Collaborative Sensor Networking Towards Real-Time Acoustical Beamforming in Free-Space and Limited Reverberence

https://escholarship.org/uc/item/7v03j6kw

Center for Embedded Network Sensing, 3(3)

Bergamo, Pierpaolo
Asgari, Shadnaz
Wang, H B
et al.

2004-05-05

Peer reviewed
Collaborative Sensor Networking Towards Real-Time Acoustical Beamforming in Free-Space and Limited Reverberance

Pierpaolo Bergamo, Shadnaz Asgari, Hanbiao Wang, Daniela Maniezzo, Len Yip, Student Member, IEEE, Ralph E. Hudson, Kung Yao, Fellow, IEEE, and Deborah Estrin, Fellow, IEEE

Abstract—Wireless sensor networks have been attracting increasing research interest given the recent advances in microelectronics, array processing, and wireless networking. Consisting of a large collection of small, wireless, low-cost, integrated sensing, computing, and communicating nodes capable of performing various demanding collaborative space-time processing tasks, wireless sensor network technology poses various unique design challenges, particularly for real-time operation. In this paper, we review the Approximate Maximum-Likelihood (AML) method for source localization and direction-of-arrival (DOA) estimations. Then, we consider the use of least-squares (LS) method applied to DOA bearing crossings to perform source localization. A novel virtual array model applicable to the AML-DOA estimation method is proposed for reverberant scenarios. Details on the wireless acoustical testbed are given. We consider the use of Compaq iPAQ 3760s, which are handheld, battery-powered device normally meant to be used as personal organizers (PDAs), as sensor nodes. The iPAQs provide a reasonable balance of cost, availability, and functionality. It has a build-in StrongARM processor, microphone, codec for acoustic acquisition and processing, and a PCMCIA bus for external IEEE 802.11b wireless cards for radio communication. The iPAQs form a distributed sensor network to perform real-time acoustical beamforming. Computational times and associated real-time processing tasks are described. Field measured results for linear, triangular, and square subarrays in free-space and reverberant scenarios are presented. These results show the effective and robust operation of the proposed algorithms and their implementations on a real-time acoustical wireless testbed.

Index Terms—Beamforming, source localization, distributed sensor network, wireless network, microphone array, time synchronization, reverberance.

1 INTRODUCTION

Recent advances in microelectronic-mechanical and wireless communication systems have allowed the construction of low-cost, low-power small sensor nodes. The challenge of sensor networks is to achieve dependable mission performance by the dynamic configuration and collaboration of these low reliable sensor nodes with limited sensing and communication capabilities. Some of the missions for these sensor networks may include the operations of detection, localization, tracking, and identification of objects which are of interest to diverse military, industrial, and civilian applications [1], [2], [3], [4], [5].

This paper, in this special issue on mission-oriented sensor networks, illustrates that the proper usage of novel array and signal processing algorithms implemented on very low-cost Commercial-Off-The-Shelf (COTS) platforms is capable of achieving quite sophisticated real-time acoustical beamforming operation for source localization. The objective of source localization is to estimate the positions of either a fixed or moving source using a passive and stationary sensor network. Various beamforming methods can be used to determine the direction-of-arrival(s) (DOA) and the location(s) of one or more acoustic source(s). Beamforming is a space-time operation in which a waveform originating from a given source but received at spatially separated sensors is combined in a time-synchronous manner. If the propagation medium preserves sufficient coherency among the received waveforms, then the beamformed waveform can provide an enhanced Signal-to-Noise-Ratio (SNR) compared to a single sensor system. Beamforming and localization are two interlinking problems and many algorithms have been proposed to tackle each problem individually and jointly (i.e., localization is often needed to achieve beamforming and some localization algorithms take the form of the beamformer). Many tutorial papers [6], [7] and books [8], [9] have dealt with beamforming and localization.

Over the past few years, a number of hardware platforms have emerged in the wireless sensor network arena, most of them involve custom-made devices. While these platforms may be energy-efficient and have small form factors, the COTS platforms are particularly attractive due to their lower cost, readily availability, and ease of use (minimal hardware/software modifications). While
the dedicated platforms are probably required for efficient real-world deployments. COTS platforms facilitate rapid application prototypes and data collection. This is particularly important in order to verify newly proposed array and signal processing algorithms in a testbed environment. In our COTS sensor platform, we use Compaq iPAQ 3760s [10] as the sensor nodes, which are handheld, battery-powered devices normally meant to be used as Personal Digital Assistants (PDA). We select the iPAQ for its compactness, reasonable battery life, Linux open-source operating system support [11], and reasonable computational resource. Each iPAQ has a built-in microphone and codec capable of a sampling rate from 8 kHz to 48 kHz for acoustical acquisition. It also supports the IEEE 802.11b wireless card for the data transmission. In order to perform coherent processing of the achieved waveforms, the signals collected from the iPAQ must be time-synchronized with respect to each other. The synchronization is achieved using Reference-Broadcast Synchronization (RBS), described by Elson et al. in [12].

