Hindawi Publishing Corporation
EURASIP Journal on Applied Signal Processing
Volume 2006, Article ID 96706, Pages 1–19
DOI 10.1155/ASP/2006/96706

An Automated Acoustic System to Monitor and Classify Birds

**C. Kwan,[1] K. C. Ho,[2] G. Mei,[1] Y. Li,[2] Z. Ren,[1] R. Xu,[1] Y. Zhang,[1] D. Lao,[1]
M. Stevenson,[1] V. Stanford,[3] and C. Rochet[3]**

[1] *Intelligent Automation, Inc., 15400 Calhoun Drive, Suite 400, Rockville, MD 20855, USA*
[2] *Department of Electrical and Computer Engineering, University of Missouri-Columbia, 349 Engineering Building West, Columbia, MO 65211, USA*
[3] *National Institute of Standards and Technology, Building 225, Room A216, Gaithersburg, MD 20899, USA*

Received 4 May 2005; Revised 3 October 2005; Accepted 11 October 2005

Recommended for Publication by Hugo Van hamme

This paper presents a novel bird monitoring and recognition system in noisy environments. The project objective is to avoid bird strikes to aircraft. First, a cost-effective *microphone dish* concept (microphone array with many concentric rings) is presented that can provide directional and accurate acquisition of bird sounds and can simultaneously pick up bird sounds from different directions. Second, direction-of-arrival (DOA) and beamforming algorithms have been developed for the circular array. Third, an efficient recognition algorithm is proposed which uses Gaussian mixture models (GMMs). The overall system is suitable for monitoring and recognition for a large number of birds. Fourth, a hardware prototype has been built and initial experiments demonstrated that the array can acquire and classify birds accurately.

1. INTRODUCTION

Collisions between aircraft and birds have become an increasing concern for both human and bird safety. More than four hundred people and over four hundred aircraft have been lost globally since 1988, according to a Federal Aviation Agency (FAA) report [1]. Thousands of birds have died due to these collisions. Bird strikes have also caused more than 2 billion dollars worth of damage each year.

There are several ways to monitor the birds near the airports. First, X-band radars are normally used for monitoring birds. One drawback is that the radar cannot distinguish between different birds even though it can monitor birds several kilometers away. Second, infrared cameras are used to monitor birds. However, cameras do not work well under bad weather conditions and cannot provide bird species information. Third, according to Dr. Willard Larkin at the Air Force Office of Scientific Research, microphone arrays are being considered for monitoring birds. The conventional arrays are linear arrays with uniform spacing. One serious drawback is that there is a cone of angular ambiguities. Moreover, no microphone array product has been produced yet.

In this research, we propose a novel circular microphone array system that includes both hardware and software for bird monitoring. This new concept can eliminate the drawbacks of linear arrays, can provide no angular ambiguities, can generate more symmetric beam patterns, and can produce more directional beams to acquire bird sounds and hence more accurate bird classification. Consequently, the technology will save both human and bird lives, and will also significantly reduce damage costs due to bird strikes.

Besides bird monitoring and recognition, the system can be applied to wildlife monitoring, endangered species monitoring in inaccessible areas, speech enhancement in communication centers, conference rooms, aircraft cockpits, cars, buses, and so forth. It can be used for security monitoring in airport terminals, and bus and train stations. The system can pick up multiple conversations from different people and at different angles. It can also be used as a front-end processor to automatic speech recognition systems. We expect that this new system will significantly increase speech quality in noisy and multispeaker environments.

Here we will present the technical details of the proposed bird monitoring system and summarize the experimental results. Some preliminary work of the proposed system has been presented in a bird monitoring workshop [2]. This paper provides a comprehensive description of the entire system, develops in details the signal processing techniques in

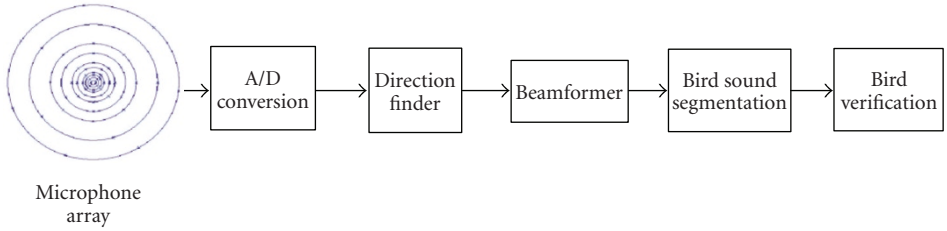


Figure 1: Proposed automated bird monitoring and recognition system.

each component, and provides more complete simulation and experimental results.

The paper is organized as follows. Section 2 gives a brief overview of the proposed system, which consists of several major parts: microphone dish and data acquisition system, direction-of-arrival (DOA) estimation algorithm, beamformer to eliminate interferences, and bird classifier. Section 3 will summarize a wideband DOA estimation algorithm and provide a comparative study between estimation results using a linear array and a circular array. A new beamforming algorithm and a comparative study between a linear array and a circular array will be summarized in Section 4. It was found that the dish array has several key advantages over the linear array, including less number of ambiguity angles, more consistent performance, better interference rejection capability, and so forth. Section 5 describes the bird classification results using GMM method. The development of a prototype microphone dish will be included in Section 6. A dish array consisting of 64 microphone elements has been developed and used to collect sound data in the laboratory and in an open space. In Section 7, experimental results will be described to demonstrate the performance of the software and hardware. Finally, conclusions will be drawn in Section 8.

2. OVERALL BIRD MONITORING SYSTEM DESCRIPTION

The circular microphone array concept for bird monitoring is novel. Based on our literature survey [3], the circular microphone array with many concentric rings has not been produced in the past. No DOA or beamforming algorithms exist for this type of arrays. Figure 1 shows the proposed bird monitoring system, which consists of a microphone dish, a data acquisition system, and software processing algorithms such as direction finder, beamformer, and bird sound classification.

Our analysis of bird sounds shows that the frequency range of bird sounds is between 100 Hz to 8 kHz. In the data acquisition part, the goal is to simultaneously acquire 64 microphone signals and digitize them with 22 kHz sampling rate. This is not an easy task. With the help of engineers from the National Institute of Standards and Technology (NIST), we were able to build a data acquisition system that can satisfy this goal.

In the direction finding part, we modified and customized the well-known multiple signal classification (MUSIC) [4] algorithm to the circular array. Our studies found that the circular array can provide more accurate and less ambiguous DOA estimations than linear arrays.

For the beamformer, new algorithms were developed specifically for the concentric circular arrays. Our algorithms can provide symmetric beam patterns, offer no angular ambiguities, and guarantee consistent residual sidelobe for all frequencies from 100 Hz to 8 kHz. In the design of the beamformer, we have systematic ways to choose the interring spacing and interelement spacing in each ring in order to achieve the above merits such as symmetric and directional beam patterns.

The bird classification was done by using GMM, which is a well-known technique and has been widely used in human speaker verification.

Once the directions of birds are determined, some bird control systems will be activated. For example, a control device that can create a loud bang in the direction of the birds could be activated to scare the birds away.

3. DOA ESTIMATION ALGORITHM FOR CIRCULAR MICROPHONE ARRAYS

3.1. DOA estimation algorithm for circular arrays

Figure 2 shows the configuration of the proposed circular array. It has M concentric rings with radii R_m, $m = 1, 2, \ldots, M$, and $R_1 < R_2 < \cdots < R_M$. Ring m has N_m elements, and the ith element in ring m has an angle of $\upsilon_{m,i} = 2\pi(i-1)/N_m$ rad. with respect to the x-axis. The signal sample received by the ith element in ring m is denoted as $x_{m,i}(n)$, where n is the time index. The azimuth angle in the x-y plane with respect to the x-axis is denoted as θ, and the elevation angle with respect to the z-axis is represented as ϕ.

