Design and Testing of Robust Acoustic Arrays for Localization and Enhancement of Several Bird Sources

*Electrical Engineering Department, UCLA, Los Angeles, CA 90095
Email: yao@ee.ucla.edu
†Computer Science Department, UCLA, Los Angeles, CA 90095
Email: hbwang@lecs.cs.ucla.edu
‡Ecology and Evolutionary Biology Department, UCLA, Los Angeles, CA 90095

Abstract—Sensor network technology can revolutionize the study of animal ecology by providing a means of non-intrusive, simultaneous, unmanned monitoring. In this paper, we investigate the design, analysis, and testing of acoustic arrays for localizing bird vocalizations of different species. The spectra of the bird waveforms affect the desired dimension of the array. Microphones are placed in a uniform circular array and are finely synchronized within a few microseconds. We apply the Approximate Maximum Likelihood (AML) method to estimate the source direction-of-arrival (DOA) and perform beamforming for signal enhancement. The crossing of the distributed DOA estimated bearings is used to localize the birds. The enhanced signals are used for training and estimation for the classification of the birds. The experimental results demonstrate the practicality and robustness of our array design.

I. INTRODUCTION

Recent emergence of distributed sensor network technology has dramatically boosted our capability to monitor the state of the complex physical world [1], [2]. The study of ecology is one of the many fields that can benefit from this technology. Ecologists have long been seeking a non-intrusive way of monitoring animal behaviors and diversity. By using acoustic sensor networks, researchers can passively capture the natural vocalization of animals and study the interaction of multiple individuals in their habitats. In this paper, we investigate the design, analysis, and testing of acoustic arrays for localizing and beamforming bird vocalizations of different species.

Acoustic sensor networks have been proposed for animal localization due to its advantage of non-intrusiveness. Most existing work on passive acoustic localization of animals is about localizing cetaceans [3]. Fewer efforts have been devoted to passive acoustic localization of animals in terrestrial habitats. One of the highlights of such early efforts is empirical studies about the accuracy of passive acoustic localization of birds in open meadow and woodland habitats [4]. In that study, the TDOA estimates between microphone pairs are first obtained through cross-correlation method, and source localization is performed using bearing crossings of hyperbolae that correspond to TDOA differences between microphone pairs. Because cross-correlation method fails when there are simultaneous sounds from multiple sources, this acoustic localization system only works when there is only single vocalization present at the moment of observation.

An automated acoustic system for bird monitoring and classification has also been proposed in the literature [5] to avoid collisions between aircrafts and birds. In their work, the wideband acoustic signal is divided into many narrowbands and then the MUSIC [6] algorithm is applied on those narrowbands. The DOA estimation for the wideband signal is generated by combining estimated results from all the narrowband components. Due to the advantages of MUSIC algorithm, this system can work with multiple simultaneous sources. Acoustic sources beamforming has also been performed after DOA information is acquired. However that system requires a very large number of microphones (range from 64 to 200), high speed CPU and huge memory. These hardware requirements result in huge power consumption and make this system not suitable for bird behavior monitoring in the wild which is far away from any power supply infrastructure.

In our design of the acoustic source localization system, we choose the Approximate-Maximum Likelihood (AML) algorithm for both DOA estimation and beamforming of wideband sources. As is discussed in previous publications [7], the AML algorithm is optimal in the ML sense when the size of the FFT frame is large. Similar to MUSIC, the AML algorithm also has the capability of estimating DOAs of multiple simultaneous acoustic sources. Comparing to other’s work [5], our system contains less microphones and power consumptions and therefore is more suitable to deploy in the birds’ natural habitats. To test the practicability and robustness of our system, we have conducted field experiments in different environments. In the experiment, bird vocalizations are played back from a computer speaker as the acoustic sources, and DOA of bird sources are estimated within each circular array. Each circular array contains 4 or 8 microphones and a Presonus Firepod recording system is used to synchronize the signals from each microphone. After DOA is estimated, source localization is performed by applying least-square fit to the bearing crossings and the waveforms of each acoustic source is reconstructed through beamforming. In this preliminary investigation, we consider only 2D source localization.