The paper is organized as follows: Numerous beamforming methods have been proposed for DOA and source localization; among these methods, the maximum likelihood (ML) technique is optimum in the minimum estimation error variance sense for large SNR. In Section 2, we summarize a practical version of the ML algorithm, denoted as the Approximate ML (AML) algorithm, for wideband source localization and DOA estimation originally proposed in [13]. This algorithm is then implemented on the iPAQs to perform the real-time operation. Section 3 considers the least-squares (LS) method as applied to DOA bearing crossings to perform source localization. This approach provides a practical way to solve the far-field source localization problem from estimated DOA values. Section 4 proposes a novel virtual array model in which the physical array, whose location is assumed to be known, is reflected along known reflecting walls to create the virtual arrays. Then, applying the LS method on the estimated DOAs of the direct as well as the reflected paths arriving at the physical and virtual arrays, the location of the source(s) in reverberant environments can also be estimated. Section 5 provides various details on the testbeds. Hardware descriptions of the iPAQ system are first given, then the time synchronization and audio server issues are considered, and, finally, some client server paradigms are discussed. In Section 6, detailed computational times related to sampling, FFT, AML, angle estimation, and bearing crossing for real-time operations are given. In Section 7, we report various AML source localization results in free-space and reverberant scenarios. We first considered configurations of triangular, linear, and square subarrays for free-space and circular subarray for reverberant case. Then, we present various real-time testbed free-space and reverberant scenario results.

At the theoretical level, this paper first shows the feasibility of using COTS iPAQs for the implementation of a real-time acoustical beamforming testbed. The relative ease (in terms of months instead of years) in getting the testbed operational and the relatively low-cost (in the low thousands of dollars) of implementation were both happily unexpected.

2 AML ALGORITHM FOR WIDEBAND DOA ESTIMATION

In this section, we derive the AML algorithm for wideband source DOA estimation. We assume the source is in the far-field of the array, wavefront arriving at the array is assumed to be planar, and only the angle of arrival can be estimated. For simplicity, we assume both the source and sensor array lie in the same plane (a 2D scenario), as shown in Fig. 1.

Let there be M wideband sources, each at an angle θm, from the array with the reference direction pointing to the east. The sensor array consists of P randomly distributed sensors, each at position \( r_p = [x_p, y_p]^T \). The sensors are assumed to be omnidirectional and have an identical response. The array centroid position is given by \( r_c = \frac{1}{P} \sum_{p=1}^{P} r_p = [x_c, y_c]^T \). We use the array centroid as the reference point and define a signal model based on the relative time-delays from this position. The relative time-delay of the mth source is given by

\[
\tau_{cp}^{(m)} = \tau_c^{(m)} - \tau_p^{(m)} = \left[ (x_c - x_p) \cos \theta_m + (y_c - y_p) \sin \theta_m \right] / v,
\]

where \( \tau_c^{(m)} \) and \( \tau_p^{(m)} \) are the absolute time-delays from the mth source to the centroid and the pth sensor, respectively, and v is the speed of propagation. In a polar coordinate system, the above relative time delay can also be expressed as

\[
\tau_{cp}^{(m)} = r_p \cos(\theta_m - \phi_p) / v,
\]

where \( r_p \) and \( \phi_p \) are the range and angle of the pth sensor with respect to the array centroid. The data received by the pth sensor at time n is then

\[
x_p(n) = S_p(n) + w_p(n),
\]

for \( n = 0, \ldots, N - 1 \), \( p = 1, \ldots, P \), and \( m = 1, \ldots, M \), where \( N \) is the length of the data vector, \( S_p(n) \) is the mth source signal arriving at the array centroid position, \( \tau_{cp}^{(m)} \) is allowed to be any real-valued number, and \( w_p \) is the zero mean white Gaussian noise with variance \( \sigma^2 \).
For ease of derivation and analysis, the received wideband signal can be transformed into the frequency domain via the DFT, where a narrowband model can be given for each frequency bin. However, the circular shift property of the DFT has an edge effect problem for the actual linear time shift. These finite effects become negligible for a sufficient long data. Here, we assume the data length $N$ is large enough to ignore the artifact caused by the finite data length. For $N$-point DFT transformation, the array data model in the frequency domain is given by

$$X(\omega_k) = D(\omega_k)S(\omega_k) + \eta(\omega_k),$$

for $k = 0, \ldots, N-1$, where the array data spectrum is $X(\omega_k) = [X_1(\omega_k), \ldots, X_P(\omega_k)]^T$, the steering matrix $D(\omega_k) = [d^{(1)}(\omega_k), \ldots, d^{(M)}(\omega_k)]$ the steering vector is given by

$$d^{(m)}(\omega_k) = [d_1^{(m)}(\omega_k), \ldots, d_P^{(m)}(\omega_k)]^T,$$

and the source spectrum is given by

$$S(\omega_k) = [S^{(1)}(\omega_k), \ldots, S^{(m)}(\omega_k)]^T.$$

The noise spectrum vector $\eta(\omega_k)$ is zero mean complex white Gaussian distributed with variance $N\sigma^2$. Note, due to the transformation to the frequency domain, $\eta(\omega_k)$ asymptotically approaches a Gaussian distribution by the central limit theorem, even if the actual time-domain noise has an arbitrary i.i.d. distribution (with bounded variance). This asymptotic property in the frequency-domain provides a more reliable noise model than the time-domain model in some practical cases. Throughout this paper, we denote superscript $^T$ as the transpose and $^{-1}$ as the complex conjugate transpose.

The AML estimator performs the data processing in the frequency domain. The maximum-likelihood estimation of the source spectrum can be given by

$$X(\omega_k) = D(\omega_k)S(\omega_k) + \eta(\omega_k),$$

for $k = 0, \ldots, N-1$, where the array data spectrum is $X(\omega_k) = [X_1(\omega_k), \ldots, X_P(\omega_k)]^T$, the steering matrix $D(\omega_k) = [d^{(1)}(\omega_k), \ldots, d^{(M)}(\omega_k)]$ the steering vector is given by

$$d^{(m)}(\omega_k) = [d_1^{(m)}(\omega_k), \ldots, d_P^{(m)}(\omega_k)]^T,$$

and the source spectrum is given by

$$S(\omega_k) = [S^{(1)}(\omega_k), \ldots, S^{(m)}(\omega_k)]^T.$$

The noise spectrum vector $\eta(\omega_k)$ is zero mean complex white Gaussian distributed with variance $N\sigma^2$. Note, due to the transformation to the frequency domain, $\eta(\omega_k)$ asymptotically approaches a Gaussian distribution by the central limit theorem, even if the actual time-domain noise has an arbitrary i.i.d. distribution (with bounded variance). This asymptotic property in the frequency-domain provides a more reliable noise model than the time-domain model in some practical cases. Throughout this paper, we denote superscript $^T$ as the transpose and $^{-1}$ as the complex conjugate transpose.

