

Acoustic Sensor Networks for Woodpecker Localization

H. Wang^a, C.E. Chen^b, A. Ali^b, S. Asgari^b, R.E. Hudson^b, K. Yao^b, D. Estrin^a, and C. Taylor^c

^aUCLA Computer Science Department ;

^bUCLA Electrical Engineering Department ;

^cUCLA Ecology and Evolutionary Biology Department

ABSTRACT

Sensor network technology can revolutionize the study of animal ecology by providing a means of non-intrusive, simultaneous monitoring of interaction among multiple animals. In this paper, we investigate design, analysis, and testing of acoustic arrays for localizing acorn woodpeckers using their vocalizations.^{1,2} Each acoustic array consists of four microphones arranged in a square. All four audio channels within the same acoustic array are finely synchronized within a few micro seconds. We apply the approximate maximum likelihood (AML) method³ to synchronized audio channels of each acoustic array for estimating the direction-of-arrival (DOA) of woodpecker vocalizations. The woodpecker location is estimated by applying least square (LS) methods to DOA bearing crossings of multiple acoustic arrays. We have revealed the critical relation between microphone spacing of acoustic arrays and robustness of beamforming of woodpecker vocalizations. Woodpecker localization experiments using robust array element spacing in different types of environments are conducted and compared. Practical issues about calibration of acoustic array orientation are also discussed.

Keywords: Acoustic Localization, Acoustic Beamforming, Sensor Networks, Robust DOA Estimation, Sensor Calibration, Ecological Monitoring, Acorn Woodpecker

1. INTRODUCTION

Recent emergence of distributed sensor network technology could dramatically boost our capability to monitor the state of the complex physical world.^{4,5} The study of ecology is one of many fields that benefit greatly from sensor network technology.^{6,7} Acoustic sensor networks can provide a non-intrusive means to observe animal behavior and diversity^{8,9} by passively listening to animal vocalizations. In addition, interaction of multiple individuals can be simultaneously monitored by acoustic sensor networks. In this paper, we present our preliminary studies about the design and implementation of acoustic array networks that can estimate locations of acorn woodpeckers using their vocalization.^{1,2}

Ecologists have long been pursuing the goal of passive acoustic localization of animals due to its obvious advantage of non-intrusiveness. Most existing work on passive acoustic localization of animals is about localization of cetaceans.¹⁰ 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.¹¹ In that study, the passive acoustic localization system consists of several microphones. Time-difference-of-arrival (TDOA) between each pair of microphones is estimated using cross-correlation between their sound waveforms. The bird location is estimated using cross-bearing of hyperbolas that correspond to TDOAs between microphone pairs. Because cross-correlation fails in TDOA estimation when there are simultaneous sounds from multiple sources, this acoustic localization system works only when there is a single vocalization present at the moment of observation. In addition, this acoustic localization system works under close human supervision. During data analysis, the user identifies section of animal vocalizations that were not masked and that were relatively free from background noise by visual inspection of spectrograms. Cross-correlation function displays are also visually inspected by the user to identify and discard possibly spurious cross-correlation maxima.

Further author information: (Send correspondence to Hanbiao Wang)
Hanbiao Wang: E-mail: hbwang@cs.ucla.edu, Telephone: 1 310 206 3925

An iPAQ-based wireless acoustic network has also been developed for frog call detection and localization in real-time.⁸ Each iPAQ is a hand-held computer with a built-in microphone. 802.11 cards are used to establish a wireless network among iPAQs. Audio data acquisition across multiple iPAQs is synchronized within a few microseconds. The frog vocalizations of interest in the acquired audio streams are automatically detected by matched filter based on spectrograms. TDOA between each pair of microphones is estimated online using cross-correlation of sound waveforms. There are multiple non-linear equations that relate TDOAs, microphone locations, sound speed, and the frog location. These non-linear equations are linearized by introducing extra unknown variables.¹² The frog location is estimated online using least square method. Due to the limit of cross-correlation-based TDOA estimation, this system also works only when there is a single vocalization present at the moment of observation.

An acoustic beamformer has been proposed to monitor birds in order to prevent collision between aircrafts and birds.¹³ The beamformer consists of a large number of microphones arranged into multiple concentric rings on a dish. DOAs of birds are estimated using multiple signal classification (MUSIC) algorithm.¹⁴ Due to advantages of MUSIC algorithm, this system can work with multiple simultaneous sound sources. Only DOAs of birds are desired. Localization of birds is not a goal of this system. In order to achieve high resolution of DOA estimation, this beamformer requires a large number of microphones. Each microphone is digitized at a sampling rate of 22 KHz. Therefore, each beamformer requires a large amount of memory and a high-speed CPU for real-time processing. The above hardware requirement will result in high power consumption. Thus this system is not suitable for bird behavior monitoring in the wild that is far away from any power supply infrastructure.

In our design of acoustic sensor networks for woodpecker localization, we choose the approximate maximum likelihood (AML) method for wideband acoustic beamforming³ to estimate DOA of woodpeckers using four synchronized audio channels of each acoustic array. As discussed in a previous publication,³ the AML algorithm is an optimal DOA estimator (optimal as the FFT size becomes large). Similar to MUSIC, AML can also estimate DOAs of multiple simultaneous acoustic sources. Since AML needs only a small number of audio channels from each acoustic array in order to estimate DOAs of woodpeckers, the requirements on memory space, CPU speed, and power consumption for data processing are much easier to meet. Thus, AML-based acoustic sensor networks are more suitable for animal monitoring in the wild. Our acoustic localization system not only estimates DOAs of woodpeckers using individual acoustic arrays, but also estimates woodpecker locations by applying least square (LS) method to DOA bearing crossings of multiple acoustic arrays. To test our acoustic array design, we have conducted several localization experiments in different environments. A computer speaker is used as a mock bird to play back the waka calls of the acorn woodpecker. Each acoustic array has four M53 microphones¹⁵ placed in a square. The waka calls are recorded synchronously by the PreSonus FIREPOD recording studio¹⁶ that connects microphones to a laptop computer. The recorded waka calls are processed offline for woodpecker DOA estimation and localization. In our preliminary investigation, we have only tested two-dimensional beamforming and localization. Other generalization are also possible.

