Circuits Syst Signal Process (2010) 29: 837–856
DOI 10.1007/s00034-010-9187-3

Acoustic Source Localization in Wireless Sensor Network

Zoran M. Saric · Dragan D. Kukolj · Nikola D. Teslic

Received: 5 October 2008 / Revised: 20 June 2009 / Published online: 27 April 2010

Abstract In this paper, we consider the problem of acoustic source localization in a wireless sensor network based on different measured signal quantities, such as the received signal strength (RSS), the angle of arrival (AOA) and the time of arrival (TOA). For each of these quantities, an appropriate weighted least squares criterion function is developed to be used for sound source localization. The weights of each criterion function take into account the decrease in the signal-to-noise ratio (SNR) with distance from the source. In addition, RSS localization algorithm proposed in this paper provides improvement of the localization accuracy for low SNR. Finally, separate criterion functions for RSS, TOA and AOA are used together to obtain minimal localization error and maximal reliability of the acoustic source localization. Simulation analysis confirms improved performance of the proposed localization algorithm.

Keywords Angle of arrival · Received signal strength · Sensor network · Time delay of arrival · Time of arrival

Z.M. Saric (✉)
RT-SP Computer Based Systems, Novi Sad, Serbia
e-mail: zoran.saric@rt-sp.com

D.D. Kukolj · N.D. Teslic
Faculty of Engineering, The University of Novi Sad, Novi Sad, Serbia

D.D. Kukolj
e-mail: dragan.kukolj@rt-rk.com

N.D. Teslic
e-mail: nikola.teslic@rt-rk.com

BIRKHÄUSER

1 Introduction

Sound source localization has been a subject of increased research efforts during the last decades. The systems for detection, localization and tracking of the acoustic sources are widely applicable in industry, defence, robotics and security. The advantage of the objects' localization by their sound radiation is that it is in essence a passive method which needs no collaboration of the object in the localization process. Moreover, localization object is not aware that it is under observation. This is important in the defence and object protection applications. Another motivation to use localization by sound is that it can be a supplement to the other methods of environmental surveillance when visibility is not good enough, or when an object is not optically visible. A typical application is gun shot detection.

The functionality of the localization systems, which is achieved using a variety of approaches, depends on the requirements and constrains of the system at hand. The existing acoustic source localization techniques are typically based on three types of physical quantities: (1) time difference of arrival TDOA [3, 4, 18] or its variation, time of arrival (TOA) [13, 17, 27], (2) angle of arrival AOA (usually named as direction of arrival) [1, 8, 20, 24, 27], and (3) received signal energy [6, 10, 21, 26]. Received signal energy is usually termed as received signal strength (RSS).

Widely used localization methods are time of arrival (TOA) and time difference of arrival (TDOA) [7, 11, 19, 27]. Both these methods provide accurate source localization in wireless sensors networks [11, 27]. When they are used in distributed localization algorithms, there are some differences between them. In a TOA based localization system, both the signal detection and the TOA estimation are performed at each sensor node, and only the TOA estimates are sent to the central processing station. On the other hand, TDOA approach requires a pair of signals from a pair of sensor nodes to be simultaneously processed. Therefore, sensor nodes have to transmit the raw data to the processing station for TDOA estimation. This significantly increases the need for communication and energy resources. TDOA and TOA both need network synchronization.

The angle of arrival (AOA) estimates can be obtained using array processing techniques [5, 8, 10, 20, 24]. Each sensor node has to be equipped with at least two sensing devices. The AOA estimation is based on time difference between the sensing device signals or its phase difference in the case of narrow band signals. In addition, array processing techniques, such as beamforming and adaptive beamforming, enhance the signal estimate that can be useful for improvement of the RSS and TOA estimates. The drawback of the AOA approach is the device and computational complexity.

Recently, with rapid development of the systems based on wireless sensor networks, the approaches using received signal strength (RSS) have become highly attractive. In contrast to TDOA and AOA estimation, RSS is comparatively much easier and less costly to obtain from the time series recordings [6, 12, 21]. In contrast to TDOA and TOA, estimation of the RSS does not need network synchronization. Additionally, RSS approach is suitable for wireless sensor networks due to limited power and energy from the battery.

In the noiseless cases, any of the three source localization approaches, if used separately, will result in a unique solution—the position of the real sound source.

The number of disturbances, such as wind blowing and ambient noise, reduce the reliability and the accuracy of the sound source localization. To improve the accuracy and the reliability of the source localization, the three quantities, TOA, AOA and RSS, can be used simultaneously. Starting from the generalized maximum likelihood (ML) approach, we define separate likelihood functions for TOA, AOA and RSS based localization. Then, we simplify them to the weight nonlinear least squares criterion (WNLS) functions. All the three-criterion functions depend of the distance between the source and sensor nodes. By taking into account that the SNR decreases with the square of the distance from the source, it will be shown that the RSS weights decrease with the fourth power of the distance, while TOA and AOA weights decrease with the square of the distance. This is the reason why the accuracy of the RSS based localization is much worse than that of the TOA and AOA approaches for long distance sources. Starting from the ML approach for localization based on RSS, TOA and AOA, we define the general WNLS criterion function as the sum of the individual criterion functions. The improved localization accuracy of the proposed localization method is experimentally established by simulations.

Different measurement quantities RSS, TOA and AOA, when used separately, provide different localization accuracy. The best localization accuracy can be obtained by TOA, and the worst by RSS. The difficulty in TOA application is its cost, synchronization requirement, and power consumption. The RSS sensor nodes are low cost, with long battery life and easy to be installed in any place. Their lower accuracy can be improved by denser placement in the field of interest. The AOA sensing stations can be used to increase the low SNR by beamforming, but they need wider space to be installed. In this paper, we show that all the three localization methods can be combined into an integral localization system utilizing the good properties of the individual localization methods.

The rest of the paper is organized as follows. The formulation of the sound localization problem is given in Sect. 2. Sections 3, 4 and 5 deal with RSS, TOA and AOA based localization approaches, respectively. The general WNLS criterion function is given in Sect. 6. Simulation results are presented in Sect. 7. The conclusions are given in Sect. 8.

2 Estimation of the Input Features

The localization system consists of M sensor nodes ($M \geq 3$) placed in the area of interest, and the central processing station (Fig. 1). Each sensor node is equipped with a microphone array with 4 microphones placed on vertices of the regular tetrahedron. Each node locally processes acquired data to obtain estimates of the three physical quantities: TOA, AOA and RSS. These estimates are transferred to the central station where the proposed localization algorithm runs.

There is bidirectional communication between the central processing station and each sensor node (Fig. 1). The central processing node sends various commands to the sensor nodes, while the sensor nodes carry out the commands and reply its state. In regular processing mode, each node processes acoustic signals to see if there is any interesting sound event. After detection of a new acoustic event, the node classifies it

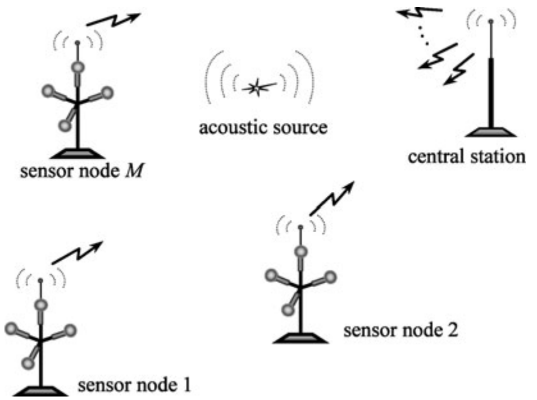


Fig. 1 Sensor nodes and central processing station in controlled area. Each sensor node is capable of estimating RSS, AOA and TOA

to some typical classes, estimates RSS, TOA and AOA quantities and sends an event alert and data to the central processing station. Central processing station uses all received data to make the decision about acoustic event by voting. If the decision is true, source position is estimated using received estimates of RSS, TOA and AOA.

Wireless sensor network design has to solve various problems such as accurate node positions, sensor node synchronization, communication protocol, sensor calibration, battery life and its exchange, and so on. All these problems are out of the scope of this paper.