A beamformer requires the direction of arrival (DOA) in terms of a particular pair of (θ, ϕ) of the source signal for beamforming in order to enhance the desired signal. The signal DOA is not known in practice and needs to be estimated. There are two challenges in this work. First, not many DOA estimation algorithms exist for circular arrays. Second, the bird signals are wideband signals. DOA estimation algorithms are normally developed for narrowband signals. This section presents the DOA estimation of a wideband source

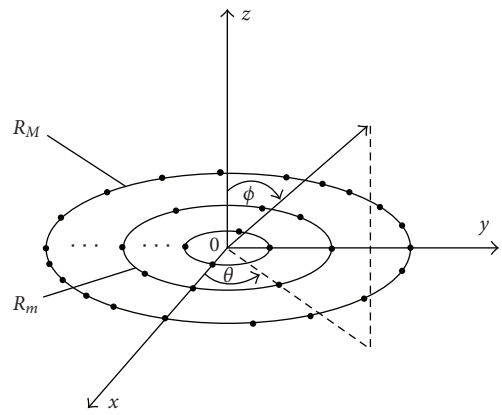


Figure 2: The proposed concentric circular microphone array.

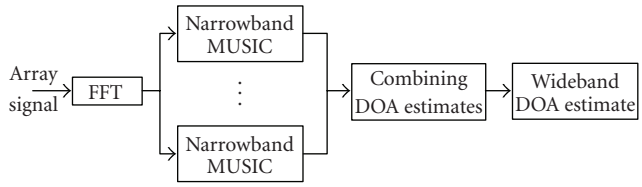


Figure 3: Block diagram of the DOA estimation of a wideband source [5].

signal (bird calls), based on the MUSIC algorithm for narrowband signal.

Multiple signal classification (MUSIC) algorithm [4] is a DOA estimation algorithm for narrowband signal. For wideband signal DOA estimation, we will divide the wideband signal into many narrowband components and then apply MUSIC on those narrowbands. The DOA estimation for the wideband signal is generated by combining estimated results from all the narrowband components. The process is shown in Figure 3. Other wideband DOA estimation techniques for linear arrays can be found in [3, 5] but they are more computationally intensive. At present, we implemented the narrowband combining technique.

As shown in Figure 3, the DOA estimation algorithm consists of the narrowband MUSIC algorithm, which is followed by a peak searching technique to obtain the DOA estimates for each frequency band, and the combination of the DOAs from different frequency bands to form the final estimate. The processing components are described below.

Figure 2 shows the geometry of a circular array. Recall that the signal received at microphone element i in ring m is $X_{m,i}(n)$. $X_{m,i}(n)$ is separated into frames of size 256 samples using a rectangular window with 50% overlap. For each frame l, fast Fourier transform (FFT) is applied to form the frequency-domain samples $X_{m,i}(k,l)$, where $k = 0, 1, \ldots, 255$ is the frequency index. Putting the frequency-domain data at index k over the array elements in ring m forms the data vector at frame l, $\mathbf{X}_m(k,l) = [X_{m,1}(k,l), X_{m,2}(k,l), \ldots, X_{m,N_m}(k,l)]^T$. The collection of the data vectors from all rings forms the overall signal vector at frequency index k, $\mathbf{X}(k,l) = [\mathbf{X}_1(k,l), \mathbf{X}_2(k,l), \ldots, \mathbf{X}_m(k,l)]^T$. The length of $\mathbf{X}(k,l)$ is equal to $K = N_1 + N_2 + \cdots + N_M$, the total number of receiving elements.

In the presence of additive noise, we have the model commonly used in array processing:

$$\mathbf{X}(k,l) = \mathbf{A}(\theta,\phi)\mathbf{S}(k,l) + \mathbf{N}(k,l), \tag{1}$$

where $\mathbf{S}(k,l) = [\mathbf{S}_1(k,l), \mathbf{S}_2(k,l), \ldots, \mathbf{S}_D(k,l)]^T$ is a $D \times 1$ vector containing the source signals spectrum component at frequency index k and frame l, D is the number of sources, and $\mathbf{A}(\theta,\phi) = [{}^1\mathbf{a}(\theta,\phi), {}^2\mathbf{a}(\theta,\phi), \ldots, {}^D\mathbf{a}(\theta,\phi)]$ is a $K \times D$ matrix whose columns are $K \times 1$ directional vectors for different sources. $\mathbf{N}(k,l)$ is a $K \times 1$ ambient noise vector. It is assumed that the noise is uncorrelated with the source signal.

(i) The narrowband MUSIC algorithm

The narrowband DOA algorithm follows the MUSIC technique. It first generates the data correlation matrix over L frames:

$$\mathbf{R}_k = \frac{1}{T}\sum_{l=1}^{T} \mathbf{X}(k,l)\mathbf{X}^H(k,l), \tag{2}$$

where T is the total number of frames used in DOA estimation and is chosen to be 100. The superscript "H" represents complex conjugate transpose. Second, eigendecomposition on $\mathbf{R}_k$ is performed, giving

$$\mathbf{R}_k = \mathbf{U}_s\mathbf{\Lambda}_s\mathbf{U}_s^H + \mathbf{U}_n\mathbf{\Lambda}_n\mathbf{U}_n^H, \tag{3}$$

where $\mathbf{U}_s$ is a matrix whose column vectors are eigenvectors spanning the signal subspace and $\mathbf{\Lambda}_s$ contains the corresponding eigenvalues; $\mathbf{U}_n$ is the matrix whose column vectors are eigenvectors spanning the noise subspace and $\mathbf{\Lambda}_n$ contains the corresponding eigenvalues. The first D largest eigenvalues compose $\mathbf{\Lambda}_s$ and the rest form $\mathbf{\Lambda}_n$, and D is the expected number of signal sources. Third, the MUSIC spatial spectrum is generated over the angles θ and ϕ according to

$$P_M(\theta,\phi) = \frac{\mathbf{a}^H(\theta,\phi)\mathbf{a}(\theta,\phi)}{\mathbf{a}^H(\theta,\phi)\widehat{\Pi}^{\perp}\mathbf{a}(\theta,\phi)}. \tag{4}$$

$\widehat{\Pi}^{\perp} = \mathbf{U}_n\mathbf{U}_n^H$ is the noise subspace matrix, $\mathbf{a}(\theta,\phi) = [\mathbf{a}_1(\theta,\phi), \mathbf{a}_2(\theta,\phi), \ldots, \mathbf{a}_M(\theta,\phi)]^T$, where $\mathbf{a}_m(\theta,\phi) = [e^{-j\gamma R_m \sin\phi\cos(\theta-\upsilon_0)}, e^{-j\gamma R_m \sin\phi\cos(\theta-\upsilon_1)}, \ldots, e^{-j\gamma R_m \sin\phi\cos(\theta-\upsilon_{N_m-1})}]^T$ is the array manifold for ring m of the circular array, and $\gamma = 2\pi F_s/L$ is the wave number with L equal to the FFT size and F_s the sampling frequency.

Figure 4 shows the MUSIC spatial spectrum obtained from 4 concentric circular arrays with a total of $K = 30$ elements, which is the same as the 4th subarray in the 64-element array configuration presented in Section 6. The source signal used are two random amplitude narrowband signals at 500 Hz, coming from ($\theta = 90^\circ$, $\phi = 70^\circ$) and ($\theta = 45^\circ$, $\phi = 60^\circ$), respectively. As shown in Figure 4, the MUSIC spectrum contains 2 peaks suggesting 2 DOAs.

