Maximum Likelihood Wireless Sensor Network Source Localization Using Acoustic Signal Energy Measurements

Xiaohong Sheng and Yu-Hen Hu *Fellow*
Department of Electrical and Computer Engineering
University of Wisconsin - Madison, WI 53706, USA,
Email: sheng@ece.wisc.edu, hu@engr.wisc.edu

Abstract

A maximum likelihood (*ML*) acoustic source location estimation method is presented. This method uses acoustic signal energy measurements taken at individual sensors of an ad hoc wireless sensor network to estimate the locations of multiple acoustic sources. Compared to existing acoustic energy based source localization methods, this proposed *ML* method delivers more accurate results and offers the enhanced capability of multiple source localization. A multi-resolution search algorithm and an expectation- maximization (*EM*) like iterative algorithm are proposed to expediate the computation of source locations. The Cramer-Rao Bound (*CRB*) of the *ML* source location estimate has been derived. When there is only a single source in the sensor field, the corresponding *CRB* formulation can be used to analyze the impacts of sensor placement to the accuracy of location estimates. Extensive simulations have been conducted. Empirically, it is observed that this proposed *ML* method consistently outperforms existing acoustic energy based source localization methods. An example applying this method to track military vehicles using real world experiment data also demonstrates the performance advantage of this proposed method over a previously proposed acoustic energy source localization method.

Index Terms

Source Localization, Sensor Network, Acoustic Energy, Acoustic Intensity, Maximum Likelihood, Nonlinear Least Square

This project is supported by DARPA under grant no. F 30602-00-2-0555

I. Introduction

The emergence of miniature low-power devices that integrate micro-sensing and actuation with on-board processing and wireless communication capabilities has stimulated great interests in developing wireless ad hoc sensor network [1], [2]. In an ad hoc network, the sensors do not form a structured formation but often are deployed randomly. Here we assume that the senosr locations are known in advance.

A wireless ad hoc sensor network often performs monitoring tasks such as detection, classification, localization and tracking of one or more targets in the sensor field. The sensors are typically battery-powered and have limited wireless communication bandwidth. Therefore, efficient collaborative signal processing algorithms that consume less energy for computation and less bandwidth for communication are needed [3].

In this paper, we focus on the task of source location estimation using passive and stationary acoustic sensors (microphones) in a wireless ad hoc sensor network. Source localization using acoustic sensors has found numerous applications. In sonar signal processing, the focus is on locating under-water acoustic sources using an array of hydrophones [4]. In video conference and multimedia human computer interface applications, microphone arrays have been developed to locate and track speakers head position in a room environment [6], [7], [8], [9], [10],[11], [12]. Acoustic signatures have also been used to estimate vehicle locations in an open-field sensor network [13], [14].

Existing acoustic source localization methods make use of three types of physical measurements: time delay of arrival (TDOA), direction of arrival (DOA) and source signal strength or energy. DOA can be estimated by exploiting the phase difference measured at receiving sensors [15], [16], [17], [18],[19] and is applicable when the acoustic source emits a coherent, narrow band signal. TDOA is suitable for broadband acoustic source localization and has been extensively investigated [5], [6] [13], [14], [20], [21]. It requires accurate measurements of the relative time delay between sensor nodes. It is known that the intensity or equivalently the energy of acoustic signal attenuates as a function of distance from the source. Using this property, recently, an energy-based acoustic source localization method has been reported [22] for locating single target in an open-terrain acoustic sensor field.

For wireless sensor network applications, energy (intensity) based acoustic features is an appropriate choice since the acoustic power emitted by targets such as moving vehicles usually varies slowly with respect to time. As such, the acoustic energy time series can be sampled at a much lower rate compared to the raw acoustic time series. Furthermore, the positions of ground moving targets need not be up-

dated too often. Therefore, little data will need to be transmitted to a data fusion center via the often congested wireless communication channels. This, in terms, will reduce the energy consumption for data transmission on individual sensor nodes, and will reduce the demand of communication bandwidth over shared wireless channels. Moreover, acoustic signal intensity measurements can also be used to detect the presence of a target. Hence there is no need to compute additional features for the task of source localization.

The method to be presented in this paper is based on a maximum likelihood estimation of both the source locations and corresponding acoustic energy readings. Compared to a previously reported method [22], this proposed method promises two significant advantages:

- The *ML* method can handle more than one targets within a sensor field. The previously reported method can handle only single target localization.
- The *ML* method yields higher accuracy in terms of source location estimates compared to the earlier methods.

In this paper, an acoustic signal energy attenuation model as a function of source-to-sensor distance is first established. Based on this model, the acoustic source localization problem is formulated as a maximum likelihood estimation problem. Two complementary methods then are proposed to solve this nonlinear optimization problem. The first method is based on a projection formulation and uses multi-resolution search to expediate computation. The second method is based on an Expectation-Maximization (*EM*) like iterative algorithm. This method has less computation complexity but may converge to a local minimum. Next, a formulation of the *Cramer-Rao Bound (CRB)* of the variance of the location estimates is derived. When there is only a single target, the corresponding *CRB* can be used to analyze the impact of sensor deployment strategy on the localization accuracy. Specifically, the formula clearly verified the conventional belief that smaller estimation error is achieved if the sensors are densely and uniformly laid out over the sensor field.

The remaining of this paper is organized as follows: In section II, the acoustic energy attenuation model is developed. In section III, the *ML* source location estimation method is derived. Both the projection and the *EM* solution methods are also presented. In section IV, the *CRB* is derived. In particular, in the single source case, explicit formulation of *CRB* is derived. This expression allows one to explore the impacts of different sensor deployment strategies to the localization accuracy. Experiments and simulations are provided in section V. Extensive simulation have been conducted to compare the performance of the proposed *ML* method to three existing energy based source localization methods. Then, time series data

samples obtained from a field experiment conducted in California during November 2001 are used to demonstrate the feasibility of using this proposed method to solve real world sensor network source localization problem.

