Energy-Based Sound Source Localization with Low Power Consumption in Wireless Sensor Networks

Fang Deng, *Member, IEEE*, Shengpan Guan, Xianghu Yue, Xiaodan Gu, Jie Chen, *Senior Member, IEEE*, Jianyao Lv, and Jiahong Li

***Abstract*—In this study, we show that the energy-based method of sound source localization can be successfully exploited for sound source localization under low power consumption conditions. Sound source localization is widely applied in battlefield environments where low power consumption is especially crucial and necessary for extending the lifespan of sensor nodes. We propose several variables that may (possibly) affect the path loss exponent. We provide data that shows that energy-based methods of sound source localization can accurately determine the appropriate path loss exponent and can improve localization accuracy. The results of our study also demonstrate that energy-based methods of sound source localization can significantly reduce localization errors by adjusting sensors' weight coefficients when ambient (background) noise exists. Our test results indicate that our method of source localization using low power consumption is consistently accurate and is able to determine sound source localization multiple times over an extended period.**

***Index Terms*—Energy-based algorithm, low power consumption, path loss exponent, sound energy attenuation model, sound source localization.**

I. Introduction

SOURCE localization has garnered a great deal of attention due to its numerous applications in communications, navigation, and target tracking [1]–[8]. There are many localization methods in existence, including the Multiple Signal Classification (MUSIC) algorithm, correlation approaches, and beamforming-based approaches [9]. While these methods can be used in sound source localization systems, they must first satisfy the conditions of low power consumption in order to extend the life of the sensor nodes used in wireless sensor networks.

The MUSIC method (first proposed in 1986) belongs to the so-called "high resolution" approach because of the sharpness (or high resolution) of its results. Applying the MUSIC method to broadband signals is somewhat difficult because computations must be conducted in an environment with extreme noise. Asano was the first to apply the MUSIC algorithm to robotics. As part of an experiment, Asano distributed an array of $N = 8$ microphones around the periphery of Jijo-2 (a robot). Asano was able to obtain realistic localizations of vocal sources by applying the narrowband method to broadband signals [10]. Such an application of the MUSIC method may be seen as "naive" because it so closely follows the mandates of the narrowband algorithm. Asano's use of this variation of the MUSIC method is known as Standard Eigen Value Decomposition-MUSIC. Another interesting variation of the MUSIC method was proposed by Nakamura *et al.* in 2011. Nakamura *et al.* proposed the Generalized Eigen Value Decomposition-MUSIC (GEVD-MUSIC) method as a way to identify noise and signal in [11].

Time difference of arrival (TDOA) is one of the correlation approaches used to determine source localization. The main function of TDOA is to estimate the time delay between a pair of microphones. Generalized cross correlation (GCC) techniques, employed in conjunction with the PhaT weighting function (GCC-PhaT) [12], [13], are by far the most commonly used methods in a robotics context (due to the high temporal resolution in TDOA estimations). For example, Kim *et al.* used a triangular 3-microphone array to infer source location from short-time observations in order to cope with the movement of the sound source (or the robot) [14]. Additionally, Badali *et al.* provide a useful evaluation of various real-time sound localization approaches using a cubical 8-microphone array in which GCC-PhaT is compared with beamforming techniques [15]. Finally, Hilsenbeck and Kirchner proposed a robust approach to the acoustic perception of the presence of people from a pair of microphones [16].

However, despite GCC-PhaT's obvious uses (detailed above), this method still has some issues. For example, GCC-PhaT only takes into account the phase of the perceived signals in the intercorrelation computation and assigns the same importance to

Manuscript received August 31, 2016; revised November 16, 2016; accepted December 7, 2016. Date of publication January 16, 2017; date of current version May 10, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61304254, Grant 61321002, and Grant 61120106010, and in part by Beijing NOVA Program xx2016B027.

F. Deng, X. Yue, X. Gu, J. Chen, J. Lv, and J. Li are with the Key Laboratory of Intelligent Control and Decision of Complex Systems, School of Automation, Beijing Institute of Technology, Beijing 100081, China (e-mail: dengfang@bit.edu.cn; yuexianghubit@163.com; x_dan_2002@163.com; chenjie@bit.edu.cn; ljy091@126.com; lijiahong.bit@gmail.com).

S. Guan is with Qingdao Topscomm Communication Co., Ltd., Qingdao 266024, China (e-mail: guanshengpan@126.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2017.2652394

each frequency. In order to solve this problem, Valin *et al.* created an alternative process which penalizes frequencies with a low signal-to-noise ratio (SNR) [17]. Furthermore, while correlation approaches have improved localization accuracy, they mainly assume planar wave fronts in order to limit algorithmic complexity.

Among all the methods rooted in signal processing, those based on beamforming are the most commonly used in robotics because of their low computational cost. Mattos and Grant used a beamforming approach in their experiments and concluded that while the beam pattern of the main lobe is thin enough for frequencies over 1 kHz, frequencies below 800 Hz cannot be exploited for localization [18]. Ultimately, the localization of conventional beam-formers is limited to low frequencies, and thus alternative beamforming methods are necessary for future work [9].

Furthermore, sound source localization remains a challenging task due to factors including environmental noise and power consumption. For sound source localization with microphone arrays, existing methods can be categorized into two groups [19]. The first group consists of more direct methods in which either microphone arrays are positioned in various locations, and seek out the peak output power of signals, or signals are analyzed for a correlation matrix (see [20]–[22]). The location with the highest output power is considered to be the sound source location. The second group encompasses two methods. The first method is based on estimating the TDOA [23], [24], for the signals received by spatially separated microphone pairs, and then calculating the source position based on the TDOA estimation and array geometry. The second method is based on the strength of the received sound energy, the results of which first define the sound energy attenuation model, and then calculate the sound source position.