In this paper, we demonstrate critical issues in designing acoustic sensor network for bird monitoring and also suggest different sizes and geometries for different bird species. Beamforming results show that we can successively separate and enhance the SNR of several simultaneous bird vocalizations. Several localization experiments under different environments are conducted and compared.

The paper is organized as follows: Section II reviews the AML algorithm. Section III discusses the design issues on the array sizes and number of microphones in each array. Section IV discusses the signal enhancement and separation through beamforming technique. Section V demonstrates the experimental results of DOA estimation, source localization in different environments, and beamforming. Section VI concludes this paper.

II. AML DOA ESTIMATOR

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 the sensor array lie in the same plane (a 2-D scenario) as shown in Fig. 1.
Let there be $M$ wideband sources, each at an angle $\theta_m$ 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 omni-directional and have 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 $m$th source is given by 

$$t_{cp}^{(m)} = t_{c}^{(m)} - t_{p}^{(m)} = [(x_c - x_p) \cos \theta_m + (y_c - y_p) \sin \theta_m] / v, \quad \text{where} \quad (t_c^{(m)}) \text{and} \ (t_p^{(m)}) \text{are the absolute time-delays from the} \ (m) \text{th} \ \text{source to the} \ (m) \text{th} \ \text{sensor and the} \ (m) \text{th} \ \text{sensor respectively, and} \ v \ \text{is the speed of propagation.}$$

In a polar coordinate system, the above relative time delay can also be expressed as $t_{cp}^{(m)} = r_p \cos(\theta_m - \phi_p) / v$, where $r_p$ and $\phi_p$ are the range and angle of the $p$th sensor with respect to the array centroid. The data received by the $p$th sensor at time $n$ is then

$$x_p(n) = \sum_{m=1}^{M} S^{(m)}(n - t_{cp}^{(m)}) + w_p(n), \quad \text{for} \ \ n = 0, ..., N - 1, \ \ p = 1, ..., P,$$

where $N$ is the length of the data vector, $S^{(m)}$ is the $m$th source signal arriving at the array centroid position, $t_{cp}^{(m)}$ is allowed to be any real-valued number, and $w_p$ is the zero mean white Gaussian noise with variance $\sigma^2$.

For the 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 $k = 0, ..., N - 1$, where the array data spectrum is $X(\omega_k) = [X_1(\omega_k), ..., X_P(\omega_k)]^T$, the steering matrix $D(\omega_k) = [d^{(1)}(\omega_k), ..., d^{(M)}(\omega_k)]$, the steering vector is given by $d^{(m)}(\omega_k) = [d_1^{(m)}(\omega_k), ..., d_P^{(m)}(\omega_k)]^T$, $d_p^{(m)} = e^{-j2\pi k t_{cp}^{(m)} / N}$, and the source spectrum is given by $S(\omega_k) = [S^{(1)}(\omega_k), ..., S^{(M)}(\omega_k)]^T$. The noise spectrum vector $\eta(k)$ 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 $^*$ as the transpose, and $H$ 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 [7]

$$\max_{\Theta, S} L(\Theta, S) = \min_{\Theta, S} \sum \limits_{k=1}^{N/2} || X(\omega_k) - D(\omega_k)S(\omega_k) ||^2,$$

which is equivalent to a nonlinear least square problem. Using the technique of separating variables [8], the AML DOA estimate can be obtained by solving the following likelihood function

$$\max \limits_{\Theta} J(\Theta) = \max \limits_{\Theta} \sum \limits_{k=1}^{N/2} \text{tr}(P(\omega_k, \Theta)R(\omega_k)),$$

where $P(\omega_k, \Theta) = D(\omega_k)D^\dagger(\omega_k)$, $D^\dagger = (D(\omega_k)^HD(\omega_k))^{-1}D(\omega_k)^H$ is the pseudo-inverse 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

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

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 eq. (4). In the multiple sources case, the computational complexity of the AML algorithm requires multidimensional search, which is much higher than the MUSIC type algorithm that requires only 1-D 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 [9].

III. DESIGN OF ACOUSTIC SENSOR ARRAYS FOR DOA ESTIMATION OF BIRD VOCALIZATIONS

Most studies of animal behavior focus on one species or just a few species. As a result, it is usually possible to restrict attention to a limited range of the sound spectrum. Restricting attention in this manner will generally permit the arrays to be designed to perform better DOA, localization and beamforming. Accordingly, for this study we have made several critical assumptions to simplify the design process.