As the system is aimed for the acoustic impulse detection, the first processing step is the signal detection. The assumption is that templates of the acoustic impulse events are a priori known and this knowledge has to be used for cross-correlation based detection. In the localization system based on TDOA estimates, the raw data has to be transferred from a sensor node to the central station, where TDOA estimates are obtained by cross-correlation. The alternative approach used in this paper is to estimate TOA on each sensor node using impulse detection algorithm. In this case, only the detected time has to be transferred to the processing station. This significantly saves communication consumption in comparison to the TDOA approach.

For the sake of simplification of the system, sensor nodes do not communicate with one another, but only with the central processing station. This communication is bidirectional. The processing starts when the central station sends commands to the sensor nodes to turn them to the operating mode. In this mode, nodes process acoustic signals locally. If some of them detect a typical acoustic signal, and classify it as a signal of interest, an alert signal is sent to the central station as well as the corresponding RSS, TOA and AOA estimates. Collected data are processed in the central station to obtain the position of the acoustic source. Hence, data transfer is activated only after the detection of the acoustic source of interest, and during periodical check and calibration of the system. The low communication requirement reduces energy consumption and saves the battery life.

Estimates of RSS, TOA and AOA will be denoted by $\hat{E}_i$, $\hat{T}_i$ and $\hat{\theta}_i$, respectively, where the index i denotes the ith sensor node. We will use the following assumptions:

- The ambient noise $n(t)$ at each sensor node has the same variance σ_n^2, ($\sigma_n^2 = E(n^2(t))$).

- Estimates $\hat{E}_i$, $\hat{T}_i$ and $\hat{\theta}_i$ will be modeled by

$$\hat{E}_i = E_i + \xi_{E_i}, \qquad \hat{T}_i = T_i + \xi_{T_i}, \qquad \hat{\theta}_i = \theta_i + \xi_{\theta_i}, \quad i = 1, \dots, M,$$

  where E_i, T_i and θ_i are true values, while ξ_{E_i}, ξ_{T_i}, ξ_{θ_i}, are zero mean Gaussian errors.
- The estimation errors ξ_{E_i}, ξ_{T_i}, ξ_{θ_i}, at the sensor nodes i and j are mutually uncorrelated.
- The speed of the sound in an acoustic environment is supposed to be known and equal to c ($c = 340$ m/s in air).

Estimated quantities are grouped into vectors

$$\hat{\mathbf{E}} = [\hat{E}_1, \dots, \hat{E}_M], \tag{1}$$

$$\hat{\mathbf{T}} = [\hat{T}_1, \dots, \hat{T}_M], \tag{2}$$

$$\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \dots, \hat{\theta}_M]. \tag{3}$$

The true source position in Cartesian coordinates will be denoted as $\boldsymbol{s} = [x_{1_s}, x_{2_s}]^T$. The location of the ith sensor node will be denoted by $\boldsymbol{x}_i = [x_{1_i}, x_{2_i}]^T$.

3 Received Signal Strength Approach

The localization based on the received signal strength uses the property that sound energy attenuates with the square of the distance from the source. Suppose that M sensor nodes are located at the known locations $\boldsymbol{x}_i$, $i = 1, \dots, M$, in outdoor environment. Signal at the ith sensors node is [21]

$$x_i(t) = \frac{x_s(t - \tau_i)}{d_i} + n_i(t), \qquad E\big(n_i(t)\big) = 0, \qquad E\big(n_i^2(t)\big) = \sigma_{n_i}^2, \tag{4}$$

where $x_s(t)$ is a zero mean signal with variance σ_s^2. The signal $x_s(t)$ represents an acoustic pulse K-samples wide. The value $d_i = \|\boldsymbol{x} - \boldsymbol{x}_i\|$ is the distance between the sound source and the ith sensor node. The variable $n_i(t)$ is additive white Gaussian noise (AWGN) $N(0, \sigma_n^2)$, while $\tau_i = d_i/c$ is the time delay. E is the expectation operator, and $\|\cdot\|$ denotes the vector 2-norm $\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2}$. In [6, 10, 26] as well as in Appendix A.1, it is shown that the measured acoustic energy $\hat{E}_i$ (actually estimated energy at the ith sensor node) can be modeled by

$$\hat{E}_i = \frac{E_s}{d_i^2} + \xi_{E_i}, \tag{5}$$

where E_s is the true energy of the sound source at a distance of 1 m, while ξ_{E_i} is the zero mean estimation error. The value $\hat{E}_i$ is an unbiased estimate of the signal energy at the ith node (see (A.3)). Its variance $\sigma_{E_i}^2 = E(\xi_{E_i}^2)$, (A.6), is

$$\sigma_{E_i}^2 = \frac{2\sigma_n^4}{K}\left(1 + \frac{2\sigma_s^2}{d_i^2 \sigma_n^2}\right). \tag{6}$$

The unknown energy E_s from (5) can be eliminated using two sensor nodes i and j by

$$d_i^2\hat{E}_i - d_j^2\hat{E}_j = \xi_{E_{i,j}}, \quad i = 1, \dots, M-1, \; j = i+1, \dots, M, \tag{7}$$

where $\xi_{E_{i,j}}$ is the zero mean random variable $\xi_{E_{i,j}} = d_i^2\xi_{E_i} - d_j^2\xi_{E_j}$. The variance of $\xi_{E_{i,j}}$ is $\sigma^2_{E_{i,j}} = E(\xi_{E_{i,j}}\xi_{E_{i,j}}) = \sigma^2_{E_i}d_i^4 + \sigma^2_{E_j}d_j^4$.

According to [6, 10], (7) can be transformed to

$$\|\boldsymbol{x} - c_{ij}\|^2 = \rho_{ij}^2 + \zeta_{ij}, \quad i = 1, \dots, M-1, \; j = i+1, \dots, M, \tag{8}$$

where c_{ij} is the center and ρ_{ij} the radius of the hyper-sphere on which the source locations lie

$$c_{ij} = \frac{\boldsymbol{x}_i - k_{ij}\boldsymbol{x}_j}{1 - k_{ij}^2}, \qquad \rho_{ij} = \frac{k_{ij}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{1 - k_{ij}^2}, \quad k_{ij} = \sqrt{\frac{\hat{E}_j}{\hat{E}_i}}. \tag{9}$$

The random variable ζ_{ij} is the transformed measurement noise which, according to [6], is given by $\zeta_{ij} = \frac{d_i^2\xi_{E_i} - d_j^2\xi_{E_j}}{\hat{E}_i - \hat{E}_j}$. Applying (8) to all available node pairs, an appropriate nonlinear least squares (NLS) criterion function can be defined [6, 10].

The problem with this approach is that the system of (8) becomes ill-conditioned when k_{ij} is near to unity (see (9)). To overcome this problem, we applied ML approach directly to the system of (7). For M sensor nodes, we can apply $N = \binom{M}{2}$ pairs of sensor nodes i, j to (7). The error vector is

$$\boldsymbol{\xi}_E = [\xi_{E_{1,2}} \cdots \xi_{E_{1,M}} \quad \xi_{E_{2,3}} \cdots \xi_{E_{2,M}} \cdots \xi_{E_{M-1,M}}]^T, \tag{10}$$

where ^{T} denotes transposition. We will assume that the error vector $\boldsymbol{\xi}_E$ is approximately a zero mean Gaussian process with conditional probability density function

$$f_{\xi_E|X}(\boldsymbol{\xi}_E|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{N}{2}}(\det(\mathbf{R}_E))^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\boldsymbol{\xi}_E^T\mathbf{R}_E^{-1}\boldsymbol{\xi}_E\right), \tag{11}$$

where $\mathbf{R}_E$ is its covariance matrix ($\mathbf{R}_E = E(\boldsymbol{\xi}_E\boldsymbol{\xi}_E^T)$). The elements of the matrix $\mathbf{R}_E$ are

$$E(\xi_{E_{i_1,i_2}} \cdot \xi_{E_{i_3,i_4}}) = \begin{cases} \sigma^2_{E_{i_1}}d_{i_1}^4 + \sigma^2_{E_{i_2}}d_{i_2}^4, & \text{if } i_1 = i_3 \wedge i_2 = i_4 \text{ (main diagonal)}, \\ \sigma^2_{E_{i_1}}d_{i_1}^4 \text{ or } \sigma^2_{E_{i_2}}d_{i_2}^4, & \text{if } i_1 = i_3 \wedge i_2 \neq i_4 \text{ or} \\ & \quad i_1 \neq i_3 \wedge i_2 = i_4, \\ -\sigma^2_{E_{i_1}}d_{i_1}^4 \text{ or } -\sigma^2_{E_{i_2}}d_{i_2}^4, & \text{if } i_1 = i_4 \wedge i_2 \neq i_3 \text{ or} \\ & \quad i_1 \neq i_4 \wedge i_2 = i_3, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