The second group is the more indirect of the two methods, and requires more steps. However, when compared to the more direct group, the indirect group does have two advantages. The first advantage is that the indirect group has a lower computational cost. The second advantage is that it uses less sampling data. Consequently, the indirect group is suitable for use in an environment that requires a low sampling rate and less sampling data [25]. Furthermore, while the TDOA-based method is very suitable for pulse signals [26], it is insufficient for real-time applications when the sound signal extends over a long period of time.

In this study, we have adopted an energy-based method in order to satisfy the requirements of real-time applications and low power consumption. We propose possible variables that may affect the path loss exponent and we provide a method of determining an appropriate path loss exponent that improves localization accuracy. Our proposed method also reduces localization errors by adjusting the weight coefficients of the sensors when ambient noise exists.

The remainder of this paper is organized as follows. In Section II, we present (in detail) the sound energy attenuation model. In Section III, we describe the localization algorithm. In Section IV, we provide an analysis of power consumption. In Section V, we provide the experimental and simulation results. Finally, we present the conclusion in Section VI.

II. Traditional Sound Energy Attenuation Model [27]

When sound travels through the air, acoustic energy is emitted omnidirectionally from the sound source. The strength of a sound source diminishes at a rate inversely proportional to the square of the distance. This formula serves as the basis of the theory of wireless sensor networks sound source localization based on energy. The traditional sound energy attenuation model is defined as

$$y_i(t) = \zeta_i \frac{S(t)}{|r_i - r(t)|^{\alpha}} + \varepsilon_i(t) \tag{1}$$

where $y_i(t)$ indicates the signal energy measured on the ith sensor, ζ_i is the gain factor of the ith acoustic sensor, $S(t)$ is the sound energy 1 m from the sound source (we can think of this as the energy of the sound source). r_i and $r(t)$ indicates the coordinates of ith sensor node and sound source at time t. Each variable is a vector with two additional variables (when in a two-dimensional (2-D) plane). As we only consider the case of 2-D planes in this study, the variables can be a set as (x, y), with $|r_i - r(t)|$ being the distance difference between the ith sensor and the sound source at time t, α being the path loss exponent ($\approx$2), $\varepsilon_i(t)$ being the cumulative effects of the modeling error of parameters ζ_i, and α, and the additive observation noise being $y_i(t)$ (for the purposes of this paper, $y_i(t)$ can be thought of as ambient noise).

III. Indoor/Outdoor Experiments About the Sound Energy Attenuation Model

The existing sound energy attenuation model (mentioned above) is not accurate. Although the path loss exponent (α) is assumed to be constant ($\approx$2) in [27], it was actually obtained using a specific point, and an omnidirectional sound source in a favorable environment (a gentle breeze and mild temperatures). The parameters below may vary in different situations, but they are all influenced (primarily) by the following factors.

1) Ambient noise will have an impact on α.
2) The ideal point (an omnidirectional sound source which does not exist in a real environment) will have an impact on α.
3) The distances between the sensors and the sound source will impact α.

In order to validate the potential impacts (described above), we conducted experiments in both indoor and outdoor environments.

Before conducting these experiments, it was necessary to solve the problem of how to find accurate sound energy. Assuming there are n sampling points, X_i is the sampling value of the ith point, and X the sampling value when there is no sound source. Energy of the sound can be defined as

$$E = \gamma \frac{1}{n} \sum_{i=1}^{n} |X_i - X|^2 \tag{2}$$

where γ is a constant. The magnitude of a continue sound wave changes continuously, as shown in Fig. 1. Therefore, it is not possible to obtain an accurate value of the sound energy when there is little sampling data. A more accurate value can be found

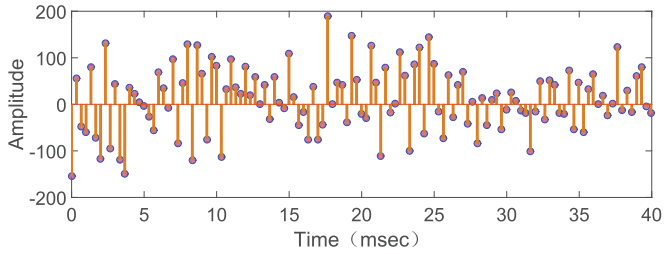

Fig. 1. Sound wave of the tank used in the experiments. In consideration of energy consumption and localization accuracy, the sampling frequency was set at approximately 3000 Hz and the sampling length was set at approximately 40 ms.

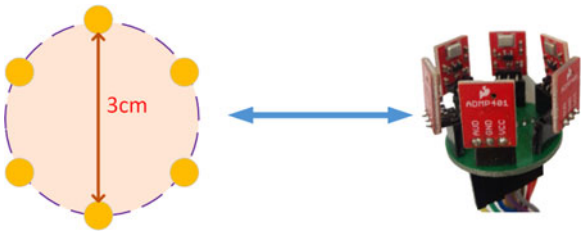

Fig. 2. Sound sensor array. Six sound sensors constitute a circular array. Sound source energy can be determined more accurately by calculating the average value of the six sound sensors.

TABLE I
INDOOR EXPERIMENTAL RESULTS

ID	0°		90°	
	α	RMSE	α	RMSE
1	1.57	0.01	1.21	0.023
2	1.50	0.02	1.26	0.03

if there is more sampling data; however, more sampling data requires more time when using the same sampling rate and does not satisfy the requirements of real-time applications. As such, it is necessary to use six sound sensors in order to find a more accurate value (by calculating the average value); The sound sensor array is shown in Fig. 2.