The AML estimator performs the data processing in the frequency domain. The maximum-likelihood estimation of the source DOA and source signals is given by the following optimization criterion [13]:

$$\max_{\Theta, S} J(\Theta, S) = \max_{\Theta, S} \sum_{k=1}^{N/2} \|X(\omega_k) - D(\omega_k)S(\omega_k)\|^2,$$

which is equivalent to a nonlinear least square problem. Using the technique of separating variables [14], the AML DOA estimate can be obtained by solving the following likelihood function:

$$\max_{\Theta} J(\Theta) = \max_{\Theta} \sum_{k=1}^{N/2} \|P(\omega_k, \Theta)X(\omega_k)\|^2$$

$$= \max_{\Theta} \sum_{k=1}^{N/2} \text{tr}(P(\omega_k, \Theta)R(\omega_k)),$$

where

$$P(\omega_k, \Theta) = D(\omega_k)D^\dagger(\omega_k), D^\dagger = (D(\omega_k)^H D(\omega_k))^{-1} D(\omega_k)^H$$

is the pseudoinverse of the steering matrix $D(\omega_k)$ and $R(\omega_k) = X(\omega_k)X(\omega_k)^H$ is the one snapshot covariance matrix. Once the AML estimate of $\Theta$ is found, the estimated source spectrum can be given by

$$S^{ML}(\omega_k) = D^\dagger(\omega_k, \Theta^{ML})X(\omega_k).$$

The AML algorithm performs signal separation by utilizing the physical separation of the sources and, for each source signal, the SINR is maximized in the ML sense. Note that no closed-form solution can be obtained in (4). In the multiple source case, the computational complexity of the AML algorithm requires multidimensional search, which is much higher than the MUSIC type algorithm that requires only 1D search. Various numerical solutions were proposed to obtain the AML estimate. These include the Alternating Projection (AP), Gauss-Newton (GN), and Conjugate-Gradient (CG). For detail derivation of these methods, see [15].

### 3 Source Localization Based on LS Estimation of DOAs

In an idealized free-space (nonreverberant) model, each sensor receives one direct path ray (i.e., wavefront) from a source. A subarray of such sensors using various methods [16], [17], and [13] can exploit the information from these rays to estimate the direction-of-arrival (DOA) of the source. If a second subarray is not colinear with the first subarray and the source, then the crossing of these two DOAs yields the location of the source. Due to uncertainties in the estimation of these two DOAs, there will be an uncertainty region about the location of the source. By using three or more noncolinear subarrays and appropriate least-squares (LS) estimation technique, the uncertainty region about the true location of the source can be reduced. This technique can be extended to two or more sources, as shown below.

Instead of describing the LS estimation technique for completely arbitrary subarray(s) and source(s) geometry, consider the simplified model of having two sources and three subarrays, as shown in Fig. 2. In this model, source $j, j = 1, 2$, is located at $(x_j, y_j)$ and subarray $k, k = 1, 2, 3$, is located at $(x_k, y_k)$. In Fig. 2, for simplicity, we set $(x_1, y_1) = (0, 0)$, $(x_2, y_2) = (x_2, 0)$, and $(x_3, y_3) = (0, 0)$. The six DOAs at these subarrays directed toward sources 1 and 2 are denoted by $\{\theta_i, i = 1, \ldots, 6\}$, as shown in this figure. In this model, we assume the locations of the two sources are unknown and need to be determined, but the locations of the three subarrays are known and the DOAs are known by some estimation method.
From trigonometry, the six DOA angles satisfy the following six equations:

\[
\begin{align*}
\tan(\theta_1) &= \frac{s_1^1}{s_1^2}, \\
\tan(\theta_2) &= \frac{s_1^2}{(x_2 - s_2^2)}, \\
\tan(\theta_3) &= \frac{s_1^3}{(x_3 - s_3^1)}, \\
\tan(\theta_4) &= \frac{s_2^1}{s_2^2}, \\
\tan(\theta_5) &= \frac{s_2^2}{(x_2 - s_2^2)}, \text{ and} \\
\tan(\theta_6) &= \frac{s_3^1}{(x_3 - s_3^1)}.
\end{align*}
\]

These six equations can be expressed as a linear system of equation \(AZ = B\), where \(A\) is a \(6 \times 4\) matrix, \(Z\) is a \(4 \times 1\) vector, and \(B\) is a \(6 \times 1\) vector defined by

\[
A = \begin{bmatrix}
-tan(\theta_1) & 0 & 1 & 0 \\
0 & -tan(\theta_4) & 0 & 1 \\
tan(\theta_2) & 0 & 1 & 0 \\
0 & -tan(\theta_5) & 0 & 1 \\
-tan(\theta_3) & 0 & 1 & 0 \\
0 & -tan(\theta_6) & 0 & 1
\end{bmatrix}, \\
Z = \begin{bmatrix}
s_1^1 \\
s_1^2 \\
s_2^1 \\
s_2^2 \\
s_3^1 \\
s_3^2
\end{bmatrix}, \\
B = \begin{bmatrix}
0 \\
-x_2 \tan(\theta_2) \\
-x_2 \tan(\theta_5) \\
x_3 \tan(\theta_3) + y_3 \\
x_3 \tan(\theta_6) + y_3
\end{bmatrix}.
\]

Fig. 3. Rays for a source and subarray in a reverberant scenario with two walls.