A major contribution of this paper is the identification of the critical relation between microphone spacing of acoustic arrays and robustness of beamforming of woodpecker vocalizations. Woodpecker localization experiments in different type of environments are conducted and compared. Practical issues about calibration of acoustic array orientation are also discussed. The rest of this paper is organized as follows. Section 2 reviews the AML algorithm. Section 3 examines the critical relation between microphone spacing of acoustic arrays and robustness of beamforming of woodpecker vocalizations. Section 4 compares results of localization experiments conducted in environments with different characteristics of ambient noise and multipath effects. Section 5 examines calibration of acoustic array orientation to improve DOA and location estimation. Section 6 outlines our future plans for three-dimensional localization of faraway acoustic source and real-time on-line processing in wireless acoustic array sensor networks. Section 7 concludes this paper.

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 2-D scenario) as shown in Fig. 1.

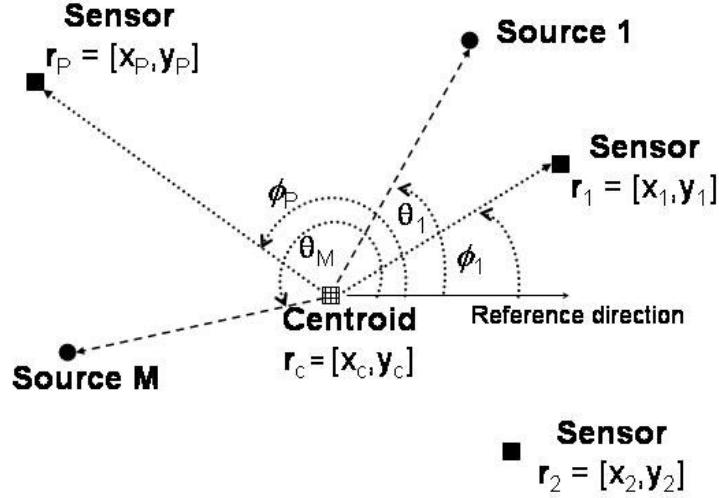


Figure 1. Far-field notations for sources, sensors, and sensor array centroid

Let there be M wideband sources, each at an angle θ_m from the array with the reference direction pointing to the east. The sensor array consists of P randomly distributed sensors, each at position $\mathbf{r}_p = [x_p, y_p]^T$. The sensors are assumed to be omni-directional and have identical response. The array centroid position is given by $\mathbf{r}_c = \frac{1}{P} \sum_{p=1}^P \mathbf{r}_p = [x_c, y_c]^T$. 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 $t_{cp}^{(m)} = t_c^{(m)} - t_p^{(m)} = [(x_c - x_p) \cos \theta_m + (y_c - y_p) \sin \theta_m]/v$, in where $t_c^{(m)}$ and $t_p^{(m)}$ are the absolute time-delays from the m th source to the centroid and the p th sensor, respectively, and v 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 ϕ_p are the range and angle of the p sensor with respect to the array centroid. The data received by the p th sensor at time n is then

$$x_p(n) = \sum_{m=1}^M S^{(m)}(n - t_{cp}^{(m)}) + w_p(n), \quad (1)$$

for $n = 0, \dots, N - 1$, $p = 1, \dots, P$, and $m = 1, \dots, M$, 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 σ^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

$$\mathbf{X}(\omega_k) = \mathbf{D}(\omega_k)\mathbf{S}(\omega_k) + \eta(\omega_k), \quad (2)$$

for $k = 0, \dots, N - 1$, where the array data spectrum is $\mathbf{X}(\omega_k) = [X_1(\omega_k), \dots, X_P(\omega_k)]^T$, the steering matrix $\mathbf{D}(\omega_k) = [\mathbf{d}^{(1)}(\omega_k), \dots, \mathbf{d}^{(M)}(\omega_k)]$, the steering vector is given by $\mathbf{d}^{(m)}(\omega_k) = [d_1^{(m)}(\omega_k), \dots, d_P^{(m)}(\omega_k)]^T$, $d_p^{(m)} = e^{-j2\pi k t_{cp}^{(m)}/N}$, and the source spectrum is given by $\mathbf{S}(\omega_k) = [S^{(1)}(\omega_k), \dots, S^{(M)}(\omega_k)]^T$. 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 T 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 criterion³

$$\max_{\Theta, \mathbf{S}} L(\Theta, \mathbf{S}) = \min_{\Theta, \mathbf{S}} \sum_{k=1}^{N/2} \|\mathbf{X}(\omega_k) - \mathbf{D}(\omega_k)\mathbf{S}(\omega_k)\|^2, \quad (3)$$

which is equivalent to a nonlinear least square problem. Using the technique of separating variables,¹⁷ the AML DOA estimate can be obtained by solving the following likelihood function

$$\max_{\Theta} J(\Theta) = \max_{\Theta} \sum_{k=1}^{N/2} \|\mathbf{P}(\omega_k, \Theta)\mathbf{X}(\omega_k)\|^2 = \max_{\Theta} \sum_{k=1}^{N/2} \text{tr}(\mathbf{P}(\omega_k, \Theta)\mathbf{R}(\omega_k)), \quad (4)$$

where $\mathbf{P}(\omega_k, \Theta) = \mathbf{D}(\omega_k)\mathbf{D}^\dagger(\omega_k)$, $\mathbf{D}^\dagger = (\mathbf{D}(\omega_k)^H\mathbf{D}(\omega_k))^{-1}\mathbf{D}(\omega_k)^H$ is the pseudo-inverse of the steering matrix $\mathbf{D}(\omega_k)$ and $\mathbf{R}(\omega_k) = \mathbf{X}(\omega_k)\mathbf{X}(\omega_k)^H$ is the one snapshot covariance matrix. Once the AML estimate of Θ is found, the estimated source spectrum can be given by

$$\hat{\mathbf{S}}^{ML}(\omega_k) = \mathbf{D}^\dagger(\omega_k, \hat{\Theta}^{ML})\mathbf{X}(\omega_k). \quad (5)$$

