Sound-based remote sensing of terrestrial animals: localization and error analysis

SHEN Shaoqing[1], GONG Peng[1,2], CHENG Xiao[3], YING Qing[1]

1. *State Key Laboratory of Remote Sensing Science*, *Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University*, *Beijing* 100101, *China*;
2. *Department of Environmental Science*, *Pallcy and Management*, *University of Calitornia*, *Berkeley*. *CA*, *USA*;
3. *College of Global Change and Earth System Science*, *Beijing Normal University*, *Beijing* 100875, *China*

Abstract: Advanced sound-based remote sensing technology for terrestrial animals could greatly enhance our ability of animals monitoring, research and protection. In this research, we designed an inexpensive bioacoustic localization system for terrestrial animals, which integrates commercial off-the-shelf recorders and wireless controllers and is much cheaper than most animal localization systems. Combined with the bioacoustic localization software we developed, the system is verified to have the ability to fulfill our requirements because its localization errors along the X and Y directions are both less than 1.69 m, although the error along the Z direction is a little bit larger. In order to assess the factors influencing the localization accuracy, we applied Monte Carlo simulation method to conduct error analysis. We found that errors, including in surveying recording stations, estimating Time Difference of Arrival (TDOA), and estimating sound velocity, will all influence the final localization accuracy. Besides, the Monte Carlo method could also be used for choosing the values of system parameters when implementing a bioacoustic localization system of terrestrial animals, such as the total number of recording stations and spatial ranges of the system.

Key words: sound-based remote sensing, bioacoustic localization, biological monitoring, error analysis, Monte Carlo

CLC number: TP79 **Document code:** A

Citation format: Shen S Q, Gong P, Cheng X and Ying Q. 2011. Sound-based remote sensing of terrestrial animals: localization and error analysis. *Journal of Remote Sensing*, **15**(6): 1255–1275

1 INTRODUCTION

Many efforts have been made to establish various monitoring systems for animals, especially endangered wildlife, in order to assess the impacts of human activities and more effectively protect the animals. Different types of information are extracted while monitoring, among which, location of animals is of special interest. This is because, in the first place, habitat is the spatial media where interactions between animals and their environment as well as interactions among different kinds of animals come into effece. The more precise the habitat is, the more effectively can people protect them. Furthermore, location information can be used to analyze animals' behaviors. For example, behavioral researchers found that individual animals in a chorus dynamically adjust spacing in order to influence behaviors such as mate choice (Jones & Ratham, 2009; Mennill, *et al*., 2006). Finally, localizing animals in time can avoid tragedies resulting from spatial collisions between mankind and animals, such as avian localization systems developed for avoiding strikes between birds and planes (Ning, *et al.*, 2010). Hence it can be seen that localization systems of animals are very useful and worth developing.

Among all kinds of localization systems of animals, the bioacoustic localization system has its specific advantages. Firstly, as a passive method, the natural movements of animals are not inhibited by bioacoustic localization. Secondly, sound is one way for animals to communicate. This ensures a bioacoustic localization system can record communication between animals while localizing, which can be further used in species analysis or behavior research and so forth (Mennill, *et al.*, 2006). Thirdly, the fact that sound travels farther in water than light does make a bioacoustic localization system the most suitable device to locate animals living in the water, such as the famous Right Whale Listening Network (Clark, *et al*., 2007). Last but not least, acoustic device is not constrained by light condition and can be used to locate animals in the night. For example, migratory movement of birds is recorded in the night, and all those recordings are important references for making migratory bird conservation plans (Farnsworth, 2005).

Bioacoustic localization systems have been used extensively to locate many marine organisms, especially large mammals in the

Received: 2010-12-06; **Accepted:** 2011-07-01

Foundation: The National High Technology Research and Development Program(863 Program)(No. 2006AA12Z112, 2009AA12200101)

First author biography: SHEN Shaoqing (1985—), female, Ph. D. candidate. Her research interests are sound-based remote sensing, bioacoustics and environmental change. E-mail: s_s_q@126.com

Corresponding author biography: GONG Peng (1965—), professor, he received his Ph. D. in Geography at the University of Waterloo in 1990. His main research interest includes land-cover/land-use mapping and change detecting, photo-ecometrics, global change and human health and technology and society. He has published more than 400 academic papers and over 7 books. E-mail: gong@irsa.ac.cn

ocean (Stafford, *et al.*,1998; Wahlberg, *et al.*, 2001, 2003; Mellinger & Clark, 2003; Clark & Clapham, 2004). This is mainly because there is no other more effective method to locate marine animals. However, the applications of bioacoustic localization systems for terrestrial animals are not well known (Mennill, *et al.*, 2006). In a classic bioacoustic localization system for terrestrial animals, multiple microphones are usually put in a place precisely and connected to a computing system via a multi-channel audio card (McGregor, *et al.*, 1997; Exadaktylos, *et al*., 2008; Jones & Ratham, 2009; Mennill, *et al.*, 2006). Albeit the single audio card guarantees time synchronization among all the channels and this system performs well for locating animals within a small area, such systems still need improvements. First, the centralized way to acquire and compute audio data render the system highly dependent on some components (*e.g.*, the audio card and the computing system). Second, wired connection often limits the relative position and distance between any two of the microphones, the audio card and the computing system, resulting in low level of flexibility in implementing and reconstructing. Also, exposed wires are vulnerable to damage from human beings, other animals and the environment.

Recent advances in Wireless Sensor Network (WSN) technology can be used to improve the current bioacoustic localization systems for terrestrial animals. The WSN technology refers to a modern information technology which integrates the sensor technology, automation technology, data transmission, data storage, data process and analysis technology (Gong, 2007). A WSN-based bioacoustic localization system for terrestrial animals consists of multiple recording nodes which cooperate to localize the animals. A single recording node is a tiny integrated system which is composed of microphone sensors, audio processors, audio storage, microprocessor, wireless communication module, and power module.

WSN technology improves traditional bioacoustic localization from the following aspects. First, the wireless connection among the recording nodes renders much more convenient implementation in the field and the dispatch of data acquisition. Computation to each recording node makes the system more robust, although this also causes another challenge: time synchronization of recording nodes. Second, WSN-based recording nodes allow embedded Global Positioning System (GPS) modules for self positioning. This makes bioacoustic localization much more automatic and convenient because the coordinates of recording nodes are traditionally obtained by using additional survey instruments manually. Third, WSN-based recording nodes also allow integrating temperature sensors and humidity sensors (or sound velocity meter instead). These sensors are used to estimate sound velocity in the air, resulting in more accurate bioacoustic localization. Fourth, the computation ability of recording nodes enables embedded bioacoustic recognition, which helps the bioacoustic localization system focus on species of interest.

This paper aims at developing a WSN-based bioacoustic localization system for terrestrial animals and discussing its accuracy. First, the architecture of the WSN-based bioacoustic localization system for terrestrial animals is introduced. Meanwhile, a prototype system including software and hardware is demonstrated. The outdoor experiments and relevant results are also included. Next we present an assessment of different factors which influence the accuracy of the bioacoustic localization system for terrestrial animals, and propose an algorithm based on the Monte Carlo method to quantitatively analyze the location errors, we conclude and discuss about the future work at last.

2 SYSTEM ARCHITECTURE

The architecture of a WSN-based bioacoustic localization

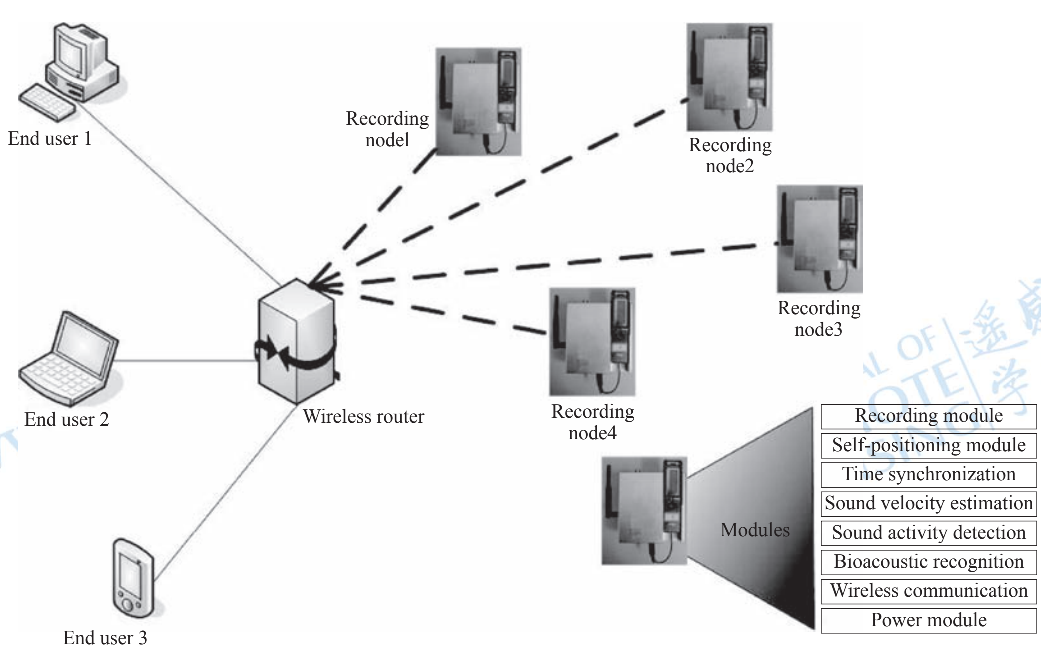


Fig. 1 The architecture of a WSN-based bioacoustic localization system for terrestrial animals

system for terrestrial animals is shown in Fig. 1. Generally, there should be eight modules in a recording node: (1) recording module; (2) self-positioning module; (3) time synchronization; (4) sound velocity estimation; (5) sound activity detection; (6) bioacoustic recognition; (7) wireless communication; (8) power module. The bioacoustic localization system for terrestrial animals has different needs from common localization systems, such as self-positioning, sound velocity estimation, and recognition of animal species (sound activity detection and bioacoustic recognition), and WSN technology could be employed to fulfill these special needs.

2.1 Self-positioning

The coordinates of recording nodes are necessary inputs for bioacoustic localization and their accuracy will influence the final localization accuracy (Section 4.1). The traditional way to survey recording nodes in the field is to use a total station in terms of several ground control points. Although this way can achieve very high accuracy, it is labor intensive when implementing, Furthermore, it is difficult to find appropriate ground control points in the field. Wireless Sensor Network allows equipping each recording node with an embedded GPS module so that each recording node can easily localize itself and send the position information through wireless network to relevant devices for later bioacoustic localization. Hence displacement and movement of recording nodes will not be a problem because recording nodes will send out the new position timely and automatically, ensuring possible of system calibration.

2.2 Sound velocity estimation

The velocity of sound is not a constant, and it depends upon temperature, barometric pressure, altitude, humidity, wind velocity. The velocity of sound is (331.45±0.05) m/s in still, dry air under standard conditions of temperature and pressure (0℃ and 760 mm of Hg pressure).

The velocity of sound increases as temperature increases in a relationship of

$$c = 331.45 \times \sqrt{1 + \frac{t}{273}} \tag{1}$$

where c refers to the velocity of sound, and t is the temperature in degrees Celsius. Accordingly, when temperature decreases with increasing altitude, the velocity of sound will also decrease as altitude increases. Meanwhile, the sound is refracted upward, away from listeners on the ground, creating an acoustic shadow in a distance from the source. Otherwise, when temperature increases with increasing altitude, the velocity of sound increases as altitude increases, and sound is refracted downward (Everest & Ken, 2009).

The velocity of sound also increases with increasing moisture, following in a rule of

$$c = 1508.528\sqrt{\frac{R_w}{M_w}} \tag{2}$$

where $R_w = \frac{7+h}{5+h}$ is the universal gas constant, and M_w=29–11h, refers to the mean molecular weight of the gas at sea level, h is defined to be equal to the fraction of molecules that are water.

If the sound travels with the wind, the velocity of sound will increase; otherwise, the velocity of sound will decrease. In addition, wind could alter the main direction of sound (Everest & Ken, 2009).

Since bioacoustic localization of terrestrial animals usually occurs outdoor and the above environmental parameters can vary all the time, nearly real-time sound velocity estimation is necessary for enhancing the localization accuracy. WSN technology could integrate temperature sensor, humidity sensor, and wind velocity sensor into the whole a system for sound velocity estimation, or directly use sound velocity meter instead.

2.3 Recognition of animal species

People always have clear goals when using a bioacoustic localization system and how to localize animals of their interest is the key for a practical bioacoustic localization system. Therefore, the step before bioacoustic localization is recognition of animal species, which could be accomplished by bioacoustic recognition. Since bioacoustic recognition of animal species is also based on the features of animal vocalizations and WSN-based recording nodes have the ability of computing, it could be possible to perform embedded bioacoustic recognition of animal species inside recording nodes before localization, resulting in a much smarter bioacoustic localization system, such as the system described in Exadaktylos (2008).

Bioacoustic recognition of animal species is essentially a pattern recognition. There are many methods, such as template matching (Anderson, *et al.*, 1996), neural network (McIlraith & Card, 1997), Gaussian mixture models (Kwan, *et al.*, 2005), support vector machine (Fagerlund, 2007), Hidden Markov Model (Kogan & Margoliash, 1998).

3 PROTOTYPE SYSTEM

According to the above architecture of a WSN-based bioacoustic localization system for terrestrial animals, we develop a prototype consisting of hardware (Section 3.2) and software (Section 3.3). We test its localization performance in the field, detailed in Section 3.4 and 3.5.