II. Acoustic Energy Attenuation Model

Let N be the number of acoustic sensors, and K be the number of acoustic sources in the sensor field. In theory [23], the intensity of an acoustic signal emitted omni-directionally from a point sound source and propagating through ground surface will attenuate at a rate that is inversely proportional to the distance from the source. In this paper, we will further assume that the acoustic intensities of the K sources will be linearly superimposed without any interaction among them. The acoustic signal received at the i^{th} sensor, $i = 1, 2, ...N$, during time interval n then can be expressed as:

$$x_i(n) = s_i(n) + \nu_i(n) \tag{1}$$

where

$$s_i(n) = \gamma_i \sum_{k=1}^{K} \frac{a_k(n - t_{ki})}{\|\boldsymbol{\rho}_k(n - t_{ki}) - \mathbf{r}_i\|} \tag{2}$$

is the acoustic intensity measured at the i^{th} sensor due to all K acoustic sources, and $\nu_i(n)$ is the background noise. In this paper, $\nu_i(n)$ is modeled as a zero-mean additive white Gaussian (AWGN) noise random variable with variance σ_n^2. $a_k(n - t_{ki})$ is the intensity of the k^{th} acoustic source measured at one meter from that source, and t_{ki} is the propagation delay of the acoustic signal from the k^{th} source to the i^{th} sensor. $a_k(n - t_{ki})$ will be modeled as a zero-mean random variable uncorrelated with each other. That is, $E\{a_k(n-t_{ki})\} = 0$ and $E\{a_k(n-t_{ki})a_j(n-t_{ki})\} = E\{a_k(n-t_{ki})\} \cdot E\{a_j(n-t_{ji})\} = 0$ if $k \neq j$. $\boldsymbol{\rho}_k$ is an unknown $p \times 1$ vector denoting the position vector of the k^{th} target; and $\mathbf{r}_i$ is a given $p \times 1$ vector denoting the position vector of the i^{th} stationary sensor. γ_i is sensor gain factor of the i^{th} acoustic sensor.

We remark that this model does not account for the potential echoes and reverberation [24] due to the presence of obstacles such as man-made walls or rocky hills. It also does not account for the potential impacts of wind direction [25] [26] and dense vegetation [27] in the sensor field. Moreover, in practice, some of the assumptions made in deriving this model may not be always true and hence may affect its accuracy. For example, the engine sound of a vehicle may not be omni-directional, and will be biased toward the side closer to the engine. The physical size of the acoustic source may be too large compared to the sensor field to be adequately modelled as a point source. In an outdoor environment, strong background noise, including wind-gust may be encountered during operation. In addition, the gain of

individual microphones will need to be calibrated to yield consistent acoustic energy readings. Some of these issues will be addressed in a future work.

Assume that $s_i(n)$ and $\nu_i(n)$ are uncorrelated such that $E\{s_i(n)\nu_i(n)\} = E\{s_i(n)\}E\{\nu_i(n)\} = 0$. Then, one may represent the acoustic energy as (setting $g_i = \gamma_i^2$ and $S_k(n - t_{ki}) = E\left[a_k^2(n - t_{ki})\right]$):

$$\begin{aligned} E\left[s_i^2(n)\right] &= \gamma_i^2 \sum_{k=1}^{K} \frac{E\left[a_k^2(n - t_{ki})\right]}{\|\boldsymbol{\rho}_k(n - t_{ki}) - \mathbf{r}_i\|^2} \\ &= g_i \sum_{k=1}^{K} \frac{S_k(n - t_{ki})}{\|\boldsymbol{\rho}_k(n - t_{ki}) - \mathbf{r}_i\|^2} \end{aligned}$$

In practice, the expectation is realized using the ensemble average over a time window $T = M/f_s$, where M is the number of sample points used for estimating the acoustic energy received by the sensor during this time interval, and f_s is the sampling frequency. Denote this average energy measurements over the time window $[t - T/2, t + T/2]$ as $y_i(t)$, it leads to

$$\begin{aligned} y_i(t) &= \frac{1}{T} \sum_{n=\frac{t-T/2}{f_s}}^{\frac{t+T/2}{f_s}} x_i^2(n) \\ &= \underbrace{\frac{1}{T} \sum_{n=\frac{t-T/2}{f_s}}^{\frac{t+T/2}{f_s}} s_i^2(n)}_{signal\ energy,\ y_s(t)} + \underbrace{\frac{1}{T} \sum_{n=\frac{t-T/2}{f_s}}^{\frac{t+T/2}{f_s}} \nu_i^2(n)}_{noise\ energy,\ y_n = \varepsilon_i(t)} \end{aligned} \tag{3}$$

Since the sources are assumed to be within a sensor network, it is safe to assume that the propagation delay t_{ki} is small enough such that $a(n - t_{ki}) \approx a(n)$ and $\boldsymbol{\rho}(n - t_{ki}) \approx \boldsymbol{\rho}(n)$. This leads to a more concise acoustic energy decay model

$$y_i(t) = y_s(t) + \varepsilon_i(t) = g_i \sum_{j=1}^{K} \frac{S_j(t)}{d_{ij}^2(t)} + \varepsilon_i(t) \tag{4}$$

where $d_{ij}(t) = \|\boldsymbol{\rho}_j(t) - \mathbf{r}_i\|$ is the Euclidean distance between the i^{th} sensor and the j^{th} source. The square of the background noise, $\nu_i^2(n)$, will have a χ^2 distribution with mean equals to $E[\nu_i^2(n)] = \sigma_i^2$ and variance equals to $2\sigma_i^4/M$. If M is sufficiently large ($M >> 30$), according to the central limit theorem, ε_i can be approximated well with a normal distribution, namely, $\varepsilon_i \sim N(\sigma_i^2, 2\sigma_i^4/M)$. For convenience, we shall denote $\mu_i = \sigma_i^2$, and $\sigma_i^2 = 2\sigma_i^4/M$ in later derivations. The validity of this energy attenuation model has been verified with a simple experiment. Details can be found in [22].

III. Maximum Likelihood Location Estimation

In this section, we will introduce the *ML* estimation with different solutions to estimate the source location. Note that the estimation is based on single frame of energy readings from different individual

sensors. Let us now define the following matrix notations. The time index t is omitted in the interests of brevity.

$$\begin{aligned}
\mathbf{Z} &= \begin{bmatrix} (y_1-\mu_1)/\sigma_1 & \dots & (y_N-\mu_N)/\sigma_N \end{bmatrix}^T \\
\mathbf{G} &= diag\begin{bmatrix} g_1/\sigma_1 & g_2/\sigma_2 & \dots & g_N/\sigma_N \end{bmatrix} \\
\mathbf{D} &= \begin{bmatrix} 1/d_{11}^2 & 1/d_{12}^2 & \dots & 1/d_{1K}^2 \\ 1/d_{21}^2 & 1/d_{22}^2 & \dots & 1/d_{2K}^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1/d_{N1}^2 & 1/d_{N2}^2 & \dots & 1/d_{NK}^2 \end{bmatrix} \\
\mathbf{S} &= \begin{bmatrix} S_1 & S_2 & \cdots & S_K \end{bmatrix}^T \\
\mathbf{H} &= \mathbf{GD} \\
\boldsymbol{\xi} &= \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \xi_N \end{bmatrix}^T
\end{aligned} \tag{5}$$

where $\xi_i = (\epsilon_i - \mu_i)/\sigma_i \sim N(0,1)$ are independent Gaussian random variables. Using these notations, equation (4) can be represented as:

$$\mathbf{Z} = \mathbf{GDS} + \boldsymbol{\xi} = \mathbf{HS} + \boldsymbol{\xi} \tag{6}$$