In our indoor experiments, our acoustic source was a tape recorder placed in a fixed location. Sensor nodes with acoustic microphones were placed at varying intervals from the energy source (1, 2, 3, 4, and 5 m). The microphones were placed approximately 20 cm above the ground, and the temperature was approximately 25 °C. The following two experiments were conducted (with each experiment being conducted two times):

1) The sound source was facing the sensor nodes, and the angle was 0°.
2) The sound source was not facing the sensor nodes, and the angle was 90°.

The indoor experimental results are shown in Table I.

In the outdoor experiments, our acoustic energy source was a tape recorder placed in a fixed location. The microphones were placed approximately 20 cm above the ground, then weather was clear (with a gentle breeze), and the temperature was approximately 24 °C. Four experiments were conducted and every experiment was conducted two times.

TABLE II
OUTDOOR EXPERIMENTAL RESULTS

ID	Distance Interval With 1 m				Distance Interval With 5 m			
	0°		90°		0°		90°	
	α	RMSE	α	RMSE	α	RMSE	α	RMSE
1	1.93	0.02	1.65	0.01	2.45	0.02	1.71	0.01
2	2.01	0.02	1.65	0.01	2.24	0.01	1.73	0.02

In the first experiment, the energy source was facing the sensor nodes (with the angle 0°); sensor nodes with acoustic microphones were placed at varying intervals from the energy source (5, 10, 15, 20, 25, and 30 m).

The second experiment employed a sound source that did not face the sensor nodes (with the angle 90°); sensor nodes with acoustic microphones were placed at varying intervals from the energy source (5, 10, 15, 20, 25, and 30 m).

The third experiment employed an energy source that faced the sensor nodes (with the angle 0°); sensor nodes with acoustic microphones were placed at varying intervals from the energy source (1, 2, 3, 4, and 5 m).

The fourth experiment employed a sound source that did not face the sensor nodes (with the angle 90°); sensor nodes with acoustic microphones were placed at varying intervals from the energy source (1, 2, 3, 4, and 5 m).

The outdoor experimental results are shown in Table II.

We recorded the time series for both acoustic sensors in both the indoor environment and the outdoor environment at a sampling rate of 3000 Hz, after which point the energy readings were computed. Given the experimental results already obtained, it can be concluded that:

When the distances and the source types are the same, the values of α will be smaller in an indoor environment than in an outdoor environment. As such, α will have different values in different environments.

When the distances and environments were the same, the values of α were greater in the sources facing the sensors than the sources not facing the sensors. Consequently, the type of sound source utilized will have an impact on α.

When the source types and environments are the same, the values of α will be greater for distances 5, 10, 15, 20, 25, and 30 m than for distances of 1, 2, 3, 4, and 5 m. Consequently, the distances between the sensors and the sound source will also impact α.

As such, we now know that the values of α will be different when the three factors (environments, source types, and distances) are different. Next, we present how best to find the appropriate value of α.

First of all, the following assumptions are made.

1) There is a single moving sound source that does not initially make a sound.
2) The coordinates of the sensor nodes are known.

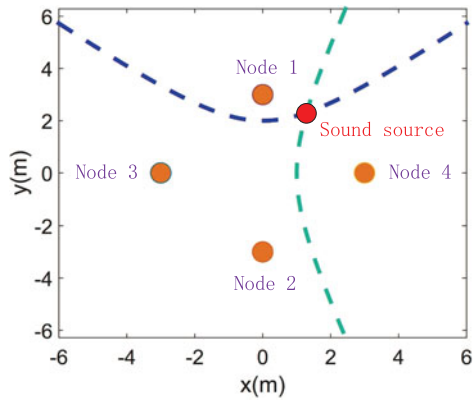

Fig. 3. Sound source localization according to the time difference of arrival.

3) At least four sensor nodes are used in the 2-D sensor field (this is presented in Section IV).

The following method was then adopted to find α.

Step 1: When the sound source has not yet made a sound, the ambient noise from each sensor node can be obtained from the sound sensor.

Step 2: When the sound source begins to make a sound, the four sensor nodes record the arrival times of sounds as t_1, t_2, t_3, t_4, and the sound energy as y_1, y_2, y_3, y_4.

Step 3: The hyperbolic can be obtained from t_1, t_2, t_3, t_4 just as shown in Fig. 3, and the position of the sound source r can be obtained.

Step 4: Now the sound source coordinates, the ambient noise from each sensor node, and the four sensor nodes coordinates are all known. From (1) α_{ij} can be defined as

$$\alpha_{ij} = \log_{\frac{|r_j - r(t)|}{|r_i - r(t)|}} \frac{y_i - \varepsilon_i}{y_j - \varepsilon_j} \tag{3}$$

where α_{ij} is the path loss exponent obtained by the ith and jth sensor nodes ($\alpha_{ij} = \alpha_{ji}$). When there are four sensor nodes, the value of α can be obtained as

$$\alpha = \frac{\alpha_{12} + \alpha_{13} + \alpha_{14} + \alpha_{23} + \alpha_{24} + \alpha_{34}}{6}. \tag{4}$$

When there are m sensor nodes, the value of α can be obtained as

$$\alpha = \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \alpha_{ij}. \tag{5}$$

This method can reduce the localization error greatly (this is validated in Part B of Section VI).