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 source case, the computational complexity of the AML algorithm requires multi-dimensional 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.¹⁸

3. ARRAY ELEMENT SPACING

In this section, we examine the critical relation between microphone spacing of acoustic arrays and robustness of beamforming of woodpecker vocalizations. We will also propose a guideline to choose proper array element spacing in order to achieve beamforming robustness.

3.1. Non-Robustness of Beamforming with Large Sensor Spacing

In beamforming of narrowband signals with wavelength λ using a linear array with sensors regularly spaced at distance 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. 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 average out no matter whether the array is linear and how large is sensor spacing.

In this paper, we are interested in beamforming of waka calls of acorn woodpeckers using a square array of four microphones. As shown in figure 2, waka calls are wideband signals. However, waka calls do not have comparable power spectral density across a wide range of frequency. Instead, the power spectral density of waka calls is more like a combination of a small number of narrowband signals centered around 3.4 KHz, 2.5 KHz, 1.7 KHz, and 2.1 KHz, in descent order of power.

Theoretically, with an ideal wideband signal, the grating lobe effect due to a single tone will not be observed. However, the non-uniformity of the power spectral contents of waka calls makes some grating lobe effects due to various frequency contents more difficult to average out in beamforming using large microphone spacing. This is confirmed by a simulation of waka call beamforming using AML as shown in figure 3(a). The square acoustic array is set to 61 cm by 61 cm in the simulation. In the simulated beam pattern, although there is no grating lobe to cause DOA estimation ambiguity, several side lobes are almost 70% tall as the main lobe. In the simulation, we have assumed no noise, interference, or multipath effects in acquired audio data.

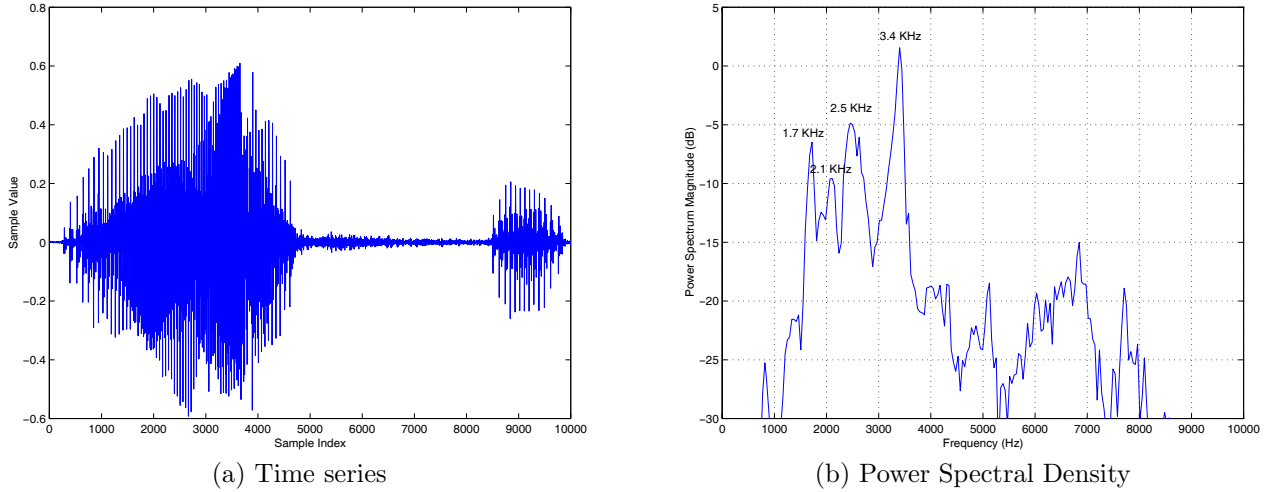


Figure 2. Waka calls of acorn woodpeckers.

In real habitats of acorn woodpeckers, 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 tall side lobes could easily cause gross DOA estimation error when side lobes grow taller than the main lobe. This is exactly what we have observed in a field experiment of AML beamforming conducted in the UCLA Botanical Garden. The experimental site is densely populated with trees and bushes. As in the simulation, the acoustic array is also 61 cm by 61 cm. A speaker to play back waka calls is at 72 degree to the acoustic array. If there were no noise or multipath effects, we should have obtained a beam pattern in the field experiment similar to that of the simulated beamforming. However, in the experiment beam pattern as shown in figure 3(b), the lobe pointing to the true DOA 72 degree shrinks dramatically while the new main lobe points to 342 degree. In this experiment, the estimated DOA 342 degree is completely false. The simulation and the experiment have shown that beamforming of waka calls using large microphone spacing is not robust due to waka calls' special spectral property. Factors such as ambient noise and multipath effects in the real environment can easily distort beam pattern and result in gross DOA estimation errors when microphone spacing is large.