The joint probability density function of $\mathbf{Z}$ then can be expressed as:

$$f(\mathbf{Z}|\boldsymbol{\theta}) = (2\pi)^{-\frac{N}{2}} exp\left\{-\frac{1}{2}(\mathbf{Z}-\mathbf{HS})^T(\mathbf{Z}-\mathbf{HS})\right\} \tag{7}$$

where

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\rho}_1^T & \boldsymbol{\rho}_2^T & \cdots & \boldsymbol{\rho}_K^T & S_1 & S_2 & \cdots & S_K \end{bmatrix}^T$$

is the vector of unknown parameters. The negative log-likelihood function is proportional to a nonlinear quadratic form:

$$\ell(\boldsymbol{\theta}) = \|\mathbf{Z} - \mathbf{GDS}\|^2 \tag{8}$$

Thus the maximum likelihood parameter estimation of $\boldsymbol{\theta}$ can be obtained by minimizing $\ell(\boldsymbol{\theta})$.

Equation (8) is a *nonlinear* least square cost function because the $\mathbf{D}$ matrix that contains NK elements is a nonlinear function of the Kp unknown source location coordinates $\{\boldsymbol{\rho_j}; 1 \le j \le K\}$. The $\mathbf{S}$ vector also contains K unknown parameters. Since there are $K(p+1)$ unknown parameters, there must be at least $K(p+1)$ or more sensors reporting acoustic energy readings to yield an unique solution.

To minimize $\ell(\boldsymbol{\theta})$, the solution must lie on a stationary point where

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial S_k} = 0, \quad \text{and} \quad \nabla_{\boldsymbol{\rho}_k} \ell(\boldsymbol{\theta}) = 0.$$

for $k = 1, 2, \cdots K$. These conditions lead to the following set of relations

$$\mathbf{S} = \mathbf{H}^{\dagger}\mathbf{Z} \tag{9}$$

$$\mathbf{B}^T\mathbf{G}(\mathbf{Z} - \mathbf{H}\mathbf{S}) = 0 \tag{10}$$

where $\mathbf{H}^{\dagger}$ is the *pseudo-inverse* of the matrix $\mathbf{H}$, and $\mathbf{B} = \left[\begin{array}{cccc} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_N \end{array} \right]$ is a $K \times N$ matrix with its k^{th} column defined as:

$$\mathbf{B}_k \stackrel{def}{=} \frac{\partial (\mathbf{D}\mathbf{S})^T}{\partial \boldsymbol{\rho}_k} = -2S_k \left[\begin{array}{ccc} \mathbf{b}_{1k}/d_{1k}^3 & \ldots & \mathbf{b}_{Nk}/d_{Nk}^3 \end{array} \right]^T \tag{11}$$

and $\mathbf{b}_{ij} = \partial d_{ij}/\partial \boldsymbol{\rho}_j = \left(\boldsymbol{\rho}_j - \mathbf{r}_i\right)/d_{ij}$ is a unit vector from the j^{th} source to the i^{th} sensor.

Equation (10) may be expressed alternatively as:

$$\begin{aligned} \nabla_{\boldsymbol{\rho}_k} \ell(\boldsymbol{\theta}) &= 2S_k \sum_{i=1}^{N} \frac{g_i}{\sigma_i} \left(\frac{\boldsymbol{\rho}_k - \mathbf{r}_k}{d_{ik}^4} \right) \left(z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^{K} \frac{S_m}{d_{im}^2} \right) \\ &= 0 \end{aligned} \tag{12}$$

Solving above equation, $\boldsymbol{\rho}_k$ can be represented as a linearly weighted sum of all sensor locations $\mathbf{r}_i$:

$$\boldsymbol{\rho}_k = \left(\sum_{i=1}^{N} \alpha_i \mathbf{r}_i \right) / \left(\sum_{i=1}^{N} \alpha_i \right) \tag{13}$$

where

$$\alpha_i = \sum_{i=1}^{N} \frac{g_i}{\sigma_i} \left(\frac{1}{d_{ik}^4} \right) \left(z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^{K} \frac{S_m}{d_{im}^2} \right)$$

We caution readers that equation (13) is *not* an explicit expression for $\boldsymbol{\rho}_k$ since the distance d_{im} is a function of $\boldsymbol{\rho}_m$ on the right hand side of this equation.

Note that equation (9) gives explicit expression of $\mathbf{S}$ as a function of $\mathbf{H}$ and $\mathbf{Z}$. However, equation (10) is an implicit relation where the unknown $\{\boldsymbol{\rho}_k\}$ appear in both sides of the equation. These two equations motivated the development of two different approaches to solve for the maximum likelihood source localization problem as described below:

A. Multi-Resolution Projection Solution

Substitute equation (9) into equation (8), the variables $\{S_k\}_{k=1}^K$ can be elimited, giving a modified negative log-likelihood function

$$\begin{aligned} & \mathbf{L}'(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_K) = \mathbf{Z^T}(\mathbf{I} - \mathbf{P_H})^{\mathbf{T}}(\mathbf{I} - \mathbf{P_H})\mathbf{Z} \\ = \quad & \mathbf{Z^T}(\mathbf{I} - \mathbf{P_H})\mathbf{Z} = \mathbf{Z^T}\mathbf{Z} - \mathbf{Z^T}\mathbf{P_H}\mathbf{Z} \end{aligned}$$

where

$$\mathbf{P}_H = \mathbf{H}(\mathbf{H^T}\mathbf{H})^{-\mathbf{1}}\mathbf{H^T} = \mathbf{U_H}\,\mathbf{U_H^T}.$$

is a projection matrix, and $\mathbf{U}_H$ is the left singular vectors of the $\mathbf{H}$ matrix. The properties $\mathbf{P}_H = \mathbf{P}_H^T$ and $\mathbf{P}_H \cdot \mathbf{P}_H = \mathbf{P}_H$ have been used during derivation. Since $\mathbf{Z^T}\mathbf{Z}$ are known (normalized) energy measurements, **minimizing** $\mathbf{L}'$ is equivalent to **maximizing**

$$\mathbf{L}\,(\boldsymbol{\rho_1}, \boldsymbol{\rho_2}, \ldots, \boldsymbol{\rho_K}) = \mathbf{Z^T}\mathbf{U_H}\mathbf{U_H^T}\mathbf{Z} = \|\mathbf{U_H^T}\mathbf{Z}\|^{\mathbf{2}} \tag{14}$$

that depends only on source coordinates $\{\boldsymbol{\rho}_{k=1}^K\}$. In a way, equation (14) is the log-likelihood function with the constraint of equation (9) imposed. As such, $\mathbf{L}$ contains K fewer unknown variables than $\ell(\boldsymbol{\theta})$.