IV. Localization Algorithm

Since the sound energy attenuation model has been obtained in a real environment, we now present the algorithm of sound source localization. Assuming there are m sensor nodes around the sound source, the sound energy detected by the ith sensor node is $y_i(t)$ and the sound source detected by the jth sensor node is $y_j(t)$. Equations (6) and (7) can be obtained from (1) as [27]

$$y_i(t) = \zeta_i \frac{S(t)}{|r(t) - r_i|^\alpha} + \varepsilon_i(t) (i \in m) \tag{6}$$

$$y_j(t) = \zeta_j \frac{S(t)}{|r(t) - r_j|^\alpha} + \varepsilon_j(t) (j \neq i, j \in m). \tag{7}$$

Equations (8) and (9) can be obtained from (6) and (7) as [27]

$$|r(t) - r_i| = \left[\frac{y_i(t) - \varepsilon_i(t)}{\zeta_i S(t)} \right]^{\frac{1}{\alpha}} \tag{8}$$

$$|r(t) - r_j| = \left[\frac{y_j(t) - \varepsilon_j(t)}{\zeta_j S(t)} \right]^{\frac{1}{\alpha}}. \tag{9}$$

When the ith sensor and the jth sensor are in the same environment, it can be assumed that the sensor nodes have the same gain factor, the relationship between the distance ratio and the energy ratio can be obtained as [27]

$$k_{ij} = \frac{|r(t) - r_i|}{|r(t) - r_j|} = \left[\frac{y_j(t) - \varepsilon_j(t)}{y_i(t) - \varepsilon_i(t)} \right]^{\frac{1}{\alpha}} \tag{10}$$

where $k_{ij} > 0$ and it is a constant, we can get its value from (10). Equation (10) can be expressed as [27]

$$\left[r(t) - \frac{r_i - k_{ij}^2 r_j}{1 - k_{ij}^2} \right]^2 = \frac{k_{ij}^2 (r_i - r_j)}{(1 - k_{ij}^2)^2}. \tag{11}$$

Let $C_{ij} = \frac{r_i - k_{ij}^2 r_j}{1 - k_{ij}^2}$, $\rho_{ij} = \frac{k_{ij} |r_i - r_j|}{1 - k_{ij}^2}$, then (11) can be expressed as [27]

$$|r(t) - C_{ij}|^2 = \rho_{ij}^2. \tag{12}$$

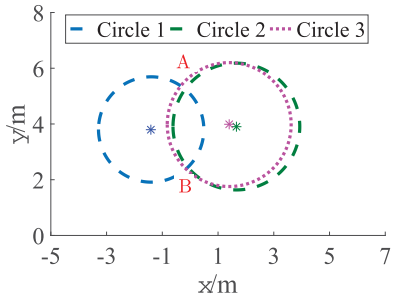

Fig. 4. Intersection areas of A and B.

Importantly, every two sensor nodes can determine a circle. There must be three circles in order to obtain a sound source position (when using three sensor nodes). However, occasionally there will be special cases (for example, Fig. 4), where we will not know whether the sound source was near the A area or the B area. In such cases, it is necessary to use at least four sensor nodes that can determine at least six circles. These circles will still not have an intersection point [28] because of ambient noise.

When there are m sensor nodes, the number of the circles is $\frac{m(m-1)}{2}$, assuming that (x_t, y_t) is the center of the tth circle, r_t is the radius of the tth circle (assuming that the tth circle

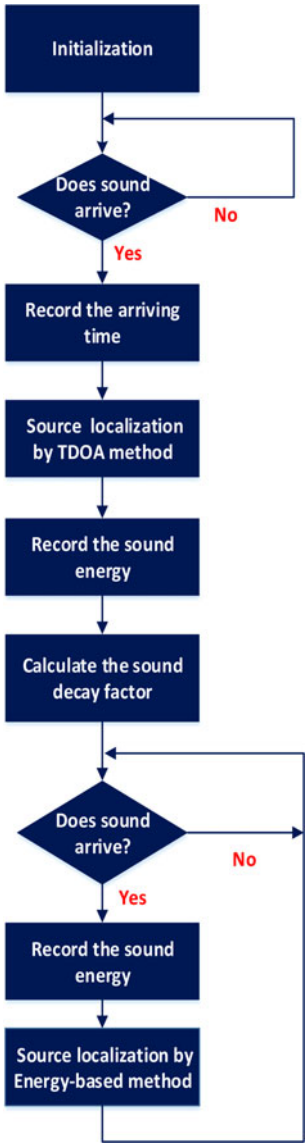

Fig. 5. Overview of the localization algorithm.

is defined by the ith and jth sensor nodes, then they can be obtained from C_{ij} and ρ_{ij} mentioned above); the sound source (x_s, y_s) can be determined when f_1 has the minimum value shown as

$$f_1 = \sum_{t=1}^{\frac{m(m-1)}{2}} |\sqrt{(x_s - x_t)^2 + (y_s - y_t)^2} - r_t|. \quad (13)$$

From (13) the sound source position can be obtained when f_1 has the minimum value. This means that all of the circles are equal. However, when ambient noise can not be ignored, the data measured by the sensor nodes closer to the sound source will be more precise, and the circles determined by these sensor nodes must have a higher weight coefficient value. So (13) can be modified as

$$f_2 = \sum_{t=1}^{\frac{m(m-1)}{2}} k_t |\sqrt{(x_s - x_t)^2 + (y_s - y_t)^2} - r_t| \quad (14)$$

TABLE III
COMPARISONS WITH OTHER METHODS

Method	AD Sampling Frequency (kHz)	Localization Times
GCC	44.1	one
MVC	51.2	one
PEOBO	102.4	one
Proposed	3	many

where k_t is the weight coefficient of the tth circle. The sound source (x_s, y_s) can be determined when f_2 has the minimum value. This can reduce the localization error by adjusting the weight coefficients (This is discussed in Part A of Section VI). An overview of the localization algorithm is depicted in Fig. 5.