From (11) the log-likelihood function $L_E(\boldsymbol{x})$ is

$$L_E(\boldsymbol{x}) = -\frac{1}{2}\log\bigl(\det(\mathbf{R}_E)\bigr) - \frac{1}{2}\boldsymbol{\xi}_E^T\mathbf{R}_E^{-1}\boldsymbol{\xi}_E. \tag{13}$$

Note that $\mathbf{R}_E$ is a function of $\boldsymbol{x}$ as its elements depend on d_i (see (12)). The maximization of the criterion function $L_E(\boldsymbol{x})$ is highly nonlinear that poses high computation burden. As the term $\log(\det(\mathbf{R}_E))$ does not influence too much the estimate for sufficiently high K, we will neglect it. Detail explanation of this approximation is given in Appendix A.2. Another approximation is to take only the diagonal elements of $\mathbf{R}_E$. Obtained weighted nonlinear least squares (WNLS) criterion function $J_E(\boldsymbol{x})$ is

$$J_E(\boldsymbol{x}) = \sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\frac{\xi_{E_{i,j}}^2}{2\sigma_{E_{ij}}^2} = \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} w_{E_{i,j}}\bigl(d_i^2\hat{E}_i - d_j^2\hat{E}_j\bigr)^2, \tag{14}$$

where $w_{E_{i,j}}$ are the weights defined by $w_{E_{i,j}} = \frac{1}{2(\sigma_{E_i}^2 d_i^4 + \sigma_{E_j}^2 d_j^4)}$, $\sigma_{E_i}^2 = \frac{2d_i^2\sigma_n^4 + 4\sigma_s^2\sigma_n^2}{K d_i^2}$.

The estimate of the source location is $\hat{s} = \arg\min_x J_E(\boldsymbol{x})$.

Note the difference between the proposed algorithm and the algorithm ER-WNLS from [6], where the weights are developed under the assumption that d_i^2 can be approximated by $d_i^2 \approx E_s/\hat{E}_i$. This approximation is valid only for high SNR [6, (10)], i.e., $d_i^2 = (E_s + d_i^2\xi_{E_i})/\hat{E}_i \approx E_s/\hat{E}_i$. For low SNR and for a long distance source, it does not hold. In these cases, the system (8) of ER-WNLS algorithm [6] becomes ill-conditioned as the measurement ratio $\hat{E}_j/\hat{E}_i$ tends to unity. A more accurate model used in the proposed RSS solution and the ill-conditioned problem of the ER-WNLS algorithm are the reasons why the proposed algorithm outperforms the ER-WNLS algorithm (see Experiments 3 and 4 in Sect. 7). It is worth noting that the weight $w_{E_{i,j}}$ decreases with the fourth order of the distances d_i and d_j.

4 Time of Arrival Approach

Each sensor node locally estimates the time of arrival (TOA). These estimates will be denoted by $\hat{T}_i$, $i = 1, \ldots, M$, and used as measurements in the localization algorithm. The estimation error is assumed to be white additive noise [12] defined by

$$\hat{T}_i = T_i + \xi_{T_i}, \quad T_i = T_s + \frac{d_i}{c}, \tag{15}$$

where T_i is the true time of arrival at the ith sensor node, T_s is the instant of the acoustic event at the source location, d_i is the distance from the source to the ith sensor node, c is the speed of sound, and ξ_{T_i} is the TOA estimation error. The estimation error ξ_{T_i} is a mean zero random variable with variance $\sigma_{T_i}^2$.

The estimation of T_i is performed simultaneously with acoustic pulse detection under the assumption that the signal $s(t)$, $t = 0, \ldots, T_s$, is a priori known on the basis of the bank of the signal templates. The maximum likelihood estimation of the sound arrival time can be obtained by finding the value of τ_d that maximizes the correlation function of the received signal $x(t)$ and the template $s(t)$ as follows

$$\hat{\tau}_d = \arg\max_{\tau} r_{xs}(\tau), \quad r_{xs}(\tau) = \frac{1}{T_s}\int_{T_s} x(t)s(t-\tau)\,dt.$$

According to the general theory of ML estimators, the lower limit for $\sigma_{T_i}^2$ in the presence of an additive white Gaussian noise is given by the Cramer–Rao lower bound (CRLB) [14, 25] as $\sigma_{T_i}^2 \geq \mathrm{CRLB(TOA)} = \frac{\sigma_n^2}{2\int_{f\,\min}^{f\,\max}(2\pi f)^2|G_{S_i}(f)|^2\,df}$, where $|G_{S_i}(f)|^2$ is the bilateral power spectral density of the useful signal at the ith sensor node. The experimental measurements of TOA error variance for sound source are given in [2]. Let us define CRLB at a distance of 1 m from the sound source as

$$C_\tau = \mathrm{CRLB}_{1m}(\mathrm{TOA}) = \frac{\sigma_n^2}{2\int_{f\,\min}^{f\,\max}(2\pi f)^2|G_0(f)|^2\,df}, \tag{16}$$

where $|G_0(f)|^2$ is the power spectral density at a range of 1 m from the source. Taking into account $|G_{S_i}(f)|^2 = |G_0(f)|^2/d_i^2$, we can approximate the variance of TOA estimate at the ith sensor node by

$$\sigma_{T_i}^2 = d_i^2 \frac{\sigma_n^2}{2\int_{f\,\min}^{f\,\max}(2\pi f)^2|G_0(f)|^2\,df} = C_\tau d_i^2. \tag{17}$$

To eliminate the unknown T_s, we have to subtract $\hat{T}_i$ and $\hat{T}_j$ of the sensor nodes i and j [12] according to

$$\Delta\hat{T}_{i,j} = \hat{T}_i - \hat{T}_j = \frac{1}{c}\Delta d_{i,j}(x) + \xi_{T_{i,j}},$$

where $\Delta d_{i,j} = d_i - d_j = \|\boldsymbol{x} - \boldsymbol{x}_i\| - \|\boldsymbol{x} - \boldsymbol{x}_j\|$ is the difference of the distances from the source to the sensor nodes i and j. Value $\xi_{T_{i,j}}$ is a zero mean estimation error $\xi_{T_{i,j}} = \Delta\hat{T}_{i,j} - \frac{1}{c}\Delta d_{i,j}(x)$. If the errors ξ_{T_i} and ξ_{T_j} are independent, the variance of $\xi_{T_{i,j}}$ is $\mathrm{var}(\xi_{T_{i,j}}) = \sigma_{T_{i,j}}^2 = \sigma_{T_i}^2 + \sigma_{T_j}^2 = C_\tau(d_i^2 + d_j^2)$.