From Fig. 2, by observation, we know source 1 is associated with DOAs \(\{\theta_1, \theta_2, \theta_3\}\) and source 2 is associated with DOAs \(\{\theta_4, \theta_5, \theta_6\}\). In practice, for the three subarrays and two sources case, we have to try \(2^3 = 8\) possible ways of associating each of the two sources with the estimated DOAs \(\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}\). In other words, we need to solve (6) for each of the eight possible sets of these DOAs. For example, suppose we solve (6), using \(\{\theta_1, \theta_2, \theta_3\}\) for source 1 and \(\{\theta_4, \theta_5, \theta_6\}\) for source 2. Then, the norm of the residual \(\epsilon = ||AZ - B||^2\) will be zero, yielding the correct source locations \((s_1^1, s_1^2)\) and \((s_2^1, s_2^2)\) when the locations of the three subarrays are known exactly, and the estimated DOAs are also estimated perfectly. On the other hand, suppose we associate DOA \(\theta_1\) with source 2 and DOA \(\theta_4\) with source 1 and solve (6) with \(\theta_1\) and \(\theta_4\) interchanged, then the residual \(\epsilon\) will be nonzero since there is no set of consistent solutions \((s_1^1, s_1^2)\) and \((s_2^1, s_2^2)\) satisfying (6). Under ideal conditions, only one of the residuals will be zero and the other residuals will yield nonzero residuals. In practice, with imperfect DOA estimations and uncertainties in the subarray locations, the solution of \((s_1^1, s_1^2)\) and \((s_2^1, s_2^2)\) associated with the minimum residual is used as the best LS criterion source locations.

Clearly, the above considered LS technique for source localization proposed for dealing with multiple real sources and multiple subarrays can equally be applied to multiple real sources and multiple virtual subarrays as proposed in Section 4. In the example to be considered in Fig. 3 with two reflecting walls, there is one source and four equivalent subarrays (one real and three virtual). As the number of real sources, reflecting surfaces, and real subarray increases, the number of equivalent sources, equivalent subarrays, and DOAs increase rapidly. The ability of each subarray to resolve these DOAs also becomes critical. We show the operation of this proposed LS method in Section 7 for various open-field and reverberant scenarios.
order of reflections grows rapidly with increasing the number of surrounding walls. For example, in a simple case of having two perpendicular walls, we have at most second order reflection and, so, for each subarray, there would be four rays (i.e., the direct path ray, the multipath ray reflected only from wall 1, the multipath ray reflected only from wall 2, and the multipath ray reflected from reflection on wall 1 and wall 2 consecutively). In the “Virtual Array Model,” we find the mirrored image of the subarray and its corresponding rays with respect to each of the surrounding walls. As the result of this imaging, we would have some virtual subarrays and also some virtual rays in addition to the real subarray and real rays. The number of virtual subarrays is also dependent on the number of surrounding walls. In this model, we assume the location of the real subarray is known and the DOAs are estimated by the AML algorithm. Our goal is to locate the source. Suppose we are in a reverberant scenario with only two walls, the left wall (wall 1) and the below wall (wall 2) and we have only one source at $(x_1, y_1)$. Fig. 3 shows the virtual subarrays and their corresponding virtual rays in such a condition. The direct path DOA at the real subarray toward the source is denoted by $\theta$. The multipath DOA which is reflected on wall 1 is denoted by $\theta_1$ and the multipath DOA which is reflected on wall 2 is denoted by $\theta_2$. Finally, the multipath DOA which is reflected from wall 1 and wall 2 consecutively is denoted by $\theta_3$. Although each subarray has four rays (four real rays for the real subarray and four virtual rays for each of the three virtual subarrays), only one consistent ray from each subarray passes through the location of the source. The consistent rays for our two wall scenario is shown in Fig. 4. Because we assumed the location of the real subarray is known, then the location of the virtual subarrays is also known. So, by estimating the set of the DOA angles $\theta, \theta_1, \theta_2, \theta_3$, we can write the linear equation for each of the consistent rays. Note that the tangent of the consistent ray for the real array, virtual array 1, virtual array 2, and virtual array 3 is $\theta, (\pi - \theta_1), (2\pi - \theta_2)$, and $(\pi - \theta_3)$, respectively. These four linear equations for consistent rays can be expressed in the matrix form of $AZ = B$, where $A$ is a $4 \times 2$ matrix, $Z$ is a $2 \times 1$ vector, and $B$ is a $4 \times 1$ vector defined by

$$A = \begin{bmatrix} -\tan(\theta) & 1 \\ \tan(\theta_1) & 1 \\ \tan(\theta_2) & 1 \\ -\tan(\theta_3) & 1 \end{bmatrix},$$

$$Z = \begin{bmatrix} s_x^1 \\ s_y^1 \end{bmatrix}, \quad B = \begin{bmatrix} y_1 - (x_1 \tan(\theta)) \\ y_1 - (x_1 \tan(\theta_1)) \\ y_1 + (x_1 \tan(\theta_2)) \\ y_1 + (x_1 \tan(\theta_3)) \end{bmatrix}. \tag{7}$$

In practice, because of imperfect DOA estimation, we need to find the location of the source by using the LS solution. From Fig. 4, by observation we know the consistent ray from real array, virtual array 1, virtual array 2, and virtual array 3 are associated with DOAs $\theta, \theta_1, \theta_2$, and $\theta_3$, respectively. In practice, we have to try all different permutations of the estimated angles ($4! = 24$ different possible ways) and choose the final solution based on the minimum residual $\varepsilon = ||AZ - B||$. Note that, based on the quadrant where the source is located relative to the sensor subarray, we need to place some constraints on the vector $Z$; otherwise, an incorrect source location can result just by selecting the minimum residual solution. The necessary constraints on $Z$ are described below based on the quadrant where the source is located relative to the subarray.