3.1 Principle and method of bioacoustic localization

There are mainly two ways to locate animals acoustically. One is based on energy-range which is used in the above mentioned “Right Whale Listening Network”. The other is the TDOA based calculations (Wahlberg, *et al.*, 2001) in which animals can be localized according to the time difference of animal vocalizations arriving at any two recordings nodes that are set up precisely at known positions. The TDOA method is the one we use in this paper. First, set up four (or more) recording nodes(Fig. 2), and survey their positions accurately. Then estimate time differences of the bioacoustic signal arriving at each recording node based on the audio recordings. Last, localize the uttering animals according to the TDOAs.

Consider a bioacoustic localization system with N recording nodes at known locations.

$$T_i = \frac{1}{c}\sqrt{(x_i - sx)^2 + (y_i - sy)^2 + (z_i - sz)^2} \qquad i=1,\cdots,n \tag{3}$$

Let vector (x_i, y_i, z_i) denotes the true coordinate of recording node i, c denotes the true sound velocity. Let (sx, sy, sz) denotes the to-be-solved coordinate of the target animal. Randomly pick up the r^{th} recording node as the reference one, then TDOA τ_{ri} between the reference node and i^{th} recording node can be calculated by

$$
\begin{aligned}
\tau_{ri} &= T_r - T_i \\
&= \frac{1}{c}\sqrt{(x_r - sx)^2 + (y_r - sy)^2 + (z_r - sz)^2} - \\
&\quad \frac{1}{c}\sqrt{(x_i - sx)^2 + (y_i - sy)^2 + (z_i - sz)^2} \\
&\quad (i=1,\cdots, n,\ i\neq r)
\end{aligned} \tag{4}
$$

Sound is assumed to travel in the air in a uniform linear motion at velocity $c^{'}$. The coordinates of the recording node can be surveyed with the results denoted by ($x^{'}_i$, $y^{'}_i$, $z^{'}_i$). The arrival time difference τ_{ri} can be estimated through the Generalized Cross Correlation (GCC) method with the results denoted by $\tau^{'}_{ri}$. Then the error in TDOA could be expressed by Eq. (4).

The source (*sx, sy, sz*) can be solved when the following objective function *F* achieves the minimum. *F* is the sum of the squared error, *i.e.*

$$
F = \sum_{i=1}^{n} f_i^2(c^{'}, \tau^{'}_{ri}, x^{'}_r, y^{'}_r, z^{'}_r, x^{'}_i, y^{'}_i, z^{'}_i, sx, sy, sz) \tag{5}
$$

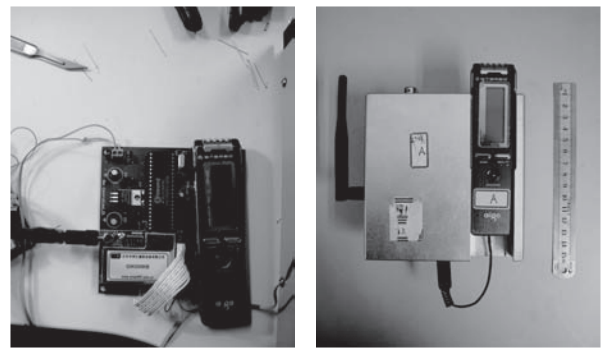
$$
\begin{aligned}
& f_i(c^{'}, \tau^{'}_{ri}, x^{'}_r, y^{'}_r, z^{'}_r, x^{'}_i, y^{'}_i, z^{'}_i, sx, sy, sz) \\
&= \tau^{'}_{ri} - \frac{1}{c^{'}}\sqrt{(x^{'}_r - sx)^2 + (y^{'}_r - sy)^2 + (z^{'}_r - sz)^2} + \\
&\quad \frac{1}{c^{'}}\sqrt{(x^{'}_i - sx)^2 + (y^{'}_i - sy)^2 + (z^{'}_i - sz)^2} \\
&\quad i=1,\cdots, n,\ i\neq r
\end{aligned} \tag{6}
$$

Errors exist in the survey value ($x^{'}_i$, $y^{'}_i$, $z^{'}_i$), the estimated TDOA $\tau^{'}_{ri}$, and the estimated sound velocity $c^{'}$. All these errors will affect the final localization accuracy, which will be discussed in Section 4.

3.2 Hardware description

We have already implemented four modules inside a recording node, including the recording module, time synchronization, wireless communication, and power module. Aigo R5588 recorder is used in the recording module. We use radio signal to time-synchronize all the recording nodes. Storage batteries are used in power systems which are 2.3 Ah, 12V.

As for the other four modules, a total station is used to survey the recording nodes. Sound activity detection and bioacoustic recognition have been implemented in the software, but not yet integrated into the recording node hardware. Sound velocity estimation will be implemented later, and the sound is assumed to travel in a linear uniform way at this stage.



(a) (b)

Fig. 2 The recording node
(a) Before encapsulation; (b) After encapsulation

3.3 Software description

We have developed the software which implements bioacoustic recognition and bioacoustic localization of animals. The software is further developed based on Audacity ([2010-12-06] http://audacity.sourceforge.net/). In our software, "double threshold algorithm" is used to detect sound activity. Then Maximum Likelihood based on Gaussian Mixture Models (GMM) is applied to recognize the animals' species according to their sounds. TDOAs are estimated by Generalized Cross Correlation (GCC) (Knapp & Carter, 1976). The Least Squares method is applied to bioacoustic localization of animals (Fig. 3 and Fig. 4).

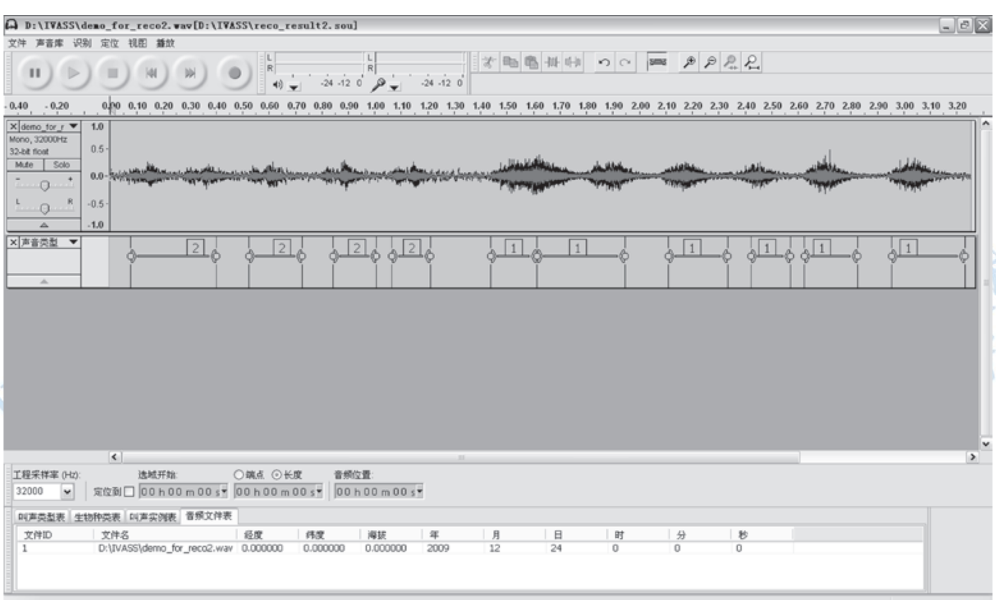


Fig. 3 The function of bioacoustic recognition in the software
(The Figure shows results of sound activity detection and bioacoustic recognition using the software. It extracts sound clips containing animals' calls annotating their starting time and stopping time, and recognizes each clip. "1" denotes the calls of "Azure-winged Magpie", and "2" denotes the calls of "Common Magpie")

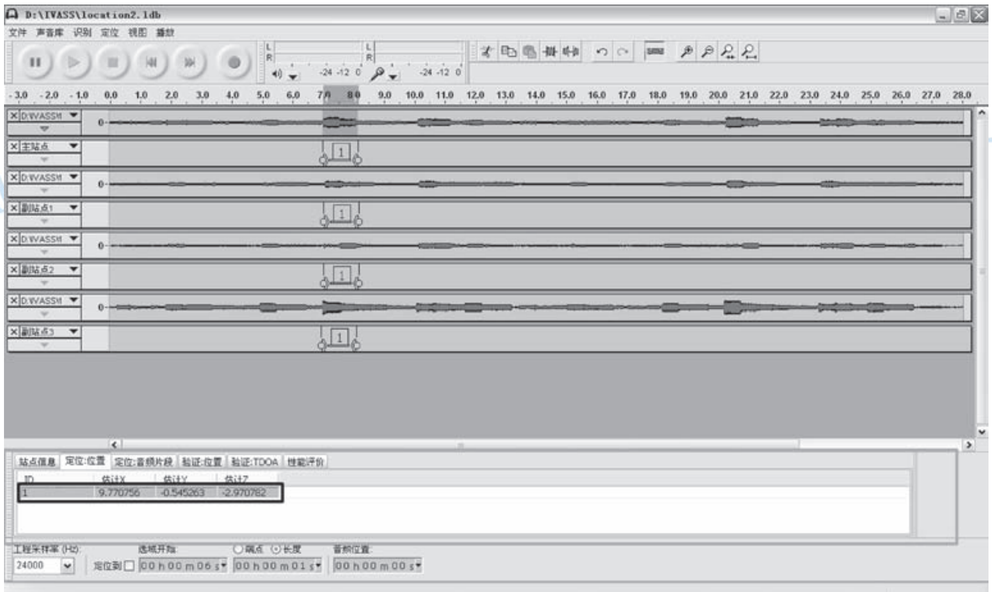


Fig. 4 The function of bioacoustic localization in the software

(Results of bioacoustic localization based on audio files from four different recording nodes are shown in the figure. Four recording nodes generate four audio files (the four waves in the figure). Sound clips containing animals' calls are detected using "sound activity detection" technology and annotated by number, such as the "1" annotation in each audio file. The estimated coordinates of the animal who is making the current sound clip is shown in the rectangle in the bottom-left corner)

3.4 Experimental design

Experiments were conducted in the field to test the performance of the prototype system (Fig. 6). The test field is the playground in front of the Institute of Remote Sensing Application, Chinese Academy of Sciences, Beijing, China. In order to lessen the impacts of environmental factors, a non-windy day was selected.

During experiments, four recording nodes were placed at the corners of the playground. The playground is about 50 meters long and 25 meters wide (Point A, B, C, and D in Fig. 5). The spatial coordinates of the recording nodes are displayed in Table 1. The four recording nodes were fixed at different heights on a pole (Fig. 6, Fig. 7).

We used a loudspeaker as an artificial sound source placed at five known positions to testify the localization accuracy of this prototype system (Point 1, 2, 3, 4, 5 shown in Fig. 5). These 5 positions were randomly selected and surveyed with a total station and a ruler.

Table 1 The positions of recording nodes and the artificial sound sources/m

Point	X	Y	Z
A	14.633	–25.728	1.817
B	–29.609	0.830	2.817
C	–14.6	25.808	3.817
D	23.986	0.029	0.874
1	–8.532	3.055	1.762
2	10.082	–0.532	0.678
3	3.280	–11.995	1.145
4	–25.192	10.486	1.988
5	23.071	–13.887	0.415

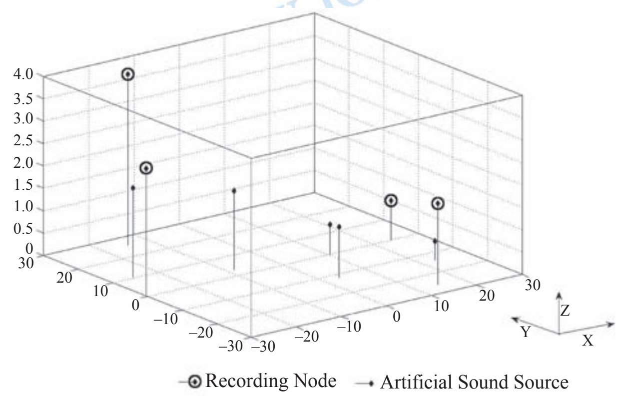


Fig. 5 Positions of recording nodes and the artificial sound sources/m

Fig. 6 The whole acoustic localization system

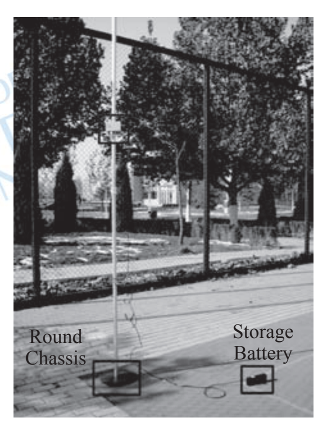


Fig. 7 A recording node at work

3.5 Results of experiments

An artificial sound source was placed at 5 different places one after another to assess the localization performance of the prototype system in the field. The surveyed coordinates of the artificial sound source and the correspondent localization results calculated by the prototype system are listed in Table 2, with our error analysis in Table 3 and Table 4, respectively.

First, localization errors both in the X axis and the Y axis are relatively small. The mean value of absolute localization errors is 0.51 m in the x axis, and 0.87 m in the y axis; the mean value of relative localization errors is 3.5% in the x axis, and 11.4% in the y axis. Since our purpose for such a system is to estimate the positions of bird species, we are satisfied with the prototype system in its localization accuracy.

Second, compared with the x and the y axes, localization errors in the z axis are relatively large. The mean value of absolute localization errors is 4.99 m in the z axis and the mean value of relative localization errors is 528.9%. In our experiments, the three-dimensional localization errors mainly result from the localization errors in the z axis, which can be drawn from the two columns: $\sqrt{(X'-X)^2+(Y'-Y)^2}$ and $\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}$ in Table 3. The maximum of the two-dimensional localization errors $\sqrt{(X'-X)^2+(Y'-Y)^2}$ is only 1.97 m. Once considering the Z axis, the maximum of the three-dimensional localization errors $\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}$ is increased to 5.99 m. The column $\frac{|Z'-Z|}{\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}}$ in Table 4 also verifies this point, in which the values are more than 90% and the maximum is 94.2%. Therefore, the location results in the z axis are unreliable. We have to enhance the localization accuracy in the z axis to improve the overall performance of the bioacoustic localization system.