Let us define the error vector for $\binom{M}{2}$ pairs of the sensor nodes as

$$\boldsymbol{\xi}_T = [\xi_{T_{1,2}} \cdots \xi_{T_{1,M}} \quad \xi_{T_{2,3}} \cdots \xi_{T_{2,M}} \cdots \xi_{T_{M-1,M}}]^T. \tag{18}$$

We will assume that the measurement error vector $\boldsymbol{\xi}_T$ is an approximately zero mean Gaussian process with conditional probability density function

$$f_{\xi_T|X}(\boldsymbol{\xi}_T|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{N}{2}}(\det(\mathbf{R}_T))^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\boldsymbol{\xi}_T^T\mathbf{R}_T^{-1}\boldsymbol{\xi}_T\right), \tag{19}$$

where N, $N = \binom{M}{2}$, is the dimension of $\boldsymbol{\xi}_T$, and $\mathbf{R}_T$ is its covariance matrix. The elements of the matrix $\mathbf{R}_T$ are

$$E(\xi_{T_{i_1,i_2}} \cdot \xi_{T_{i_3,i_4}})$$

$$
= \begin{cases} C_\tau(d_{i_1}^2 + d_{i_2}^2), & \text{if } i_1 = i_3 \wedge i_2 = i_4 \text{ (main diagonal)}, \\ C_\tau d_{i_1}^2 \text{ or } C_\tau d_{i_2}^2, & \text{if } i_1 = i_3 \wedge i_2 \neq i_4 \text{ or } i_1 \neq i_3 \wedge i_2 = i_4, \\ -C_\tau d_{i_1}^2 \text{ or } -C_\tau d_{i_2}^2, & \text{if } i_1 = i_4 \wedge i_2 \neq i_3 \text{ or } i_1 \neq i_4 \wedge i_2 = i_3, \\ 0, & \text{otherwise.} \end{cases} \tag{20}
$$

From (19), the log-likelihood function $L_\tau(\boldsymbol{x})$ is

$$
L_T(x) = -\frac{1}{2}\log\bigl(\det(\mathbf{R}_T)\bigr) - \frac{1}{2}\boldsymbol{\xi}_T^T \mathbf{R}_T^{-1} \boldsymbol{\xi}_T. \tag{21}
$$

Note that $\mathbf{R}_T$ is a function of $\boldsymbol{x}$ as its elements depend on d_i, $d_i = \|\boldsymbol{x} - \boldsymbol{x}_i\|$ (see (20)). The maximization of the criterion function $L_T(\boldsymbol{x})$ is a highly nonlinear process that presents a high computation burden. As the term's $\log(\det(\mathbf{R}_T))$ influence becomes insignificant to the estimate for high SNR, we can neglect it. The proof for this simplification is given in Appendix A.3. Another approximation is to take into account only the diagonal elements of $\mathbf{R}_T$. The final criterion function is the weighted sum of errors,

$$
J_T(\boldsymbol{x}) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \frac{\xi_{T_{i,j}}^2}{2\sigma_{T_{i,j}}^2} = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} w_{\tau_{i,j}} \left(\Delta\hat{T}_{i,j} - \frac{1}{c}\Delta d_{i,j}(\boldsymbol{x}) \right)^2, \tag{22}
$$

where the weights are$w_{\tau_{i,j}} = \frac{1}{2C_\tau(d_i^2+d_j^2)}$. The estimate of the source location is $\hat{s} = \arg\min_x J_T(\boldsymbol{x})$. It is worth noting that the weight $w_{\tau_{i,j}}$ decreases with the second order of the distances d_i and d_j.

5 Angle of Arrival Approach

This approach assumes that each sensor node is equipped with a sensor array which is capable of estimating the angle of arrival (AOA). Measured, actually estimated, AOA can be described by [12],

$$
\hat{\theta}_i = \theta_i + \xi_{\theta_i}, \quad i = 1, \dots, M, \tag{23}
$$

where $\hat{\theta}_i$ is the measured AOA, θ_i is the true AOA at the ith sensor node, and

$$
\xi_{\theta_i} = \hat{\theta}_i - \theta_i, \tag{24}
$$

is the estimation error. The estimation error ξ_{θ_i} is usually modeled as a zero mean Gaussian random process with variance $\sigma_{\theta_i}^2$. For the circular array-case, the Cramer–Rao lower bound of the variance $\sigma_{\theta_i}^2$ [5] is

$$
\sigma_{\theta_i}^2 \geq \mathrm{CRLB}_{\mathrm{circular}}(\mathrm{AOA}) = \frac{\phi_{\mathrm{BW}}^2}{\mathrm{SNR}_{\mathrm{array}}}, \tag{25}
$$

where ϕ_{BW} is the *effective beamwidth* [5] defined by $\phi_{\mathrm{BW}} = \frac{c}{\pi\rho k_{\mathrm{nrwms}}}$, $k_{\mathrm{nrwms}} = \frac{2}{N}\sqrt{\frac{\sum_{f=1}^{N/2} f^2|S_s(f)|^2}{\sum_{f=1}^{N/2}|S_s(f)|^2}}$, $S_s(f)$ being the $N/2$ points DFT of the source signal, k_{nrwms} the *normalized root weighted mean squared source frequency*, and ρ the radius of the sensor array. The denominator in (25) is the array SNR defined by [5] $\mathrm{SNR}_{\mathrm{array}} = \frac{R}{N\sigma_n^2}\sum_{f=1}^{N/2}|S_s(f)|^2$, where R is the number of sensor devices (microphones). Taking into account that the signal strength at the ith sensor node is inversely proportional to the distance from the source, $S_s(f) \sim \frac{1}{d_i}$, the variance of ξ_{θ_i} can be approximated by

$$\sigma_{\theta_i}^2 \approx C_\theta d_i^2, \quad C_\theta = \frac{\phi_{\mathrm{BW}}^2}{\mathrm{SNR}_{\mathrm{array}}}, \tag{26}$$

where d_i is the distance between the source and the ith sensor node, and C_θ is the CRLB of AOA for the SNR at a distance of 1 m from the source. Assuming that the AOA estimates at different sensor nodes are independent, the conditional joint probability density of all the node estimates is

$$f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{M}{2}}(\prod_{i=1}^{M}\sigma_{\theta_i}^2)^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\sum_{i=1}^{M}\frac{\xi_{\theta_i}^2}{\sigma_{\theta_i}^2}\right), \tag{27}$$

where $\boldsymbol{\xi}_\theta$

$$\boldsymbol{\xi}_\theta = [\xi_{\theta_1} \quad \xi_{\theta_2} \quad \cdots \quad \xi_{\theta_M}]^T, \tag{28}$$

is the error vector. Taking into account (24) and (26), $f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x})$ is given by

$$f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x}) = \frac{1}{(2\pi C_\theta)^{\frac{M}{2}}\prod_{i=1}^{M}d_i}\exp\left(-\frac{1}{2C_\theta}\sum_{i=1}^{M}\frac{(\hat{\theta}_i - \theta_i(\boldsymbol{x}))^2}{d_i^2}\right),$$
$$\theta_i(\boldsymbol{x}) = \mathrm{arctg}\left(\frac{x_2 - x_{2i}}{x_1 - x_{1i}}\right),$$

where (x_1, x_2) are the Cartesian coordinates of the source position $\boldsymbol{x}$, and (x_{1i}, x_{2i}) are the Cartesian coordinates of the ith sensor node. If we neglect the constants that do not influence the estimate, the log-likelihood function is

$$\begin{aligned} L_\theta(\boldsymbol{x}) &= -\sum_{i=1}^{M}\log(d_i) - \frac{1}{2C_\theta}\sum_{i=1}^{M}\frac{(\hat{\theta}_i - \theta_i(\boldsymbol{x}))^2}{d_i^2} \\ &= -\sum_{i=1}^{M}\log(d_i) - \frac{\mathrm{SNR}_{\mathrm{array}}}{2\phi_{\mathrm{BW}}^2}\sum_{i=1}^{M}\frac{(\hat{\theta}_i - \theta_i(\boldsymbol{x}))^2}{d_i^2}. \end{aligned} \tag{29}$$

For sufficiently high $\mathrm{SNR}_{\mathrm{array}}$, taking into account (26), the second term becomes more important then the first one, and hence can be neglected. Thus, the log-likelihood function becomes a weighted nonlinear least squares criterion function given by

$$J_\theta(\boldsymbol{x}) = \sum_{i=1}^{M} w_{\theta i}\big(\theta_i - \theta_i(\boldsymbol{x})\big)^2, \tag{30}$$

where $w_{\theta_i} = \frac{1}{2C_\theta d_i^2}$. The estimate of the source location is $\hat{s} = \arg\min_x J_\theta(\boldsymbol{x})$. Note that the weight w_{θ_i} decreases with the square of the distance d_i.

6 Multi-objective Source Localization

In Sects. 3, 4 and 5, we used different single quantities to locate the acoustic source. In this section, we will prove that we can increase the localization accuracy by using TOA, AOA and RSS quantities together. The proof will be performed using the CRLB of the localization error. Then we will define a multi-objective localization criterion for joint TOA, AOA and RSS quantities.