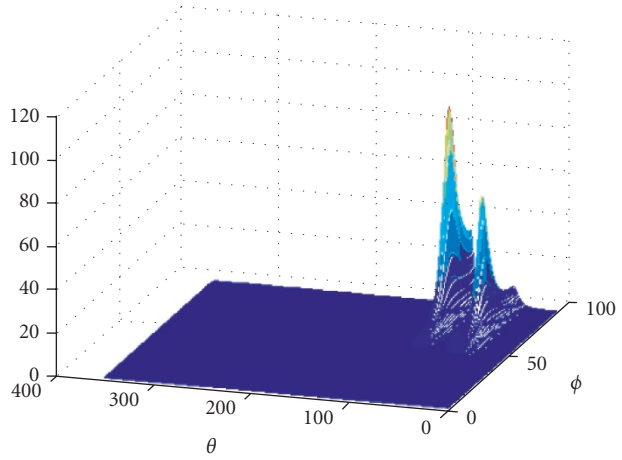


Figure 4: Narrowband MUSIC spectrum for 2 DOAs (500 Hz narrowband signal).

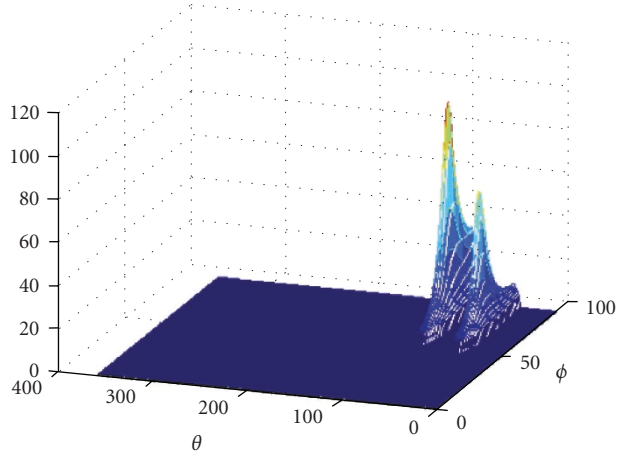


Figure 5: Narrowband MUSIC spectrum with small peaks from noise removed.

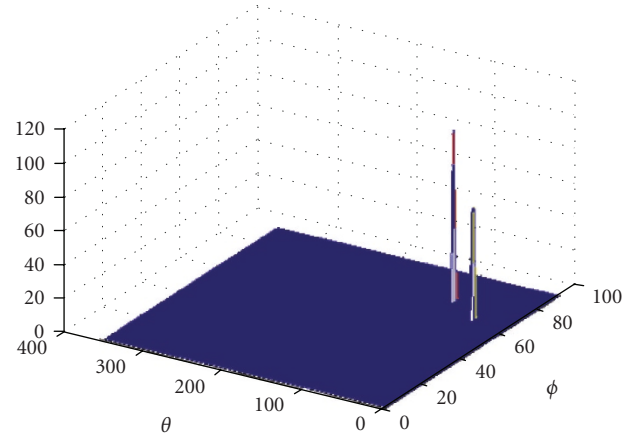


Figure 6: Narrowband MUSIC spectrum with only local maxima and minima.

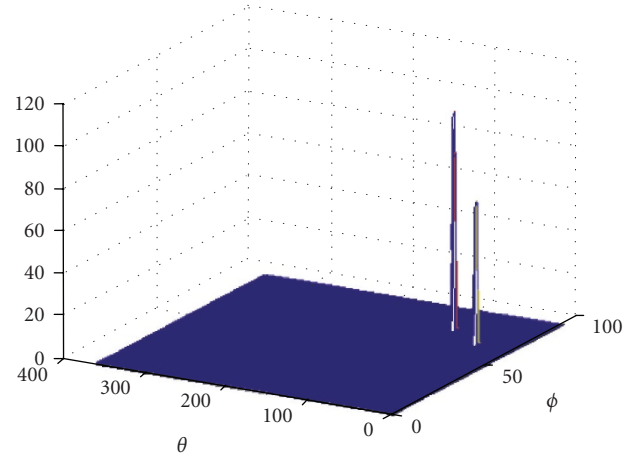


Figure 7: Narrowband MUSIC spectrum with only local maxima.

(ii) Two-dimensional peak searching algorithm

After the MUSIC spatial spectrum is obtained, the next task is to identify the location of those peaks in the spectrum which correspond to the DOAs. We use the MUSIC spectrum in Figure 4 as an example to illustrate the 2D peak searching algorithm as described below.

(1) A noise floor threshold is chosen to remove small local maxima. The MUSIC spectrum with small peaks removed is shown in Figure 5. The threshold is chosen experimentally by observing the floor level of the MUSIC spectrum. Other criteria should be used to enable automatic processing later.

(2) The 1st derivatives of $P(\theta, \phi)$ along θ and ϕ are computed. The zero-crossing locations of $dP(\theta, \phi)/d\theta$ and $dP(\theta, \phi)/d\phi$ are recorded. Regions of $P(\theta, \phi)$ around those zero-crossing points correspond to local minima and local maxima are kept for further processing. Other regions are removed. Figure 6 shows such a processed MUSIC spectrum. Note that local minima do not occur in this simulation case.

(3) After Step 2, the remaining regions contain both local maxima and minima. Among those regions, only local maxima have negative 2nd derivatives. Thus 2nd derivatives of $P(\theta, \phi)$ along θ and ϕ are computed. Only regions with both $d^2P(\theta, \phi)/d^2\theta < 0$ and $d^2P(\theta, \phi)/d^2\phi < 0$ are kept. Figure 7 shows the local maxima after this process.

Due to numerical precision problem, some peaks' locations may be lost in this step. Thus a smearing of those locations picked out by the 2nd derivatives condition is necessary. The smearing is done by enlarging the regions picked out by 1 more point in all directions.

(4) In this last step, the D peaks corresponding to the D DOAs are picked out. This is simply done by sequentially finding the largest D values in the remaining regions of $P(\theta, \phi)$. After the 1st peak is identified, a small region surrounding the 1st peak will be excluded from the remaining searches, and so on for the 2nd, ..., $(D-1)$th peaks. This is to ensure that smaller peaks instead of regions around larger peaks can be identified.

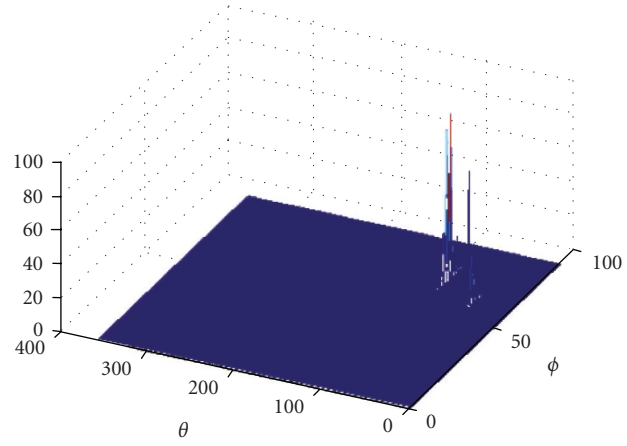


Figure 8: Combined narrowband DOA estimates.

(iii) Combining narrowband DOA estimation results to form the final DOA estimates

Using the circular array composed of 4 rings with about 30 elements, the narrowband DOA estimation results have bias, especially in the ϕ direction. We found out that when we use windowing to compute FFT of the array signal, the spectrum smearing of windowing will introduce bias in the result. To avoid smearing, longer window is preferred, and this also suggests that a larger number of spectral components generally give smaller bias in estimation result. Based on this observation, the estimated results from the narrowband MUSIC are combined in a way by taking their spectrum energy into consideration.

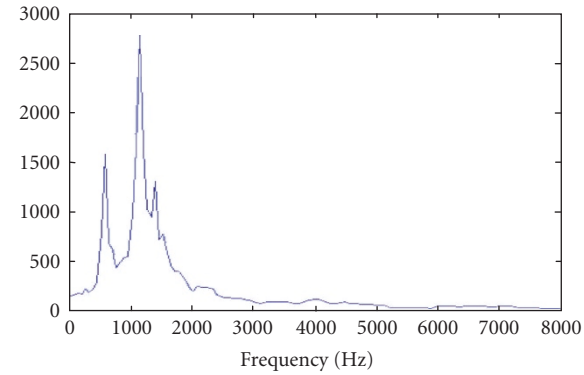
The peak value in the MUSIC spectrum will be associated with an estimated DOA as its confidence value. A histogram is generated to combine the narrowband DOA estimates using the confidence values of the estimated narrowband DOAs, and it is shown in Figure 8.