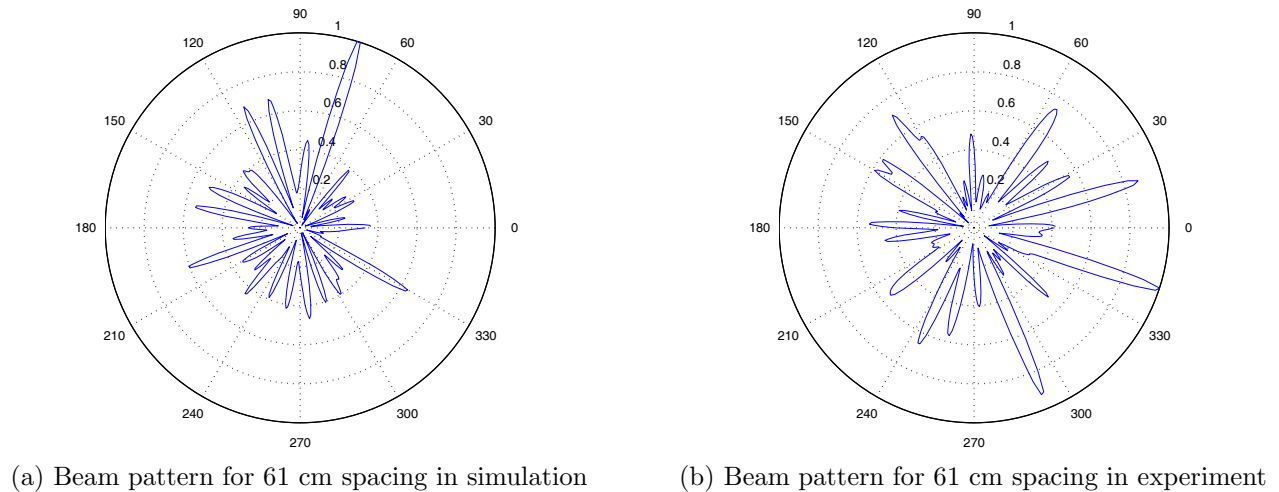


Figure 3. Beam pattern in AML when waka calls arrive at 72 degree. Function in polar plot is J defined in equation 4.

3.2. Sensor Spacing for Robust Beamforming

As we have seen, the non-robustness of waka call beamforming using large sensor spacing is caused by large side lobes in beam pattern. One way to solve this non-robustness problem is to remove side lobes in beam pattern by using smaller sensor spacing. We can follow the half-wavelength sensor spacing guideline of narrowband beamforming for choice of sensor spacing in order to remove side lobe in beam pattern of waka calls. As shown in figure 2, We can model a waka call 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 side lobe in beam pattern of the combined signal, we have to remove side lobe in beam pattern of each component narrowband signals. Follow 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.

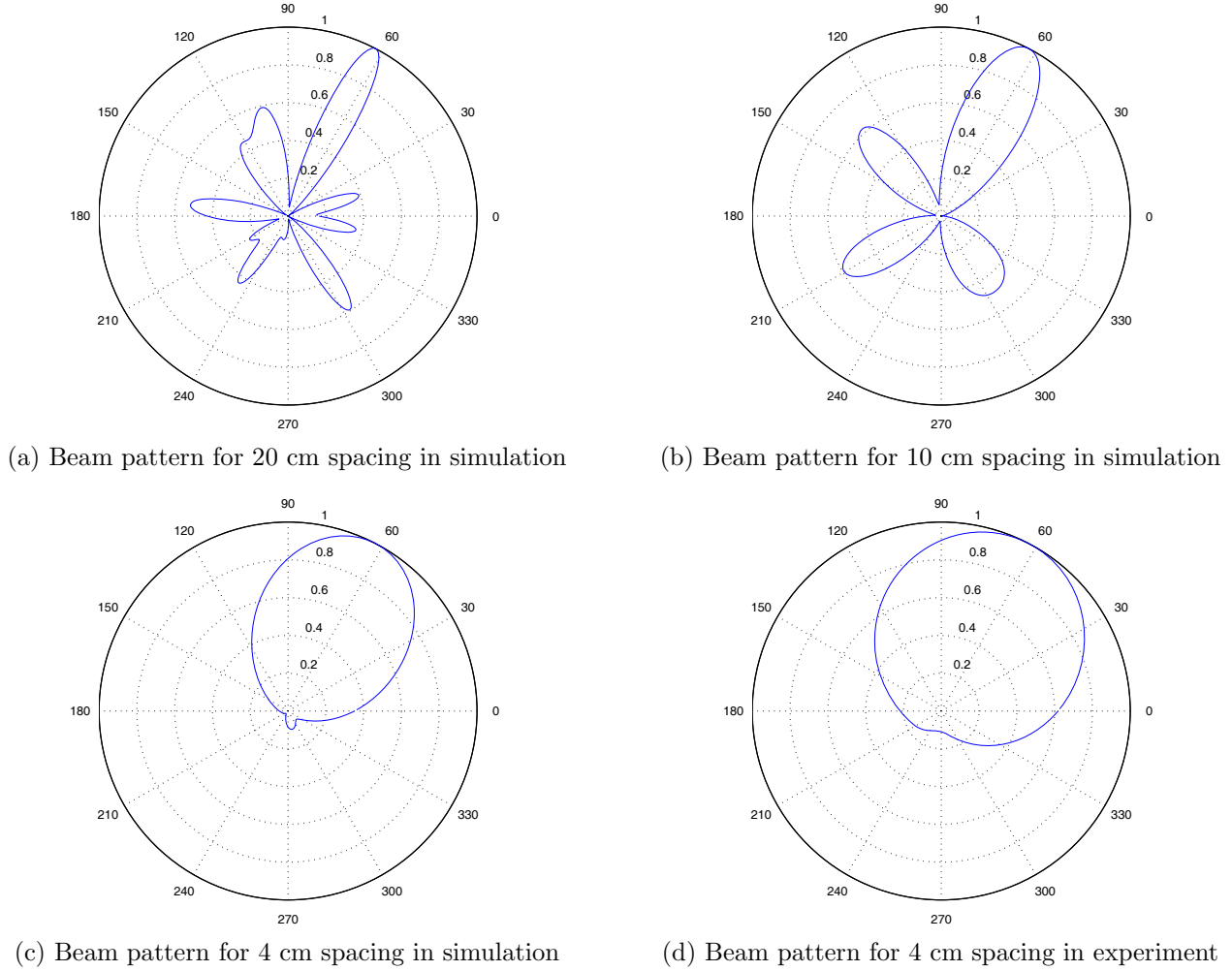


Figure 4. Beam patterns in AML when waka calls at 61 degree. Function in polar plot is J defined in equation 4