- As shown in Fig. 2, high percentage of the energy density falls within a band; hence we can safely filter the unwanted frequency outside the band to remove noise or possible interferer.
- We choose to use 2D array geometry since their regional location is far more important than locating them in space.
- We choose the Uniform Circular Array (UCA) configurations to ensure the beampattern as symmetric as possible with source angle since we do not have prior information of the source location.

The above assumptions lead us with two design criteria: the number of sensors and the array size.
A. Review of Linear Array Design for Narrowband Signals

In beamforming of narrowband signals with wavelength $\lambda$ using a linear array with sensors regularly spaced at an inter-sensor spacing of $a$, grating lobes will appear in the beam pattern and cause ambiguity in DOA estimation when $a > \lambda/2$. This is spatial aliasing of narrowband beamforming. On the other hand, if $a < \lambda/2$, there is no grating lobe, but the main lobe will become large, hence we are losing spatial resolution. From these two properties, we know that there is a tradeoff between the spatial resolution and the grating lobes by varying the size of the array.

If we fix the inter-sensor spacing of $a$ to be $\lambda/2$ and increase the number of sensors in a linear array, the gratings lobe will not appear, but the mainlobe becomes thinner. This means we gain spatial resolution without having grating lobes that pose ambiguity in DOA estimation. Note that this is the same as increasing the spacing distance of the original array and then adding microphones in between the arrays to make $a = \lambda/2$. We just noted that increasing the array size will reduce the mainlobe width, but increase the grating lobe. Then adding the microphones causes the reduction of grating lobe. The latter can be understood as adding more sampling points in space to remove the spatial aliasing problem.

B. Array Size and Number of Sensors Design for Bird Vocalizations

Wideband signals have non-zero power spectral density across a range of frequencies. In wideband beamforming, the grating lobes for different frequencies are in different locations and thus averaged out whether the array is linear and how large is the sensor spacing. In this ideal wideband case, the grating lobe effect due to a single tone will not be observed due to the averaging nature; however, the non-uniformity of the power spectral contents of bird vocalizations makes some grating lobe effects due to various frequency contents more difficult to average out in beamforming using large microphone spacing. This is confirmed in [10].

In Acorn Woodpecker (Melanerpes formicivorus) and Mexican Antthrush (Formicarius moniliger) vocalizations, both bird vocalizations are wideband signals as shown in Fig. 2. However, these vocalizations do not have comparable power spectral density across a wide range of frequency. Instead, the power spectral density is more like a combination of a small number of narrowband signals. In Woodpecker vocalization, they centered around 3.4 KHz, 2.5 KHz, 1.7 KHz, and 2.1 KHz, in descent order of power, and in Antthrush vocalization they is only one peak at 2.0 KHz.

![Power spectral density](image)

(a) Acorn Woodpecker.  (b) Mexican Antthrush.

Fig. 2. Power spectral density.

In real habitats, there are always additional ambient noise and multipath effects that can cause acoustic signals to degrade. Signal degradation can distort lobe heights and lobe positions in beam pattern. Such beam pattern with large side lobes could easily cause gross DOA estimation error when side lobes grow larger than the main lobe as observed in [10]. One way to solve this problem is to follow the half-wavelength guideline of narrowband beamforming for choice of sensor spacing. As shown in Fig. 2, we can model a Woodpecker vocalization as combination of 4 narrowband signals centered at 3.4 KHz, 2.5 KHz, 1.7 KHz, and 2.1 KHz. In order to completely remove the side lobe in the beam pattern of the combined signal, we have to remove the side lobe of each component’s beampattern. From the half-wavelength sensor spacing guideline, there is no side lobe when sensor spacing $a < (345 \text{ m/s} / 3.4 \text{ KHz}) / 2$. That is, $a < 5 \text{ cm}$. To be conservative, we choose 4 cm microphone spacing for our acoustic array design, and in radius $r = 4/\sqrt{2} = 2.83 \text{ cm}$. Another way to solve this problem is to put a constraint on the grating lobe peak to be at most 20% of the mainlobe and find the optimum array size by means of simulation. To do this, we simulate the beampattern with different sizes and select the largest array size that satisfy the constraint. The Mexican Antthrush optimum array size is performed this way as seen in Fig. 3d. The Antthrush can also be modeled as a narrowband signal centered at 2.0 KHz and by using the half-wavelength sensor spacing guideline, we have $a < 8.6 \text{ cm}$, which translates to array with radius $r = 6.1 \text{ cm}$. Fig. 3c shows that it still contains sidelobes that are greater than 20% of the mainlobe. Combining the two methods, we can use the first method to provide good initial estimate, and the second method to search the optimum array size.