Table 2 The true and estimated coordinates of artificial sound sources/m

Source ID	X	X'	Y	Y'	Z	Z'
1	–8.53	–8.56	3.06	2.41	1.76	7.71
2	10.08	9.77	–0.53	–0.55	0.68	–2.97
3	3.28	3.47	–12.00	–13.18	1.15	–4.43
4	–25.19	–24.17	10.49	12.18	1.99	–3.53
5	23.07	24.05	–13.89	–14.71	0.42	–3.81

Note: X, Y, Z denotes true coordinates; X', Y', Z' denotes estimated coordinates

Table 3 Absolute localization error/m

Source ID	|X'–X|	|Y'–Y|	|Z'–Z|	$\sqrt{(X'-X)^2+(Y'-Y)^2}$	$\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}$
1	0.03	0.65	5.95	0.65	5.99
2	0.31	0.02	3.65	0.31	3.66
3	0.19	1.18	5.58	1.20	5.71
4	1.02	1.69	5.52	1.97	5.86
5	0.98	0.82	4.23	1.28	4.42
Max.	1.02	1.69	5.95	1.97	5.99
Min.	0.03	0.02	3.65	0.31	3.66
Mean	0.51	0.87	4.99	1.08	5.13

Table 4 Relative localization error and localization error ratio/%

Source ID	$\left\|\frac{X'-X}{X}\right\|$	$\left\|\frac{Y'-Y}{Y}\right\|$	$\left\|\frac{Z'-Z}{Z}\right\|$	$\frac{\|Z'-Z\|}{\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}}$
1	0.4	21.2	338.1	99.4
2	3.1	3.8	536.8	99.6
3	5.8	9.8	485.2	97.8
4	4.0	16.1	277.4	94.2
5	4.2	5.9	1007.1	95.7
Max.	5.8	21.2	1007.1	99.6
Min.	0.4	3.8	277.4	94.2
Mean	3.5	11.4	528.9	97.3

4 ERROR ANALYSIS

Localization error is induced by errors in the input variables that are used for locating the sound source, such as the sound velocity in the medium, time difference of arrival (TDOA), and the coordinates of recording nodes. The aim of error analysis is to estimate localization error according to errors in the input variables. Since bioacoustic localization is essentially a nonlinear optimization problem, it is complex to assess error propagation with an analytical method. Therefore, we used Monte Carlo simulation method to perform error analysis.

The Monte Carlo method is a statistical simulation algorithm. Localization error can be estimated by simulating the bioacoustic localization procedure in tens of thousands times with a computer. We implemented the algorithm in MatLab (Mathworks Inc., Natick, MA).

A MINNA (Wahlberg, *et al.*, 2001) system is an acoustic localization system that consists of the MINimum Number of receiver Array (*i.e.* recording node in this paper) required to find the source location. The smallest number of recording nodes for a three-dimensional acoustic location system is 4. If there are more recording nodes than what is needed for the MINNA solution, the system is an ODA (OverDetermined Array) (Wahlberg, *et al.*, 2001). In MINNA systems the number of recording nodes is just sufficient to localize the sound source, and there are no extra recording nodes available to decrease the impact of errors in the input variables. However an ODA has extra recording node to decrease the impact of errors in the input variables through Least Square method (Wahlberg, *et al.*, 2001). Therefore, an ODA usually has smaller localization error than a MINNA, which can be clearly demonstrated in the following sections.

Algorithm 1 Monte Carlo method-based error analysis ([L_x,L_y,L_z,] denotes the space range where the recording nodes and the sound source should be ; "*time_resolution*" is the precision of TDOA; "*N*" means the total number of recording nodes. All these parameters are used to depict the characteristics of a bioacoustic localization system.)

Step 1 Initialize. Users specify spatial_range ([L_x,L_y,L_z,]), "*time_resolution*", "*N*", the total times of simulation "*n_iters*", and the threshold for empirical cumulative probability "*emp_cum_pro_th*". Let "*iter*" = 0.

Step 2 Find out true values for all input variables. Within the cube of which the center is [0,0,0], the length is L_x, the width is L_y and the height is L_z, randomly select the position of the sound source, denoted by $[S_x,S_y,S_z]$ which are the true value for the position of the sound source. Then randomly select the places for all recording nodes, denoted by $[R_{xi},R_{yi},R_{zi}]$, i=1,2,⋯, N which is the true value for the position of the i^{th} recording node. The true value for sound velocity c is set to 340 m/s. Randomly select one recording node as the reference one, then the true TDOA can be calculated out, denoted by τ_{ri}, i=1,2,⋯, N, $i≠r$ where r is the index of the reference recording node.

Step 3 Find out the estimated values for all input variables. The errors in surveying recording nodes, estimating TDOA, estimating sound velocity are assumed as independent random variables with their expectation equals to 0. Therefore, the estimated values for all input variables ($[R'_{xi}, R'_{yi}, R'_{zi}]$, $i=1,2,\cdots,N$, c', $\tau'_{ri}, i=1,2,\cdots,N, i\neq r$) can be simulated by adding some random errors to their true value. Then the position of the sound source can be estimated, denoted by $[S'_x,S'_y,S'_z]$, using the simulated values of all input variables.

Step 4 Calculate localization error. Different types of acoustic localization error can be calculated in terms of $[S_x,S_y,S_z]$ and $[S'_x,S'_y,S'_z]$.

Step 5 *iter* = *iter* + 1. If *iter* < *n_iters*, go back to **Step 2**, otherwise, to **Step 6**.

Step 6 Many instances for each type of the acoustic localization error have been obtained. The value at which the empirical cumulative probability of some kind of localization error becomes no less than "*emp_cum_pro_th*" is used to indicate the magnitude of this kind of localization error.

4.1 Error in surveying recording nodes

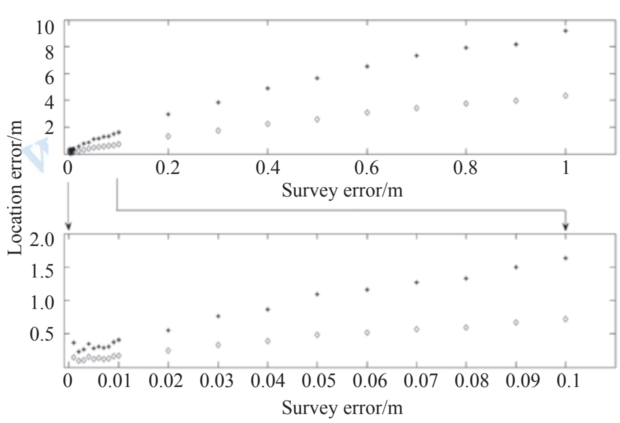
Error always exists in surveying. The survey error often obeys the normal distribution (*i.e.* Gaussian distribution). That is, along the i^{th} axis of the j^{th} recording node, the survey error obeys

$$N(0,(\sigma_{i,j}^{R})^2),\ i=x,y,z,\ j=1,2,\cdots,N \tag{7}$$

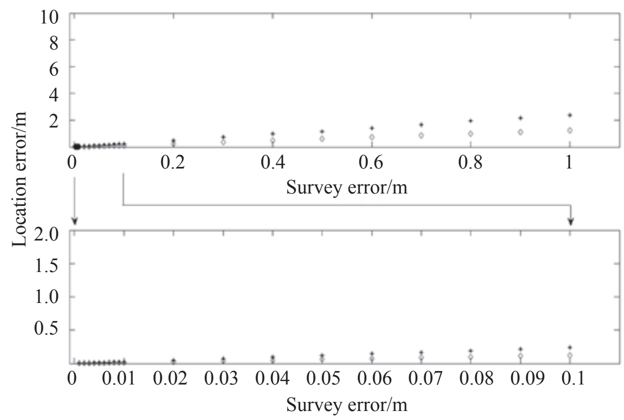
where $\sigma_{i,j}^{R}$ represents the standard deviation of the error. Suppose that survey errors in all recording nodes are independent and identical, then

$$N(0,(\sigma_{i,j}^{R})^2),\ i=x,y,z,\ j=1,2,\cdots,N \tag{8}$$

is simplified to $N(0,(\sigma_i^R)^2), i=x,y,z$. If survey errors are assumed independent and identically distributed along three different axes, then $N(0,(\sigma_i^R)^2), i=x,y,z$ is simplified to $N(0,(\sigma^R)^2)$. Simulation results of Algorithm 1 in Fig. 8 show how the localization error increases with the error in surveying recording nodes.



(a)



(b)

Fig. 8 The impact of survey error on location error
(a) MINNA, N = 4; (b) ODA, N = 10
(Each * point denotes the simulated total location error; each ◊ point denotes the simulated absolute location error along X axis (Y, Z are omitted because they are the same as X). Survey error σ^R is increasing from 0.001 m to 1 m with changing step: for [0.001 m, 0.01 m], the step is 0.00 1m; for (0.01 m, 0.1 m], the step is 0.01 m; for (0.1 m, 1 m], the step is 0.1 m. Here, *n_iters* is 10000; spatial range $[L_x,L_y,L_z]$ is [30 m, 30 m, 30 m]; sound velocity is assumed to be 340 m/s. It is assumed that there is not TDOA estimation error or sound velocity estimation error)

4.2 Error in estimating TDOA

Because the bioacoustic localization system is a discrete digital system and the windowing technology is used for signal processing. There is always estimation error between the true TDOA and an estimated TDOA. Considering a bioacoustic localization system with an audio sampling rate denoted by sample_rate(Hz) and window step in windowing technology denoted by win_step, the smallest time unit which the system could represent is win_step/sample_rate (unit: s). So the estimated TDOA must be a multiple of the smallest time unit. However, the true TDOA value is continuous and hence not necessary to be a multiple of the smallest time unit. Therefore the estimation error due to time resolution of an acoustic location system is inevitable and we call it "inherent TDOA estimation error". We take the smallest time unit as the time resolution of a bioacoustic localization system, denoted by time_resolution (s), *i.e.*

$$time\_resolution=win\_step\ /\ sample\_rate \tag{9}$$

According to the above analysis, the inherent TDOA estimation error should obey uniform probability distribution as U(–*time_resolution*/2, *time_resolution*/2). Simulation results of Algorithm 1 in Fig. 9 show how the acoustic location error increases with the time resolution of a bioacoustic localization system.

Besides the inherent TDOA estimation error, noises could render inaccurate TDOA estimation and error in synchronization of all the recording nodes can also contribute error in the TDOA estimation. Obviously, the order of magnitude of the other two errors is higher than that of the inherent TDOA estimation error, so the influence of the inherent TDOA estimation can be ignored if the other two errors exist. A normal probability distribution is used to generally describe error in estimating TDOA, denoted by $N(0, (\sigma^R)^2)$. Simulation results of Algorithm 1 in Fig. 10 show how the acoustic location error increases with the error in estimating TDOA.

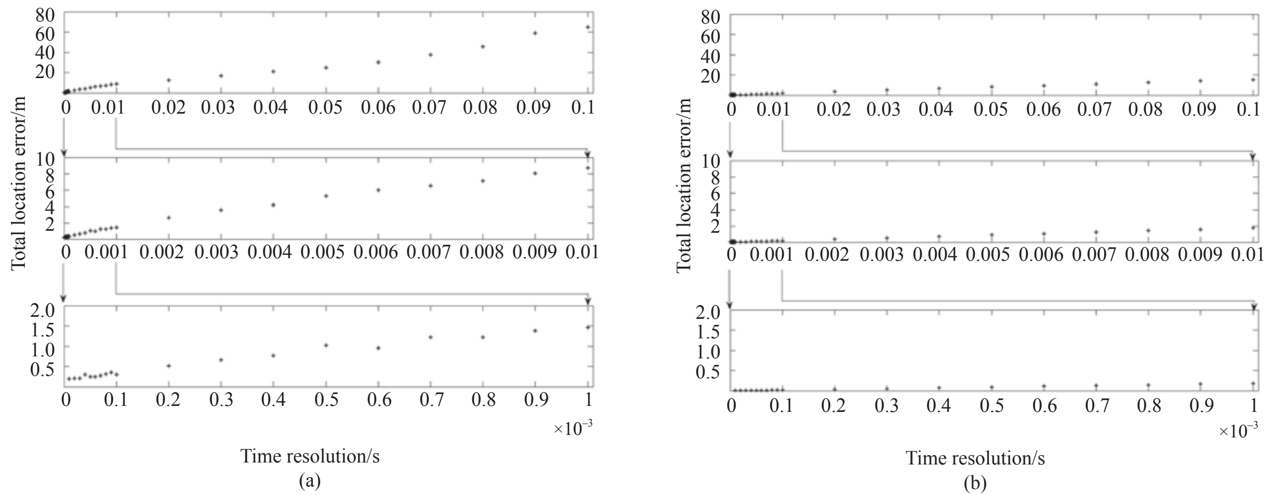


Fig. 9 The impact of inherent TDOA estimation error on location error

(a) MINNA, $N = 4$; (b) ODA, $N = 10$

(Each * point denotes the simulated total location error. Time resolution is between 0.00001 s and 0.1 s with a varying step. For [0.00001 s, 0.0001 s], the step is 0.00001 s; for (0.0001 s, 0.001 s], the step is 0.0001 s; for (0.001 s, 0.01 s], the step is 0.001 s; for (0.01 s, 0.1 s], the step is 0.01 s. Here, *n_iters* is 10000; spatial range $[L_x,L_y,L_z]$ is [30 m, 30 m, 30 m]; sound velocity is assumed to be 340 m/s. It is assumed that there is not survey error or sound velocity estimation error. It's also assumed that besides time resolution there are no other sources for TDOA estimation error)