To verify the robustness of small microphone spacing, we have conducted simulations and experiments using several different microphone spacing in the acoustic array. Beam patterns for different microphone spacing are shown in figure 4. The acoustic source is at 61 degree to the acoustic array. As shown in figure 4 (a-b), simulations have indicated that several side lobes are still as tall as 60% of the main lobe when the microphone spacing is 20 cm and 10 cm. The number of side lobes is reduced as the microphone spacing decreases. As shown in figure 4 (c), simulations have indicated that side lobes almost completely disappear when microphone spacing is reduced

to 4 cm. To verify that the beam pattern of a single main lobe with 4 cm microphone spacing is robust, we have conducted localization experiments using 4 cm by 4 cm square acoustic arrays on the same experimental site in the UCLA Botanical Garden that were used for the localization experiment using 61 cm by 61 cm square acoustic arrays described in figure 3(b). A speaker to play back waka calls is at 61 degree to the acoustic array. As shown in figure 4(d), a beam pattern of acoustic signals acquired in the experiment has only a single main lobe, which is consistent with the simulated beam pattern in figure 4(c). The beam pattern in the experiment for the 4 cm by 4 cm arrays is fatter than the beam pattern in simulations for the same microphone spacing because the acoustic signals in experiments could have been distorted by noise, interference, and multi-path effects. Since there is no side lobe in the beam pattern in the experiment, the DOA estimate will not be completely incorrect even though the main lobe could be severely distorted by noise, interference, and multipath effects. Compared to the beam pattern of 61 cm by 61 cm square arrays shown in figure 3, these beam patterns of 4 cm by 4 cm square arrays are much more robust.

In our design of acoustic arrays for woodpecker localization, we use four microphones in each array because more microphones in an array demand larger memory, faster CPU, and more power consumption that can not be easily met in the wild. Since an acoustic array is a square with one microphone placed at each corner, a smaller microphone spacing results in a smaller array aperture. A smaller array aperture also results in a larger 3-dB-width of the main lobe in beam pattern. A fatter main lobe in beam pattern directly translate into a coarser DOA estimation resolution. Therefore, in our acoustic array design, a smaller microphone spacing causes a coarser DOA estimation resolution. So sensor spacing plays a critical role in the tradeoff between DOA estimation robustness and resolution. We believe the robustness of DOA estimation is of first order importance. Without robustness, DOA estimation can be totally false and have a gross error ranging from 0 to 360 degrees. On the other hand, we can achieve a finer resolution of DOA estimation and localization through fusion of multiple coarse but robust DOA estimates as shown in section 4.

4. LOCALIZATION EXPERIMENTS

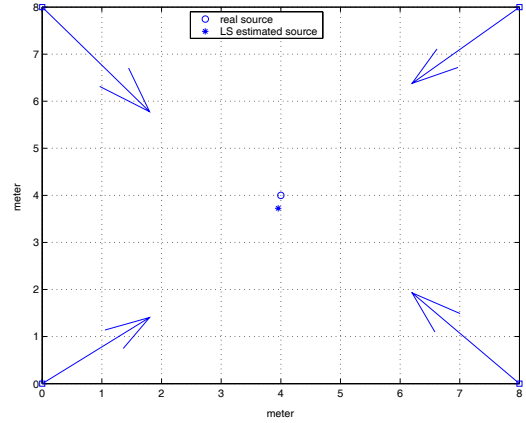
In this section, we present results of localization experiments in several environments with very different characteristics of noise, interference, and multipath effects. In these experiments, each acoustic array is a 4 cm by 4 cm square with one microphone on each corner. Both the speaker for playback and acoustic arrays for data acquisition are raised about 70 cm above the ground. Each acoustic array estimates the DOA of acoustic source relative to the local coordinate system of the acoustic array. The orientation of the acoustic array is measured relative to the global coordinate system using compass. We first translate DOA estimates of each array's local coordinate system into the global coordinate system using acoustic array orientation measurements. We then estimate the acoustic source location in the global coordinate system by applying LS method to the cross-bearing of DOA estimates in the global coordinate system.

To test the effectiveness of acoustic localization in relatively open environments such as grassland habitats we have conducted experiments in a small parking lot on the campus of the Buckley School in Sherman Oaks, California. The speaker, the acoustic arrays, the parking lot, and the surrounding area are shown in figure 5(a). Four acoustic arrays are deployed at corners of an 8 m by 8 m square. The speaker is placed at the center of the square. As shown in figure 5(b), DOA estimates by acoustic arrays in the global coordinate system are indicated by arrows. The estimated location of the acoustic source is about 28 cm off the true location of the acoustic source.

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 also very windy during the experiments in the Science Courtyard. The speaker, the acoustic arrays, the courtyard, and the surrounding area are shown in figure 6(a). Four acoustic arrays are deployed at corners of an 4 m by 4 m square. The speaker is placed at the center of the square. As shown in figure 6(b), DOA estimates by acoustic arrays in the global coordinate system are indicated by arrows. 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.