Step 1: System initialization. It mainly includes sampling rate settings, communication settings, time synchronization, etc.

Step 2: Calculate the sound decay factor when sound arriving.

Step 3: Calculate the coordinates of the sound source.

V. COMPARISONS WITH OTHER METHODS

In this section we compare our method with other existing algorithms, including the method of maximizing the GCC function [19], the method of minimum variance cepstrum (MVC) [26], and the method of peak estimation of beam former output (PEOBO) [29], as shown in Table III.

Our method has two distinct advantages. First, our sampling frequency is approximately 3 kHz, a much lower sampling frequency compared to other methods. A sampling frequency of 3 kHz demonstrates strong localization accuracy, and reduces the demand for hardware (i.e., the higher the sampling frequency, the higher the hardware requirements need). Significantly, the other three methods (previously mentioned) do not work at all when the sampling frequency is near 3 kHz, whereas our method achieved sound source localization multiple times over an extended period of time at 3 kHz. Second, the other methods can only complete one localization, whereas our method can complete multiple localizations. Thus, sounds from slow-moving targets cannot be detected by the other methods.

VI. EXPERIMENTS AND SIMULATIONS

A. Simulations

In this section, we discuss how the accuracy of an energy-based source localization method is affected by the weight coefficients k_i when the SNRs have different values (SNR is measured 1 m away from the source). For this purpose, four sensors were placed in a 30-m by 30-m sensor field with the following coordinates: (0, 0), (0, 30), (30, 0), (30, 30). The location of the target was assumed to be within this sensor field. It was also assumed that the gain factor $\zeta_i = 1$ and the path loss exponent $= 2$. One thousand independent simulations were performed and the average error and standard deviation (STD) of the results (in both x and y directions) were computed. In order to better understand their impact, we devised five different modes of weight coefficients: (1, 1, 1, 1, 1, 1), (2, 1, 1, 1, 1, 1), (2, 2, 1, 1, 1, 1), (3, 2, 1, 1, 1, 1), and (4, 2, 1, 1, 1, 1). The results are displayed in Fig. 6.

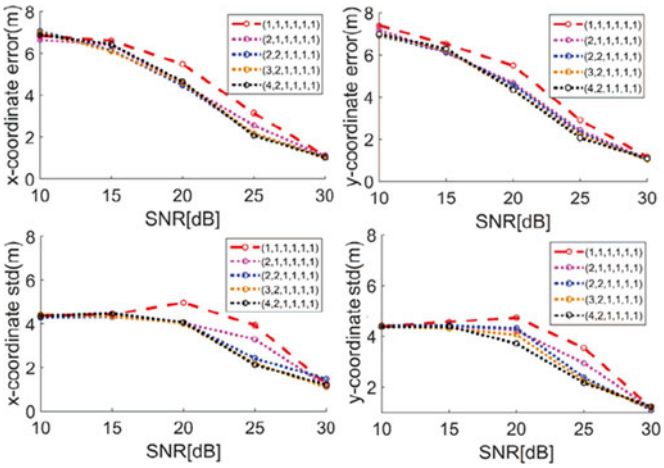

Fig. 6. Error and STD of sound source localization due to background noise.

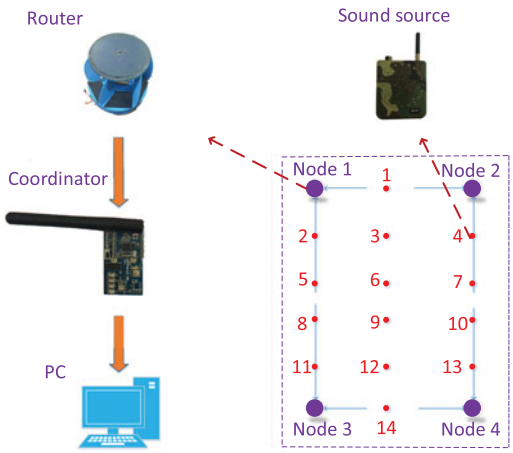

Fig. 7. Experimental environment. When the sound source emits a sound wave, the routers send the sound information to the coordinator, and the coordinator then sends the information to the PC.

TABLE IV
VALUES OF α IN THE EXPERIMENTS

Number	α							
1–8	1.14	1.09	1.05	1.43	0.95	1.31	1.36	1.00
9–16	1.30	1.23	1.48	1.25	1.33	1.48	1.60	1.43
17–24	1.90	1.46	1.44	1.22	1.35	1.34	1.15	1.77
25–32	1.18	1.46	1.49	0.97	1.39	1.14	1.86	1.59
33–40	1.50	1.85	1.60	2.16	1.38	1.54	1.08	1.44

From the results, it can be seen that adjusting the weight coefficients reduced localization errors by nearly 20% (when SNR was between 15 and 25 dB)

B. Experiments

In this section, we discuss the accuracy of the path loss exponent α calculated by our new method. We also provide an error rate comparison between the traditional method ($\alpha \approx 2$) and our new method.