After obtaining the histogram of DOA estimates from different frequency components, the 2D peak searching algorithm described earlier is used again to Figure 8 to yield the final wideband DOA estimate.

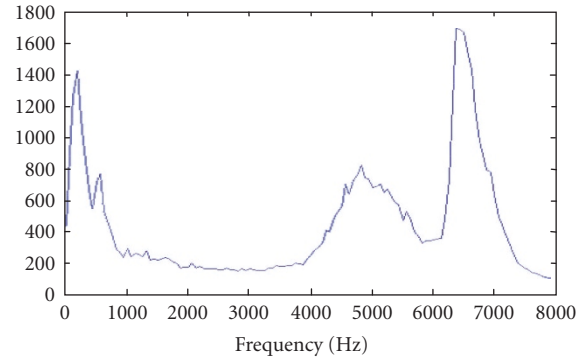
3.2. Statistical performance of the wideband DOA estimation algorithm

We used 2 bird sound files as the sources and generated the received array signals. One bird sound is Canada Goose located in the far field from the direction ($\theta = 90^\circ$, $\phi = 70^\circ$) and the other is Chip Sparrow also in the far field from the direction ($\theta = 45^\circ$, $\phi = 60^\circ$). The two sources have the same energy level. The power spectra of those 2 bird sounds are shown in Figure 9. The ambient noise level with respect to any one of the signals is -5 dB, 0 dB, and 5 dB, respectively, to create three scenarios.

Due to limitation in computational capacity, narrowband MUSIC is only performed for every other frequency



(a) Canada Goose



(b) Chip Sparrow

Figure 9: Spectrum of bird sounds.

index from 300 Hz up to 8 kHz. The narrowband MUSIC spatial spectrum is generated in the precision of 1° along θ, ϕ. 50 independent ensemble runs are conducted to generate the bias, variance, and MSE for the algorithm.

The statistical performance of the wideband DOA estimation technique is listed in Tables 1 and 2.

(1) Source 1: Canada Goose, true DOA ($\theta = 90^\circ$, $\phi = 70^\circ$).

(2) Source 2: Chip Sparrow, true DOA ($\theta = 45^\circ$, $\phi = 60^\circ$).

From 1 and 2 Tables, one can see that the algorithm gives very accurate DOA estimates under the SNRs used in the experiment. Further observation reveals the following.

(1) Bias, variance, and MSE all increase when the SNR decreases.
(2) Bias, variance, and MSE in θ are smaller than those of ϕ.
(3) Comparing the 2 signals, Chip Sparrow sound yields a slightly better performance. This may be due to several factors, such as the spectral content of the signal.

In short, the DOA estimation results are quite satisfactory and accurate enough for use in beamforming algorithm.

Table 1: Statistical performance of the proposed wideband DOA estimation technique for a given DOA.

SNR(dB)	Bias (deg.)		Variance (deg.)		MSE (deg.)	
	theta	phi	theta	phi	theta	phi
−5	−0.0200	0.8600	0.2996	2.4004	0.3000	3.1400
0	0.0400	0.3800	0.0784	2.4756	0.0800	2.6200
5	0	0.4	0	1.4400	0	1.6000

Table 2: Statistical performance of the proposed wideband DOA estimation technique for a given DOA.

SNR(dB)	Bias (deg.)		Variance (deg.)		MSE (deg.)	
	theta	phi	theta	phi	theta	phi
−5	0.0800	−1.0400	0.0736	0.1184	0.0800	1.2000
0	0	−1.0400	0	0.0384	0	1.1200
5	0	−1.0200	0	0.0196	0	1.0600

3.3. Comparison with DOA estimation results using linear array

The DOA ambiguity set of a linear array is a cone around the linear array. Thus it cannot be used to estimate the direction of a coming signal in 3D space. To illustrate the advantage of using a circular array instead of a linear array in DOA estimation, the MUSIC spectrum generated by an 11 element with half-wavelength spacing linear array is shown in Figure 10. There is only one narrowband signal at 500 Hz coming from $(\theta = 45^\circ, \phi = 60^\circ)$. The SNR is 3 dB. Although there is only one signal, there are two stripes of spectrum peaks, corresponding to the ambiguity set of a cone around the linear array. It is clear that for linear array it is not possible to yield an accurate DOA estimate without ambiguity.

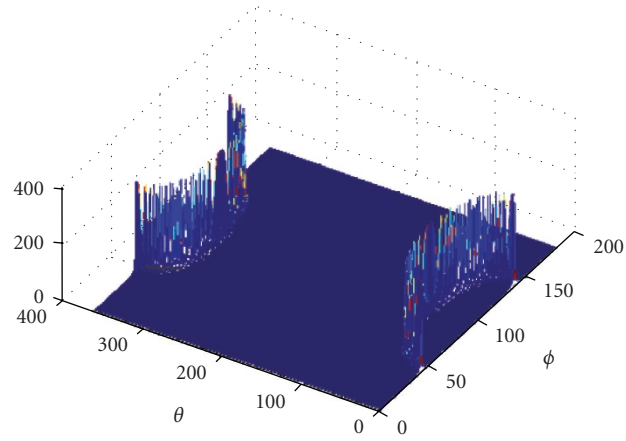


Figure 10: MUSIC spectrum for a linear array.

4. BEAMFORMING ALGORITHM FOR THE PROPOSED CONCENTRIC CIRCULAR ARRAY

The section will first present the beamforming algorithm for the concentric circular array shown in Figure 2. A compound ring structure is then described to make efficient use of array elements. This section closes with a comparison of the performance between the proposed concentric circular array and a linear array. For explanation purpose, we will first consider a narrowband input. For wideband inputs, the same procedures will be duplicated for multiple bands [6].

4.1. Beamforming algorithm

The output of the proposed beamformer is

$$z(n) = \sum_{m=1}^{M} w_m \sum_{i=1}^{N_m} h_{m,i} x_{m,i}(n), \tag{5}$$

where $x_{(m,i)}(n)$ is the received signal in microphone element i of ring m, $h_{m,i}$ is the intraring weights, and w_m is the interring weights. The proposed beamformer fixes the intrarings to be the delay-and-sum weights,

$$h_{m,i} = \frac{1}{N_m} e^{j\gamma R_m \sin\theta_o \cos(\phi_o - \upsilon_{m,i})}, \quad i = 1, 2, \ldots, N_m, \tag{6}$$

where (θ_o, ϕ_o) is the DOA of the desired signal. The novelty of the proposed beamformer is to select the interring weights to approximate a desired array pattern as illustrated below.

When we choose the intraring weights according to (6), the array pattern of ring m is

$$F_m(\theta, \phi) = \frac{1}{N_m} \sum_{i=1}^{N_m} e^{j\gamma R_m [\sin\theta_o \cos(\phi_o - \upsilon_{m,i}) - \sin\theta \cos(\phi - \upsilon_{m,i})]}. \tag{7}$$

Equation (7) can be expressed in terms of Bessel functions as [7]

$$\begin{aligned} F_m(\theta, \phi) &= J_o(\gamma R_m \rho) + 2 \sum_{q=1}^{\infty} j^{qN} J_{qN}(\gamma R_m \rho) \cos(qN\xi) \\ &\approx J_o(\gamma R_m \rho), \end{aligned} \tag{8}$$

where $J_n(\bullet)$ is the nth-order Bessel function of the first kind,

$$\rho = \sqrt{(\sin\phi\cos\theta - \sin\phi_o\cos\theta_o)^2 + (\sin\phi\sin\theta - \sin\phi_o\sin\theta_o)^2},$$
$$\xi = \arccos\left(\frac{\sin\phi\cos\theta - \sin\phi_o\cos\theta_o}{\rho}\right). \tag{9}$$

The approximation in the second line of (8) is becoming more accurate as the number of receiving elements in the ring increases. Since the beamformer output is the weighted sum of the outputs from the individual rings, the overall array pattern is

$$F(\theta,\phi) = \sum_{m=1}^{M} w_m F_m(\theta,\phi) \approx \sum_{m=1}^{M} w_m J_o(\gamma R_m \rho). \tag{10}$$

We now focus on the design of the intraring weights w_m to achieve a certain desirable beam pattern.