We will assume that the measurements, actually the errors of the TOA, AOA and RSS estimates, are statistically independent. Under this condition, we can factorize the joint conditional probability density function $f_{\hat{E},\hat{T},\hat{\theta}|X}(\hat{E},\hat{T},\hat{\theta}|\boldsymbol{x})$ by

$$f_{\hat{E},\hat{T},\hat{\theta}|X}(\hat{E},\hat{T},\hat{\theta}|\boldsymbol{x}) = f_{\zeta|X}(\boldsymbol{\zeta}|\boldsymbol{x}) = f_{\xi_E|X}(\boldsymbol{\xi}_E|\boldsymbol{x}) f_{\xi_T|X}(\boldsymbol{\xi}_T|\boldsymbol{x}) f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x}), \tag{31}$$

where $f_{\xi_E|X}(\boldsymbol{\xi}_E|\boldsymbol{x})$, $f_{\xi_T|X}(\boldsymbol{\xi}_T|\boldsymbol{x})$, $f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x})$ are the conditional probability densities of the RSS (11), TOA (19), and AOA (27), respectively. The variables $\boldsymbol{\xi}_E$, $\boldsymbol{\xi}_T$ and $\boldsymbol{\xi}_\theta$ are the RSS, TOA and AOA error vectors respectively defined by (10), (18), (28). The vector $\boldsymbol{\zeta} = [\boldsymbol{\xi}_E^T \quad \boldsymbol{\xi}_T^T \quad \boldsymbol{\xi}_\theta^T]^T$ is the associated error vector. Taking into account (31), the CRLB for the localization error is [9, 23]

$$\begin{aligned}
&\mathrm{CRLB}(E,T,\theta) \\
&\quad = \frac{1}{E\{[\frac{\partial}{\partial x}\log f_{\zeta|X}(\boldsymbol{\zeta}|\boldsymbol{x})]^2\}^{-1}} \\
&\quad = E\left\{\frac{1}{[\frac{\partial}{\partial x}\log f_{\xi_E|X}(\boldsymbol{\xi}_E|\boldsymbol{x}) + \frac{\partial}{\partial x}\log f_{\xi_T|X}(\boldsymbol{\xi}_T|\boldsymbol{x}) + \frac{\partial}{\partial x}\log f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x})]^2}\right\}.
\end{aligned} \tag{32}$$

Each of the three terms in the denominator of (32) has a negative sign. Hence, the following inequality is valid

$$\begin{aligned}
&E\left\{\frac{1}{[\frac{\partial}{\partial x}\log f_{\xi_E|X}(\boldsymbol{\xi}_E|\boldsymbol{x}) + \frac{\partial}{\partial x}\log f_{\xi_T|X}(\boldsymbol{\xi}_T|\boldsymbol{x}) + \frac{\partial}{\partial x}\log f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x})]^2}\right\} \\
&\quad \le \min\left\{E\left[\frac{1}{[\frac{\partial}{\partial x}\log f_{\xi_E|X}(\boldsymbol{\xi}_E|\boldsymbol{x})]^2}\right], E\left[\frac{1}{[\frac{\partial}{\partial x}\log f_{\xi_T|X}(\boldsymbol{\xi}_T|\boldsymbol{x})]^2}\right],\right. \\
&\qquad\qquad \left. E\left[\frac{1}{[\frac{\partial}{\partial x}\log f_{\xi_\theta|X}(\boldsymbol{\xi}_\theta|\boldsymbol{x})]^2}\right]\right\}.
\end{aligned} \tag{33}$$

From (32) and (33), the inequality

$$\mathrm{CRLB}(E,T,\theta) \leq \min\{\mathrm{CRLB}(E), \mathrm{CRLB}(T), \mathrm{CRLB}(\theta)\} \tag{34}$$

follows. From (34), we conclude that TOA, AOA and RSS used together provide a lower localization error than each of these quantities when used separately. The contribution of each quantity E, T and θ to the localization accuracy depends on its CRLB. If one of them has a much higher CRLB then the other two, its contribution to the localization accuracy is negligible and may be excluded from the criterion function. Now, we can develop a multi-objective criterion function for joint TOA, AOA and RSS based localization. Let us define the vector $\hat{\mathbf{Y}}$ of the input estimates by $\hat{\mathbf{Y}} = [\hat{\mathbf{E}}^T \quad \hat{\mathbf{T}}^T \quad \hat{\theta}^T]^T$, where $\hat{\mathbf{E}}$, $\hat{\mathbf{T}}$, $\hat{\theta}$ are estimates of the corresponding quantities defined by (1), (2) and (3), respectively. Using (31) and (13), (21), (29), the joined log-likelihood function is

$$\begin{aligned} L(\boldsymbol{x}) &= \log f_{\zeta|X}(\boldsymbol{\zeta}|\boldsymbol{x}) \\ &= -\frac{1}{2}\log\bigl(\det(\mathbf{R}_E)\bigr) - \frac{1}{2}\boldsymbol{\xi}_E^T\mathbf{R}_E^{-1}\boldsymbol{\xi}_E - \frac{1}{2}\log\bigl(\det(\mathbf{R}_T)\bigr) - \frac{1}{2}\boldsymbol{\xi}_T^T\mathbf{R}_T^{-1}\boldsymbol{\xi}_T \\ &\quad - \sum_{i=1}^{M}\log(d_i) - \frac{1}{2C_\theta}\sum_{i=1}^{M}\frac{(\hat{\theta}_i - \theta_i(\boldsymbol{x}))^2}{d_i^2}. \end{aligned}$$

As $L(\boldsymbol{x})$ is rather complicated to minimize, we will define the weighted sum criterion function $J(\boldsymbol{x})$ using (14), (22) and (30) by $J_{\mathrm{CF}}(\boldsymbol{x}) = J_E(\boldsymbol{x}) + J_T(\boldsymbol{x}) + J_\theta(\boldsymbol{x})$. The estimate of the source location is $\hat{s} = \arg\min_x J_{\mathrm{CF}}(x)$.

The criterion function $J_{\mathrm{CF}}(\boldsymbol{x})$ can be minimized by various optimization methods such as the Nelder–Mead (simplex) direct search (DS) [15], exhaustive search (ES), multi-resolution (MR) search, and gradient descent (GD) search algorithm. Another and most promising optimization method is the particle swarm optimization algorithm [16]. The benefit of this algorithm is its stability and low computational burden as it does not need gradient calculation.

Note that RSS, TOA and AOA do not equally contribute to the localization accuracy. It was shown that the TOA and AOA weights $w_{\tau_{i,j}}$, w_{θ_i} decrease with the square of the distance (see (22), (30)), while the RSS weight $w_{E_{i,j}}$ reduces with the fourth order of the distances d_i and d_j. Secondly, RSS estimates are more sensitive to the obstacles, multi-path propagation or wind flow. To overcome this problem, we can use additional sensor nodes to measure only RSS. The denser placement of RSS sensor unit-nodes will considerably increase the accuracy of the localization system. The modular structure of the criterion function $J_{\mathrm{CF}}(\boldsymbol{x})$ allows us to use different number of RSS, TOA and AOA measurements. From a practical point of view, these additional nodes will not significantly increase the price of the system because RSS sensors are cheap and have low power consumption.

7 Simulation Results

The performance of the proposed sound localization approach was tested by simulation analysis realized by Matlab/Simulink tools. The signal model is defined by (4),

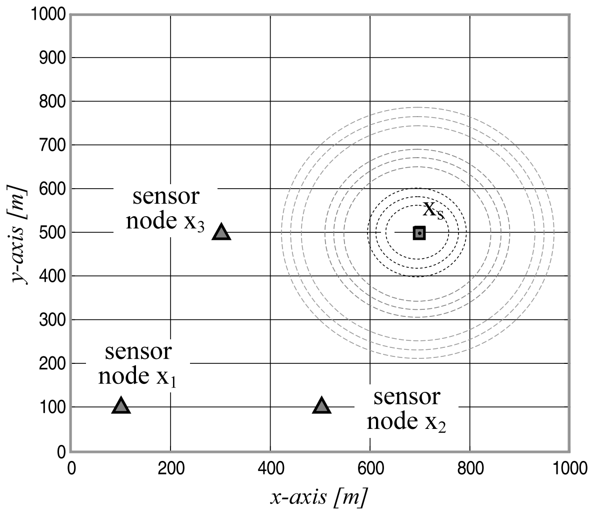


Fig. 2 The configuration of the sensor nodes x_1, x_2, x_3 and sound source x_s in the first experiment

(15) and (23). The purpose of the first experiment was to analyze root mean square localization errors (RMSE) in terms of the signal-to-noise ratio for the following estimation approaches:

(1) Approach based on RSS;
(2) Approach based on AOA;
(3) Approach based on TOA;
(4) The combination of the RSS, AOA and TOA approaches denoted as 3 CF.