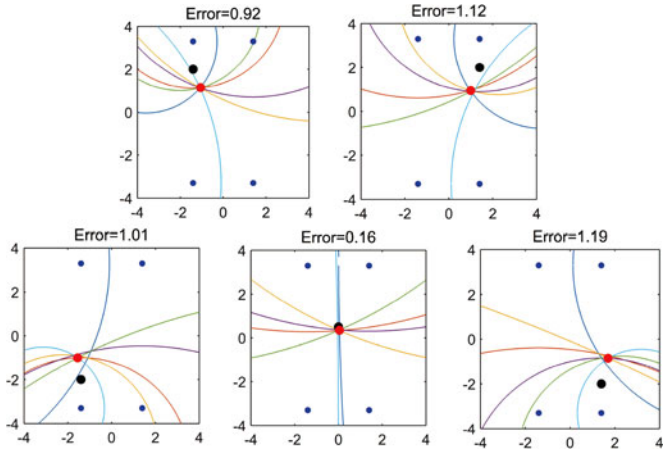

Fig. 8. Sound source localization results of the traditional method (with $\alpha = 2$) when SNR $\approx \infty$. The blue points represent the sensor nodes, the black points represent the real sound source positions, and the red points represent the localization positions.

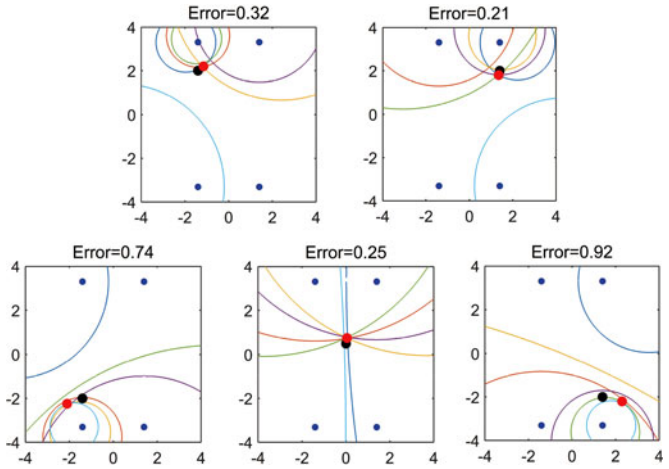

Fig. 9. Sound source localization results of the new method (with α calculated according to the environment) when SNR $\approx \infty$. The blue points represent the sensor nodes, the black points represent the real sound source positions, and the red points represent the localization positions.

At least three sensor nodes are needed to localize a sound source in a 2-D plane. We used four sensor nodes in our experiments in order to improve precision. We arranged the four nodes into a rectangle, which is the optimal arrangement of sensor nodes [30]. Naturally, the geometric arrangement of the node network can have an impact on the results of any experiment; however, this impact is inconsequential as long as the four nodes are not arranged in a straight line. In our study, we found that localization errors were larger outside of the rectangle formation than inside of the rectangle formation. Thus, using a rectangular geometric formation for networks reduces localization errors.

As shown in Fig. 7, our experiments were performed in a quiet room using five wireless sensor nodes (including one coordinator node and four router nodes). Every sensor node had a microphone array (with six sound sensors). Our acoustic energy source was a tape recorder placed in a fixed location (facing up). The AD sampling was 13 bits and the sampling frequency was

TABLE V
EXPERIMENTAL RESULTS

Number	Source Coordinate	Localization Error of Proposed Method (m)				Localization Error of Traditional Method (m)			
		SNR = ∞	SNR = 30 db	SNR = 20 db	SNR = 10 db	SNR = ∞	SNR = 30 db	SNR = 20 db	SNR = 10 db
1	(0, 3.3)	0.25	0.15	0.18	0.57	1.81	1.86	1.81	1.80
2	(−1.4, 2)	0.32	0.16	0.51	1.01	0.92	0.89	0.90	1.79
3	(0, 2)	0.45	0.25	0.41	0.76	0.92	0.86	0.86	0.78
4	(1.4, 2)	0.21	0.42	0.22	2.12	1.12	1.09	1.14	2.02
5	(−1.4, 0.5)	0.33	0.86	1.43	0.64	0.21	0.35	0.43	1.70
6	(0, 0.5)	0.25	0.36	0.32	2.85	0.16	0.11	0.30	1.17
7	(1.4, 0.5)	0.36	0.88	0.35	0.10	0.81	0.81	1.05	1.90
8	(−1.4, −0.5)	0.07	0.25	0.72	1.02	0.68	0.67	0.69	0.40
9	(0, −0.5)	0.30	1.01	0.30	0.78	0.4	0.45	0.40	0.72
10	(1.4, −0.5)	0.49	0.61	0.84	2.15	1.01	1.12	1.15	2.34
11	(−1.4, −2)	0.74	0.11	1.09	1.03	1.01	1.07	1.21	1.15
12	(0, −2)	0.51	0.27	0.61	1.99	1.30	1.27	1.21	1.46
13	(1.4, −2)	0.92	0.46	0.67	0.36	1.19	1.22	1.05	1.72
14	(0, −3.3)	0.65	0.50	0.60	0.65	2.10	2.10	2.15	1.50
	Average error	0.42	0.45	0.59	1.15	0.98	0.99	1.03	1.46

3000 Hz. Two hundred sampling points were collected and used to calculate the sound energy. Four router nodes were arranged in a rectangle 6.6 m long and 2.8 m wide. The coordinates of the four sensor nodes were (−1.4, −3.3), (−1.4, 3.3), (1.4, 3.3), and (1.4, −3.3). The sensor nodes recorded when the sounds arrived. The temperature at the time of the experiments was approximately 25 °C and the sound speed was about 346 m/s. The accuracy of the path loss exponent depended largely on two factors:

1) accuracy of the sound energy measured by the sensors;
2) measurement accuracy of the arrival times.

For the first factor, we obtained accurate sound energy values by using six sound sensors (as previously described). For the second factor, time synchronization was used every few seconds in order to improve accuracy. The time synchronization method was simple: the coordinator node sent a message to all router nodes every few seconds and the router nodes reset their timers after receiving this message. The time synchronization error was no more than 1 ms.