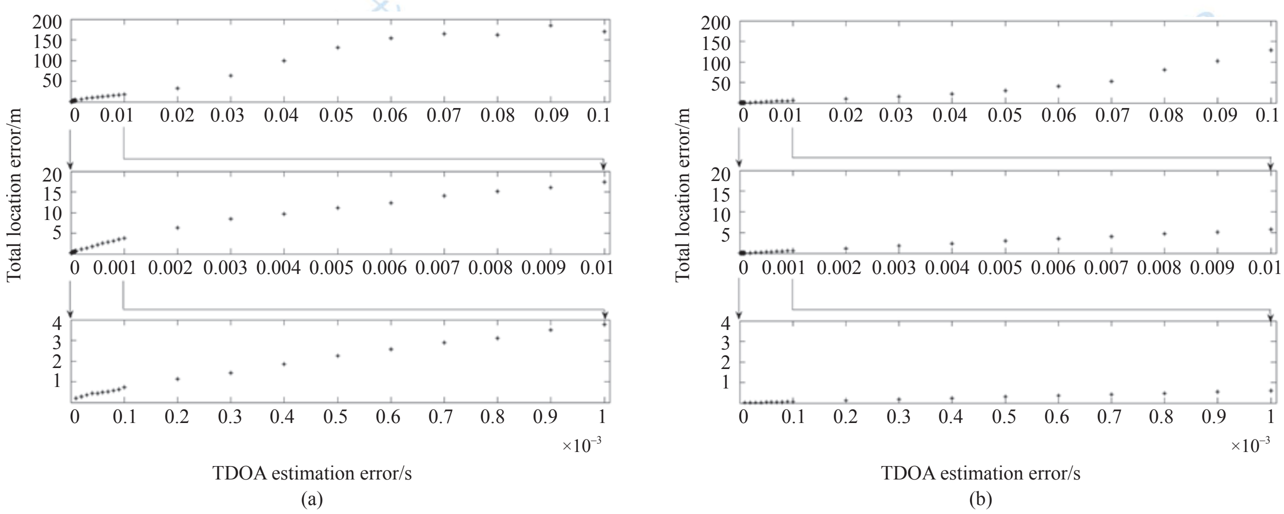


Fig. 10 The impact of TDOA estimation error on location error

(a) MINNA, $N = 4$; (b) ODA, $N = 10$

(Each * point denotes the simulated total location error. TDOA estimation error is between 0.00001 s and 0.1 s with a varying step. For [0.00001 s, 0.0001 s], the step is 0.00001s; for (0.0001 s, 0.001 s], the step is 0.0001 s; for (0.001 s, 0.01 s], the step is 0.001 s; for (0.01 s, 0.1 s], the step is 0.01 s. Here, *n_iters* is 10000; spatial range $[L_x,L_y,L_z]$ is [30 m, 30 m, 30 m]; sound velocity is assumed to be 340 m/s. It is assumed that there is not survey error or sound velocity estimation error)

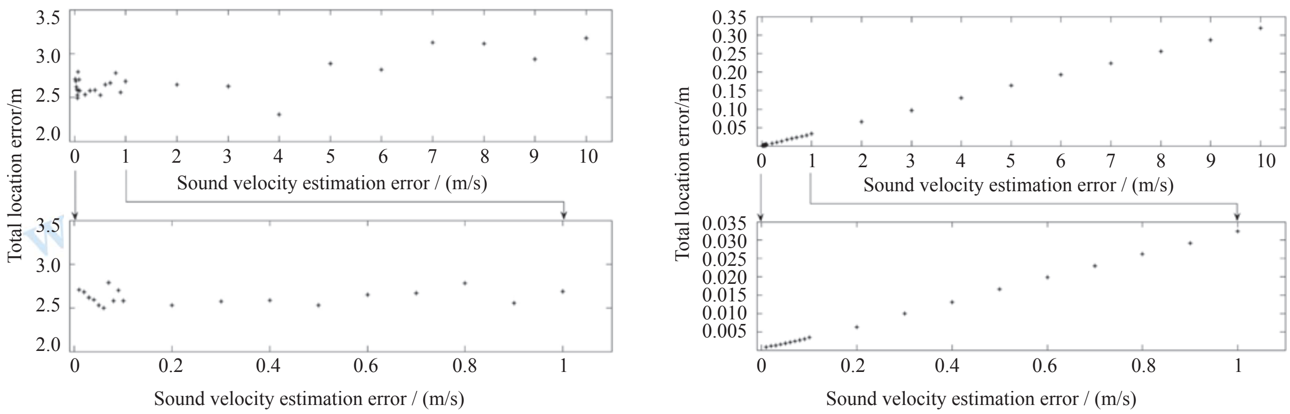
4.3 Error in estimating sound velocity

A normal probability distribution $N(0,(\sigma^s)^2)$ is used to generally describe error in estimating sound velocity. Simulation results of Algorithm 1 in Fig. 11 show how the acoustic location error increases with the error in estimating sound velocity.

4.4 Appropriate value for the total number of recording nodes

The total number of recording nodes should be decided when implementing a bioacoustic localization system. On one hand, the more recording nodes there are, the closer to theoretical location accuracy the acoustic location system can approach (Fig. 12). On the other hand, more recording nodes will cost more. Therefore, a balance should be kept between accuracy and cost when deciding the total number of recording nodes.

Algorithm 1 provides an easy way for estimating the localization accuracy of a bioacoustic localization system when varying the total number of recording nodes (Fig. 12). Combined with cost situation, it is easy to decide the appropriate total number of recording nodes.



(a) (b)

Fig. 11 The impact of sound velocity estimation error on location error
(a) MINNA, N=4; (b) ODA, N=10
(Each * point denotes the simulated total location error. Sound velocity estimation error is between 0.01 m/s and 10 m/s with a varying step. For [0.01 m/s, 0.1 m/s], the step is 0.01 m/s; for (0.1 m/s, 1 m/s], the step is 0.1 m/s; for (1 m/s, 10 m/s], the step is 1 m/s. Here, *n_iters* is 10000; spatial range [L_x,L_y,L_z] is [30 m, 30 m, 30 m]. It is assumed that there is not survey error or TDOA estimation error)

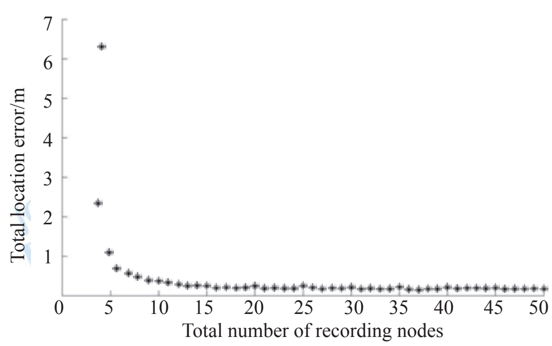


Fig. 12 The impact of the total number of recording nodes on location error
(Each * point denotes the simulated total location error. The total number of recording nodes is increasing from 4 to 50 with **step 1**. Here, *n_iters* is 100; spatial range [L_x, L_y, L_z] is [30 m, 30 m, 30 m]; survey error along any single axis is 0.1 m; time resolution is 0.001 s; sound velocity is 340 m/s. It is assumed that there is not sound velocity estimation error)

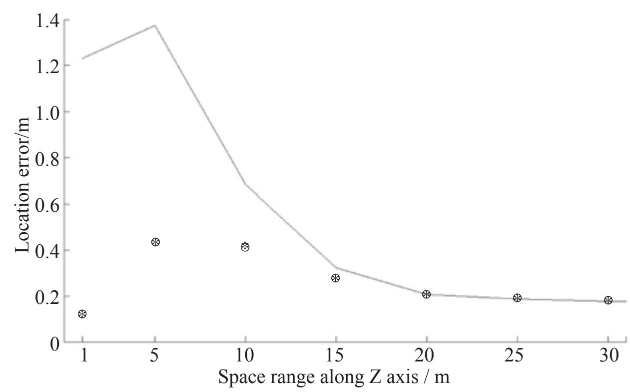


Fig. 13 The impact of space range along Z axis on location error
(Each * point denotes absolute location error along X axis, each o point denotes absolute location error along Y axis and the curve denotes absolute location error along Z axis. Space range along Z axis is between 1 m and 70 m with step 5 m. Here, space range along both X axis and Y axis is kept 30 m. The survey error along any single axis is 0.01 m. TDOA estimation error and sound velocity estimation error are not considered here. The sound velocity is assumed to be 340 m/s. *n_iters* is 10000)

4.5 Reason for lower accuracy in the Z axis

From the results of the field experiment, we can see that the location error along the Z axis is much larger than that along the X and Y axes. Using Algorithm 1, we find that the real reason is that relative to that in the X and Y axes, the space range along the Z axis are too small. Seen from Fig. 13, location error could be decreased by increasing space range along Z axis.

5 CONCLUSION

In this research, we presented a WSN-based bioacoustic localization prototype system of terrestrial animals. In outdoor experiments, the system is verified to have the ability to fulfill our requirements. In order to assess the factors influencing the localization accuracy and their impacts, we applied Monte Carlo simulation method to conduct the error analysis. Results show that the localization accuracy decreases with increasing errors in surveying recording nodes, estimating TDOA and sound velocity. The localization accuracy could be enhanced by increasing the number of recording nodes.

The prototype system is the beginning step from where we will continue to develop smarter localization systems for terrestrial animals. Wireless Sensor Network will demonstrate its powerful ability in biological monitoring since it totally changes the way people interacting with the physical world. Modules of self-survey, sound velocity estimation and bioacoustic recognition will be added into our system in future.

Acknowledgements: Thanks Yang Xirong in State Key Laboratory of Remote Sensing Science for his technical support in this research.

REFERENCES

Anderson S E, Dave A S and Margoliash D. 1996. Template-based automatic recognition of birdsong syllables from continuous recordings. *Journal of Acoustical Society of America*, **100**(2): 1209–1219 DOI: 10.1121/1.415968

Bohn D A. 1987. Environmental effects on the speed of sound. *Audio Engineering Society Convention* 83, New York: 16–19

Clark C W and Clapham P J. 2004. Acoustic monitoring on a humpback whale (*Megaptera novaeangliae*) feeding ground shows continual singing into late spring. *Proceedings of the Royal Society B-Biological Sciences*, **271**(1543): 1051–1057 DOI: 10.1098/rspb.2004.2699

Clark C W, Gillespie D, Nowacek D P and Parks S E. 2007. Listening to their world: acoustics for monitoring and protecting right whales in an urbanized ocean // Kraus S D and Rolland R M, eds. The Urban Whale: North Atlantic Right Whales at the Crossroads. Boston: Harvard University Press: 333–357

Everest F A and Ken C P. 2009. The master handbook of acoustics. New York: McGraw–Hill

Exadaktylos V, Silva M, Ferrari S, Guarino M, Taylor C J, Aerts J M and Berckmans D. 2008. Time-series analysis for online recognition and localization of sick pig (*Sus scrofa*) cough sounds. *Journal of Acoustical Society of America*, **124**(6): 3803–3809 DOI: 10.1121/1.2998780

Fagerlund S. 2007. Bird species recognition using support vector machines. *EURASIP Journal on Applied Signal Processing*, **2007**(1): 64 DOI: 10.1155/2007/38637

Farnsworth A. 2005. Flight calls and their value for future ornithological studies and conservation research. *The Auk*, **122**(3): 733–746 DOI: 10.1642/0004-8038

Gong P. 2007. Wireless sensor network as a new ground remote sensing technology for environmental monitoring. *Journal of Remote Sensing*, **11**(4): 545–551

Hardy H C, Telefair D and Pielemeier W H. 1942. The velocity of sound in air. *Journal of Acoustical Society of America*, **13**(3): 226–233 DOI: 10.1121/1.1916169

Jones D L and Ratnam R. 2009. Blind location and separation of callers in a natural chorus using a microphone array. *Journal of Acoustical Society of America*, **126**(2): 895–910 DOI: 10.1121/1.3158924

Knapp C H and Carter G C. 1976. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics Speech and Signal Processing*, **24**(4): 320–327 DOI: 10.1109/TASSP.1976.1162830

Kogan J A and Margoliash D. 1998. Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study. *Journal of Acoustical Society of America*, **103**(4): 2185–2196 DOI: 10.1121/1.421364

Kwan C, Ho K C, Mei G, Li Y, Ren Z, Xu R, Zhang Y, Lao D, Stevenson M, Stanford V and Rochet C. 2006. An automated acoustic system to monitor and classify birds. *EURASIP Journal on Applied Signal Processing*, **2006**(3): 1–19 DOI: 10.1155/ASP/2006/9670

McGregor P K, Dabelsteen T, Clark C W, Bower J L, Tavares J P and Holland J. 1997. Accuracy of a passive acoustic location system: empirical studies in terrestrial habitats. *Ethology Ecology and Evolution,* **9**(3): 269–286 DOI: 10.1080/08927014.1997.9522887

McIlraith A L and Card H C. 1997. A comparison of backpropagation and statistical classifiers for bird identification. *IEEE International Conference on Neural Networks*, **1**: 100–104 DOI: 10.1109/ICNN.1997.611645

Mellinger D K and Clark C W. 2003. Blue whale (*Balaenoptera musculus*) sounds from the North Atlantic. *Journal of Acoustical Society of America*, **114**(2): 1108–1119 DOI: 10.1121/1.1593066

Mennill D J, Burt J M, Fristrup K M and Vehrencamp S L. 2006. Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. *Journal of Acoustical Society of America*, **119**(5): 2832–2839 DOI: 10.1121/1.2184988

Ning H S, Chen W S, Mao X and Li J. 2010. Bird-aircraft strike avoidance radar. *IEEE Aerospace and Electronic Systems Magazine*, **25**(1): 19–28 DOI: 10.1109/MAES.2010.5442150

Rabiner L R and Sambur M R. 1975. An algorithm for determining the endpoint for isolated utterances. *The Bell System Technical Journal*, **54**(2): 297–315