**Case 1.** The source is located in the first quadrant relative to the subarray. For this case, we have $0 \leq \theta < \pi/2$, $\pi/2 \leq \theta_1 < \pi$, $3\pi/2 \leq \theta_2 < 2\pi$, $\pi \leq \theta_3 < 3\pi/2$. Thus, the constraint on $Z$ is described as

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} < Z.$$ 

**Case 2.** The source is located in the second quadrant relative to the subarray. For this case, we have $\pi/2 \leq \theta < \pi$, $\pi/2 \leq \theta_1 < \pi$, $\pi \leq \theta_2 < 3\pi/2$, $\pi \leq \theta_3 < 3\pi/2$. Thus, the constraint on $Z$ is described as

$$\begin{bmatrix} 0 \\ y_1 \end{bmatrix} < Z \leq \begin{bmatrix} x_1 \\ +\infty \end{bmatrix}. $$

**Case 3.** The source is located in the third quadrant relative to the subarray. For this case, we have $\pi \leq \theta < 3\pi/2$, $\pi \leq \theta_1 < 3\pi/2$, $\pi \leq \theta_2 < 3\pi/2$, $\pi \leq \theta_3 < 3\pi/2$. Thus, the constraint on $Z$ is described as

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} < Z \leq \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}. $$

**Case 4.** The source is located in the fourth quadrant relative to the subarray. For this case, we have $3\pi/2 \leq \theta < 2\pi$, $\pi \leq \theta_1 < 3\pi/2$, $3\pi/2 \leq \theta_2 < 2\pi$, $\pi \leq \theta_3 < 3\pi/2$. Thus, the constraint on $Z$ is described as
To solve this LS problem with the box boundary constraints, we use quadratic programming. So, for each permutation, we first determine the case number based on the values of four estimated angles, then we apply the corresponding constraints on Z and find the solution by quadratic programming. Sometimes, the imperfection in DOA estimation is such that none of the permutations falls in any of the above cases. In that situation, we use only three of the four estimated angles instead of the four angles and use the same method. We have \( \binom{4}{3} = 4 \) different ways for selecting three angles from the four angles. For each of those permutations, matrix \( A \) would be \( 3 \times 2 \) and vector \( B \) would be \( 3 \times 1 \). Again, the final solution is obtained based on minimum residual criterion. Even if, by considering three angles, it was not possible to find a solution, we have to use two of the four estimated angles with the same method. In that case, \( A \) would be a \( 2 \times 2 \) matrix and \( B \) would be a \( 2 \times 1 \) vector and we would have \( \binom{4}{2} = 6 \) different permutations. So, by this approach, if at least two of the four estimated angles are accurate, we can still locate the source.

5 Description of the Testbed

For low cost and readily availability, we chose COTS iPAQs 3670s [10] with external 11Mbits Orinoco PCMCIA cards to form the integrated sensor nodes. Each has a modest (206 MHz Intel StrongARM-1110) CPU and a moderate memory space (64 MB RAM and 16 MB ROM) for data processing. It has a built-in microphone and a built-in audio codec for acoustic data acquisition (8-48 kHz signed 16-bit integer). In addition, it has a 950 mAh Lithium polymer battery that lasts about 2.5 hours with the wireless communication on. We installed the open-source GNU/Linux Familiar distribution [11] on iPAQs. Together, these features make iPAQ 3760 a convenient platform for developing the acoustic wireless sensor network.

In our testbed, nodes are organized into subarrays. Each subarray forms an independent unit for beamforming. Because of the coherent nature of beamforming, the audio codecs of iPAQs (in the same subarray, at least) must either share the same time reference or be able to convert their time references back and forth in a fine grain. We use the Reference Broadcast Synchronization (RBS) protocol [12] to estimate the clock offsets among the iPAQ CPUs. The reference packets are periodically broadcast through the IEEE 802.11b channels by iPAQs in rotation. By comparing the receiving time of the same reference packet, the clock offsets among CPUs of the receiving iPAQs are estimated. Our testbed achieved a time synchronization accuracy of a few microseconds within each beamforming subarray.

The low-end consumer-grade audio codecs on iPAQ 3760s have large nondeterministic latencies when they are asked to start recording. Simply to invoke the recording command at the same time on all iPAQs does not guarantee obtaining audio data starting from the same time even if all iPAQ CPUs are perfectly synchronized. A solution to this problem is provided by the “audio server” [19]. The audio server is a process that continuously reads samples from the audio codecs, timestamps the audio samples, buffers the most recent 10 seconds of data, and makes that buffer available to user applications through a library function.

In a subarray, although all nodes carry out the acoustic data acquisition tasks, only one node, the Master Node, is designated for the beamforming computation. The data acquisition nodes are modeled as data servers, and the Master Node is modeled as a client that requests acoustic data from the data servers in its subarray. The Master Node first picks a starting time of data in terms of its own clock, then converts it to the time stamps in terms of the local clocks of the data acquisition nodes. The Master Node sends a request with the starting timestamp and the data length to a designated TCP port of the data acquisition nodes. Each data acquisition node continuously listens to the designated TCP port for incoming data requests. When receiving such a request, the data acquisition node will fetch the corresponding data segment from the audio server and then sends the data back to the Master Node. The requested data can be either sound samples or FFT. In the testbed, Master Node requests FFT. Finally, after the Master Node receives all requested data segments, it can perform the beamforming. If many arrays are considered, they need to exchange the estimated angle to perform the triangulation algorithm. In order to do that, one of the Master Nodes, is elected to be Central Node and it is supposed to receive the estimated angle from the other Master Nodes and to perform the triangulation.