Given any real-valued function $g(y)$ continuous in $[0,1]$, it can be expressed as a Fourier-Bessel series as [8]

$$g(y) = \sum_{m=1}^{\infty} A_m J_o(\delta_m y), \quad 0 < y < 1, \tag{11}$$

where δ_m is the mth zero of $J_o(\bullet)$ arranged in ascending order. The coefficients A_m are given by

$$A_m = \frac{2}{J_1^2(\delta_m)} \int_0^1 \tau g(\tau) J_o(\delta_m \tau)\, d\tau. \tag{12}$$

Comparing (10) and (11), and establishing the mapping relationship

$$\rho = 2y, \quad y \in \left[0, \frac{1+\sin\theta_o}{2}\right], \tag{13}$$

we are able to approximate any desirable beam pattern $g(y)$ by choosing the ring radius as

$$R_m = \frac{\delta_m}{2\gamma} \tag{14}$$

and the interring weights as

$$w_m = A_m. \tag{15}$$

Equation (14) fixes the array structure and (15) provides the weights to combine the outputs from different rings. There is truncation error resulted from limiting the number of summation terms up to M in (11). The truncation error is not significant as the coefficient values A_m decrease as m increases. In any case, the number of rings M can be chosen such that the amount of trunction error is within certain tolerable limit.

Figure 11 shows a design example where the desired array pattern is chosen to be a Chebyshev function with -25 dB sidelobe level. The number of rings is 4, and the numbers of elements of the rings, starting from the ring, are 6, 10, 14, and 18. It is clear that the proposed design method is able to approximate the desired array pattern well.

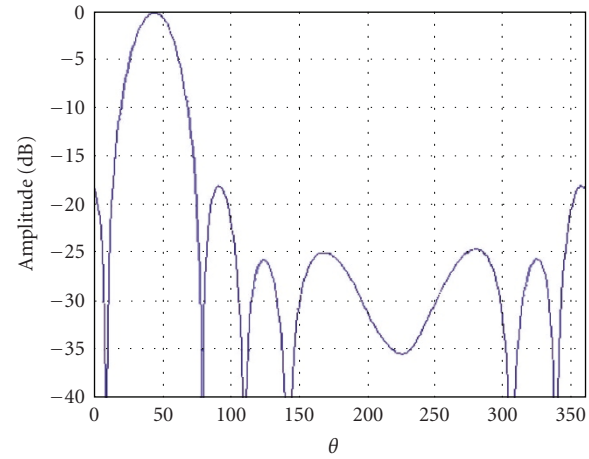


Figure 11: The beam pattern of the proposed circular array at 1 kHz.

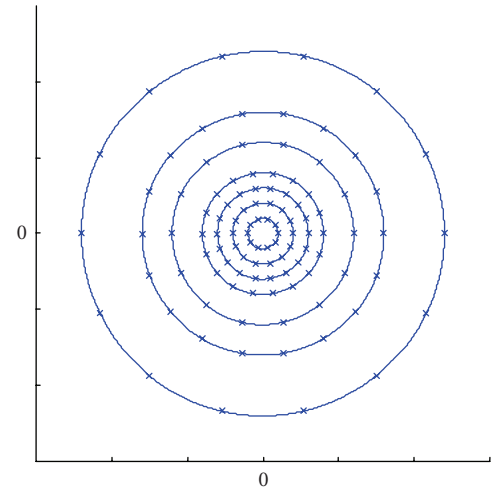


Figure 12: The proposed circular array configuration.

The above discussion is for a narrowband input. For a wideband input, we first separate the incoming data into frames, and apply FFT to decompose the input signal into narrowband components. The above design procedure is then applied at different narrowband components and the resultant output is obtained through inverse FFT. The proposed design method above can also achieve frequency invariant beam pattern and the details can be found in [7].

4.2. Compound ring structure

In the bird monitoring system, we have designed a concentric circular array that has 7 rings and 102 elements. The radius of the array is about 0.5 m which is very compact. Figure 12 is the array structure. One novelty of the proposed design is that the circular array can perform wideband beamforming, and the compound ring approach is utilized to make efficient use of array elements. In the compound ring structure, some rings are shared by several frequency bands and

Table 3: Grouping of rings into different subarrays for broadband beamforming.

	Approx. operating frequency range	Number of rings	Number of elements in each ring
Subarray 1	250 Hz–700 Hz	3	[6, 10, 14]
Subarray 2	700 Hz–1.5 kHz	4	[6, 10, 14, 18]
Subarray 3	1.5 kHz–3.5 kHz	4	[6, 10, 14, 18]
Subarray 4	3.5 kHz–8 kHz	2	[10, 14]

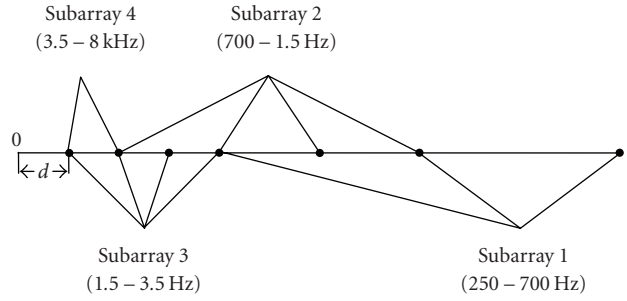


Figure 13: Grouping of the rings in the four subarrays.

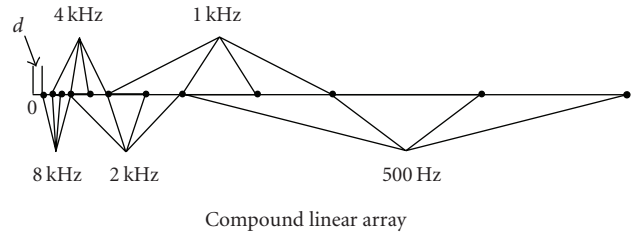


Figure 14: Configuration of compound linear array.

therefore resulting in savings in array elements. The proposed compound ring structure has 4 operating frequency bands as listed in the second column of Table 3. The third column in the table shows the number of rings in each band and the fourth column is the number of elements in each ring for the frequency band considered. The grouping of the rings for different bands is shown in Figure 13. The minimum separation between two rings is

$$d = \frac{1}{4}\frac{\delta_4 \lambda_{2\,\mathrm{kHz}}}{4\pi} = 0.0402\,\mathrm{m}. \tag{16}$$

The largest radius, and hence the size of the array, is

$$12d = 0.4827\,\mathrm{m}. \tag{17}$$

The details in deriving (16) and (17) are available in [6]. Although (14) fixes the radius of the rings, interpolation technique [9] is used to relax this constraint. Because of reusing array elements in different subarrays, the total number of elements is 10 + 14 + 14 + 18 + 14 + 18 + 14 = 102.

In general, the larger the number of rings in a subarray, the larger will be the attenuation in the ambient noise level. The power spectral density of birds has higher energy from 700 Hz–4 kHz. That is why subarrays 2 and 3 have 4 rings to provide larger attenuation to the noise.