Stafford K M, Fox C G and Clark D S. 1998. Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean. *Journal of Acoustical Society of America*, **104**(6): 3616–3625 DOI: 10.1121/1.423944

Wahlberg M, Mohl B and Madsen P T. 2001. Estimating source position accuracy of a large-aperture hydrophone array for bioacoustics. *Journal of Acoustical Society of America*, **109**(1): 397–406 DOI:10.1121/1.1329619

陆生动物声音遥感：定位与误差分析

沈少青[1]，宫鹏[1,2]，程晓[3]，应清[1]

1. 遥感科学国家重点实验室 中国科学院遥感应用研究所、北京师范大学，北京 100101；

2. 加州大学伯克利分校 环境科学、政策和管理系，美国 加利福尼亚；

3. 北京师范大学 全球变化与地球系统科学研究院，北京 100875

摘　要：声音遥感技术可以极大地提高人们对动物的监测、研究和保护能力。本研究设计了一套廉价的陆生动物声音定位系统。该系统集成了市场上现成的录音笔和无线控制设备，价格明显低于许多动物定位系统。使用自主研发的声音定位软件，在X、Y方向上的绝对定位误差最大为1.69 m，可满足大部分动物定位需求；Z方向的定位误差偏大，结果尚不理想。为了揭示影响声音定位系统精度的各种因素及其影响，采用蒙特卡罗方法进行分析，发现录音站点的位置测量误差、到达时间差的估计误差和声速的估计误差均会影响最终的声音定位精度。此外，该方法还可用于在声音定位系统布设阶段确定系统参数的合理数值，包括录音节点总个数和空间尺度。

关键词：声音遥感，生物声音定位，生物监测，误差分析，蒙特卡罗方法

中图分类号：TP79　　**文献标志码：**A

引用格式：沈少青，宫鹏，程晓，应清. 2011. 陆生动物声音遥感：定位与误差分析. 遥感学报，**15**(6): 1255–1275
Shen S Q, Gong P, Cheng X and Ying Q. 2011. Sound-based remote sensing of terrestrial animals: localization and error analysis. *Journal of Remote Sensing*, **15**(6): 1255–1275

1 引　言

人们一直在努力建立各种生物(尤其是濒危野生动物)监测系统，以便评估人类活动对其产生的影响，从而达到更有效保护生物的目的。对于监测时获取的各类信息，生物的位置信息备受人们的关注。这是因为，首先，生物的位置信息可以用来确定生物的栖息地范围。栖息地是生物与环境发生相互作用以及与其他生物发生相互作用的空间场所。越准确的栖息地信息，越能帮助人们有效地保护生物。其次，生物的位置信息可用于分析生物的行为模式。例如生物行为学家研究发现鸟类个体在对唱时空间位置的改变会影响其他的行为，如配偶的选择(Jones和Ratham，2009; Mennill 等, 2006)。再次，获取生物的实时位置信息可避免人与其他生物发生空间冲突，从而避免悲剧的发生，例如发展飞鸟定位系统用来避免飞鸟与飞机相撞的悲剧(Ning 等, 2010)。由此可见，生物定位系统的研发十分必要。

声音定位系统，作为生物定位系统的一种，有着独特的优点。首先，作为一种被动式的定位系统，它不会干扰生物的自然活动；其次，声音是生物交流的方式之一。定位系统在定位的同时，还能记录下生物之间交流的信息(Mennill 等, 2006)，留待其他分析，如生物种类信息分析、行为分析等，一举多得；再次，由于在水中声音比光传播得远得多，所以声音定位系统是定位生活在江河湖海中的生物的最理想设备。比如有名的“露脊鲸监听网络(Right Whale Listening Network)”(Clark 等, 2007)；最后，录音设备不受光照条件的限制，所以它能用于定位夜间活动的生物。例如，监测候鸟的夜间迁徙活动，这些记录是制定候鸟保护计划的重要依据(Farnsworth, 2005)。

声音定位系统已经被广泛地应用于海洋动物的定位，尤其是海洋中的大型哺乳动物(Stafford 等,1998; Wahlberg 等, 2001; Mellinger和Clark, 2003; Clark和

收稿日期：2010-12-06；**修订日期：**2011-07-01

基金项目：国家高技术研究发展计划(863计划)(编号: 2006AA12Z112，2009AA12200101)

第一作者简介：沈少青(1985—　)，女，博士研究生，研究方向为声学遥感、生物声学与环境变化。E-mail: s_s_q@126.com。

通信作者简介：宫鹏(1965—　)，男，研究员，主要研究方向为利用各类遥感技术监测从局部到全球尺度的环境变化、对环境变化进行建模，并在环境与健康领域开展应用。E-mail: gong@irsa.ac.cn。

Clapham, 2004)。这主要是因为没有比声音定位更有效的方法来定位海洋动物。然而将声音定位系统用于定位陆生动物，却是比较少的(Mennill 等, 2006)。目前典型的陆生动物声音定位系统是将多个麦克风精确地摆放到空间各处，通过一个多通道声音采集卡连接到一台计算设备进行二维定位(McGregor 等, 1997；Exadaktylos 等，2008；Jones和Ratham，2009；Mennill 等，2006)。尽管单个声音采集卡保证了多通道录音的时间同步，并且这种定位系统足够用于解决小范围简单环境中的生物声音定位，但是仍存在不少缺陷。首先，集中式的声音采集和计算方式导致系统高度依赖单个部件(声音采集卡和计算设备)，使得系统不够稳健。其次，有线连接方式限制了麦克风、声卡及计算设备的摆放位置和距离，不利于布设系统和改造系统。并且，外露的信号线还容易受到人、其他生物及环境的破坏。

蓬勃发展起来的无线传感器网络技术可用于克服上述陆生动物声音定位系统的缺陷。无线传感器网络技术是指将传感器技术、自动控制技术、数据网络传输、储存、处理与分析技术集成的现代信息技术(宫鹏，2007)。在一套基于无线传感器网络技术的陆生动物声音定位系统中，多个录音节点协同定位发出叫声的生物。单个录音节点是一个功能完整的系统，具有麦克风传感器、音频采集处理器、音频存储器、微处理器、无线通信设备和供电模块。

无线传感器网络技术可以从下述方面改进传统的陆生动物声音定位系统。第一，录音节点之间的无线连接方式使得野外的系统布设变得简单。新系统将数据采集和计算任务分派给每个录音节点，这种分布式的特点提高了系统的稳健性，虽然这也引入了录音节点之间的时间同步问题。第二，录音节点的坐标是声音定位的必要参数，GPS模块可以嵌入到录音节点中以定位节点自身的位置，这省去了传统声音定位流程中人工测量录音节点的步骤，使得整个流程变得更加自动便捷。第三，还可以将温度传感器和湿度传感器也装载到录音节点上。这些传感器可用于估计声音在空气中的实际传播速度，以提高声音定位的精度。第四，录音节点的计算能力使得它能够根据采集的音频进行生物种类识别。如果在声音定位前先进行生物种类识别，就可以让整套声音定位系统更加专注于某类动物的定位。

本文的目的就是设计出一套基于无线传感器网络的陆生动物声音定位系统，并探讨这套声音定位系统的定位精度和误差特性。首先介绍了一套基于无线传感器网络技术的陆生动物声音定位系统的总体架构，同时展示了一套软硬件齐全的原型系统；然后利用该套原型系统进行了野外声音定位实验并分析实验结果；接下来总结了影响陆生动物声音定位系统定位精度的各种因素，并采用蒙特卡罗方法来定量分析定位误差；最后全文总结和展望未来的工作。

2 系统架构

基于无线传感器网络技术的陆生动物三维声音

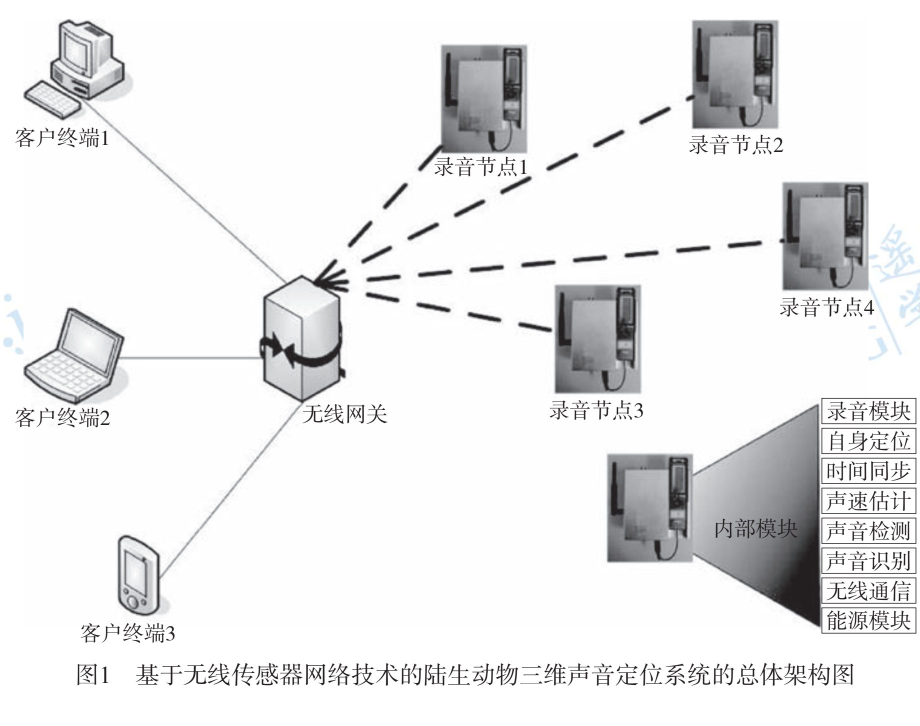


图1 基于无线传感器网络技术的陆生动物三维声音定位系统的总体架构图

定位系统的总体架构详见图1。可以看出，录音节点包含8大模块：(1)录音模块；(2)自身定位；(3)时间同步；(4)声速估计；(5)声音检测；(6)声音识别；(7)无线通信；(8)能源模块。用于定位陆生动物的三维声音定位系统有着与普通定位系统不同的应用需求，例如自身定位、声速估计和生物种类识别(声音检测与声音识别)，无线传感器网络技术可以用来满足这些特别的应用需求。

2.1 自身定位

录音节点的位置坐标是声音定位的必要输入参数，而且录音节点的位置测量精度会影响最终定位的精度(第4.1节)。在野外测量录音节点的传统方式是根据大地控制点利用全站仪来进行测量。尽管这种方法具有较高的精度，但它操作起来比较繁琐，并且在野外常常难以找到合适的大地控制点。无线传感器网络技术允许为每个录音节点装载GPS模块，这样它们就可以方便地进行自身定位，并通过无线网络将坐标发给需要这些信息的设备，以便后续的声音定位。录音节点的滑动或者移动将不会带来麻烦，因为它们会将新的位置信息及时自动地广播出去，这有利于系统校正。

2.2 声速估计

声音在空气中的传播速度不是固定不变的，而是随着温度、高度、湿度和风速等环境参数的变化而改变。当温度是0℃，大气压强为760 mm汞柱标准大气压强时，声音在静止干燥的空气中的传播速度是(331.45 ± 0.05) m/s (Hardy 等，1942)。

随着温度的升高，声速会增大(Bohn, 1987)。声速c与温度t之间的关系式为

$$c = 331.45 \times \sqrt{1 + \frac{t}{273}} \tag{1}$$

式中，c的单位是m/s，而t的单位是℃。所以，当气温随着高度增加而降低时，声音的传播速度随着高度增加而减小，声音射线就会向天空弯曲；反之，当气温随着高度增加而增大时，声音的传播速度就会随着高度增加而增加，声波射线就会向地面弯曲，给人的听觉就是“声音在下沉”(Everest和Ken，2009)。

随着湿度的增大，声速也会增大(Bohn，1987)。声速c与湿度h之间的关系式为

$$c = 1508.528\sqrt{\frac{R_w}{M_w}} \tag{2}$$

式中，$R_w = \dfrac{7+h}{5+h}$是随湿度变化的通用气体常数；M_w=29–11h是随湿度变化的平均气体分子量；h代表湿度，是水气在空气中的比重。

声音顺风传播，速度增大；逆风传播，速度减小。除此以外，风还能改变声音传播的主方向(Everest和Ken，2009)。

由于动物声音定位一般都是在野外进行的，而野外环境中上述参数均在不断变化，因此实时声速估计有助于提高动物声音定位的精度。无线传感器网络技术允许集成温度计、湿度计和风速计到声音定位系统中来估计声速，也可以直接集成声速仪来获得声速测量值。

2.3 生物种类识别

生物定位往往具有目的性，如何准确地把用户感兴趣的生物种类定位出来，这是生物定位系统实用化的关键。所以生物定位之前关键的一步是生物种类的识别。这可以通过生物声音识别来实现。既然生物种类的声音识别也是基于声音特征的，而基于无线传感器网络技术的录音节点具有较强的计算能力，生物种类的声音识别就可以在录音节点内完成，从而形成一个更加智能的生物声音定位系统(Exadaktylos 等，2008)。

生物种类的声音识别问题本质上是模式识别的问题。它有很多种方法，常见的有模板匹配(Anderson 等，1996)、神经网络(McIlraith和Card，1997)、高斯混合模型(Kwan 等，2006)、支持向量机(Fagerlund，2007)、马尔科夫模型(Kogan和Margoliash，1998)等。

3 原型系统

根据上述基于无线传感器网络的陆生动物声音三维定位系统的总体架构，我们研发了一套软硬件齐全的原型系统(第3.2节和3.3节)，并在野外测试了该原型系统的定位精度，详见于第3.4节和3.5节。