![Beampatterns](image)

(a) Woodpecker, $r = 7.07 \text{ cm}$.  (b) Woodpecker, $r = 2.83 \text{ cm}$.  (c) Antthrush, $r = 6.10 \text{ cm}$.  (d) Antthrush, $r = 4.24 \text{ cm}$.

Fig. 3. Simulated 4 sensor array beampattern with radius $r$ and source at 60 degrees.

Carrying our intuition from the linear array, we can increase the number of sensors to gain spatial resolution and keep the sidelobe peak within a constraint. This can also be seen as a two-step process. The first step is to increase the size of the array to reduce the mainlobe width as seen in Fig. 3a-b; the next step is to add microphones to remove spatial aliasing in the beampattern. To keep the radius of the array constant, we can think the 4 element array as points in a circle, and adding more sensors is essentially adding points in the circle.
as shown in Fig. 4b. As in the case of 4 sensors, the relationship between sensors distances and frequency is not as simple as the half-wavelength, but we can use the simulation method to find the optimum array size.

![Diagram of sensor array](image)

(a) Woodpecker, r = 7.07. (b) Diagram of sensor array.

Fig. 4. Simulated 8 sensor array beampattern with radius r and source at 60 degrees.

We also perform the array design analysis on several other bird vocalizations to confirm that the optimum array sizes are different for different species, and the result is tabulated in table I.

<table>
<thead>
<tr>
<th>Bird Name</th>
<th>4 sensor array radius</th>
<th>8 sensor array radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acorn Woodpecker (Melanerpes formicivorus)</td>
<td>2.83 cm</td>
<td>7.07 cm</td>
</tr>
<tr>
<td>Male Dusky Antbird (Ceracomacra tyrannina)</td>
<td>3.18 cm</td>
<td>6.36 cm</td>
</tr>
<tr>
<td>Female Dusky Antbird (Ceracomacra tyrannina)</td>
<td>3.18 cm</td>
<td>6.36 cm</td>
</tr>
<tr>
<td>Mexican Antthrush (Formicarius moniliger)</td>
<td>4.24 cm</td>
<td>8.48 cm</td>
</tr>
<tr>
<td>Barred Antshrike (Thamnophilus doliatus)</td>
<td>4.95 cm</td>
<td>10.61 cm</td>
</tr>
<tr>
<td>Great Antshrike (Tara major)</td>
<td>7.07 cm</td>
<td>14.85 cm</td>
</tr>
</tbody>
</table>

**TABLE I**

**OPTIMUM ARRAY SIZE FOR VARIOUS BIRDS.**

Adding more microphones allows us to use a larger array with the same peak main lobe to side lobe ratio, and thus better resolution. However, more microphones also demand larger memory, faster CPU, and more power consumption and may not be easily met in the field. The analysis presented in this section provides excellent guidelines to trade-off between resolution and hardware requirements.

IV. SIGNAL SEPARATION AND ENHANCEMENT THROUGH BEAMFORMING

Beamforming is a procedure of coherently combining the received signals from spatial distributed sensor nodes. It provides an array gain that increases the output SNR in proportion to the number of sensor elements. For the past few decades, many DOA estimation and beamforming algorithms have been proposed for the framework of narrow band sources in the far-field of the sensor array. There are other applications of interest such as acoustic and seismic signals in which the sources are often considered as wideband. For these applications, algorithms for wideband beamforming are required. Wideband beamforming can be applied in either time-domain or frequency domain. For time-domain beamformers, TDOA information needs to be acquired first, and spatial filtering is applied to linearly combine signals from each sensor. For frequency-domain beamformers, the array is steered to the estimated DOA direction so that the signals are coherently combined for each frequency bin. Our AML beamformer falls in this category, and is implemented by eq. (5).