Figure 11 is a typical beam pattern of the proposed circular array at 1 kHz. A main advantage of the proposed design is that it provides close to a fixed level of residue sidelobes of about −25 dB.

4.3. Beampattern comparison with linear array

For comparison purpose, a compound linear array that has the same number of array elements as the proposed circular array (102 elements) is used. The compound linear array composes of 5 subarrays operating at frequency ranges around 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz, respectively. Each subarray contains 34 elements. Half of the elements from a subarray of higher frequency will be reused in the following lower frequency subarray. Thus total number of elements is 34 + 17 ∗ 4 = 102. *Error! Reference source not found* (Figure 14) shows a compound linear array with 5 subarrays and 4 elements within each subarray. (Subarrays with as much as 34 elements are difficult to show.)

The smallest distance between two array elements is

$$d = \frac{\lambda_{8\,\mathrm{kHz}}}{2} = 0.0214\,\mathrm{m}. \tag{18}$$

The size of the 102-element compound linear array is

$$(34 - 1) \times \frac{\lambda_{500\,\mathrm{Hz}}}{2} = 11.32\,\mathrm{m}, \tag{19}$$

which is very large.

Because of the compound array structure, the beam pattern for different center frequency is same. A 3D beam pattern for one of the subarray is shown in Figure 15, the DOA is assumed to be ($\theta = 45^\circ$, $\phi = 45^\circ$). A linear array has an ambiguity region that appears as a cone.

The compound ring array used is the one described earlier. It has 7 rings and contains 102 elements. The array diameter is about 1 m. The 3D beam pattern for one of the subarray is shown in Figure 16, the DOA in assumed to be ($\theta = 45^\circ$, $\phi = 45^\circ$). Here only two ambiguity angles appear: one is above and the other is below the microphone plane.

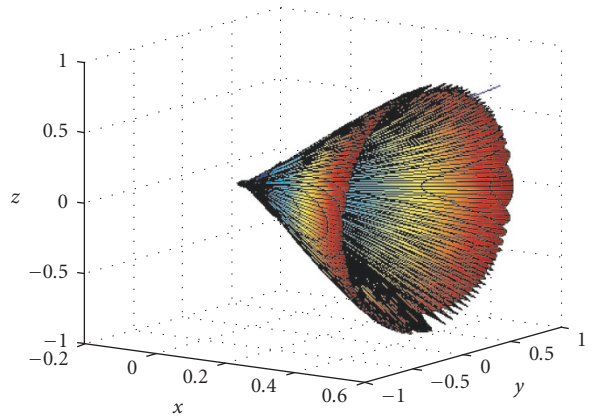


Figure 15: 3D beam pattern of compound linear array ($\theta_0 = \pi/4$, $\phi_0 = \pi/4$).

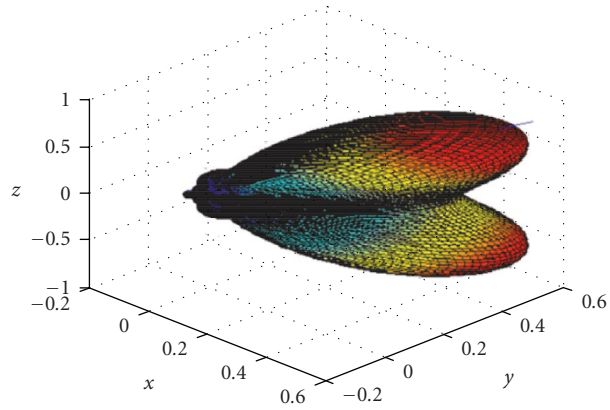


Figure 16: 3D beam pattern for ring array. ($\theta_0 = \pi/4$, $\phi_0 = \pi/4$).

Since the bird monitoring application requires the monitoring of the half-space above the microphone array, there is no angular ambiguity.

4.4. Comparison of directional interference rejection between a linear array and a circular array

The arrays used in the examples are subarrays from the previous compound arrays in Sections 4.2 and 4.3.

Linear array configuration

Here we used 34 equally spaced elements, operating at 1 kHz signal. Details of the array are described in Section 4.3. The beam pattern at 1 kHz is shown in Figure 17.

Circular array configuration

It is the subarray 2 described in Section 4.2 that operates between 700 Hz and 1.5 kHz. This subarray consists of 4 rings with a total of 48 elements. The weights are selected to achieve −20 dB sidelobe level for 1 kHz signal. The array

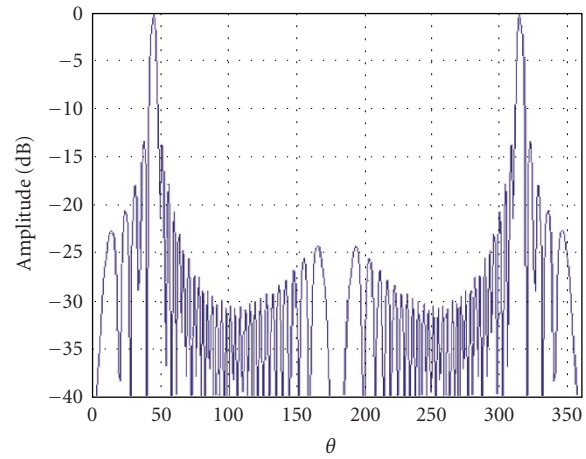


Figure 17: Beampattern for a linear array.

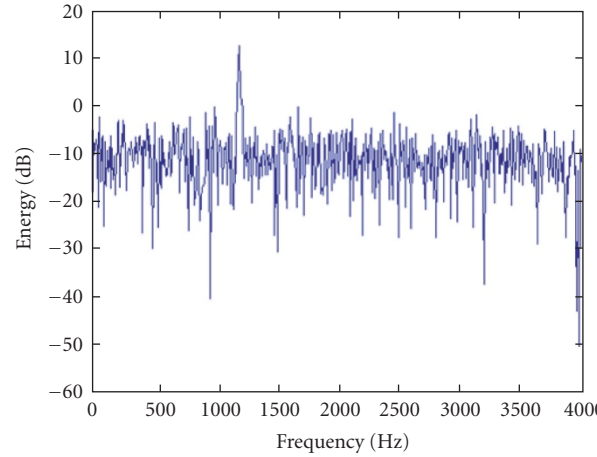


Figure 18: Received signal in one channel. Interference signal is coming from a DOA in the ambiguity set.

pattern is similar to that shown in Figure 11 with 5 dB higher sidelobe level but narrower main-beam width.

Here we assume that the interference signal is in the ambiguity set of the linear array. The DOAs and SIR and SNR are given by

(i) signal source 1 kHz signal with DOA ($\theta_0 = \pi/4$, $\phi_0 = \pi/2$);
(ii) interference 1200 Hz signal with DOA ($\theta_0 = 0$, $\phi_0 = 0.6301$);
(iii) signal-to-interference ratio SIR = −15 dB;
(iv) signal-to-ambient-noise ratio SNR = 0 dB.

(1) Figure 18 shows the signal received in one array element. The source signal is hidden in noise, and only the 1200 Hz interference is visible.

(2) Figure 19 shows the linear array output. It can be seen that both the 1200 Hz interference and 1 kHz signal are strengthened, but the 1200 Hz interference is still about 15 dB stronger than the 1 kHz signal.

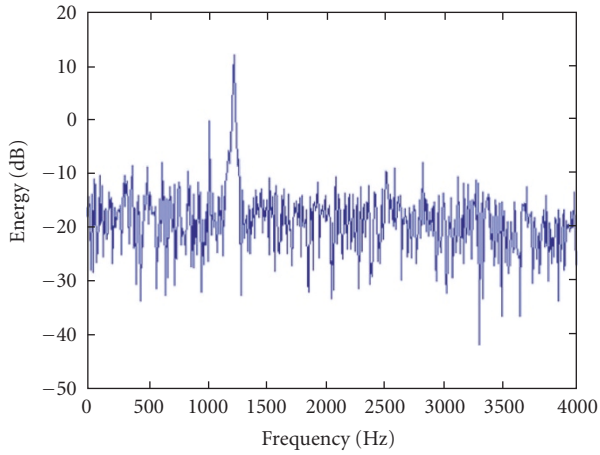


Figure 19: Output of the linear array. Interference signal is coming from a DOA in the ambiguity set.