The values of α (the real value was about 1.23) are shown in Table IV and contain data from 40 experiments. The average value is 1.1, which is close to the real value.

We selected 14 points to be localized in the area within the rectangle. We measured each point nine times, and then calculated the average error of each localization point. Figs. 8 and 9 show localization results for when there is almost no noise when using the traditional method ($\alpha \approx 2$) and the new method (α calculated according to the environment). An error comparison for different SNRs is shown in Table V.

From the results it can be concluded that our new method reduced localization errors more than the traditional method ($\alpha \approx 2$) by nearly 60%.

C. Power Consumption Test

Power consumption is divided into three categories: AD sampling, wireless sensor communication, and hardware (hardware includes the sensor nodes and the sound sensors). As such, power consumption is reduced by adopting a low sampling rate.

TABLE VI
POWER CONSUMPTION

Each Part	Power Consumption (mW)
AD sampling	16
Communication	15
Sensor node	20
Sound sensor	3
Total	54

Thus, communications and sampling data are reduced using low-power hardware. Moreover, sensors enter a dormant state when there is no sound source (in order to conserve power consumption).

The localization algorithm mentioned above is suitable for low sampling rates and requires few sample points. However, it can significantly conserve the AD sampling's power consumption. Additionally, wireless sensor communication power consumption is minimized as a result of a reduction in the sampling data transferred between nodes. In our experimental tests, the total power consumption of the system is approximately 54 mW, which is under the low power consumption requirement (about 100 mW). The maximum power consumption of each portion is shown in Table VI.

VII. CONCLUSION

In this paper, a sound source localization method based on sound source energy was developed under low power consumption conditions. At first, several variables that may affect path loss exponent were proposed, then the method to find the appropriate path loss exponent was given. To demonstrate its effectiveness, some experiments were conducted. The results showed that the energy-based sound source localization method can determine the appropriate path loss exponent accurately and reduce the localization error by around 60% compared to the traditional method. From the simulation results, it also can be seen that in

the proposed method localization error rates are significantly reduced by adjusting the sensors' weight coefficients when ambient noise exists. Above all, the proposed method demonstrates some advantages over the traditional method including better accuracy, satisfied real-time application requirements, and lower power consumption. The future work will extend this algorithm to switched wireless sensor networks by using the effective switching technique [31].

References

[1] T. Rappaport, J. Reed, and B. Woerner, "Position location using wireless communications on highways of the future," *IEEE Commun. Mag.*, vol. 34, no. 10, pp. 33–41, Oct. 1996.

[2] H. Wymeersch, J. Lien, and M. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, vol. 97, no. 2, pp. 427–450, Feb. 2009.

[3] M. Win *et al.*, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.

[4] A. Conti, M. Guerra, D. Dardari, N. Decarli, and M. Win, "Network experimentation for cooperative localization," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 467–475, Feb. 2012.

[5] J. Wang, Q. H. Gao, Y. Yu, H. Y. Wang, and M. L. Jin, "Toward robust indoor localization based on Bayesian filter using chirp-spread-spectrum ranging," *IEEE Trans. Ind. Electron.*, vol. 59, no. 3, pp. 1622–1629, Mar. 2012.

[6] B. Wang, S. L. Zhou, W. Y. Liu, and Y. J. Mo, "Indoor localization based on curve fitting and location search using received signal strength," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 572–582, Jan. 2015.

[7] H. Song, V. Shin, and M. Jeon, "Mobile node localization using fusion prediction-based interacting multiple model in cricket sensor network," *IEEE Trans. Ind. Electron.*, vol. 59, no. 11, pp. 4349–4359, Nov. 2012.

[8] J. Fu, T. Y. Chai, C. Y. Su, and Y. Jin, "Motion/force tracking control of nonholonomic mechanical systems via combining cascaded design and backstepping," *Automatica*, vol. 49, no. 12, pp. 3282–3286, Apr. 2013.

[9] S. Argentieri, P. Danes, and P. Soueres, "A survey on sound source localization in robotics: From binaural to array processing methods," *Comput. Speech Language*, vol. 34, no. 1, pp. 87–112, Nov. 2015.

[10] F. Asano, H. Asoh, and T. Matsui, "Sound source localization and signal separation for office robot 'jijo-2'," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst.*, Nov. 1999, pp. 243–248.

[11] K. Nakamura, K. Nakadai, F. Asano, and G. Ince, "Intelligent sound source localization and its application to multimodal human tracking," in *Proc. IEEE/RSJ Intl. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 143–148.

[12] B. Kwon, Y. Park, and Y. Sik Park, "Analysis of the GCC-phat technique for multiple sources," in *Proc. IEEE Int. Conf. Control Autom. Syst.*, Oct. 2010, pp. 2070–2073.

[13] X. Wan and Z. Wu, "Sound source localization based on discrimination of cross-correlation functions," *Appl. Acoust.*, vol. 74, no. 1, pp. 28–37, Jan. 2013.

[14] C. T. Kim, T. Y. Choi, B. Choi, and J. J. Lee, "Robust estimation of sound direction for robot interface," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2008, pp. 3475–3480.

[15] A. Badali, J. M. Valin, F. Michaud, and P. Aarabi, "Evaluating real-time audio localization algorithms for artificial audition in robotics," in *Proc. IEEE/RSJ Intl. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 2033–2038.

[16] B. Hilsenbeck and N. Kirchner, "Listening for people: Exploiting the spectral structure of speech to robustly perceive the presence of people," in *Proc. IEEE/RSJ Intl. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 2903–2909.