6 Computational Times

In this section, we present the computational and transferring times required by the iPAQs to perform various steps of the AML algorithm. See Table 1.

### 6.1 Sampling Time (ST)

In our testbed, we acquired 256 samples from each sensor at 48kHz; the time required to accomplish this operation is \( ST = 256 \times 1/48kHz = 5.33ms \).

### 6.2 FFT Time (FT)

The time required for FFT takes about \( FT = 80ms \). It was performed using an FFT with floating-point variables utilizing the floating-point emulator. However, iPAQs obtains integer samples. Nevertheless, we chose to use a floating-point algorithm. Up to now, we chose accuracy

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Time Name</th>
<th>1 iPAQ</th>
<th>Subarray</th>
<th>Master N.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>5.33ms</td>
<td>5.33ms</td>
<td>/</td>
</tr>
<tr>
<td>FT</td>
<td>80ms</td>
<td>80ms</td>
<td>/</td>
</tr>
<tr>
<td>DT</td>
<td>3ms</td>
<td>12ms</td>
<td>parametric</td>
</tr>
<tr>
<td>AT</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>AS</td>
<td>/</td>
<td>1ms</td>
<td>/</td>
</tr>
<tr>
<td>CT</td>
<td>/</td>
<td>/</td>
<td>1ms</td>
</tr>
</tbody>
</table>
instead of computational efficiency. In the final version of the testbed, the FFT code will use integer code. It should be remarked that each sensor computes the FFT at the same time.

6.3 FFT Data Transferring (DT)

After the computation of the FFT, each sensor transfers it to the master node. Obviously, the required time is proportional to the data length. When the 256 raw samples are applied to the FFT, the result is 256 complex variables. Each complex value is represented by two float variables of six bytes each. In future implementations, we will only transmit the positive frequency components since the FFT of a real sequence has symmetry about DC. In the current implementation, data length is about 3kB. It takes about 3kB/(11Mb/s) \(\approx 3\) ms with a wireless IEEE 802.11b connection.

While the FFTs are performed at the same time by each sensor of the subarray, the data transferring to the master node is sequential.

6.4 AML Time (AT)

The actual time required for the AML operation is affected basically by three parameters. The first is the number of sources \(M\) we need to localize at the same time. As previously discussed, AML can estimate more than one DOA angle. The second parameter that affects \(AT\) is the number of bins \(N_B\) that is the number of the most active bins considered for the AML. We used \(N_B = 20, 30, 40,\) and \(50\). The third parameter is the search angle step size \(\Delta_a\). The smaller the minimum step size considered, the longer the computational time. In this paper, we used \(\Delta_a = 1, 2, 3, 4,\) and \(5\) degrees.

Figs. 5, 6, 7, and 8 show some of our earlier measured system performance results. They show the computation time as a function of the angle precision for different values of \(N_B\). So far, we used a floating-point-based code in the most part of the code. Fig. 5 shows interesting results. We can see that the delay due to the AML computation for one source is equal to 4.0 seconds for \(N_B = 20\) bins and angle accuracy equal to 5 degrees. With such settings, the system is not very accurate if the source has a large number of frequency bands. We argue that it could be reasonable if the source has some (few) characteristic frequencies.

6.5 Angle Sending Time (AS)

When the master nodes have performed the AML algorithm, they send the estimated angle(s) to the central node. The Angle Sending time (AS) denoted by \(AS \approx 1\) ms is the sum of all the transmissions. It is only performed over the wireless channel.

6.6 Bearing Crossing Time (CT)

The central node is assumed to know the location of the subarrays, so it can estimate the location of the the source(s) by crossing the estimated DOAs. The bearing crossing Time (CT) takes about \(CT \approx 1\) ms.

6.7 Some Example of the Entire Computation

Our preliminary testbed can localize one audio source at least in 4.1 seconds with 5 degrees of angle search step size and 20 bins. In order to have greater accuracy and/or to localize a greater number of sources, the code needs to be

---

Fig. 5. Computation time as a function of search angle step size (with \(N_B = 20\)).

Fig. 6. Computation time as a function of search angle step size (with \(N_B = 30\)).

Fig. 7. Computation time as a function of search angle step size (with \(N_B = 40\)).
modified for fixed point computations. We plan to do this in the future.

7 RESULTS

In order to verify the AML algorithm and the wireless acoustical testbed, we perform several outdoor measurements with different sources and array configurations in free-space and reverberant scenarios. In the following, we show different array and subarray geometries. The first three cases are for free-space and the last case is for a reverberant scenario. We use only 256 samples (sampling rate of 48kHz, 16 bits per sample). The choice of 256 samples is to not overload the processor of each master node and to run the system as fast as possible. For each geometry, localization results are reported.

7.1 Triangular Subarray Configuration

Fig. 9 shows the configurations of the subarrays we consider for the far-field bearing crossing source localization. The subarrays are deployed on plastic tables of the same height (at 0.50m from the floor). The first measurement set is conducted with triangular sensor subarrays as shown at Fig. 9. The distance of the sensors is 0.30m. A low frequency source recording of a tracked vehicle sound is played at four different source location, S1, . . . , S4, sequentially.