We caution the reader that although $\mathbf{L}$ is expressed as a quadratic form, maximizing it still requires the solution of a *nonlinear* least square problem as $\mathbf{U}_H$ is still a nonlinear function of the source locations $\{\boldsymbol{\rho}_k\}_{k=1}^K$.

1) Multi-resolution (MR) search: A straight-forward method to find a solution that maximizes $\mathbf{L}$ is exhaustive search. However, the computation cost is extremely high especially when there are multiple sources. For example, let there be K sources and q grid points to be searched in each dimension. Then the total number of search points with a p-dimensional sensor field will be equal to q^{pK}. While the computation complexity may be feasible for a desktop computer, it is likely to be overly excessive for low power sensor nodes with limited computing capabilities.

This exponentially growing computation complexity can be mitigated with the use of multi-resolution (MR) search method. Among several choices, a logrithmic MR search strategy will examine only w points in each dimension per iteration where $q = w^m$, m being a positive integer. Hence, in each iteration, only w^{pK} grid points needs to be searched. Then, another iteration of search will be confined in the neighborhood of the current best solution by sub-dividing the coarser mesh around the current solution into w sub-divisions, and perform search. After m iterations, the MR method will search at a grid size equal to that of exhaustive search. However, the total search points will be $m \cdot w^{pK}$ rather than $q^{pK} = w^{mpK}$. To appreciate the amount of saving, let $q = 128 = 2^7 = w^m$, and $p = 2$. Using exhaustive

search, the total search points will be 2^{14K}. Using MR search, the total number of search points will be $7 \cdot 2^{2K}$. For $K = 1$, the difference is 4096 points versus 28 points. For $K = 2$, the difference is 268435456 points versus 112 points. Obviously, due to the coarser search grid at earlier iterations, the MR method may be trapped in a local minimum and yields an inferior solution.

B. Expectation-Maximization Iterative Solution

Equations (9) and (13) can be used together to yield an iterative solution similar to the expectation-maximization algorithm. With this solution, we assume the source intensity vector **S** is the *missed* data rather than unknown parameter. We initiate the K unknown source location ρ_k at the beginning. During the iteration, we expect the *missed* data according to equation (9). And then we maximize the log-likelihood function using equation (13) to get the updated estimate ρ_k. The iteration keeps on going until it convergence. This EM algorithm has much less computation complexity compared to the projection solution. However, it will trap into a local minimum. Hence, they may be applied to refine the MR search results.

C. Comparison with Other Energy Based Source Localization Methods

There are other energy based acoustic source localization methods besides the maximum likelihood method presented in this paper. We now summarize these algorithms and exploit their relationship to the proposed ML method.

1) Closest Point Approach (CPA) Methods: The closest point approach (CPA) is a nevigation term that describes the closest position of two objects moving along non-intersecting straight lines. Here, we borrow the term to refer to a nearest neighbor localization method: *Identify a sensor with largest acoustic energy measurement:*

$$i^* = \arg\max_i y_i$$

Assign the source location to be the sensor location

$$\boldsymbol{\rho} = \mathbf{r}_{i^*} \tag{15}$$

In general, the CPA method is suitable for single source situation. When there are multiple sources, the algorithm must first identify all local maxima among all sensor acoustic energy readings. As a matter of fact, the CPA algorithm can be deduced from equation (13) by setting $K = 1$. Specifically, let $d_i = d_{ik}$ when $K = 1$, and set $d_{i^*} \to 0$. Then equation (13) becomes equation (15). Another implementation detail is that the actual source location can not be overlap with the sensor location. Hence, it is often chosen as a location that is nearest to the sensor with largest energy reading.

D. Energy Ratios Source Localization, Nonlinear Least Square (ER-NLS) and Least Square (ER-LS) Formulations

When there is only a single source ($K = 1$) within the sensor field, the $\mathbf{H}$ matrix become a vector

$$\mathbf{H} = \left[\frac{g_1}{\sigma_1\, d_1^2}, \frac{g_2}{\sigma_2\, d_2^2}, \cdots, \frac{g_n}{\sigma_n\, d_n^2}\right]^T,$$

As such, $\mathbf{U_H} = \mathbf{H}/\|\mathbf{H}\|$ will be a unit vector. If each entry of the $\mathbf{U}_H$ vector *were* an independent variable, then an obviously solution to maximize the modified cost function $\mathbf{L}$ in equation (14) would be $\mathbf{U_H} = \mathbf{c} \cdot \mathbf{Z}$ where c is a proportional constant. Equivalently, for $i = 1, \cdots, N$,

$$\frac{y_i - \mu_i}{\sigma_i} = c \cdot \frac{g_i}{\sigma_i d_i^2} = c \cdot \frac{g_i}{\sigma_i \|\boldsymbol{\rho} - \mathbf{r_i}\|^{\mathbf{2}}} \tag{16}$$

where $\boldsymbol{\rho}$ denotes the source location. Note that although there are N equalities, there are actually $p{+}1 < N$ unknowns, including the proportional constant c.

One way to solve these set of nonlinear equalities is to first eliminate the unknown constant c by computing the energy ratio φ_{ij} of the i^{th} and the j^{th} sensors as follows:

$$\varphi_{ij} = \left[\frac{(y_i - \mu_i)/(y_j - \mu_j)}{g_i/g_j}\right]^{-1/2} = \frac{\|\boldsymbol{\rho} - \mathbf{r}_i\|}{\|\boldsymbol{\rho} - \mathbf{r}_j\|} \tag{17}$$

By sorting the calibrated energy readings $(y_i - \mu_i)/g_i$, for $0 < \varphi_{ij} \neq 1$, all the possible source coordinates $\boldsymbol{\rho}$ that satisfy equation (17) reside on a p-dimensional hyper-sphere described by the equation:

$$\|\boldsymbol{\rho} - \mathbf{c}_{ij}\|^2 = \zeta_{ij}^2 \tag{18}$$

where the center $\mathbf{c}_{ij}$ and the radius ζ_{ij} of this hyper-sphere associated with sensor i and j are given by:

$$\mathbf{c}_{ij} = \frac{\mathbf{r}_i - \varphi_{ij}^2 \mathbf{r}_j}{1 - \varphi_{ij}^2},\ \zeta_{ij} = \frac{\varphi_{ij}\|\mathbf{r}_i - \mathbf{r}_j\|^2}{1 - \varphi_{ij}^2} \tag{19}$$