[17] J. M. Valin, F. Michaud, and J. Rouat, "Robust 3d localization and tracking of sound sources using beamforming and particle filtering," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal*, May 2006, vol. 4, pp. IV–IV.

[18] L. Mattos and E. Grant, "Passive sonar applications: Target tracking and navigation of an autonomous robot," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2004, vol. 5, pp. 4265–4270.

[19] Y. L. Zhao, X. Chen, and B. Wang, "Real-time sound source localization using hybrid framework," *Appl. Acoust.*, vol. 74, no. 12, pp. 1367–1373, Dec. 2013.

[20] J. Wierzbicki, P. Malecki, and J. Wiciak, "Localization of the sound source with the use of the first-order ambisconic microphone," *Acta Phys. Pol. A*, vol. 123, no. 6, pp. 1114–1117, Jun. 2013.

[21] T. P. P. A. Gauthier and A. Berry, "Inverse problem with beamforming regularization matrix applied to sound source localization in closed wind-tunnel using microphone array," *J. Sound Vib.*, vol. 333, no. 25, pp. 6858–6868, Dec. 2014.

[22] Z. G. Chu, Y. Yang, and Y. S. He, "Deconvolution for three-dimensional acoustic source identification based on spherical harmonics beamforming," *J. Sound Vib.*, vol. 344, pp. 484–502, Mar. 2015.

[23] P. Bestagini, M. Compagnoni, F. Antonacci, A. Sarti, and S. Tubaro, "TDOA-based acoustic source localization in the space-range reference frame," *Multidimensional Syst. Signal Process.*, vol. 25, no. 2, pp. 337–359, Apr. 2014.

[24] Y. Wang and K. C. Ho, "TDOA source localization in the presence of synchronization clock bias and sensor position errors," *IEEE Trans. Signal Process*, vol. 61, no. 18, pp. 4532–4544, Sep. 2013.

[25] P. Julian, A. G. Andreou, L. Riddle, S. Shamma, D. H. Goldberg, and G. Cauwenberghs, "A comparative study of sound localization algorithms for energy aware sensor network nodes," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 4, pp. 640–648, Apr. 2004.

[26] C. S. Park, J. H. Jeon, and Y. H. Kim, "Localization of a sound source in a noisy environment by hyperbolic curves in quefrency domain," *J. Sound Vib.*, vol. 333, no. 21, pp. 5630–5640, Oct. 2014.

[27] D. Li and Y. H. Hu, "Energy-based collaborative source localization using acoustic microsensor array," *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 4, pp. 321–337, Mar. 2013.

[28] H. S. He, J. Lu, L. F. Wu, and X. J. Qiu, "Time delay estimation via non-mutual information among multiple microphones," *Appl. Acoust.*, vol. 74, no. 8, pp. 1033–1036, Aug. 2013.

[29] D. H. Seo, J. W. Choi, and Y. H. Kim, "Impulsive sound source localization using peak and RMS estimation of the time-domain beamformer output," *Mech. Syst. Signal Process.*, vol. 49, no. 1, pp. 95–105, Dec. 2014.

[30] Y. Chen, J. A. Francisco, and R. P. Martin, "A practical approach to landmark deployment for indoor localization," in *Proc. IEEE Conf. Sensors, Mesh Ad Hoc Commun. Netw.*, Sep. 2006, vol. 1, pp. 365–373.

[31] J. Fu, R. C. Ma, and T. Y. Chai, "Global finite-time stabilization of a class of switched nonlinear systems with the powers of positive odd rational numbers," *Automatica*, vol. 54, no. 4, pp. 360–373, Dec. 2015.

Fang Deng (M'15) received the B.E. and Ph.D. degrees in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2004 and 2009, respectively.

He is currently an Associate Professor with the School of Automation, Beijing Institute of Technology. His current research interests include nonlinear estimation, fault diagnosis, control of renewable energy resources, and wireless sensor networks.

Shengpan Guan received the B.E. degree in automation from Ocean University of China, Qingdao, China, in 2014, and the M.E. degree in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2016.

He is currently with Qingdao Topscomm Communication Co., Ltd., Qingdao. His research interests include sound source localization and wireless sensor networks.

Xianghu Yue received the B.E. degree in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2016, where he is currently working toward the M.E. degree in the School of Automation.

His current research interests include sound source localization and pattern recognition.

Xiaodan Gu received the B.E. and M.S. degrees in communication engineering and information system from PLA University of Science and Technology, Nanjing, China, in 2003 and 2006, respectively. She is currently working toward the Ph.D. degree in control science and engineering in the School of Automation, Beijing Institute of Technology, Beijing, China.

Her current research interests include information security, fault diagnosis, and sensor networks.

Jie Chen (SM'12) received the B.S., M.S., and Ph.D. degrees in control science and engineering from Beijing Institute of Technology, Beijing, China, in 1986, 1996, and 2001, respectively.

He is currently a Professor with the School of Automation, Beijing Institute of Technology. His current research interests include complex systems, multi-agent systems, multi-objective optimization and decision, constrained nonlinear control, and optimization methods.

Jianyao Lv received the B.E. degree in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2014, where he is currently working toward the M.E. degree in control science and engineering.

His current research interests include pattern recognition and intelligent signal processing.

Jiahong Li received the B.E. degree in electrical engineering and control from Beijing Institute of Technology, Beijing, China, in 2011, where he is currently working toward the Ph.D. degree under the supervision of Prof. Jie Chen in the School of Automation.

His research interests include multi-sensor data fusion, estimation of noise statistics, nonlinear state estimation, and wireless sensor networks.