In a single source scenario, a time-domain beamformer and our AML beamformer can work equally well given that the TDOA/DOA estimates are fairly accurate. However, in multiple sources scenario, especially for the case when the spectra of the multiple sources overlap, the time-domain beamformer often performs poorly. In our application of collecting acoustic data for bird classifications, there are all sorts of potential acoustic sources in the forest, and beamforming cannot be performed without an interference cancellation mechanism. The proposed AML beamformer has superior performance under this scenario since it has the capability of beamform multiple sources by exploiting the information of spatial separation among sources. When the algorithm performs beamforming toward each DOA, the signal of interest is enhanced while the interference from other sources are eliminated by placing nulls at all the other sources’ DOA directions. In the following paragraph, we present the simulation results of separating two acoustic sources [7]. Generalization to cases of more than two sources is straightforward.

Consider two acoustic sources located in the far-field of the sensor array. One is a Woodpecker lying at 60 degree, and the other is a male Dusky Antbird (Cercomacra tyrannina) lying at 180 degree, all with respect to the array. Fig. 5 shows typical bird vocalizations of these two species. The two bird vocalizations are normalized to have the same power, and the simulation is performed under SNR=10dB. In this simulation, a circular array of radius 7.07 cm with 8 microphones evenly placed around the perimeter. The design of the array has taken the spectrum of two bird sources into account and conforms to the analysis discussed in section III. The two bird sources generates its own vocalizations at the same time, and our objective is to reconstruct the original bird vocalizations through beamforming from the mixed waveform (Fig.6a).

![Time domain waveforms and power spectrum of two bird vocalizations](image)

(a) Acorn Woodpecker. (b) Dusky Antbird.

Fig. 5. Time domain waveforms and power spectrum of two bird vocalizations.

In order to steer our array toward the direction of two acoustic sources, DOA of each bird needs to be estimated first. We follow the AML DOA estimation procedure described in section II. Here we show the DOA estimate of 2D search for better illustration. Other computational efficient variants of AML algorithm converge to the same DOA results. Fig.6b shows the maximum-likelihood metric parameterized by two DOAs. The brightest spot indicates the maximum-likelihood estimates of DOAs are 60 and 180 degrees.

We then apply frequency beamforming and steer the circular array toward these two DOAs. High pass filtering is also applied after the beamformers to remove low frequency noise components. Simulation results show excellent signal separation (Fig. 7).
V. EXPERIMENTAL RESULTS

In our experiments, we use Presonus Firepod 8 channels A/D to collect the data and record them into a hard drive. This Firepod can sample at 44100 Hz, and two Firepods can be synchronized to support 16 channels simultaneously. The synchronization error is less than 65 µs which is neglectable in our applications.

A. Beamforming Experiments

In this subsection, we present our beamforming design effectiveness with experiment at UCLA Science Courtyard. We use a computer speaker located at 60 degree and spaced 10 m from the array (Fig. 9). Fig. 8a shows the Woodpecker beampattern using 8 sensors with radius 7.07 cm with DOA = 59 degree. The mainlobe is very similar to the simulated experiment in Fig. 4a, but the sidelobe is increased. In Fig. 8b, we show the Mexican Antthrush bird vocalization with 4 sensors, array radius 4.94 cm and DOA = 66 degree. This array size is slightly greater than the optimum simulated size listed at table I, but the sidelobe is tremendously larger and close to the mainlobe. This confirms the necessity to be very conservative with the sidelobe peak.

B. Beamforming Experiments

To demonstrate the effectiveness and practicality of our beamforming algorithm in separating multiple simultaneous bird vocalizations, we conducted the following experiment in UCLA Science Courtyard. The settings of the experiment is similar to what we used in simulations of Section IV. We use two computer speakers which are placed 10 meter away from our 8-element circular array. One speaker generates a typical Woodpecker vocalization from 60 degree, and the other generates a typical Dusky Antbird vocalization from 180 degree. In this experiment, we got DOA1 = 59 degree for Woodpecker source, and DOA2 = 187 degree for Dusky Antbird, as is shown in Fig. 10a. Based on this DOA estimate, we steer our array toward the source bearings and performed our beamforming algorithm. Fig. 10b shows the received waveform of a single microphone after removing low frequency out-of-band noise and Fig.11 shows the beamforming results. The experimental results show that the signals from two acoustic sources are successfully separated. We also observe there is little signal coupling from Woodpecker vocalizations to Dusky Antbird vocalization, which is due to the fact that we have a larger DOA error in that direction.