Three sensor nodes were used for localization of the source. The positions of the source and sensor nodes in a 2D area are shown in Fig. 2. An acoustic source x_s was emitting a sequence of acoustic impulses of $T_s = 6.3$ ms in width. The signal-to-noise ratio at a 1 m distance was 62, 68, 74, 80, and 86 dB. 100 measurement sets were generated for each particular value of SNR. The measurements of RSS, AOA and TOA were generated by (4), (15) and (23), respectively. The measurement errors ξ_{E_i}, $\xi_{T_{i,j}}$ and ξ_{θ_i} were generated using a random number generator with Gaussian distributions $N(0,\sigma_{E_i}^2)$, $N(0,\sigma_{T_i}^2)$ and $N(0,\sigma_{\theta_i}^2)$, respectively. Their variances, $\sigma_{E_i}^2$, $\sigma_{T_i}^2$ and $\sigma_{\theta_i}^2$, were calculated by (6), (17) and (25), respectively. In (6), we used $K = 50$.

The calculation of C_τ in (16) was based on the assumption that the signal is bandlimited around $f_c = 600$ Hz. Under this assumption, (16) was simplified [17, 27] to

$$C_\tau = \frac{1}{8\pi^2 T_s f_c^2 \mathrm{SNR}}.$$

The variance $\sigma_{\theta_i}^2$ of the AOA estimate was calculated assuming $R = 3$ sensing devices in a circular array (26). The forth microphone at the top of the tetrahedron was inactive because the localization was in the 2D area. The radius ρ of the array was 0.3 m . The criterion functions $J_E(\boldsymbol{x})$, $J_\tau(\boldsymbol{x})$, $J_\theta(\boldsymbol{x})$ and $J_{\mathrm{CF}}(\boldsymbol{x})$ were optimized by gradient descent search algorithm.

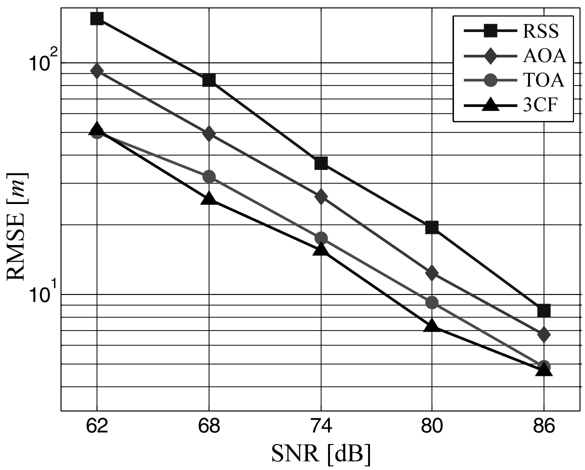


Fig. 3 The localization RMSE in term of SNR. The localization is based on the objective functions: RSS, AOA, TOA and their combination (3CF). The reference value of SNR was measured at a distance of 1 m from the source

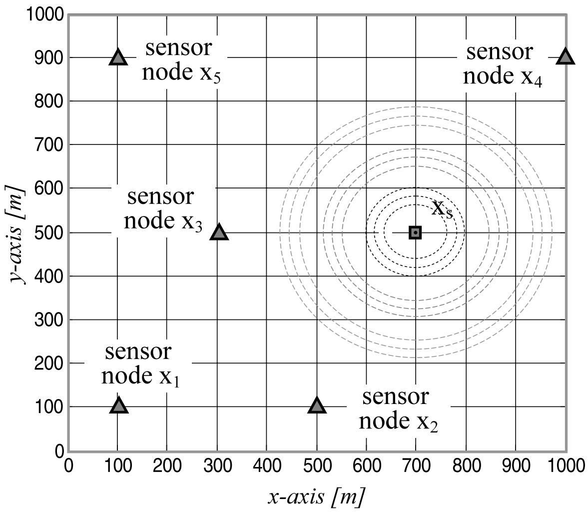


Fig. 4 Configuration of the sensor nodes $x_1, \dots, x_5$ and sound source x_s in the second experiment. Sensor nodes x_1, x_2 and x_3 measure TOA and RSS, while sensor nodes x_4 and x_5 measure only RSS

Localization errors for each approach (RSS, AOA, TOA and 3CF) are depicted in Fig. 3. As expected, all the root mean square errors linearly decay with increasing SNR. The error is maximal for the localization based on RSS. AOA and TOA provide lower RMSE than RSS. The error obtained by the combination of the RSS, AOA and TOA criterion functions is the lowest one. Although the difference in performance of TOA and 3CF is small in this example, the performance of 3CF is better when RSS, TOA and DOA errors are close to one another. In addition, some unpredictable disturbances like reflections may increase the TOA localization error. In these circumstances, 3CF provides a considerable improvement both in the accuracy and the reliability of the localization.

In the second experiment, we want to show that the additional sensor nodes, capable of measuring only RSS, can considerably increase the accuracy of the localization system. In this experiment, two additional sensor nodes were placed at positions x_4

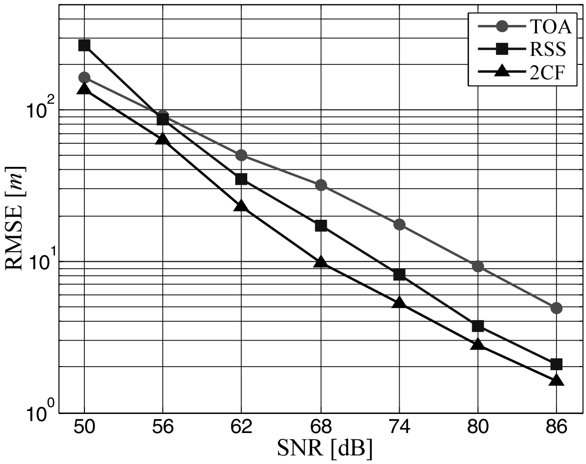


Fig. 5 The localization RMSE in terms of SNR for the scenario depicted in Fig. 4. The localization is based on the objective functions: RSS, TOA and their combination (2CF). The reference value of SNR was measured at a distance of 1 m from the source

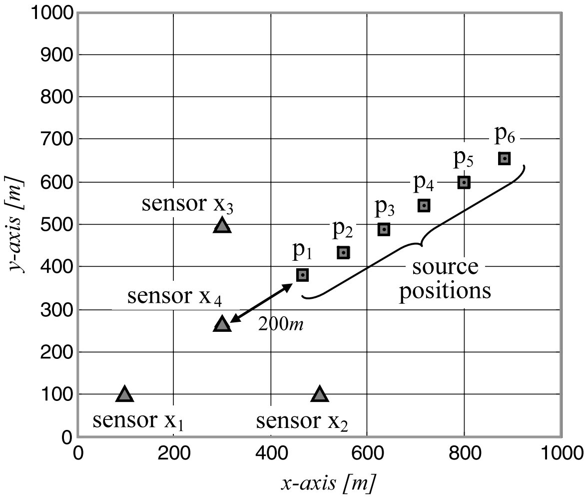


Fig. 6 The source and the sensors positions in the second experiment

(1000, 900) and x_5 (100, 900) (Fig. 4). The sensor nodes x_1, x_2 and x_3 are capable of measuring both TOA and RSS, while the sensor nodes x_4 and x_5 are capable of measuring only RSS. The localization errors for each approach, RSS with nodes $x_1, \dots, x_5$, TOA with nodes $x_1, \dots, x_3$, and the combination of them 2CF are depicted in Fig. 5. RSS with five sensors provides lower localization error than TOA with three sensors except for 50 dB. The combination of the RSS and TOA criterion functions 2CF provides the lowest localization error.