If $\varphi_{ij} = 1$, the solution of equation (17) form a hyper-plane between $\mathbf{r}_i$ and $\mathbf{r}_j$, i.e.:

$$\boldsymbol{\rho}(t)\boldsymbol{\iota}_{ij} = \tau_{ij} \tag{20}$$

where $\boldsymbol{\iota}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $\tau_{ij} = |\mathbf{r}_i|^2 - |\mathbf{r}_j|^2/2$. With all the energy ratios computed, the target location can be solved by minimizing a nonlienar least square cost function:

$$J(\boldsymbol{\rho}) = \sum_{l_1=1}^{L_1} \left(|\boldsymbol{\rho} - \mathbf{c}_{l_1}| - \zeta_{l_1}\right)^2 + \sum_{l=1_2}^{L_2} \left(\boldsymbol{\iota}_{l_2}^T \boldsymbol{\rho} - \tau_{l_2}\right)^2 \tag{21}$$

where L_1 and L_2 are the number of hyper-spheres and the number of hyper-planes respectively. We call it the energy-ratio nonlinear least square (ER-NLS) method.

Since every pair of hyper-spheres (with double indices ij replaced by a single index m for the brevity of notations) $\|\boldsymbol{\rho} - \mathbf{c}_m\|^2 = \zeta_m^2$ and $\|\boldsymbol{\rho} - \mathbf{c}_n\|^2 = \zeta_n^2$, a hyper plane can be determined by eliminating the common terms:

$$(\mathbf{c}_m - \mathbf{c}_n)^T \boldsymbol{\rho} = \left((\|\mathbf{c}_m\|^2 - \|\mathbf{c}_n\|^2) - (\zeta_m^2 - \zeta_n^2) \right) /2 \tag{22}$$

Combining equations (22) with (20), the source location can be solved using a least square solution without lengthy nonlinear optimization search. We call it the energy-ratio least square (ER-LS) method.

The ER-NLS and ER-LS methods have been reported previously in [22]. They are summarized here for the convenience of readers.

IV. Cramer-Rao Bounds (CRB)

CRB is a theoretical lower bound of the variance of an unbiased parameter estimate. It is defined as the inverse of the *Fisher information Matrix*:

$$\begin{aligned} \mathbf{J} &= -E\left[\frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln\, f(\mathbf{Z}|\boldsymbol{\theta}) \right)^T \right] \\ &= E\left\{ \left[\frac{\partial \ln f(\mathbf{Z}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \left[\frac{\partial \ln f(\mathbf{Z}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^T \right\} \end{aligned} \tag{23}$$

Substitute equation (7) into equation (23), one has

$$\mathbf{J} = \frac{\partial (\mathbf{DS})^T}{\partial \boldsymbol{\theta}} \mathbf{G}^\mathbf{T} \mathbf{G} \frac{\partial (\mathbf{DS})}{\partial \boldsymbol{\theta}^T} \tag{24}$$

where $\partial(\mathbf{DS})/\partial \boldsymbol{\theta}^T = \begin{bmatrix} \mathbf{B} & \mathbf{D} \end{bmatrix}$ since $\partial(\mathbf{DS})/\partial \mathbf{S}^T = \mathbf{D}$, and $\partial(\mathbf{DS})^T/\partial \boldsymbol{\rho}_j = \mathbf{B}_j$ as defined in equation (11). Hence

$$\mathbf{J} = \begin{bmatrix} \mathbf{B}^\mathbf{T} \\ \mathbf{D}^\mathbf{T} \end{bmatrix} \mathbf{G}^\mathbf{T} \mathbf{G} \begin{bmatrix} \mathbf{B} & \mathbf{D} \end{bmatrix} \tag{25}$$

The lower bound of the variance of the source location estimates $var\left(\widehat{\rho_{i_j}}\right)$ can be expressed as

$$var\left(\widehat{\rho_{i_j}}\right) \geq \left(J^{-1}\right)_{(i-1)p+j,(i-1)p+j}$$

$1 \leq i \leq K$, and $1 \leq j \leq p$.

A. CRB for the Single Source Case

For convenience, we will focus on the single source ($K = 1$) and two-dimensional sensor field ($p = 2$) situation. Furthermore, we set $g_i = g$ and $\sigma_i = \sigma$ for $i = 1, ..., n$. Then, the Fisher Information matrix

becomes

$$\begin{aligned}\mathbf{J} &= \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & J_{22} \end{bmatrix} \\ &= \frac{g^2}{\sigma^2}\begin{bmatrix} 4s^2 \cdot \sum_{i=1}^{n} \mathbf{b}_i \mathbf{b}_i^T / d_i^6 & -2s \sum_{i=1}^{n} \mathbf{b}_i / d_i^5 \\ -2s \sum_{i=1}^{n} \mathbf{b}_i / d_i^5 & \sum_{i=1}^{n} 1/d_i^4 \end{bmatrix}\end{aligned}$$

Define $\mathbf{E} = \mathbf{J}_{11} - \mathbf{J}_{12}\mathbf{J}_{12}^T / J_{22}$, it can be factored as:

$$\begin{aligned}\mathbf{E} &= \kappa \begin{bmatrix} \mathbf{L_X^T} \\ \mathbf{L_Y^T} \end{bmatrix} \begin{bmatrix} \mathbf{L_X} & \mathbf{L_Y} \end{bmatrix} \\ &= \kappa \begin{bmatrix} |\mathbf{L_X}|^2 & |\mathbf{L_X}| \cdot |\mathbf{L_Y}| \cos\beta \\ |\mathbf{L_X}| \cdot |\mathbf{L_Y}| \cos\beta & |\mathbf{L_Y}|^2 \end{bmatrix}\end{aligned} \tag{26}$$

where $\kappa = 4s^2g^2/\sigma^2$, $\mathbf{L_X}$ and $\mathbf{L_Y}$ are both $N \times 1$ vectors, and $\beta = \cos^{-1}\left(\mathbf{L_X^T L_Y} / |\mathbf{L_X}|\ |\mathbf{L_Y}|\right)$. The j^{th} column of the $2 \times N$ matrix $\begin{bmatrix} \mathbf{L_X} & \mathbf{L_Y} \end{bmatrix}^T$ can be expressed as:

$$\begin{aligned}\begin{bmatrix} \mathbf{L_X} & \mathbf{L_Y} \end{bmatrix}_j^T &= \frac{\mathbf{b}_j}{d_j^3} - \frac{1/d_j^2}{\sum_{i=1}^{N} \left(1/d_i^2\right)^2} \sum_{i=1}^{N} \left(\frac{1}{d_i^2}\right)\left(\frac{\mathbf{b}_i}{d_i^3}\right) \\ &= \mathbf{b}_j/d_j^3 - \overline{\mathbf{b}}/d_j^2\end{aligned} \tag{27}$$

where $\overline{\mathbf{b}}$ is a weighted average of the individual (weighted) source-to-sensor direction vector.