C. Localization Experiments

In this subsection, we present the results of localization experiments in several environments with very different noise, interference,
and multipath characteristics. In these experiments, each acoustic array is a 4-element circular array with radius 2.8 cm, and we deploy four circular arrays for source localization. A typical Woodpecker vocalization is used as our acoustic source and is playbacked from a computer speaker. Both the speaker and acoustic arrays for data acquisition are raised about 70 cm above the ground. Each acoustic array estimates the DOA of acoustic source, and the source location is determined by applying the LS method to the cross-bearing of four DOA estimates.

To test the effectiveness of our acoustic localization system under different environments, we have conducted experiments in the following sites:

- Small parking lot at Buckley School in Sherman Oaks, California.
- Science Courtyard at UCLA.
- Woody hill at Buckley School in Sherman Oaks, California.

1) Small parking lot at Buckley School: This site is chosen to study the performance of our localization system in a relatively open environment such as grassland habitats. The speaker, the acoustic arrays, the parking lot, and the surrounding area are shown in Fig. 12(a). Four acoustic arrays are deployed at corners of an 8 m by 8 m square, and the speaker is placed at the center of the square. DOA estimates are indicated by arrows and are shown in Fig. 12(b). The estimated location of the acoustic source is about 28 cm off the true location of the acoustic source.

2) Science courtyard at UCLA: During the experiment in the parking lot of the Buckley School, it was relatively quiet. To test effectiveness of acoustic localization in relatively open environments but with strong ambient noise, we have conducted experiments in the Science Courtyard at UCLA. The Science Courtyard is surrounded by Boelter Hall, Math Building, and Young Hall from three directions. There is strong persistent noise from ventilation systems of Boelter Hall and Young Hall. In addition, it was often very windy during the experiments in the Science Courtyard. The speaker, the acoustic arrays, the courtyard, and the surrounding area are shown in Fig. 13(a). Four acoustic arrays are deployed at corners of an 4 m by 4 m square, and a speaker is placed at the center of the square. DOA estimates are indicated by arrows and are shown in Fig. 13(b). Although there is strong ambient noise, beamforming using 4cm by 4cm square arrays is very robust, and the estimated location of the acoustic source is about 13 cm off the true location of the acoustic source.

The large localization error is largely due to the large DOA error of acoustic array 4 at the upper right corner, which is about 18 degrees off the true DOA while all other array have DOA error less than 8 degrees. We think that the large DOA error of array 4 might largely be due to the false orientation measurement of array 4 that is used to convert DOA estimate from the local coordinate system to the global coordinate system. A large measurement error of array orientation is possible in experiments on the hill of the Buckley School because the terrain is rough and we rely on visual inspection to control the alignment between the compass and small acoustic arrays. In [10], we discussed calibration issues in the acoustic array orientation and observed significant localization accuracy improvements through simple array calibration mechanism.

VI. CONCLUSION

In this paper, we propose an acoustic sensor network design for bird localization. We present methods to design acoustic array beam pattern for robust DOA estimation. We also suggested array geometries that are suitable for different bird species. By using frequency domain beamforming of AML algorithm, we are able to enhance the SNR of our source signal, and exploit the spatial information to separate simultaneous bird vocalizations. Field experiments have been conducted under different environment to study the
The performance of our localization algorithm. The experimental results verify the robustness of the design and suggest its practicality for bird monitoring application.

ACKNOWLEDGMENT

This work is partially supported by NSF CENS program under Cooperative Agreement CCR-0121778, NSF grant EF-0410438, ARO-MURI PSU Contract 50126, and UC-Discovery grant sponsored by ST Microelectronics. We appreciate the assistance of Kathy Griffith, Joe Wise, and Chih-Kai Chen, in conducting various experiments and Alex Kirschel and Yuan Yao in supplying us with the bird vocalization files. We also want to express our appreciation to the Buckley School in Sherman Oaks, California for providing their facility for two of our experiments.

REFERENCES