In the third experiment, only RSS was used for source localization. Three different RSS based algorithms were compared:

(1) Energy ratio based nonlinear least squares algorithm (ER-NLS) [10, 22];
(2) Energy ratio based weighted nonlinear least squares algorithm (ER-WNLS) [6];
(3) Weighted nonlinear least squares algorithm (WNLS) proposed in this paper.

Fig. 7 The localization RMSE in terms of distance between the source and sensor x_4. The reference value of SNR was measured at a distance of 1 m from the source

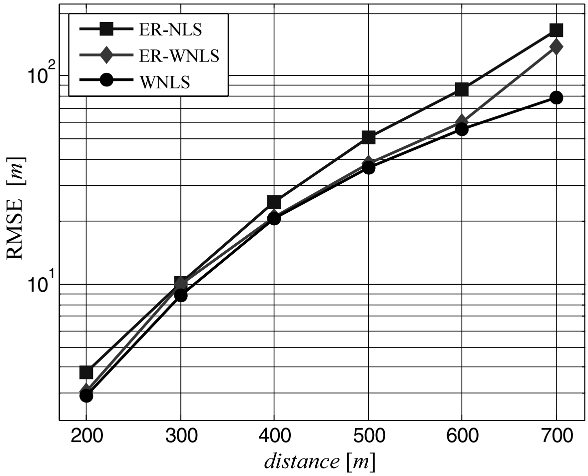
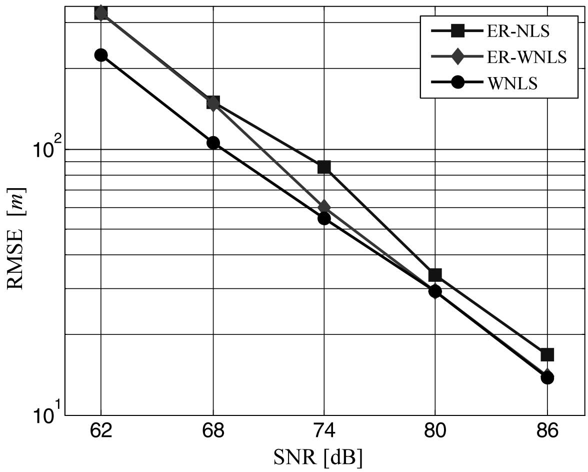


Fig. 8 The localization RMSE in terms of the input SNR. The reference value of SNR was measured at a distance of 1 m from the source



The positions of the acoustic source and three sensor nodes are shown in Fig. 6. Node positions are marked by x_1, x_2 and x_3. The point x_4 is the reference point for source distance measuring. The source position was varied from p_1 to p_6. The distance between the source and the sensor x_4 took values 200, 300, 400, 500, 600, and 700 m. The signal-to-noise ratio was 74 dB at the reference distance of 1 m from the source. At each source position, 100 independent RSS measurements were generated. Criterion functions were optimized by multi-resolution search algorithm. Root mean square errors of the source localization for algorithms ER-NLS, ER-WNLS and WNLS are shown in Fig. 7. It can be seen that the algorithms ER-WNLS and WNLS outperform the ER-NLS algorithm. Similar relation between ER-NLS and ER-WNLS was obtained in [6]. The algorithms WNLS and ER-WNLS have similar performance, but WNLS is a little better for distances 200, 300, 400, 500, and 600 m. For the distance of 700 m, WNLS is much better than ER-WNLS. The reason for this is that System (8) of ER-WNLS algorithm becomes ill-conditioned when measurement ratios $\hat{E}_j/\hat{E}_i$ for some of the sensor pairs are close to unity. An Ill-conditioned problem is particularly pronounced for low SNR and long distance sources.

In the last experiment, source position was fixed to $p_5 = (800, 600)$, while SNR was varied from 62 to 86 dB. The root mean square errors are shown in Fig. 8. The conclusions about the performance of the algorithms are similar to those of the previous experiment. The algorithms ER-WNLS and WNLS outperform the ER-NLS algorithm, while the performance of the proposed WNLS is better than the ER-WNLS algorithm at low SNR.

8 Conclusion

In this paper, the quantities RSS, TOA and AOA are used for the acoustic source localization. Localization algorithms based on single physical values RSS, TOA and AOA were derived from the general maximum likelihood estimation approach. To reduce the computational burden, likelihood functions were approximated by weighted least squares criterion functions. Weights in criterion functions take into account the decay of the sound power with distance from the acoustic source and the decrease of the SNR at sensors positions. Finally, the criterion functions for RSS, TOA and AOA are used together to obtain minimal localization error and maximal reliability of the sound localization.

The well known geometrical approach for RSS based source localization [6] is improved by applying maximum likelihood criterion to the original model of noisy measurements. This criterion function is based on the more accurate signal model. The performance of the proposed RSS algorithm is improved particularly for low SNR and for the long distance case. There are two reasons for the better performance of the proposed algorithm. Firstly, the proposed RSS algorithm is well conditioned, while the geometrical approach is ill-conditioned when an energy ratio for a sensor pair tends to unity. Secondly, the weights in the ML criterion function of the proposed approach provide better approximation of this function at low SNR, as stated at the end of Sect. 3. Simulations have shown good performance of the proposed algorithm.

Although the proposed localization system is aimed at the localization of the impulse acoustic sources, it can be also used for the localization of other acoustic sources, such as vehicle tracking. The potential applications of the proposed sound localization system are environmental monitoring, security support, mobile robot navigation, battlefield surveillance, and so on.

Acknowledgements This work was partially supported by Ministry of Science and Technology Development of the Republic Serbia under Grant number OI-161003 and by Hungarian Development Agency under the TÁMOP-4.2.2/08/1/2008-0008 program.

Appendix

A.1 Energy Model

Suppose the source emits a zero mean random signal $x_s(t)$ during the time interval of length T. If the sampling frequency is f_s, the signal will be represented by $K = Tf_s$

samples. The measurements of $x_s(t)$ correspond to the reference distance of 1 m from the source. Under the assumption that the noise is additive, the signal at the ith sensor node is

$$\tilde{x}_i(t) = x_i(t) + n_i(t), \quad x_i(t) = \frac{x_s(t-\tau_i)}{d_i}, \tag{A.1}$$

where d_i is the distance between the source and the ith sensor node, $x_i(t)$ is the delayed and attenuated signal at the ith sensor node, $n_i(t)$ is white Gaussian noise $N(0, \sigma_n^2)$, and τ_i, $\tau_i = d_i/c$, is the time delay. The signal energy at the ith sensor node is modeled by

$$\tilde{E}_i(t) = \frac{E_s(t-\tau_i)}{d_i^2} + \xi_{E_i}(t), \tag{A.2}$$

where $\tilde{E}_i(t)$ is the K-samples based energy estimate defined by

$$\tilde{E}_i(t) = \frac{1}{K} \sum_{l=t-K/2+1}^{t+K/2} \tilde{x}_i^2(l) - \sigma_n^2, \tag{A.3}$$

and $E_s(t-\tau_i)$ is the delayed signal energy

$$E_s(t-\tau_i) = \frac{1}{K} \sum_{l=t-\tau_i-K/2+1}^{t-\tau_i+K/2} x_s^2(l). \tag{A.4}$$

The estimation error $\xi_{E_i}(t)$ is

$$\xi_{E_i}(t) = \tilde{E}_i(t) - \frac{E_s(t-\tau_i)}{d_i^2} = \frac{1}{K} \sum_{l=t-K/2+1}^{t+K/2} \left(\left(n^2(l) - \sigma_n^2\right) + 2x_i(l)n(l)\right). \tag{A.5}$$

Under the assumption that $x_s(t)$ and $n_i(t)$ are uncorrelated, the estimate $\tilde{E}_i(t)$ is unbiased as $E(\xi_{E_i}(t)) = 0$. From (A.5), the variance of $\xi_{E_i}(t)$ is

$$\sigma_{E_i}^2 = E\left(\xi_{E_i}^2(t)\right) = \frac{2\sigma_n^4}{K}\left(1 + \frac{2\sigma_s^2}{d_i^2\sigma_n^2}\right), \tag{A.6}$$

where σ_s^2, $\sigma_s^2 = E(x_s^2(t))$, is the variance of the signal during source activity.