(a) Parking lot

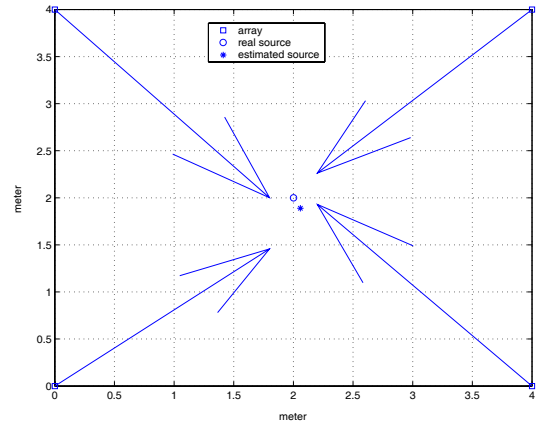


(b) Localization Result

Figure 5. Localization experiment in the parking lot of the Buckley School using 4cm by 4cm square acoustic arrays



(a) Science Courtyard



(b) Localization Result

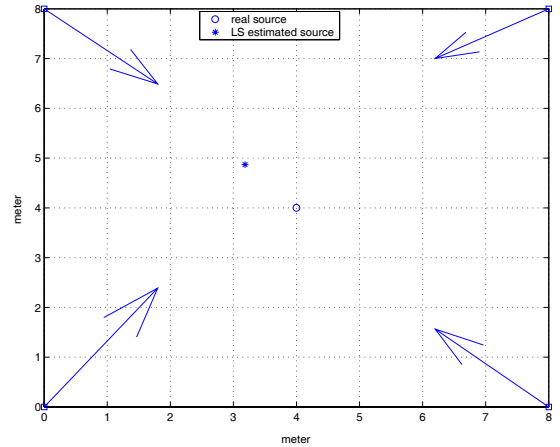
Figure 6. Localization experiment in the Science Courtyard at UCLA using 4cm by 4cm square acoustic arrays

To test effectiveness of acoustic localization in less open environments such as woodland habitats, we have conducted experiments in the woods on the hill in the Buckley School in Sherman Oaks, California. The speaker, the acoustic arrays, the hill, and the woods are shown in figure 7(a). The experimental site is on a hill slope of about 30 degree elevation and is completely covered by dense tree crowns about a few meters tall. The multipath effects in the woods are much sever than those in the parking lot. Four acoustic arrays are deployed at corners of an 8 m by 8 m square. The speaker is placed at the center of the square. As shown in figure 7(b), DOA estimates of acoustic arrays in the global coordinate system are indicated by arrows. The localization error of experiments in woods is about 119 cm, which is much higher than localization error 28 cm of experiments in the parking lot. The large localization error is largely due to the large DOA error of acoustic array 4 at the upper right corner. In the global coordinate system, DOA estimate of array 4 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 wrong 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 section 5, we will discuss calibration

of acoustic array orientation. We will see significant localization accuracy improvements due to acoustic array orientation calibration.



(a) Hill



(b) Localization Result

Figure 7. Localization experiment on the hill of the Buckley School using 4cm by 4cm square acoustic arrays

Figure 8 shows the beam patterns of the 4cm by 4 cm acoustic arrays in the experiment on the hill of the Buckley School. We can clearly see that there is a single main lobe in each beam pattern pointing to the right DOA direction although strong multi-path effects and other noise and interference have distorted the beam pattern to some extents. These beam patterns of the acoustic localization experiments in woods again verify that 4 cm microphone spacing for beamforming is very robust.

5. CALIBRATION OF ARRAY ORIENTATION

As we have pointed out in section 4, the DOA estimates relative to each acoustic array’s local coordinate system needs to be translated into a global coordinate system using acoustic array orientation measurements before DOA estimates are combined to estimate the acoustic source location. Thus the accuracy of the acoustic array orientation measurements is as important as DOA estimation accuracy to acoustic source localization accuracy. As we have seen in the localization experiment shown in figure 7, there could be large errors in array orientation measurements because acoustic arrays are small and the experimental site has a rough terrain. Large acoustic array orientation errors will finally result in a large acoustic source localization error. In this section, we describe to calibrate acoustic array orientation in order to improve localization accuracy.

To investigate the effectiveness of array orientation calibration, we have conducted several acoustic localization experiments in the UCLA Botanical Garden that is densely populated with trees and bushes. Each acoustic array has four microphones placed in a 4 cm by 4 cm square. After we deployed two acoustic arrays on a small trail about 70 cm above the ground, we also placed a speaker as a landmark in the same line as these arrays. Since we know that the speaker is at 180 degree to both arrays in the global coordinate system, the speaker DOA estimated by each acoustic array relative to the local coordinate system of the array indicates the acoustic array orientation in the global coordinate system. Although we have made our best efforts to use a compass to give each acoustic array a zero degree orientation relative to the global coordinate system, estimated DOAs of the landmark speaker indicate that the left array and the right array have 1 degree and 10 degree orientation respectively. This calibration method achieves more accurate array orientation measurements than the compass.