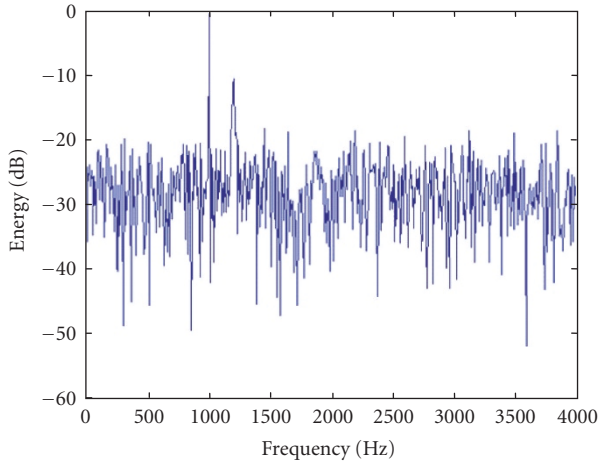


Figure 20: Output of the ring array. Interference signal is coming from a DOA in the ambiguity set.

(3) Figure 20 shows the output of a circular array with −20 dB desired sidelobe level. The target 1 kHz signal is strengthened and becomes obvious. The 1200 Hz interference had about −20 dB attenuation.

(4) Figure 21 shows the output of a circular array with a null placed in the DOA of the 1200 Hz interference. The 1200 Hz signal is completely eliminated.

Based on the above comparisons, we concluded the following.

(i) Circular array has an ambiguity set of direction of arrival (DOA) of only 2 directions, while linear array has a larger ambiguity set of (DOA) which is cone.

(ii) The beam pattern of circular array can be rotated to arbitrary direction in the x-y plane without suffering great fluctuation. This is not the case for linear array.

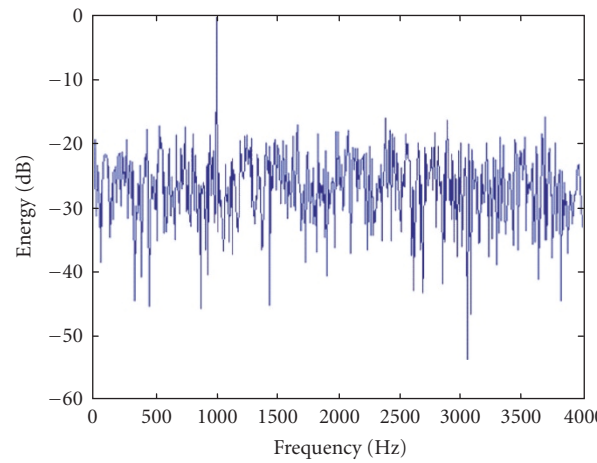


Figure 21: Output of ring array with null at the DOA of 1200 Hz interference. Interference signal is coming from a DOA in the ambiguity set. A null is created in the direction of the interference signal.

(iii) Compound linear array is incapable of attenuating directional interference in the DOA ambiguity set, circular array has much less ambiguity set, thus it can remove the directional interference in most cases linear array fails.

5. BIRD CLASSIFICATION ALGORITHM USING GMM

According to the evaluations done by National Institute of Standards and Technology (NIST) engineers [10], GMM has been proven to be quite useful in speaker verification applications. The birds have similar spectrum as humans. The individual component densities of a multimodal density may model many underlying sets of acoustic classes. A linear combination of Gaussian basis functions is capable of representing a large class of sample distributions.

The bird classification consists of two major steps: (1) preprocessing the extract features; (2) applying GMM models to classify different birds.

5.1. Preprocessing to extract features of birds

To identify the bird species, the algorithm we have been using is to first extract the feature vectors from the bird sound data, then match these feature vectors with GMMs, each trained specifically for each bird class. The difference between the probabilities is compared to a preset threshold to decide if a given bird sound belongs to a specific bird class.

The feature extraction subsystem can be best described by Figure 22. This architecture has been implemented for human speaker verification [10, 11]. The bird sound spectrum lies between a few hundred Hz to 8 kHz and is quite similar to that of human's.

The purpose of feature extraction is to convert each frame of bird sound into a sequence of feature vectors. In our system, we use cepstral coefficients derived from a mel-frequency filter bank to represent a short-term bird sound

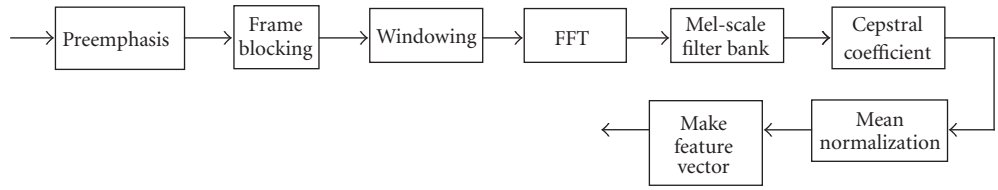


Figure 22: Preprocessing steps in the feature extraction subsystem.

spectrum. The digital bird sound data is first preprocessed (preemphasized, set to overlapped frames and windowed) and then mel-frequency cepstral coefficient analysis is applied. Currently we have 13 mel-frequency cepstral coefficients in each feature vector to describe the bird sound within a short period of time (~10 milliseconds), in which the bird sound is quasistationary. The first-order and second-order differences of these coefficients can also be included in the feature vector. So there are 39 feature elements in total that we can use to build feature vectors. Tests were performed with different features and the results showed that the feature vector with the 13 mel-frequency cepstral coefficients can achieve the best classification result.

5.2. Gaussian mixture model for birds

The GMM for birds is a probabilistic model by which the distribution of data is modeled as a linear combination of several multivariate Gaussian densities. There are two motivations for using Gaussian mixture densities as a representation of bird identity [12]. The first is the intuitive notion that the individual components densities of a multimodal density, like the GMM, may model some underlying set of acoustic classes. The second motivation is the empirical observation that a linear combination of Gaussian basis function is capable of representing a large class of sample distribution. The GMM is usually trained with the expectation-maximization (EM) algorithm to maximize the likelihood of the observation data from an individual class.

For bird classification, each bird is represented by a model. The GMM can have several different forms depending on the choice of covariance matrices. The model can have one covariance matrix per Gaussian component (nodal covariance), one covariance matrix for all Gaussian components in a speaker model (grand covariance), or a single covariance matrix shared by all speaker models (global covariance). The covariance can also be full or diagonal.

A free Matlab toolbox called Netlab was used to perform the GMM model estimation. The toolbox was developed by Ian T. Nabney at Aston University in the UK and it provides many useful Matlab functions for speech processing.

For bird classification, a group of S bird classes is represented by GMM's $\lambda_1, \lambda_2, \ldots, \lambda_S$. The objective is to find the bird model, which has the maximum a posteriori probability for a given observation sequence. Assuming equally likely birds, the classification is quite simple. In practice, for a given airport, the bird species information can be obtained easily. Hence, the *a priori* probability information can be incorporated accordingly.

5.3. Bird classification performance using GMM

Bird database

Table 4 summarizes the bird sound file numbers and the feature numbers that can be extracted from these sound files. The sound files were purchased from Cornell University. These bird species are quite dangerous to aircraft because of their masses and sizes.

Classification performance

Effect of feature window lengths on classification performance

We fixed the number of Gaussian mixture functions to 20, and tested the classification of bird sound segments with different feature window lengths. For example, a bird sound segment of 0.7 seconds can produce 70 features in the feature window. Even-numbered feature windows were used for training of GMMs and odd-numbered feature windows were used for testing.