A.2 The Influence of $\log(\det(\mathbf{R}_E))$ Term on Estimation of the Acoustic Source Position by Signal Strength

Taking into account (6) and (12), the covariance matrix $\mathbf{R}_E$ for the signal strength estimates based on K samples is

$$\mathbf{R}_E = \frac{1}{K^N}\mathbf{R}_{E1}, \tag{A.7}$$

where $\mathbf{R}_{E1}$ is the covariance matrix for the signal strength estimated by single sample, and N is the order of the matrix $\mathbf{R}_{E1}$. Substituting (A.7) into (13), the log-likelihood function becomes

$$L_E(\boldsymbol{x}) = \frac{1}{2}\log\bigl(K^N\bigr) - \frac{1}{2}\log\bigl(\det(\mathbf{R}_{E1})\bigr) - \frac{1}{2}K^N\boldsymbol{\xi}_E^T\mathbf{R}_{E1}^{-1}\boldsymbol{\xi}_E. \tag{A.8}$$

The first term of (A.8) is constant and does not influence the estimate of the source position. The second term does not depend on K. The third term increases with K according to the factor K^N. Then, from (A.8) we can conclude that the second term becomes negligible compared to the third term and can be neglected for sufficiently large K.

A.3 The Influence of $\log(\det(\mathbf{R}_T))$ Term on Estimation of the Acoustic Source Position by TOA

From (16), (17) and (20), the covariance matrix $\mathbf{R}_T$ is proportional with CRLB constant C_τ

$$\mathbf{R}_T = C_\tau^N\bar{\mathbf{R}}_T, \tag{A.9}$$

where $\bar{\mathbf{R}}_T$ is a conditional covariance matrix for the unit value of the CRLB, and N is the order of the matrix $\bar{\mathbf{R}}_T$. Taking into account (A.9) and (21), the log-likelihood function is

$$L_T(\boldsymbol{x}) = -\frac{1}{2}\log\bigl(C_\tau^N\bigr) - \frac{1}{2}\log\bigl(\det(\bar{\mathbf{R}}_T)\bigr) - \frac{1}{2C_\tau^N}\boldsymbol{\xi}_T^T\bar{\mathbf{R}}_T^{-1}\boldsymbol{\xi}_T. \tag{A.10}$$

The first term does not depend on the source position and can be ignored. For high SNR, the noise variance σ_n^2 is small compared to the signal power, and CRLB factor C_τ reduces according to (17). The third term of (A.10) becomes more significant than the second term, which can be neglected. Hence, the likelihood function can be approximated by the third term

$$L_T(\boldsymbol{x}) \approx -\frac{1}{2}\boldsymbol{\xi}_T^T\mathbf{R}_T^{-1}\boldsymbol{\xi}_T. \tag{A.11}$$

References

1. P. Aarabi, The fusion of distributed microphone arrays for sound localization. EURASIP J. Appl. Sign. Process., Spec. Issue Sens. Netw. **4**, 338–347 (2003)
2. J.N. Asha, R.L. Mosesb, Acoustic time delay estimation and sensor network self-localization: experimental results. J. Acoust. Soc. Am. **118**(2), 841–850 (2005)
3. M.S. Brandstein, J.E. Adcock, H.F. Silverman, A localization-error-based method for microphone array design, in *Proc. IEEE Int. Conf. Acoustic., Speech, Signal Processing*, Atlanta, GA, (1996), pp. 901–904
4. M.S. Brandstein, J.E. Adcock, H.F. Silverman, A closed-form location estimation for use with room environment microphone array. IEEE Trans. Speech Audio Process. **5**, 45–50 (1997)
5. J.C. Chen, K. Yao, R.E. Hudson, Acoustic source localization and beamforming: theory and practice. EURASIP J. Appl. Sign. Process., Spec. Issue Sens. Netw. **4**, 359–370 (2003)

6. K. Deng, Z. Liu, Weighted least-squares solutions for energy-based collaborative source localization using acoustic array. IJCSNS Int. J. Comput. Sci. Netw. Secur. **7**(1), 159–165 (2007)
7. J.H. DiBiase, H.F. Silverman, M.S. Brandstein, Robust localization in reverberant room, in *Microphone Arrays*, ed. by M.S. Brandstein, D. Ward (Springer, New York, 2001), pp. 157–180
8. N. Iwakiri, T. Kobayashi, Ultra-wideband time-of-arrival and angle-of-arrival estimation using transformation between frequency and time domain signals. J. Commun. **3**(1), 12–19 (2008)
9. B. Kovačević, Ž. Đurović, *Fundamentals of Stochastic Signals, Systems and Estimation Theory with Worked Examples* (Academic Mind, Belgrade, 1999), ISBN 86-7466-007-x
10. D. Li, Y.H. Hu, Energy-based collaborative source localization using acoustic microsensor array. EURASIP J. Appl. Sign. Process., Spec. Issue Sens. Netw. **4**, 321–337 (2003)
11. L. Mailaender, On the geolocation bounds for round-trip time-of-arrival and all non-line-of-sight channels. EURASIP J. Adv. Sign. Process. **2008**, 1–10
12. G. Mao, B. Fidan, B.D.O. Anderson, Wireless sensor network localization techniques. Comput. Netw. **51**, 2529–2553 (2007)
13. R.L. Moses, D. Krishnamurthy, R.M. Patterson, A self-localization method for wireless sensor networks. EURASIP J. Appl. Sign. Process., Spec. Issue Sens. Netw. **4**, 348–358 (2003)
14. L.D. Nardis, M.-G.D. Benedetto, Joint communication, ranging, and positioning in low data-rate UWB networks, in *Joint 2nd Workshop on Positioning, Navigation and Communication 2005, (WPNC'05) & 1st Ultra-Wideband Expert Talk (UET'05)*, March 17 2005, Hannover, Germany, pp. 191–200
15. J.A. Nelder, R. Mead, A simplex method for function minimization. Comput. J. **7**, 308–313 (1965)
16. M.M. Noel, P.P. Joshi, T.C. Jannett, Improved maximum likelihood estimation of target position in wireless sensor networks using particle swarm optimization, in *Proceedings of the Third International Conference on Information Technology: New Generations (ITNG'06)* (IEEE Computer Society, Los Alamitos, 2006), pp. 274–279
17. N. Patwari, J.N. Ash, S. Kyperountas, A.O. Hero III, R.L. Moses, N.S. Correal, Locating the nodes. IEEE Sig. Process. Mag. 54–69 (2005)
18. C.W. Reed, R.E. Hudson, K. Yao, Direct joint source localization and propagation speed estimation, in *Proc. IEEE Int. Conf. Acoustic., Speech, Signal Processing*, Phoenix, AZ (1999), pp. 1169–1172
19. R.A. Saeed, S. Khatun, B.M. Ali, M.A. Khazani, Performance of ultra-wideband time-of-arrival estimation enhanced with synchronization scheme. ECTI Trans. Electr. Eng. Electron. Commun. **4**(1), 78–84 (2006)
20. R.O. Schmidt, Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**(3), 276–280 (1986)
21. X. Sheng, Y.-H. Hu, Energy based acoustic source localization, in *Information Processing in Sensor Networks*, ed. by F. Zhao, L. Guibas (Springer, Berlin, 2003), pp. 285–300
22. X. Sheng, Y.H. Hu, Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. IEEE Trans. Sign. Process. **53**(1), 44–53 (2005)
23. H.L.V. Trees, *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York, 2002)
24. H.L.V. Trees, *Optimum Array Processing—Part IV* (Wiley, New York, 2002)
25. H. Urkowitz, *Signal Theory and Random Processes* (Artech House, Norwood, 1983)
26. X. Wang, D. Bi, L. Ding, S. Wang, Agent collaborative target localization and classification in wireless sensor networks. Sensors **7**, 1359–1386 (2007)
27. K. Yao, F. Lorenzelli, Localization in sensor networks. ST J. Res. **4**(1), 80–95 (2007). Wireless sensor networks, University of California, Los Angeles