Using the matrix inversion formula, the lower bound of the variance of the ML source location estimates can be found in terms of the CRB as:

$$Var(\hat{\rho_x}) \geq \{\mathbf{E}^{-1}\}_{11} = \frac{1}{\kappa} \frac{1}{|\mathbf{L_X}|^2(1 - cos^2\beta)} \tag{28}$$

$$Var(\hat{\rho_y}) \geq \{\mathbf{E}^{-1}\}_{22} = \frac{1}{\kappa} \frac{1}{|\mathbf{L_Y}|^2(1 - cos^2\beta)} \tag{29}$$

Due to the weight factors of $1/d_i^3$, sensors that are close to the source will have a much higher impact on the CRB than those far away. Specifically, the fact that the CRBs are the diagonal elements of $\mathbf{E}^{-1}$ implies that the $\mathbf{E}$ matrix must be strongly non-singular. That in turn implies that at least two of the sensors must be close enough to the source so that the $[\mathbf{L_X}\ \mathbf{L_Y}]^T$ matrix will have full column rank $(= p)$.

To reduce the CRB and hence to improve the accuracy of the source localization results, three approaches may be taken: (i) To keep κ large, (ii) To increase both $|\mathbf{L_X}|^2$ and $|\mathbf{L_Y}|^2$, and (iii) to let $\beta \rightarrow 90^o$. Since source energy emission level s and background noise level σ can not be controlled,

to increase κ, the sensor gain g needs to be increased. To increase the magnitude of both $|\mathbf{L_X}|^2$ and $|\mathbf{L_Y}|^2$, the set of sensor target distance $\{d_i; 1 \leq i \leq N\}$ must be reduced. This implies that there must be sufficient number of sensors close to the acoustic source whenevern the source is within the sensor field. Since the target may be moving within the sensor field, this requirement translates to the requirement of dense and uniform sensor deployment within the sensor field. As for (iii), to make $\beta \rightarrow 90^o$, it would require the sensors within the sensor field be laid out in a particular formation. Nonetheless, this requirement is not suitable for an ad hoc sensor network application where sensor deployment is assumed to have no specific formation or structure.

V. Performance Evaluation

A. Performance Comparison using Simulation – Single Source Case

Extensive simulation runs have been conducted to compare the performance of the proposed *ML* energy based source localization algorithm to other energy based source localization algorithms. We use equation (4) to generate the acoustic energy readings of a two-dimensional ($p = 2$) sensor field of size 100 meters by 100 meters. Since existing localization algorithms are all for single target situation, we set $K = 1$. The source location and the sensor locations are randomly chosen from within the sensor field in each run. The source energy is set at $S = 5000$, and the background noise level is set at $\sigma_i = 1$ for all sensors in the sensor field. Note that although the SNR is 37dB at the source location, the actual SNR at different sensors depends very much on the sensor to source distance. For example, for a sensor that is 50 meters away from the source, its SNR is merely $10 \times \log_{10}(5000/50^2) = 3dB$. The energy variation $\epsilon_i(t)$ is modeled as a Gaussian random variable $N(\sigma_i^2, \sigma_i^4/M)$ with $M = 100$.

2000 repeated trials were conducted. In each trial, all four energy based acoustic source localization methods, ML, CPA, ER-NLS, and ER-LS will be used to locate the source location, and the error is recorded. For the CPA method, the target location is the location of the maximum energy sensor added by a Gaussian random variable with zero mean and unit variance. For the ML and ER-NLS methods, exhaustive search is used at a resolution of 5 meters by 5 meters per search grid. Three different sensor densities have been used: 4, 10, 25. In the last case, there is approximately one sensor in every 20 by 20 meter cells in the sensor field.

From simulation results, we observe that the mean values of these methods do not show any statistically significant bias, and hence yield unbiased estimates. Furthermore, the estimation error in different dimensions are also uncorrelated. In addition,

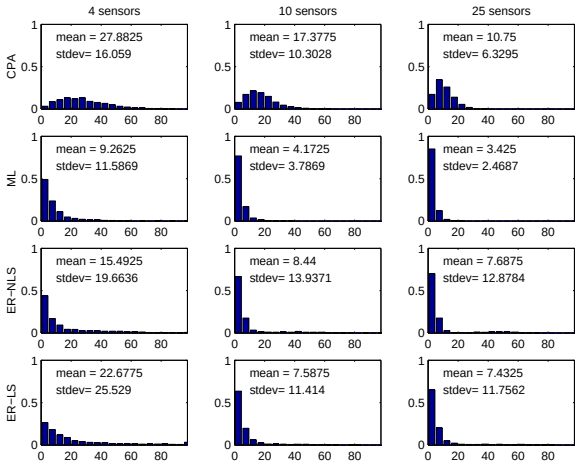

Fig. 1. Comparison of four acoustic energy-based source localization algorithms

- The *ML* method consistently outperforms all other three existing methods in terms of minimum estimation error variances and minimum mean estimation error.
- The CPA method benefits most when the number of sensors increases. Its rankings jump from the last the the second best as a result.
- With both the 10-sensor or the 25-sensor cases, the least square formulation (ER-LS) out-performs the computation-wise more expensive nonlinear least square formulation (ER-NLS).

Another way to analyze the simulation results is to examine the distribution of the magnitude of location estimation error. The results are summarized in Figure 1. In this figure, each column represent results obtained from a particular method. Each row represents results from a particular sensor density. The histograms of the magnitudes of the localization error are plotted using a bin of 5 meters increment. Since the histogram can be regarded as an approximation of the probability density function, the mean and standard deviation of the magnitude of the localization error are calculated and listed in each figure.

