Actual Coordinates
Tag 1	(-1.7,2.7) (4.4,4.6) (3.6,-4.4)	(-1.5,2.9) (4.2,4.5) (3.8,-4.2)
Tag 2	(11.3,4.9) (5.6,0.9) (15.5,-0.2)	(11.2,5.0) (5.5,1.0) (15.2,-0.1)

(c) Additionally, I found a volunteer who took one tag and move so I could test the appearance of the system when tags are moving. The system operated normally (see Figure 11).

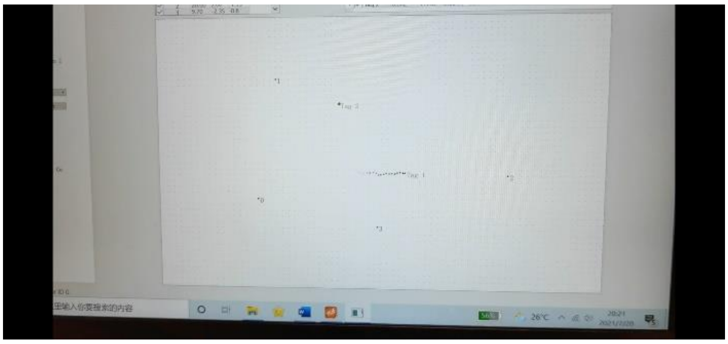


Figure 11.

(d) Error reduction

Through experiments, I found some methods to reduce error.

i. The way of setting base stations has a big influence on measuring accuracy. Base stations cannot be set close to the wall, or close to the ground. Too close to the plane will produce a larger multipath effect, resulting in a larger error. The appropriate height placement is 1.5m to ground or higher, and the appropriate height difference between Base Station A3 and other basements is 50 cm.

ii. UWB modules cannot be placed near the metal. Metal may have interference with the transmission of the signal.

iii. Put the base stations precisely in the location whose coordinates were entered in the software in advance. This is the most important point because every tiny error in the base station locating will cause a big error in tag measuring.

iv. When the tag is put in the line from a base station to another, the error will be bigger.

v. About the data on the z-axis, the heights of both tags have a big difference from the actual heights, but the height difference between tags is accurate. So, the algorithm to adjust height measuring is needed, but the error of measuring is acceptable.

The new base stations' locations and test result are as follow. The coordinates of base stations are: A0(0.00,0.00,1.60), A1(2.00,0.00,1.60), A2(0.00,2.00,1.60), A3(2.00,2.00,2.10) (see Figure 12).

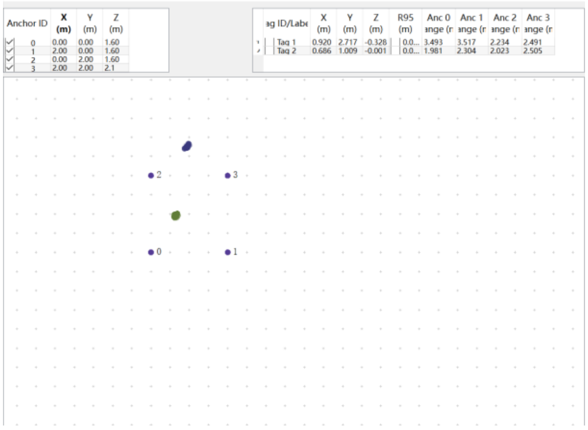


Figure 12.

	Measuring (Coordinates)	Actual (Coordinates)
Tag 1	(1.24m,1.20m) (0.48m,1.54m)	(1.26m,1.23m) (0.50m,1.50 m)
Tag 2	(0.72m,0.91m) (1.03m,0.61m)	(0.67m,0.90m) (1.00m,0.60m)
Height difference	23cm	25cm

We can see that the error is reduced to lower than 5 cm, reach to our estimate.

4.Discussion

Overall, the project explored indoor positioning schemes and algorithms used and built a prototype indoor positioning system. Through experiments, multiple methods were used to reduce error to an estimated level. Future iterations of the system may be enlarging the measuring area and keeping the error at a low level at the same time. A space like the UST football field is a possibility. To keep the accuracy remains the same while increasing the measuring area, improving signal transmission power and improving existing algorithms, and adjust related parameters may be the ways to reach our new target.