3.1 声音定位原理

主要有两种方法应用于动物的声音定位。一种是基于能量范围，曾经应用于上述露脊鲸监听网络。另一种是基于到达时间差的计算 (Wahlberg 等，

2001)，该方法利用声音到达任意两个录音节点的时间差来定位发出该叫声的动物。本文采用后一种方法。首先，在动物经常出没的地方放置4个(或者更多个)录音节点(图2)，精确测量出各个录音节点的三维坐标。然后根据各个录音节点录得的音频，估计出目标声音到达各个录音节点的时间差。最后根据这些到达时间差，就可以估计出动物的位置。

假设已知声音的传播速度c和n个录音节点的位置，其中第i个录音机的位置记为(x_i, y_i, z_i)。另设发出叫声的动物的位置为(sx, sy, sz)，它是待求的。根据“传播时间=距离/速度”，可以得到n个方程：

$$T_i = \frac{1}{c}\sqrt{(x_i - sx)^2 + (y_i - sy)^2 + (z_i - sz)^2} \quad i=1,\cdots,n \tag{3}$$

若以第r个录音机为参考录音机，根据上述n个方程，可得到以下$(n-1)$个到达时间差方程：

$$\begin{aligned}\tau_{ri} &= T_r - T_i \\ &= \frac{1}{c}\sqrt{(x_r - sx)^2 + (y_r - sy)^2 + (z_r - sz)^2} - \\ &\quad \frac{1}{c}\sqrt{(x_i - sx)^2 + (y_i - sy)^2 + (z_i - sz)^2} \\ &\quad (i=1,\cdots,n,\ i\neq r)\end{aligned} \tag{4}$$

τ_{ri}是同一声源分别传播到第r个录音机与第i个录音机的到达时间差。这些到达时间差是可以根据实际的音频记录估计出来的，常用的到达时间差估计方法就是广义互相关方法。

利用最小二乘法来估计动物的位置，即声源的估计位置(sx_i', sy_i', sz_i')，它等于下述目标函数F取最小值时(sx, sy, sz)的值：

$$F = \sum_{i=1}^{n} f_i^2(c', \tau_{ri}', x_r', y_r', z_r', x_i', y_i', z_i', sx, sy, sz) \tag{5}$$

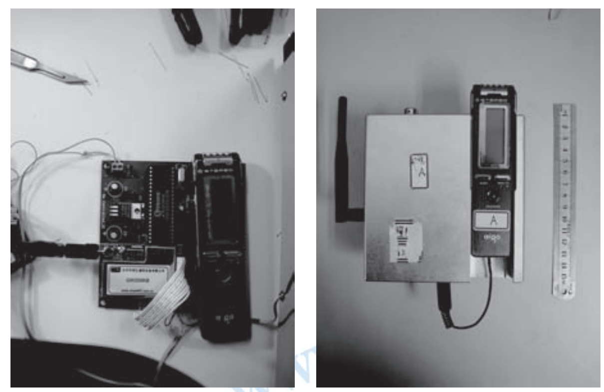
$$\begin{aligned}& f_i(c', \tau_{ri}', x_r', y_r', z_r', x_i', y_i', z_i', sx, sy, sz) \\ &= \tau_{ri}' - \frac{1}{c'}\sqrt{(x_r' - sx)^2 + (y_r' - sy)^2 + (z_r' - sz)^2} + \\ &\quad \frac{1}{c'}\sqrt{(x_i' - sx)^2 + (y_i' - sy)^2 + (z_i' - sz)^2} \\ &\quad i=1,\cdots,n,\ i\neq r\end{aligned} \tag{6}$$

式中，(x_i', y_i', z_i')是带有误差的位置测量值，τ_{ri}'是带有误差的时间差估计值，c'是带有误差的声音传播速度估计值。这些误差都会影响最终的定位精度。

3.2　硬件描述

目前我们研发出来的录音节点已经集成了录音模块、时间同步、无线通信和能源模块。录音模块借用了aigoR5588录音笔；时间同步则采用无线电信号；能源模块是奥特多品牌蓄电池，2.3 Ah，12 V。

至于其他4大模块，自身定位是借助了全站仪，GPS定位是我们下一步工作的目标；声音检测、声音识别已经实现在自行研发的软件中了，但尚未嵌入到录音节点中；声速估计是下一步工作的重要目标，目前仍然假设声音匀速直线传播。



(a)　　(b)

图2　录音节点

(a) 封装前；(b) 封装后

3.3　软件描述

我们研发出了一套生物声音处理软件，集生物识别和生物定位于一体。该软件以开源音频处理软件为基础进行了二次开发：Audacity([2010-12-06]http://audacity.sourceforge.net/)。该软件中，声音检测采用的是基于短时能量和短时过零率的双阈值算法，声音识别采用的是基于混合高斯模型的最大似然法。声音定位方面，采用广义互相关算法(Knapp和Carter, 1976)进行到达时间差的估计，并用最小二乘法估计出声源的三维位置。我们计划在未来添加更多的算法，以提高生物识别和生物定位的精度(图3、4)。

3.4　实验设计

我们在室外对这套原型系统进行了实验，以考察它的真实定位能力。实验地点是中国科学院遥感应用研究所的操场，实验情景如图6、7所示。为了减小环境因素对声音传播的干扰，我们特意选择了一个风速

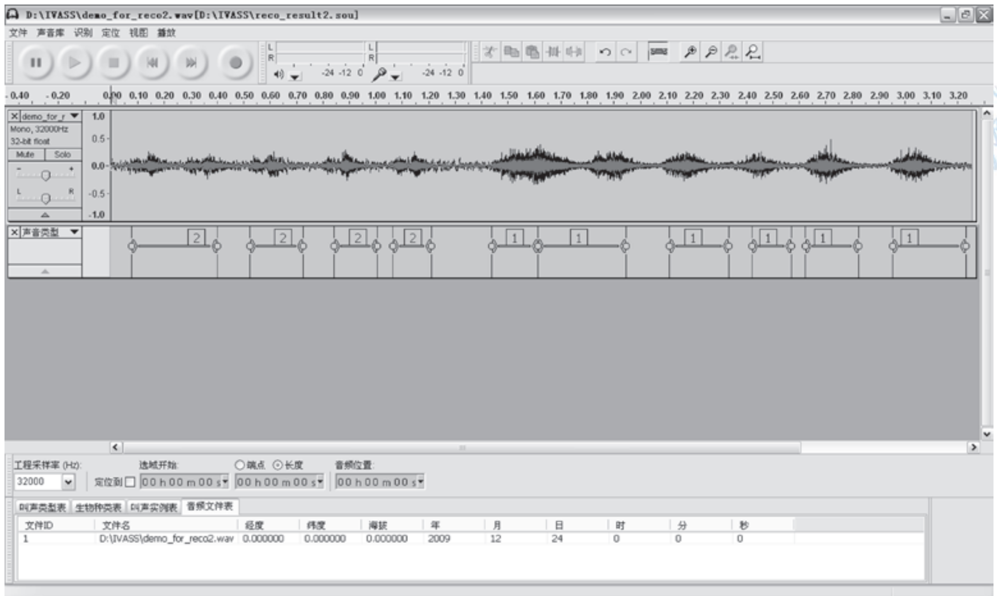


图3 生物声音识别功能的软件运行结果图

(该软件对一段音频进行了声音检测和声音识别。它从一段连续的音频中把可能包含生物声音的片段一个一个提取并标出起始时间和终止时间，并对每个片段进行了生物种类的识别。图中“1”代表了灰喜鹊的叫声，“2”代表了喜鹊的叫声)

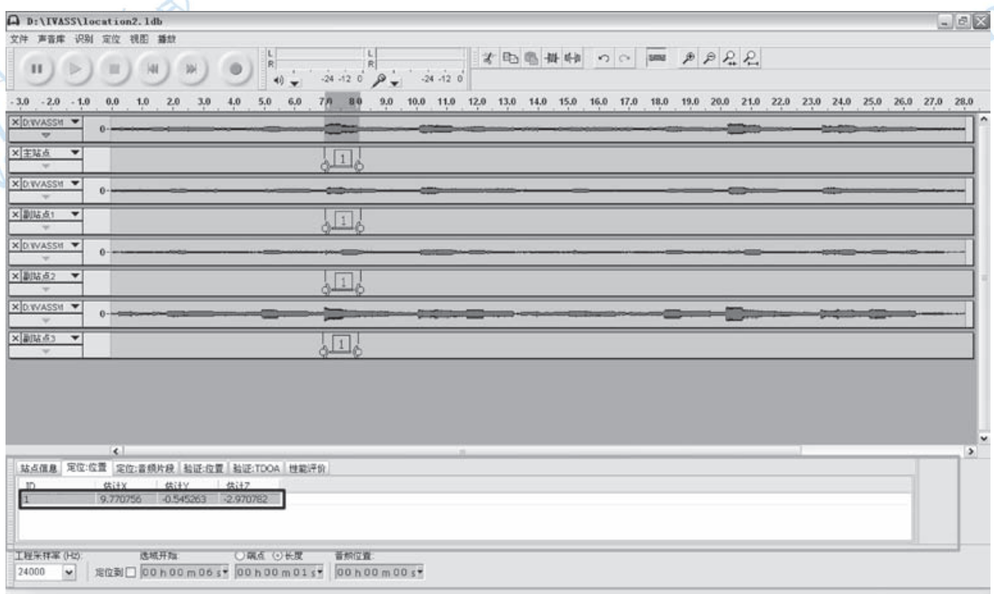


图4 生物声音定位功能的软件运行结果图

(该图展示了以4个录音节点录制的音频为数据源进行声音定位的结果。4个录音节点产生了4个音频文件，分别对应于图中的4个波形图。包含动物叫声的声音片断由软件检测并标记出来，如本图中的标记“1”。界面左下面方框里面显示的就是软件依据当前生物声音片段估计出的动物的三维坐标)

很小的晴天进行实验。

实验中，4个录音节点放在长约50 m、宽约25 m的操场的4个角点，即示意图5中的A、B、C和D4点，其坐标见表1。这4个录音节点分别固定于4杆子的不同高度。这4个杆子通过与4个铁盘相接直立于地面，详见图6、7。

为了验证原型系统的定位精度，我们在已测得的三维位置上放置人工声源，即示意图5中的1、2、3、4、5点，其坐标见表。这5个点是随机选择的。人工声源是一个喇叭，它的真实位置也是通过全站仪和卷尺测量出来的。我们通过比较人工声源的估计值和测量值，来评价这套定位系统。

表1　录音节点和人工声源的三维坐标/m

点名	X	Y	Z
A	14.633	−25.728	1.817
B	−29.609	0.830	2.817
C	−14.6	25.808	3.817
D	23.986	0.029	0.874
1	−8.532	3.055	1.762
2	10.082	−0.532	0.678
3	3.280	−11.995	1.145
4	−25.192	10.486	1.988
5	23.071	−13.887	0.415

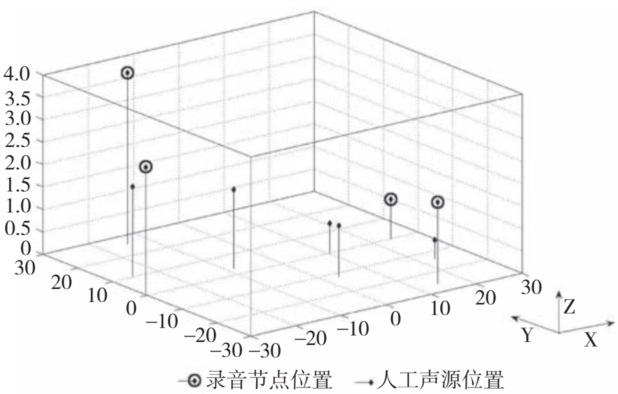


图5　录音节点和人工声源的位置示意图/m

图6　整套系统

图7　正在工作中的1个录音节点

3.5　实验结果

野外实验中我们设置了5个人工声源位置来验证动物声音定位算法。这5个人工声源的测量位置和利用原型系统定位出来的估计位置详见表2。我们对这组结果进行了分析，详见表3和表4。

首先，该套声音定位系统在X、Y方向上的误差都较小。X方向绝对误差的平均值为0.51 m，Y方向绝对误差的平均值为0.87 m；X方向相对误差的平均值为3.5%，Y方向相对误差的平均值为11.4%。目前我们需要利用这套陆生动物声音三维定位系统来研究鸟类的栖息地范围，X,Y方向的定位精度已经满足需求。

其次，与X、Y方向相比，动物声音定位算法在Z方向上的误差较大。绝对误差平均值为4.99 m，相对误差平均值竟然达到了528.9%。本次实验中，动物声音定位算法的三维位置误差主要来源于Z轴的误差。观察表3“绝对误差”中的$\sqrt{(X'-X)^2+(Y'-Y)^2}$和$\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}$这两列就可以直观地得出结论。二维位置误差$\sqrt{(X'-X)^2+(Y'-Y)^2}$的最大值只有1.97 m。而一旦将Z轴也考虑进来时，三维位置误差$\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}$的最大值就增大到了5.99 m。我们通过表4中的比重误差$\frac{|Z'-Z|}{\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}}$这一列也可以很清楚看见Z轴的误差对整个三维位置误差的影响，均大于90%，最小值也为94.2%。可以认为本实验估计出来的Z值是不可靠的。若要提高动物声音定位算法的精度，必须提高Z轴方向的定位精度。

表2　人工声源的三维坐标测量值与估计值/m

人工声源	X	X'	Y	Y'	Z	Z'
1	−8.53	−8.56	3.06	2.41	1.76	7.71
2	10.08	9.77	−0.53	−0.55	0.68	−2.97
3	3.28	3.47	−12.00	−13.18	1.15	−4.43
4	−25.19	−24.17	10.49	12.18	1.99	−3.53
5	23.07	24.05	−13.89	−14.71	0.42	−3.81