Calibration of acoustic array orientation can greatly improve accuracy of acoustic source localization. Using the same deployment of acoustic arrays, we have conducted two localization experiments. In one experiment, we placed the speaker about 2 m away from the trail. The location of the speaker by the side of the trail is estimated first without using calibrated array orientation and then using calibrated array orientation as shown

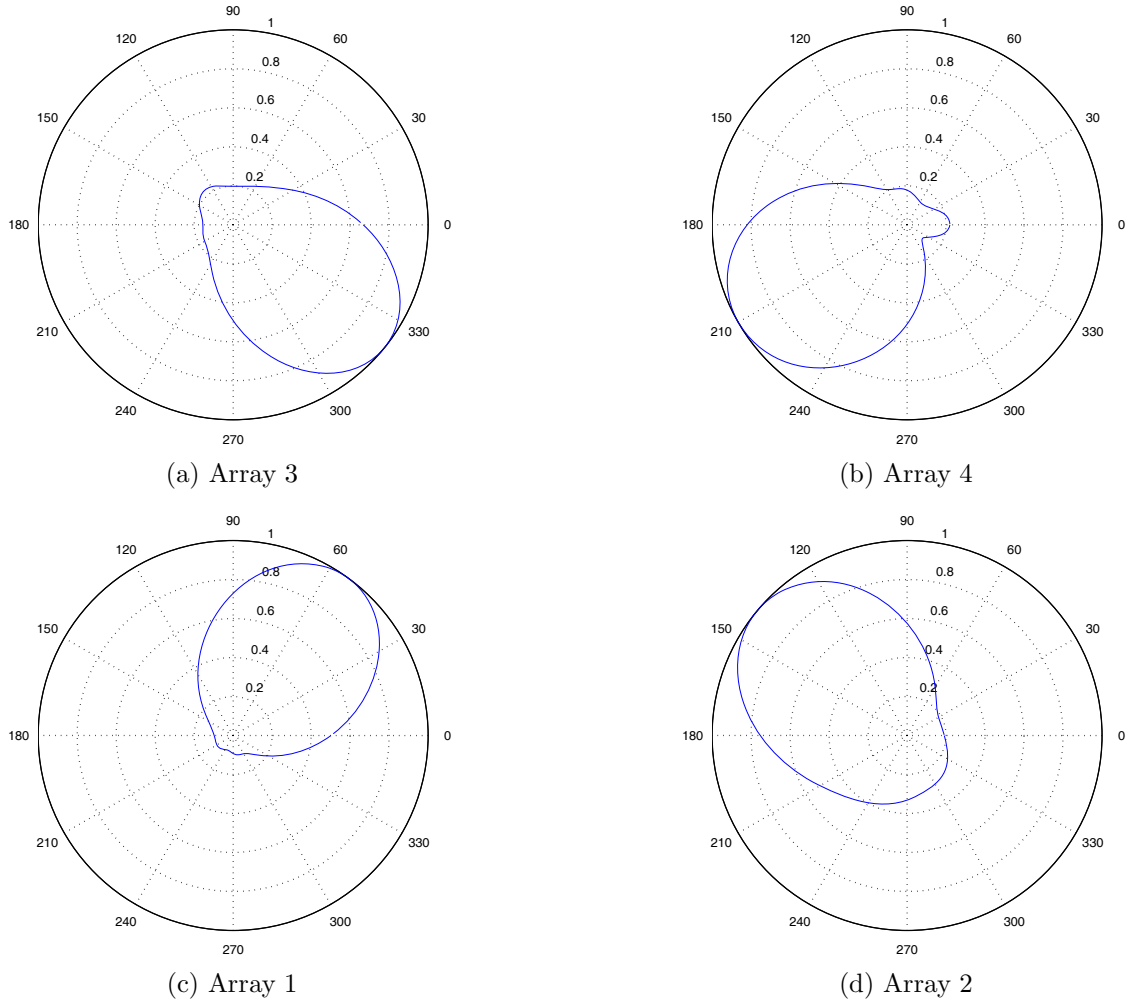
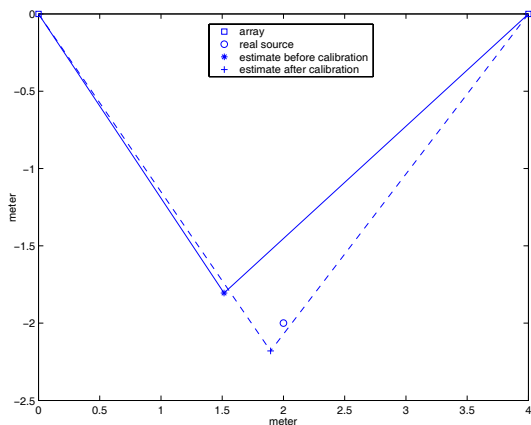


Figure 8. Beam patterns of 4cm by 4cm square acoustic arrays in localization experiment on the hill of the Buckley School. Function in polar plot is J defined in equation 4.

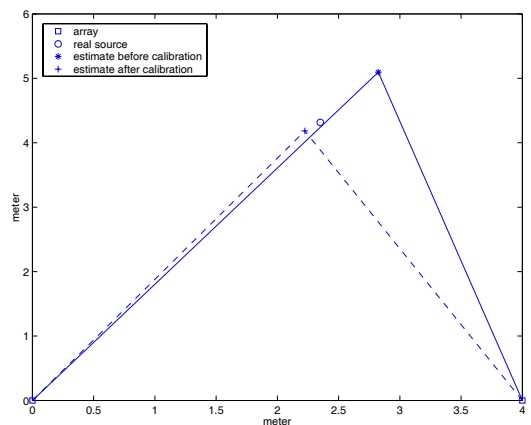
in figure 9(a). The localization error without calibration is 52 cm while the localization error with calibration is only 21 cm. In the other experiment, we placed the speaker on a tree about 5 m away from the trail. The location of the speaker on a tree is estimated first without using calibrated array orientation and then using calibrated array orientation as shown in figure 9(b). The localization error without calibration is 91 cm while the localization error with calibration is only 18 cm. Therefore, calibration of acoustic array orientation is very effective and of great importance in practice in order to improve localization accuracy.