B. Performance Analysis: Two Sources Case

We have also performed simulation to examine the performance of the proposed algorithm when there are two targets present in the sensor field. Again, we assume a sensor field of 100 by 100 meter squares. Three different source separations are used: 50 meters, 20 meters, and 10 meters. These pre-defined source locations are further perturbed randomly during simulation over a ± 5 meter square area. The background noise parameters are $\mu = 1$ and $\sigma = 0.1$. Both source intensity at 1 meter distance are defined as 5000 units. Three sensor network configurations are used: 6, 12, and 30 sensors. For each source location and

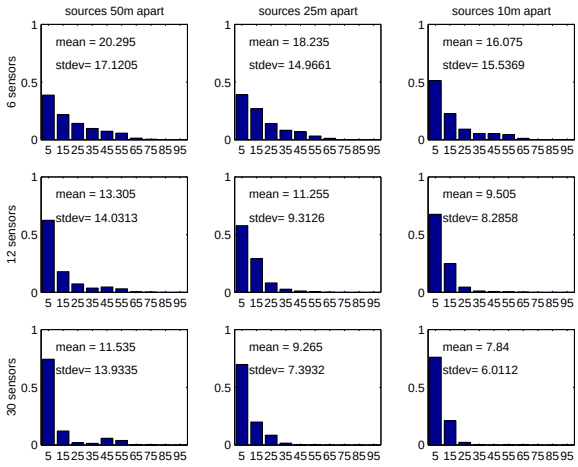

Fig. 2. Localization error of first target of the two targets

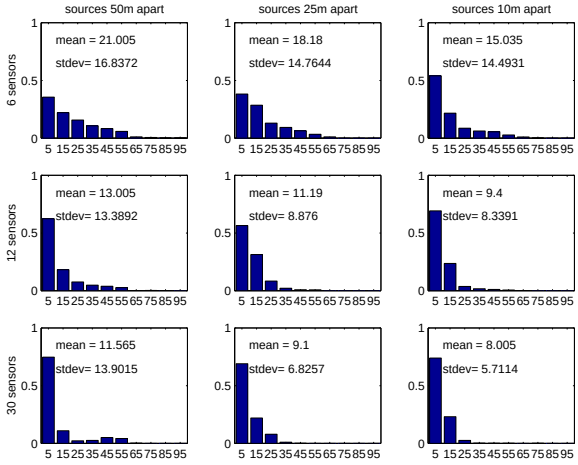

Fig. 3. Localization error of second target of the two targets

each sensor configuration, 2000 simulations are performed. The mean and variance are computed. Since there are two targets, the localized results will return two positions. We assume the target association is done properly so that the location estimation error is the minimum of the two possible assignments. The histogram of the location estimation error of each source are listed in figures 2 and 3.

C. Experiments:Application to Moving Vehicle Localization

An application of the proposed *EBSL* methods to locate moving vehicle using distributed microphone sensor nodes will be reported in this section. First, we describe the overall system. Then the experiment results will be presented.

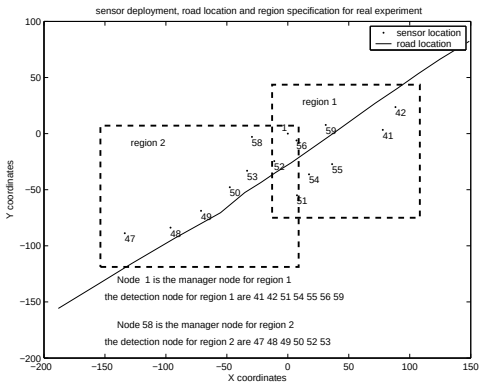

Fig. 4. sensor deployment, road coordinate and region specification for experiments

1) Sensor Network System: In November 2001, a field experiment sponsored by the DARPA ITO SensIT project has been carried out at 29 Palms California, USA. Custom-made prototype sensor nodes are laid out along side a road. Each sensor node is equipped with acoustic, seismic, and polorized infrared sensors, a 16-bit micro-processor, and radio transceiver and modem. It is powered by external car battery. During experiment, military vehicles such as AAV (Amphibious Assault Vehicle), DW (dragon wagon) were driven through the road, and sensors sampled the corresponding multi-modality data. The acoustic signal is sampled at 4960 Hz at 16-bit resolution. The set of data segments reported below are taken from the acoustic signatures of a single AAV travelling from east to west along the road during a time period of approximately 2 minutes. The sensor field is partitioned into two regions with some details given in Figure 4.

2) Acoustic Energy-based Localization Experiment: The energy reading collected from all sensor nodes within the region at the same 0.75 second time interval were used for acoustic source localization at the manager node. Fig. 5 shows the AAV ground truth and the localization results based on the *ML* algorithm with MR projection solution and ER-NLS algorithm. The ground truth is obtained by interpolating an on-board GPS recordings that recorded a position fix every 15 seconds. To use multi-resolution search, 3 grid sizes of 4, 2 and 1 The localization results are summarized in Figures 5 and 6. From figure 6, it is clear that the *ML* method out-perform the ER-NLS method.

The data sampled during this experiment is very noisy. During the experiments, Gusty wind often blew directly into the microphone, creating isolated energy spikes in some of the sensors. Many microphones are not properly calibrated, and the gain factors g_i estimated from the time series manually, and can be grossly inaccurate. The ground truth is not necessarily correct either due to the lack of differential GPS

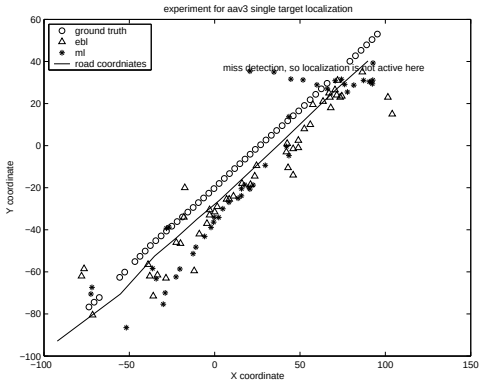

Fig. 5. AAV ground truth and localization estimation results based on ML algorithm with projection solution and NLS algorithm (MR search is used, grid size is 4*4, 2*2 and 1*1. Estimation results look bias from the ground-truth, see discussion for reasoning)

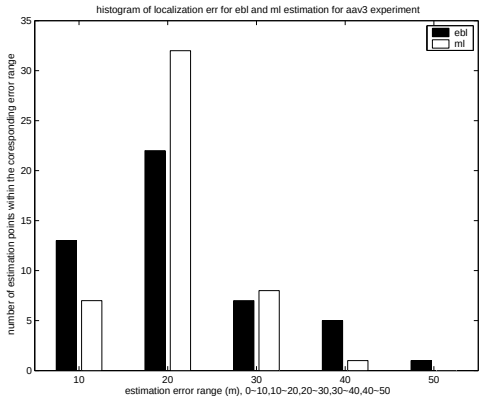

Fig. 6. Estimation error histogram for AAV experiment data

settings and the lack of long term averaging at each position fix. This is evident from figure 5 that the GPS marking is consistently off the road.

Hence, the superior performance of the proposed *ML* source localization method in this experiment clearly demonstrate the feasibility of applying this method to handle real world data.

VI. Conclusion

In this paper, a novel *ML* source localization method is presented. This method promises superior performance, multiple source localization, and is easy to compute. Besides algorithm derivation, the *CRB* of this algorithm has also been reported. Extensive simulations show that the proposed algorithm consistently out-performs other existing energy-based localization methods.

Future works include parameter sensitivity analysis, and sequential Bayesian estimation.