注：X，Y，Z表示三维坐标真实值；X'，Y'，Z'表示三维坐标估计值

表3　定位实验的绝对误差/m

人工声源	|X'−X|	|Y'−Y|	|Z'−Z|	$\sqrt{(X'-X)^2+(Y'-Y)^2}$	$\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}$
1	0.03	0.65	5.95	0.65	5.99
2	0.31	0.02	3.65	0.31	3.66
3	0.19	1.18	5.58	1.20	5.71
4	1.02	1.69	5.52	1.97	5.86
5	0.98	0.82	4.23	1.28	4.42
最大值	1.02	1.69	5.95	1.97	5.99
最小值	0.03	0.02	3.65	0.31	3.66
平均值	0.51	0.87	4.99	1.08	5.13

表4　定位实验的相对误差和比重误差/%

人工声源	$\left\|\frac{X'-X}{X}\right\|$	$\left\|\frac{Y'-Y}{Y}\right\|$	$\left\|\frac{Z'-Z}{Z}\right\|$	$\frac{\|Z'-Z\|}{\sqrt{(X'-X)^2+(Y'-Y)^2+(Z'-Z)^2}}$
1	0.4	21.2	338.1	99.4
2	3.1	3.8	536.8	99.6
3	5.8	9.8	485.2	97.8
4	4.0	16.1	277.4	94.2
5	4.2	5.9	1007.1	95.7
最大值	5.8	21.2	1007.1	99.6
最小值	0.4	3.8	277.4	94.2
平均值	3.5	11.4	528.9	97.3

4　误差分析

声音定位误差是由那些用于计算声源位置的参数所具有的不确定性引起的。这些参数包括录音节点的位置，到达时间差和声音在介质中的传播速度等。参数的不确定性是在测量或者估计这些参数时产生的，也称为参数误差。这里误差分析的目的就是根据参数误差来推测声音定位的误差。由于声音定位本质上是非线性的优化问题，很难采用解析方法来推导出声音定位的误差传播公式，所以这里我们采用蒙特卡罗方法来进行误差分析。

蒙特卡罗方法是统计模拟算法。给定待分析参数的误差分布，利用计算机大量模拟声音定位过程，就可以估计出声音定位误差。我们利用MatLab软件实现了该算法。

如果一个声音定位系统的录音节点个数恰好是定位声源所需的最少个数，这样的声音定位系统被称为最少接收器阵列 (MINimum Number of receiver Array)。对于三维声音定位系统，最少的录音节点个数为4。如果一套声音定位系统的录音节点总个数超过了最少接收器阵列所需，这样的系统被称为超定阵列 (OverDetermined Array) (Wahlberg 等, 2001)。在最少接收器阵列中，录音节点的个数刚好足够用来定位声源，没有多余的录音节点可以用来减少各种测量误差和估计误差的影响。然而一套超定阵列却有多余的录音节点，采用最小二乘法就可以减少各种测量误差和估计误差的影响 (Wahlberg 等, 2001)。所以通常超定阵列的定位误差小于最少接收器阵列的定位误差，这在以下的章节中得到了验证。

算法1 基于蒙特卡罗方法的声音定位误差分析算法($[L_x, L_y, L_z]$指定录音节点和声源的三维空间范围；“*time_resolution*”是TDOA的精度，也是系统的时间分辨率；“N”代表录音节点的总个数；所有这些参数都是用来描述一套声音定位系统的特性。)

步骤1 初始化。用户指定三维空间范围$[L_x, L_y, L_z]$，系统时间分辨率*time_resolution*，录音节点的总个数N，声音定位过程模拟的总次数*n_iters*，经验累计概率的阈值*emp_cum_pro_th*。令*iters*=0。

步骤2 找出所有参数的真实值。在以[0,0,0]为中心、长为L_x、宽为L_y、高为L_z的立方体内，随机产生一个点作为声源，其坐标为声源位置的真实值，以$[S_x,S_y,S_z]$表示。同样，随机产生N个点作为录音节点，以$[R_{xi},R_{yi},R_{zi}]$, i=1, 2, …, N表示第i个录音节点位置的真实值。假设声音传播速度的真实值为340 m/s。随机选择第r个录音节点作为参考录音节点，到达时间差的真实值就可以计算出来，以τ_{ri}, i=1, 2,…, N, $i\neq r$来表示。

步骤3 找出所有参数的估计值。录音节点位置的测量误差、到达时间差的估计误差和声速的估计误差都可以假设成独立的、满足某种概率分布且以0为期望的随机变量。所以录音节点的位置估计值$[R'_{xi},R'_{yi},R'_{zi}]$, i=1, 2, …, N，到达时间差的估计值τ'_{ri}, i=1, 2, …, N, $i\neq r$和声速的估计值c可以通过其真实值和某种随机误差之和来表示。然后利用录音节点位置的估计值、到达时间差的估计值和声速的估计值，求解出声源位置，记为$[S'_x,S'_y,S'_z]$。

步骤4 利用$[S_x,S_y,S_z]$和$[S'_x,S'_y,S'_z]$就可以计算出各种类型声音定位误差的值。

步骤5 $iter = iter + 1$。如果$iter < n_iters$，返回**步骤2**；否则，跳到**步骤6**。

步骤6 运行至此，对于每一种声音定位误差类型，都产生了大量的数值。对于某种定位误差类型，当其经验累积概率增大到不小于用户定义的某个阈值*emp_cum_pro_th*时，其所对应的数值用来表征该类型定位误差的大小。

4.1　录音节点的测量误差

测量必定存在误差。测量误差常常满足正态分布。也就是说，沿着第j个录音节点的第i个方向，测量误差的概率分布函数为高斯分布：

$$N(0,(\sigma_{i,j}^R)^2),\ i=x,y,z,\ j=1,2,\cdots,N \tag{7}$$

式中，$\sigma_{i,j}^R$是测量误差的标准差。假设所有录音节点测量误差是独立同分布的，那么

$$N(0,(\sigma_{i,j}^R)^2),\ i=x,y,z,\ j=1,2,\cdots,N$$

可以简化为$N(0,(\sigma_i^R)^2), i=x,y,z$。假设测量误差在3个方向上也是独立同分布的，那么

$N(0,(\sigma_i^R)^2), i=x,y,z$可进一步简化为$N(0,(\sigma^R)^2)$。图8展示了在这种假设条件下，声音定位误差如何随着录音节点测量误差的增大而增大。

4.2　到达时间差的估计误差

到达时间差的估计值往往与真实值之间有一定的误差。这是由声音定位系统是离散数字系统这一本质以及采用了加窗信号处理技术所决定的。假设一套声音定位系统的采样率为*sample_rate*(单位：Hz)，加窗信号处理的窗口步长为*win_step*，那么它所能精确到的最小时间间隔为*win_step* / *sample_rate*(单位：s)。可见，到达时间差的估计值必定是这个最小时间间隔的倍数。但是，到达时间差的真实值是连续的，并不一定是最小时间间隔的倍数，所以到达时间差的估计值与真实值之间往往存在误差，这种误差称为“到达时间差的固有估计误差”。我们称该最小时间间隔为声音定位系统的时间分辨率，记为*time_resolution* (单位：s)。即：

$$time\_resolution = win\_step \,/\, sample\_rate \qquad (8)$$

根据上述分析，到达时间差的固有估计误差应满足均匀分布U(–*time_resolution*/2，*time_resolution*/2)。图9显示了定位误差如何随着系统时间分辨率的增大而增大。

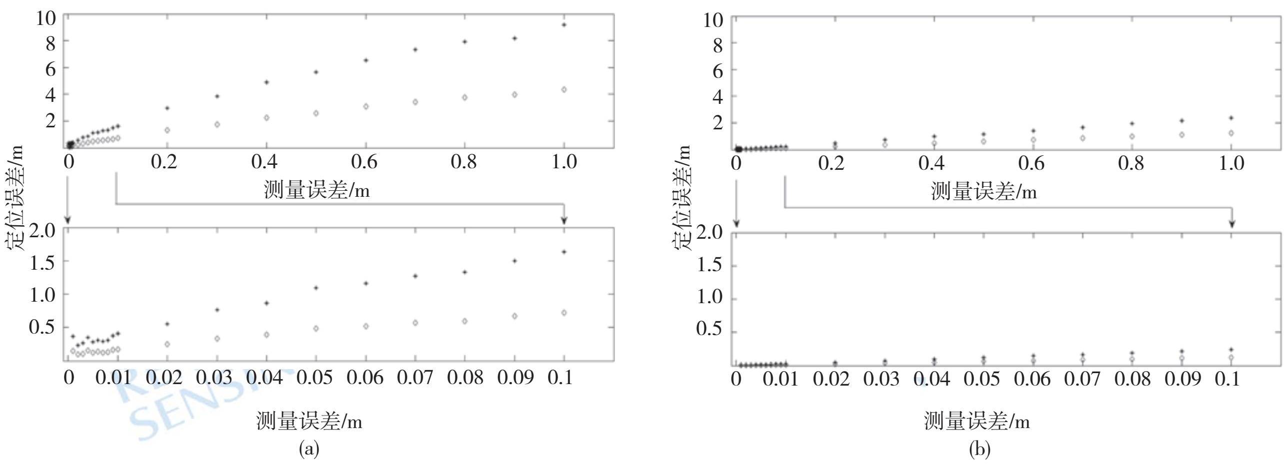


图8　测量误差对定位误差的影响

(a) 最少接收器阵列, *N*=4; (b) 超定阵列, *N*=10

(每一个*点代表定位总误差的模拟值，每一个o点表示沿X轴的绝对定位误差的模拟值 (Y轴和Z轴与X轴情况类似)。测量误差σ^R以变步长从0.001 m增大到1 m：对于[0.001 m, 0.01 m]，步长是0.001 m；对于(0.01 m, 0.1 m]，步长是0.01 m；对于(0.1 m, 1 m]，步长是0.1 m。这里，模拟次数为10000；空间范围[L_x,L_y,L_z]为[30 m, 30 m, 30 m]；声音传播速度假设为340 m/s。这里假设不存在TDOA的估计误差和声音传播速度的估计误差。)

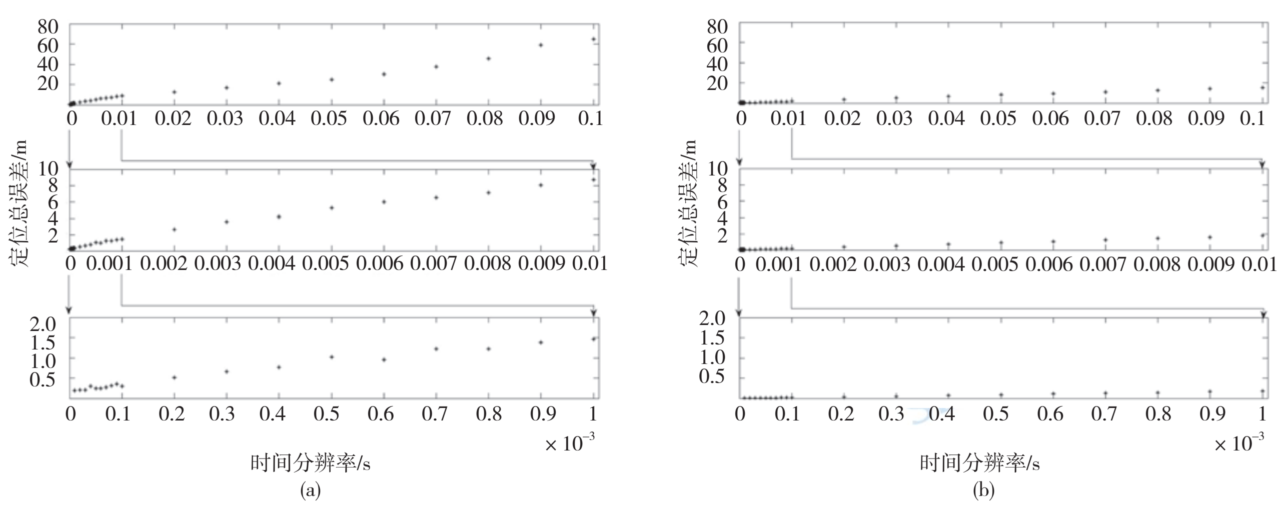


图9　时间分辨率引起的固有TDOA估计误差对定位误差的影响

(a) 最少接收器阵列，*N* = 4; (b) 超定阵列，*N* = 10

(每一个*点代表定位总误差的模拟值。时间分辨率以变步长从0.00001 s增大到0.1 s：对于[0.00001 s, 0.0001 s]，步长是0.00001 s；对于(0.0001 s, 0.001 s]，步长是0.0001 s；对于(0.001 s, 0.01 s]，步长是0.001 s；对于(0.01 s, 0.1 s]，步长是0.01 s。这里，模拟次数为10000；空间范围[L_x,L_y,L_z]为[30 m, 30 m, 30 m]；声音传播速度假设为340 m/s。这里假设不存在测量误差和声音传播速度的估计误差)