6. FUTURE WORK

In the future, we would like to extend our current work in three directions. First, we would like to extend AML to three-dimensional beamforming that simultaneously estimates both azimuth and zenith of acoustic source. We would like to investigate the design of effective acoustic array configuration for robust three-dimensional beamforming. Second, we would like to investigate localization of acoustic sources that are further away from acoustic arrays than the acoustic waves travel distance from the speaker to acoustic arrays in our previous experiments. Third, we plan to build a real-time woodpecker localization system using Intel Stargate¹⁹ platform. Each acoustic array connects to a Stargate node through a VXpocket440²⁰ sound card. Multiple Stargate nodes form a wireless network through 802.11 flash cards. Stargate nodes will tightly synchronize acoustic data



(a) Acoustic Source on trail side



(b) Acoustic source on tree

Figure 9. Localization experiments using 4cm by 4cm square acoustic arrays with and without calibration in the UCLA Botanical Garden. After calibration, DOA estimates of the left array are rotated anti-clockwise by 1 degree, and DOA estimates of the right array are rotated anti-clock wise by 10 degrees.

acquisition across multiple channels of the same acoustic array. Software modules on each Stargate node will detect woodpecker vocalizations and estimate woodpecker DOA using AML beamforming. Multiple Stargate nodes collaboratively estimate woodpecker location by combining their DOA estimates.

7. CONCLUSION

Acorn woodpecker waka calls can be modeled as a combination of a small number of narrowband signals that have well separated frequencies. For acoustic sources such as woodpecker vocalizations, we have revealed the critical relation between microphone spacing of acoustic arrays and robustness of beamforming. We have proposed a guideline to choose proper microphone spacing in order to achieve beamforming robustness. To verify our choice of microphone spacing for robust beamforming, we have conducted acoustic localization experiments in environments with different characteristics of noise, interference, and multipath effects. We have recognized that accurate DOA estimation alone can not guarantee accurate localization in practice. Acoustic array orientation error propagates directly into localization error when multiple acoustic arrays' DOA estimates are combined to determine acoustic source location. We have proposed an approach to calibrate acoustic array orientation. Localization experiments have verified that such calibration greatly improves localization accuracy.

ACKNOWLEDGMENTS

This work is partially supported by NSF CENS program under Cooperative Agreement CCR-0121778, NSF grant EF-0410438, AROD-MURI PSU Contract 50126, and UC-Discovery grant sponsored by ST Microelectronics. We appreciate the assistance of Kathy Griffith, Joe Wise, Chih-Kai Chen, and Hyunggon Park in conducting various experiments. 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

1. W. D. Koenig and D. L. Mummee, *Population Ecology of the Cooperative Breeding of Acorn Woodpecker*, Princeton University Press, Princeton, New Jersey, 1987.
2. W. D. Koenig, P. B. Stacey, M. T. Stanbeck, and R. L. Mumme, "Acorn woodpecker," *Birds of North America* **194**, pp. 1–24, 1995.
3. J. Chen, R. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE T. Signal Proces.* **50**, pp. 1843–1854, August 2002.

4. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks* **38**, pp. 393–442, March 2002.
5. D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *IEEE Computer* **37**, pp. 41–49, Aug. 2004.
6. W. group of Infrastructure for Biology at Regional to Continental Scale (IBRCS), "Ibrcs white paper: Rationale, blueprint, and expectations for the national ecological observatory network (neon)," *American Institute of Biological Sciences* .
7. K. Delin, S. Jackson, D. Johnson, S. Burleigh, R. Woodrow, J. McAuley, J. Dohm, F. Ip, T. Ferre, D. Rucker, and V. Baker, "Environmental studies with the sensor web: Principles and practice," *Sensors* **5**, pp. 103–117, 2005.
8. H. Wang, J. Elson, L. Girod, D. Estrin, and K. Yao, "Target classification and localization in habitat monitoring," in *Proc. ICASSP*, (Hong Kong, China), April 2003.
9. W. Hu, V. Tran, N. Bulusu, C. Chou, S. Jha, and A. Taylor, "The design and evaluation of a hybrid sensor network for cane-toad monitoring," in *IPSN/SPOTS 2005*, (Los Angeles, CA), April 2005.
10. L. Freitag and P. Tyack, "Passive acoustic localization of the atlantic bottlenose dolphin using whistles and ecolocation clicks," *J. Acoust. Soc. Am* **93**(4), pp. 2197–2205, 1993.
11. P. K. McGregor, T. Dabelsteen, C. W. Clark, J. L. Bower, J. P. Tavares, and I. Holland, "Accuracy of passive acoustic location system: empirical studies in terrestrial habitats," *Ethology, Ecology, and Evolution* **9**, pp. 269–286, 1997.
12. T. Tung, K. Yao, C. Reed, R. Hudson, D. Chen, and J. Chen, "Source localization and time delay estimation using constrained least squares and best path smoothing," in *Proc. SPIE'99*, **3807**, pp. 220–223, July 1999.
13. C. Kwan, K. Ho, G. Mei, Y. Li, Z. Ren, R. Xu, G. Zhao, M. Stevenson, V. Stanford, and C. Rochet, "An automated acoustic system to monitor and classify birds," in *Proc. of Bird Strike Committee Canada Conference*, 2003.
14. R. Schmidt, *A signal subspace approach to multiple emitter location and spectral estimation*. Thesis of doctor of philosophy, Stanford University, 1981.
15. "M53 microphone, http://www.linearx.com/products/microphones/m53/m53_1.htm."
16. "Presonus firepod, <http://www.presonus.com/firepod.html>."
17. I. Ziskine and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoustics, Speech, and Signal Processing* **36**, pp. 1553–1560, Oct. 1988.
18. L. Yip, J. Chen, R. Hudson, and K. Yao, "Numerical implementation of the aml algorithm for wideband doa estimation," in *Proc. SPIE*, pp. 164–172, Aug. 2003.
19. "Stargate, <http://platformx.sourceforge.net/links/resource.html>."
20. "Digigram vxpocket 440 pcmcia sound card, <http://www.digigram.com/products>."