For the experimental setting depicted in Fig. 9 (relative to each subarray), the DOA of the source is independently estimated in each subarray and the bearing crossing is used to obtain the location estimate. Fig. 10 shows results at the four locations. We note better results are clearly obtained when the source is inside the convex hull of the overall array.

7.2 Linear Subarray Configuration

In the configuration of Fig. 11, we place linear subarray with three sensors positioned 0.3m far away from each other. The tables are at the same height as that of the triangular subarray setup. Three bearing estimates locate six different locations of the source, S1, . . . , S6, which play recorded organ music, with 2kHz bandwidth and 1.75kHz central frequency.

Fig. 12 shows the one-snapshot results of the two algorithms at the six locations. The RMS error calculation shows similar performance of the first experiment, which
demonstrates that both the AML algorithm and the TDOA-CLS algorithm (proposed in [17]) can locate the source.

7.3 Square Subarray Configuration
The last free-space experimental setting is depicted in Fig. 13, where four square subarrays each with four iPAQs form a single network. As before, the subarrays are set at 0.5 m above the floor on plastic table. Two speakers, one playing the vehicle sound and the other one playing the music sound simultaneously, are placed inside the convex hull of the overall array. Fig. 14 shows the picture of one of the subarray used in this experiment. Fig. 15 shows the results when only speaker 1 is played. When both sources are playing simultaneously, we have the results shown in Fig. 16. Comparing Figs. 15 and 16 we still have quite good results for localizing two sources, but not as good as the performance with only a single source.

Note, as the number of subarray element increases, the localization accuracy of the results reported above improves, which agrees with the Cramer-Rao bound analysis reported in [20].

7.4 Reverberant Scenario (Circular Subarray Configuration)
For the reverberant scenario, we use the configuration in Fig. 17. We consider two perpendicular walls, the left wall (wall 1) and the floor (wall 2). We assume that both walls have a reflection coefficient of 0.9 in amplitude. This means that the walls are quite reflective.

The subarray is composed of eight sensors in a circle with radius $r = 0.5\text{m}$. The sensors are positioned with angles $(0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4)$ on the circle, respectively. Note that, while the other experiments were conducted in a plane parallel to the floor, in the reverberant scenario, the plane is perpendicular to the floor. Since the microphone of iPAQs is built-in inside the body and it is not external, the main body of the iPAQs would block the bouncing ray from the floor. So, we had to conduct the reverberant experiments in a plane perpendicular to the floor.
For the simulation, the source is placed in the same plane as the subarray, but at different angles relative to it. The centroid of the subarray is at \(\frac{1}{2} \times 3\) m and the distance between source and subarray is \(2\sqrt{3}\) m. The different angles that we have used for placing the source relative to the subarray are 0, \(\frac{4\pi}{6}\), and \(\frac{4\pi}{3}\) radians.

For the reverberant experiments, we considered eight different positions of the subarray and the source, but the centroid of the subarray is always at \(h = 1.3\) m from the ground. Fig. 18 shows the circular subarray that we used for our experiments. As discussed previously, four rays are considered here. The first is the direct ray. The second is the ray that bounces off the floor. The third is the ray that bounces off the wall (the subarray stands between the source and a wall) and the fourth is the ray that bounces off the floor and then on the wall before reaching the subarray.

The source we used is the tracked vehicular source. For these parameters, simulation and testbed results are shown below.

### 7.4.1 Reverberant Simulation Results

For each of the three cases, we estimated the location of the source for different SNRs. The number of iterations for each case is 10. The results are plotted in Figs. 19, 20, and 21, respectively. From these figures, we note that our proposed method can locate the source accurately most of the time. We also note that, by decreasing the SNR, the accuracy of the source localization is decreased as expected.
7.4.2 Reverberant Real-Time Testbed Results

Now, we consider eight reverberant real-time testbed results. For each of the eight cases, the results are for 40 iterations. For Experiments 1, 2, and 3, the localization results are shown in Fig. 22. Here, we place the centroid of the subarray at $d_c = 2.5m$ away from wall 1 and $h = 1.3m$ high. The source is $d_s = 2.5m$ away from the centroid of the subarray at $1m$, $1.3m$, and $1.7m$ high, respectively, for Experiments 1, 2, and 3. The accuracies of all three experiments are quite good.

For Experiments 4, 5, and 6, the localization results are shown in Fig. 23. We kept the centroid of the subarray at the same location of the previous experiments ($d_c = 2.5m$ and $h = 1.3m$). The source is moved further away to $d_c = 3.5m$ from the centroid and at $1m$, $1.3m$, and $1.7m$ high, respectively, for Experiments 4, 5, and 6. The accuracies of all three experiments are slightly worsened but are still quite good.
In Experiments 7 and 8, with results shown in Fig. 24, we moved the subarray further away (at $d_c = 3.5m$) from wall 1, keeping the height the same as before ($h = 1.3m$), but placed the locations of the source the same as those of Experiments 5 and 6, respectively ($d_s = 2.5m$ at 1.3m and 1.7m high, respectively). The error dispersions now are even lower than those in Experiments 5 and 6. Again, these results are reasonable since the subarray encounters fewer reverberations.

### 7.4.3 Comments on the Experimental Results

The results presented above demonstrated the effectiveness of our proposed AML algorithm and testbed. It should be noticed that the testbed is working quite well even in the controlled limited reverberant scenario, while the radius of the subarray is quite large, but the distances from the source to the sensors are quite small. In practice, the sensing area clearly needs to be much larger than that considered here. However, to perform controlled reverberation experiments with two (or more) large perpendicular walls will be difficult. At least for the distances considered here, no calibration was performed on these low-cost microphones in the iPAQs, which demonstrated the robustness of the algorithms. For larger distances, we may need to use better quality microphones and possibly need to calibrate/equalize these microphones.