REFERENCES

[1] D. Estrin, Culler, D., Pister, K., and Sukhatme, G.,: Connecting the Physical World with Pervasive Networks, IEEE Pervasive Computing, vol. 1, Issue 1, 2002, pp. 59-69.

[2] C. Savarese, Rabaey, J. M., and Reutel, J.: Localization in distributed Ad-hoc wireless sensor networks, Proc. ICASSP'2001, Salt Lake City, UT, (2001), 2037-2040

[3] D. Li, Wong, K.D., Hu, Y. H., Sayeed, A. M.,: Detection, classification, and tracking of targets. IEEE Signal Processing Magazine, **19**, (2002), 17–29

[4] Tolstory, A.,: Matched-Field Processing for underwater acoustics, Sigapore: World Scientific, (1993)

[5] Brandstein, M., and Silverman, H.,: A localization-error-based method for microphone-array design, Proc. ICASSP'96, Atlanta, GA, (1996), 901-904

[6] Brandstein, M. S., Adcock, J. E., and Silverman,H. F.,: A closed form location estimator for use with room environment microphone arrays, IEEE Trans. Speech and Audio Processing, vol. 5, (1997), 45-50

[7] Huang, J., Ohnishi, N., and Sugie, N.,: Sound localization in reverberant enviornment based on model of the precedence effect, IEEE. Trans. Instrumentation and Measurement, **46**, (1997), 842-846

[8] Omologo, M., and Svaizer, P.,: Acoustic source location in noisy and reverberant environment using CSP analysis, Proc. ICASSP'96, Atlanta, GA, (1996), 921-924

[9] Wang, H., and Chu, P., Voice source localization for automatic camera pointing system in videoconferencing, ICASSP-97., **1**, (1997), 187 -190

[10] Suyama, K., Takahashi, K., and Hirabayashi, R.,: A robust technique for sound source localization in consideration of room capacity, presented at Proc. IEEE Workshop on Appl. Sig. Proc. to Audio and Acoust., New Paltz, NY, (2001) 63-66

[11] Weng, J., and Guentchev, K. Y.,: Three-dimensional sound localization form a compact non-coplanar array of microphones using three-based learning, J. Acoust. Soc. Am., **110**, (2001) 311-323

[12] Griebel, S. M., and Brandstein, M. S.,: Microphone array source localization using realizable delay vectors, IEEE Workshop on Appl. Sig. Proc. to Audio and Acoust., New Paltz, NY, (2001). 71-74

[13] Yao, K., Hudson, R. E., Reed, C. W., Chen, D., and Lorenzelli, F.,: Blind beamforming on a randomly distributed sensor array system, IEEE J. Selected areas in communications, **16** (1998) 1555–1567

[14] Reed, C.W., Hudson, R., and Yao,K.,: Direct joint source localization and propagation speed estimation. In Proc. ICASSP'99, Phoenix, AZ, (1999) 1169-1172

[15] Haykin, S.,: Array Signal Processing, Englewood-Cliffs, NJ: Prentice-Hall, (1985)

[16] Taff, L. G.,: Target localization from bearings-only observations, IEEE Trans. Aerosp. Electron., **3**, issue 1, (1997) 2-10

[17] Oshman, Y., and Davidson, P., Optimization of observer trajectories for bearings-only target localization, IEEE Trans. Aerosp. Electron., **35**, issue 3, (1999), 892-902

[18] Kaplan, K. M., Le, Q., and Molnar, P.,: Maximum likelihood methods for bearings-only target localization, Proc IEEE ICASSP, **5**, (2001), 3001-3004

[19] Carter G. C.,: Coherence and Time Delay Estimation, IEEE Press, 1993.

[20] Special issue on time-delay estimation, IEEE Trans. ASSP **29**, (1981)

[21] Smith, J.O., and Abel, J.S.,: Closed form least square source location estimation from range difference measurements. IEEE Trans. ASSP **35** (1987) 1661–1669

[22] D. Li, and Y. H. Hu,: Energy Based Collaborative Source Localization Using Acoustic Micro-Sensor Array, J. EUROSIP Applied Signal Processing, 2002 (to appear)

[23] Kinsler, L.E., et al.,: Fundamentals of Acoustics. NY, NY: John, Wiley, and Sons, Inc., 1982

[24] K. Srodecki,: Evaluation of the reverberation decay quality in rooms using the autocorrelation function and the cepstrum analysis, Acustica, vol. 80, pp. 216-25, 1994.

[25] Y. L. Li, M. J. White, and S. J. Franke,: New fast field programs for anisotropic sound propagation through an atmosphere with a wind velocity profile, Journal of the Acoustical Society of America, vol. 95, pp. 718-26, 1994.

[26] E. M. Salomons,: Downwind propagation of sound in an atmosphere with a realistic sound-speed profile: a semianalytical ray model, Journal of the Acoustical Society of America, vol. 95, pp. 2425-36, 1994.

[27] T. Watanabe and S. Yamada,: Sound attenuation through absorption by vegetation, Journal of the Acoustical Society of Japan, vol. 17, pp. 175-82, 1996.

Xiaohong Sheng Xiaohong Sheng received B. Eng degree from Zhejiang University, China, and M. Eng degree from NanYang Technological University, Singapore in 1998. From 1998 to 1999, she worked for the Audio group at R&D, Creative Technology Ltd, Singapore. She received MSEE degree at the department of electrical and computer engineering, University of Wisconsin-Madison in 2001. She is currently a Ph.D. Candidate in the same department. Her research interests include Collaborative Signal Processing in wireless sensor network, Multi-Media Signal Processing and Statistical Signal Processing in Communication.

Yu Hen Hu (M'83-SM'88-F'99) Yu Hen Hu received BSEE degree from National Taiwan University, Taiwan, ROC in 1976. He received MS and PHD degree both in electrical engineering from University of Southern California, Los Angeles, CA in 1980 and 1982 respectively. Currently, he is a professor at the electrical and computer engineering department of the University of Wisconsin - Madison, WI, USA. Previously, he has been with the electrical engineering department of the Southern Methodist University, Dallas, TX.

Dr. Hu's research interests include multi-media signal processing, design methodology and implementation of signal processing algorithms and systems, and neural network signal processing. He published more than 200 journal and conference papers, edited two books Programmable Digital Signal Processors, and Handbook of Neural Network Signal Processing.

Dr. Hu is a fellow of IEEE. He served as associate editors for IEEE Transactions on Signal Processing, IEEE Signal Processing letters, Journal of VLSI signal processing, European Journal of Applied Signal Processing. He served as secretary of IEEE signal processing society, board of governors of IEEE neural network council, chair of IEEE signal processing society, neural network signal processing technical committee.