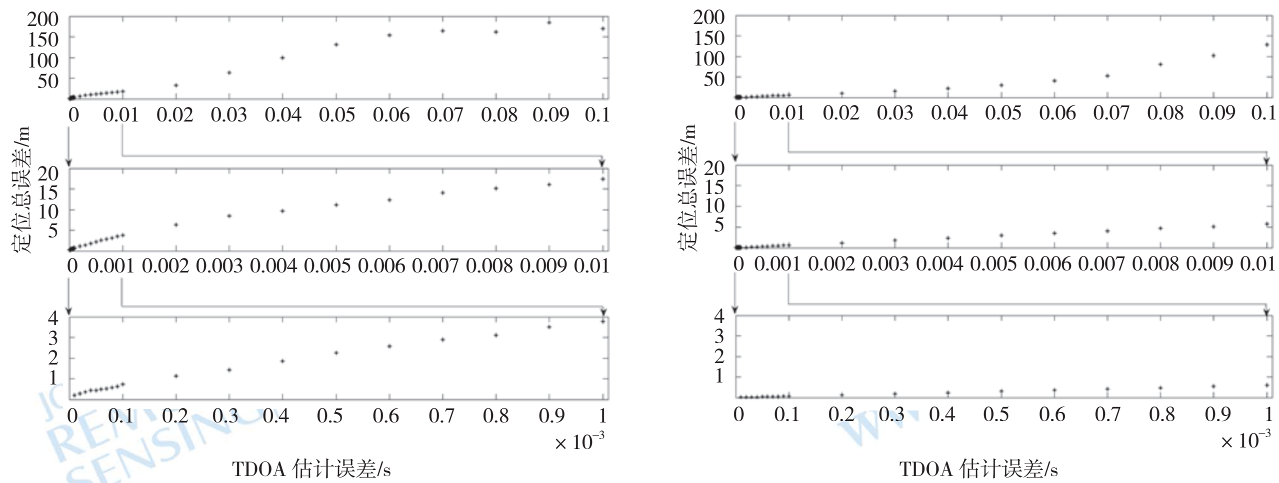
除了固有估计误差外，噪声会导致不精确的到达时间差估计，而录音节点之间的时间同步误差也是到达时间差估计误差的来源之一。显然，这两类到达时间差的估计误差的数量级大于固有估计误差，所以当存在这两类到达时间差的估计误差时，到达时间差的固有估计误差可以忽略不计。通常采用正态分布$N(0, (\sigma^R)^2)$来描述到达时间差的一般估计误差。图10显示了定位误差随着到达时间差的一般估计误差的增大而增大。

4.3 声速的估计误差

声音在空气中随着温度、高度、湿度和风速等环境参数的变化而发生改变。假设声速的估计误差满足正态分布$N(0, (\sigma^s)^2)$，图11展示了定位误差如何随着声速估计误差的增大而增大。

4.4 合理的录音节点总个数

一方面，录音节点越多声音定位系统的定位精度

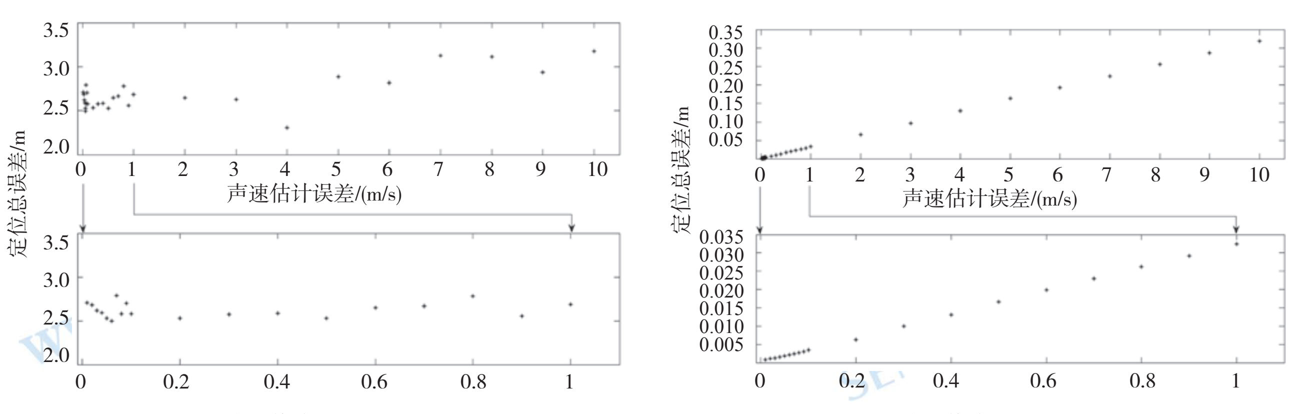


(a) (b)

图10 TDOA估计误差对定位误差的影响

(a) 最少接收器阵列, $N = 4$; (b) 超定阵列, $N = 10$

(每一个*点代表定位总误差的模拟值。TDOA的估计误差以变步长从0.00001 s增大到0.1 s：对于[0.00001 s,0.0001 s]，步长是0.00001 s；对于(0.0001 s, 0.001 s]，步长是0.0001 s；对于(0.001 s, 0.01 s]，步长是0.001 s；对于(0.01 s, 0.1 s]，步长是0.01 s。这里，模拟次数为10000；空间范围[L_x,L_y,L_z]为[30 m, 30 m, 30 m]；声音传播速度假设为340 m/s。这里假设不存在测量误差和声音传播速度的估计误差)



(a) (b)

图11 声音速度的估计误差对定位误差的影响

(a) 最少接收器阵列, N=4; (b) 超定阵列, N=10

(每一个*点代表定位总误差的模拟值。声音速度的估计误差以变步长从0.01 m/ s增大到10 m/s：对于[0.01 m/s, 0.1 m/s]，步长是0.01 m/s；对于(0.1 m/s, 1 m/s]，步长是0.1 m/s；对于(1 m/s, 10 m/s]，步长是1 m/s。这里，模拟次数为10000；空间范围[L_x,L_y,L_z]为[30 m, 30 m, 30 m]；声音传播速度假设为340 m/s。这里假设不存在测量误差和TDOA的估计误差)

越接近理论值(图12)；另一方面，录音节点越多也意味着成本越大。所以在确定录音节点的总个数时，应综合考虑精度和成本。

算法1可用来估计不同录音节点总个数下声音定位系统的定位精度(图12)。综合考虑成本预算，算法1可以辅助决策者确定合理的录音节点总个数。

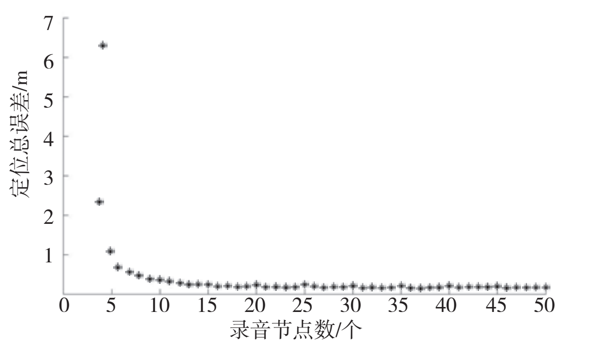


图12 录音节点个数对定位误差的影响

(每个*点表示定位总误差的估计值。录音节点的总个数从4增大到50，步长为1。这里，循环次数为100；空间范围为[30 m, 30 m, 30 m]；任一轴的测量误差为0.1 m；时间分辨率为0.001 s；声音传播速度为340 m/s。这里假设不存在声音速度的估计误差)

4.5 Z轴定位精度低的原因

从野外实验结果，我们可以看出Z轴定位误差比X轴和Y轴都大很多。利用算法1，我们发现其原因是Z轴的空间范围远小于X，Y轴的空间范围，定位误差随着Z轴的空间范围的增大而逐渐缩小(图13)。

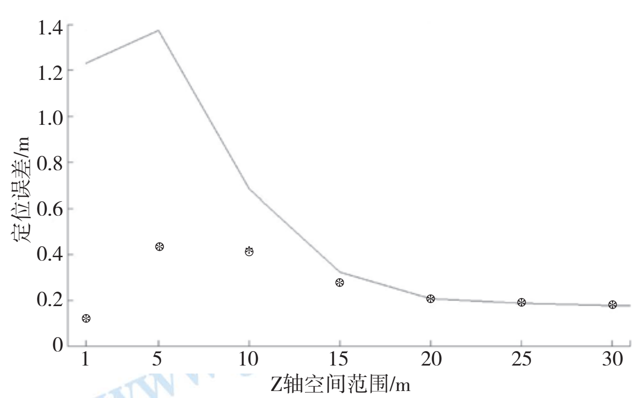


图13 Z轴空间范围对定位误差的影响

(每个*点代表X轴的绝对定位误差，每个○点代表Y轴的绝对定位误差，而曲线则代表Z轴的绝对定位误差。沿Z轴的空间范围从1 m增大到30 m，以5 m为步长。这里，X轴和Y轴的空间范围保持在30 m。沿任一轴的测量误差大小为0.01 m，不考虑TDOA的估计误差和声音速度的估计误差。声音传播速度假设伟340 m/s。模拟次数为10000)

5 总结与讨论

本文展示了一套基于无线传感器网络技术的陆生动物声音定位原型系统。户外实验证实了这套原型系统满足我们当前对生物定位的精度需求。为了揭示影响动物声音定位系统定位精度的各种因素并评价其影响，采用了蒙特卡罗方法进行分析。统计模拟结果显示，定位精度会随着录音节点自身测量误差的增大、到达时间差的估计误差的增大、声速的估计误差的增大而降低。在成本允许的情况下适当增大录音节点的总个数，有利于提高声音定位精度。

这套原型系统是我们开发更加智能的陆生动物定位系统的起点。无线传感器网络技术将在生物监测领域发挥强大的作用，因为它从根本上改变了人与物理世界交互的方式。未来，我们将集成GPS模块来实现录音节点的自身定位和录音节点之间的时间同步。另外，还会考虑复杂的声音传播模型，从而更准确地估计声音在野外环境中的传播。

志 谢 本研究得到遥感科学国家重点实验室杨习荣高级工程师的技术帮助，谨致谢意。

REFERENCES

Anderson S E, Dave A S and Margoliash D. 1996. Template-based automatic recognition of birdsong syllables from continuous recordings. *Journal of Acoustical Society of America*, **100**(2): 1209–1219 DOI: 10.1121/1.415968

Bohn D A. 1987. Environmental effects on the speed of sound. *Audio Engineering Society Convention* 83, New York: 16–19

Clark C W and Clapham P J. 2004. Acoustic monitoring on a humpback whale (*Megaptera novaeangliae*) feeding ground shows continual singing into late spring. *Proceedings of the Royal Society B-Biological Sciences*, **271**(1543): 1051–1057 DOI: 10.1098/rspb.2004.2699

Clark C W, Gillespie D, Nowacek D P and Parks S E. 2007. Listening to their world: acoustics for monitoring and protecting right whales in an urbanized ocean // Kraus S D and Rolland R M, eds. The Urban Whale: North Atlantic Right Whales at the Crossroads. Boston: Harvard University Press: 333–357

Everest F A and Ken C P. 2009. The master handbook of acoustics. New York: McGraw–Hill

Exadaktylos V, Silva M, Ferrari S, Guarino M, Taylor C J, Aerts J M and Berckmans D. 2008. Time-series analysis for online recognition and localization of sick pig (*Sus scrofa*) cough sounds. *Journal of Acoustical Society of America*, **124**(6): 3803–3809

DOI: 10.1121/1.2998780

Fagerlund S. 2007. Bird species recognition using support vector machines. *EURASIP Journal on Applied Signal Processing*, **2007**(1): 64 DOI: 10.1155/2007/38637

Farnsworth A. 2005. Flight calls and their value for future ornithological studies and conservation research. *The Auk*, **122**(3): 733–746 DOI: 10.1642/0004-8038

Gong P. 2007. Wireless sensor network as a new ground remote sensing technology for environmental monitoring. *Journal of Remote Sensing*, **11**(4): 545–551

Hardy H C, Telefair D and Pielemeier W H. 1942. The velocity of sound in air. *Journal of Acoustical Society of America*, **13**(3): 226–233 DOI: 10.1121/1.1916169

Jones D L and Ratnam R. 2009. Blind location and separation of callers in a natural chorus using a microphone array. *Journal of Acoustical Society of America*, **126**(2): 895–910 DOI: 10.1121/1.3158924

Knapp C H and Carter G C. 1976. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics Speech and Signal Processing*, **24**(4): 320–327 DOI: 10.1109/TASSP.1976.1162830

Kogan J A and Margoliash D. 1998. Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study. *Journal of Acoustical Society of America*, **103**(4): 2185–2196 DOI: 10.1121/1.421364

Kwan C, Ho K C, Mei G, Li Y, Ren Z, Xu R, Zhang Y, Lao D, Stevenson M, Stanford V and Rochet C. 2006. An automated acoustic system to monitor and classify birds. *EURASIP Journal on Applied Signal Processing*, **2006**(3): 1–19 DOI: 10.1155/ASP/2006/9670

McGregor P K, Dabelsteen T, Clark C W, Bower J L, Tavares J P and Holland J. 1997. Accuracy of a passive acoustic location system: empirical studies in terrestrial habitats. *Ethology Ecology and Evolution*, **9**(3): 269–286 DOI: 10.1080/08927014.1997.9522887

McIlraith A L and Card H C. 1997. A comparison of backpropagation and statistical classifiers for bird identification. *IEEE International Conference on Neural Networks*, **1**: 100–104 DOI: 10.1109/ICNN.1997.611645

Mellinger D K and Clark C W. 2003. Blue whale (*Balaenoptera musculus*) sounds from the North Atlantic. *Journal of Acoustical Society of America*, **114**(2): 1108–1119 DOI: 10.1121/1.1593066

Mennill D J, Burt J M, Fristrup K M and Vehrencamp S L. 2006. Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. *Journal of Acoustical Society of America*, **119**(5): 2832–2839 DOI: 10.1121/1.2184988

Ning H S, Chen W S, Mao X and Li J. 2010. Bird-aircraft strike avoidance radar. *IEEE Aerospace and Electronic Systems Magazine*, **25**(1): 19–28 DOI: 10.1109/MAES.2010.5442150

Rabiner L R and Sambur M R. 1975. An algorithm for determining the endpoint for isolated utterances. *The Bell System Technical Journal*, **54**(2): 297–315

Stafford K M, Fox C G and Clark D S. 1998. Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean. *Journal of Acoustical Society of America*, **104**(6): 3616–3625 DOI: 10.1121/1.423944

Wahlberg M, Mohl B and Madsen P T. 2001. Estimating source position accuracy of a large-aperture hydrophone array for bioacoustics. *Journal of Acoustical Society of America*, **109**(1): 397–406 DOI:10.1121/1.1329619

附中文参考文献

宫鹏．2007．环境监测中无线传感器网络地面遥感新技术．遥感学报，**11**(4): 545–551