**8 Conclusions**

In this paper, we first presented the AML algorithm for DOA estimation. Then, we showed the LS method applied to the estimated angles can be used for source localization in free-space. Upon introducing a novel virtual array model, source localization in the reverberant case was also shown to be feasible. Then, details on the testbed hardware and processing algorithms were given. Finally, extensive results on subarray configurations and measured results for free-space and controlled limited reverberant scenarios were presented.

The above results on the algorithms and the testbed clearly demonstrated their proper real-time acoustical beamforming operations in limited scenarios. In order to extend their operations to longer ranges and more realistic reverberant conditions with multiple sources, we need to solve various additional fundamental and practical issues. As mentioned in the last section, in order to increase the range of the array, higher quality microphones with tighter control of the parameters of these microphones need to be used. To operate in a more general reverberant condition, knowledge of the parameters of the dominant reflecting rays needs to be estimated by adaptive probing of the environment. In order to increase the throughput of the system, more capable floating-point processors (such as those in DSP processors) may need to be used since some of the existing throughput bottlenecks have been in the software emulation of floating-point operations in the iPAQs. Whether these needed requirements for real-time acoustical beamforming can be obtained from existing latest single unit iPAQ-like devices, instead of using multiunit systems (with discrete microphones, codecs, DSPs, radios, etc.), remains unclear and poses a challenge for future work in this area.
ACKNOWLEDGMENTS

This work is partially supported by the US National Science Foundation CENS program, the Intel Corporation, and AROD-MURI PSU Contract 50126.

REFERENCES


Pierpaolo Bergamo received the PhD degree in information engineering in 2004 from the University of Ferrara, Italy. He was a visiting researcher at the Electrical Engineering Department at the University of California at Los Angeles (UCLA) with Professor Kung Yao. He is currently a postdoctoral researcher at UCLA in the Computer Science Department with Professor Mario Gerla. His research interests include ad hoc wireless networks and sensor networks, MAC protocols, routing protocols, and security issues.

Shadnaz Asgari received the BS degree in electrical engineering from Sharif University of Technology in 1999. Since 2002, she has been a graduate student at the University of California, Los Angeles (UCLA) where she is currently doing research under Professor Yao’s supervision. Her areas of interests include array signal processing, source localization, and beamforming in sensor networks.

Hanbiao Wang received the BS degree in geophysics from the University of Science and Technology of China (USTC), Hefei, China, in 1994, the MS degree in geophysics and space physics and the MS degree in computer science from University of California, Los Angeles (UCLA) in 2000 and 2002, respectively. He is currently a PhD candidate in the Computer Science Department at UCLA and will receive the PhD degree in June 2005. His dissertation is on scalable wireless sensor networks for collaborative acoustic monitoring. His research interests include wireless network, distributed system, information fusion, and statistical signal processing.

Daniela Maniezzo received the PhD degree in information engineering in 2004 from the University of Ferrara, Italy. In 2002, she joined the research group of Professor Kung Yao at the Electrical Engineering Department, University of California at Los Angeles (UCLA). She is currently a postdoctoral researcher at UCLA in the Computer Science Department with Professor Mario Gerla. Her current research interests include the design of MAC and routing protocols for wireless ad hoc and sensor networks as well as power control, radio channel model, and propagation environment problems.

Len Yip received the BS degree and MS degree in physics from Zhongshan University, Guangzhou, China, in 1993 and 1996, respectively. He received the MS degree in electrical engineering from the University of California at Los Angeles in 2000. He is currently pursuing the PhD degree in the Department of Electrical Engineering, the University of California at Los Angeles. His research interests include sensor array processing and statistical signal processing. He is a student member of the IEEE.

Ralph E. Hudson received the BS degree in electrical engineering from the University of California at Berkeley. He received the PhD degree in optimal estimation and control from the US Naval Postgraduate School, Monterey, California. Then, he served as a lieutenant commander defining requirements, evaluating proposals, and accelerating acquisition of electronic warfare systems at the US Naval Air Systems Command, Arlington, Virginia. Afterward, he was a chief scientist at Hughes Aircraft Company, El Segundo, California (1973-1993), designing air to ground radar systems for U-2, F-15E, and B-2 aircrafts. Since 1993, he has been a research associate at the University of California at Los Angeles performing research in signal processing in acoustics and seismic sensor arrays and spatial diversity for wireless data transmission.
Kung Yao received the BS, MA and Ph.D. degrees, all in electrical engineering, from Princeton University, Princeton, New Jersey. He is a professor in the Electrical Engineering Department at the University of California at Los Angeles. His research and professional interests include sensor system, digital communication theory and systems, fading channel statistics, digital and array processing, and systolic algorithms and architectures. Dr. Yao received the IEEE Signal Processing Society’s 1993 Senior Award in VLSI Signal Processing. He is a fellow of the IEEE.

Deborah Estrin received the BS degree in electrical engineering and computer science from the University of California at Berkeley (1980), and the PhD degree in computer science from the Massachusetts Institute of Technology (1985). She is a professor of computer science at the University of California at Los Angeles and director of the Center for Embedded Networked Sensing (CENS), a US National Science Foundation Science and Technology Center awarded in 2002. She was on the faculty of computer science at the University of Southern California from 1986 through mid-2000. Dr. Estrin has been instrumental in defining the national research agenda for wireless sensor networks, first chairing a 1998 US DARPA ISAT study and then a 2001 NRC study. Her research group develops algorithms and systems to support robust networks of physically-embedded devices, with a particular focus on environmental monitoring. She is a fellow of the ACM, the IEEE, and the AAAS.

For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.