As we can see from Figure 23, increasing the feature window length will improve the classification performance by decreasing the error rate. This means a longer bird sound segment can bring a more accurate classification result.

Effect of number of Gaussian functions on classification performance

We fixed the length of feature window to $w = 20$ (a 0.2-second segment of bird sound), and tested the classification of bird sound segments with different numbers of Gaussian functions. Even-numbered feature windows were used for training of GMMs and odd-numbered feature windows were used for testing.

As we can see from Figure 24, increasing the Gaussian mixture function number will generally improve the classification performance by decreasing the error rate. But the tradeoff here is that a large number of Gaussian mixture functions will increase the computational complexity, especially with a large training set. Also, it will cause over fitting problem, which will decrease the performance of the noisy data.

Table 4: Summary of bird sound files and feature numbers for each bird class.

Bird species	Number of bird sound files	MFCC feature numbers	Class number
Mourning Dove	104	51723	1
Herring Gull	78	34241	2
Ring-billed Gull	152	60147	3
Laughing Gull	80	35783	4
Franklins Gull	34	16140	5
Red-tailed Hawk	318	122223	6
Sharp-shinned Hawk	104	36429	7
Canada Goose	83	43765	8
Snow Goose	61	26662	9
Mallard	82	36137	10
Sandhill Crane	451	178482	11

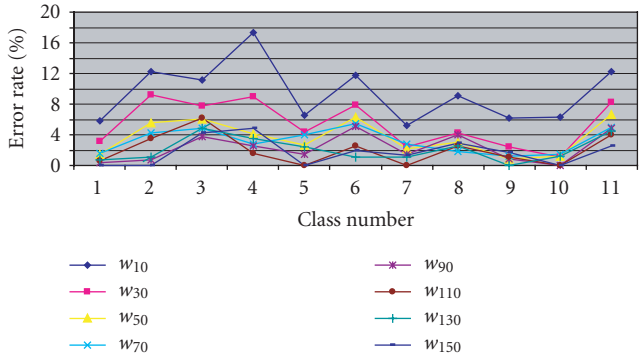


Figure 23: Error rate of 11 classes with different feature window length.

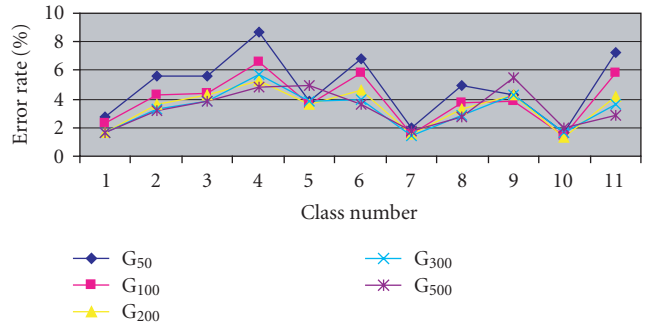


Figure 24: Error rate of 11 classes with different Gaussian mixture function number.

Effects of different training and testing methods

In the above two tests, we used the odd-number feature window in each file for training, and even-number feature window in each file for testing. Here we randomly pick up some feature windows for training and use the rest feature windows for testing. We have the following two cases.

(a) Half of the features extracted from each file were randomly selected to train the GMM model, and the rest half were used for testing. The dimension of the features was 13, and the Gaussian mixture number was 10. All the 11 classes of bird files were used to perform the test, and there were totally 30 runs. Figure 25 shows the mean error rate of the 30 runs, and the error rate range calculated by mean and standard deviation.

The results from the 30 runs are very similar. The mean error rates for all the 11 classes are small, and the standard deviations are small.

(b) Half of the files for each class were randomly selected to extract features for training and the remaining files were used to extract features for testing. The dimension of the feature is 13, and the Gaussian mixture number was 10. All the 11 classes of bird files were used to perform the test, and there were totally 20 runs. Figure 26 shows the mean error rate of the 20 runs, and the error rate range calculated by mean and standard deviation.

The result is not as good as in case (a), which is reasonable because of the inconsistency between different files. Also, the standard deviation is larger.

Effect of different SNRs on classification performance

We examined the SNR of the bird sound files and evaluated the classification performance for each SNR category. The standard deviation of the bird sound signal in each frame (10 milliseconds) was calculated for each bird file, and two times the minimum of this standard deviation is selected as the threshold to separate the "noise" and the "signal." We assumed that the noise energy is not changing much within each bird sound file. So for each bird sound file, we calculated a mean noise energy. But since the signal energy actually changes a lot within each bird file, we then associated a signal energy with each feature frame so that with each feature window, we can calculate the corresponding SNR for that feature

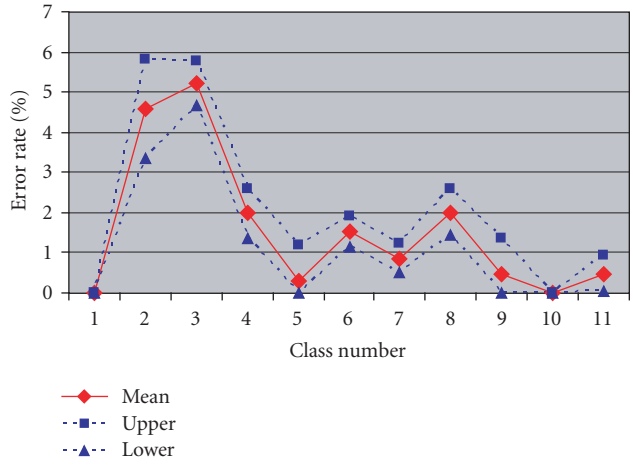


FIGURE 25: Test result of using half features randomly selected from each file for training, and the rest half for testing (30 runs).

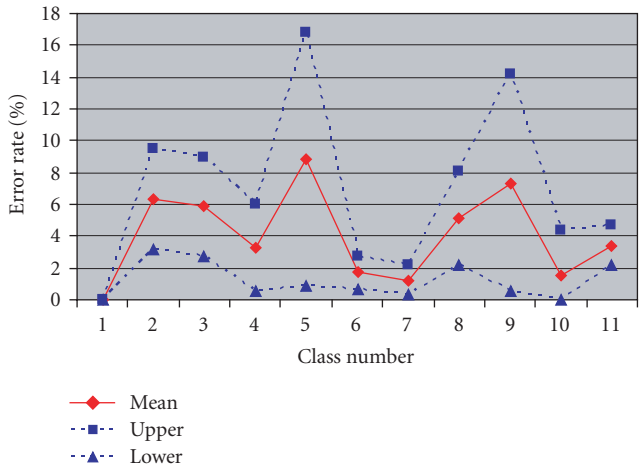


FIGURE 26: Test result of using half randomly selected files for training, and the rest half files for testing (20 runs).

FIGURE 27: Total error rates for all 11 classes under different SNR conditions.

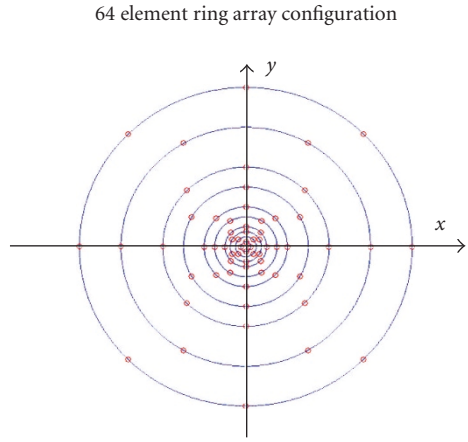


FIGURE 28: Circular array configuration.

window. The feature window length is 20, which corresponds to a 0.2-second sound segment. The Gaussian mixture function number is 200. Figure 27 plots the error rates for all the 11 classes in each different SNR category. We can see that the classification error rate is decreasing when SNR is increasing (with one exception: 15 < SNR < 20). This exception was likely the result of evaluating the SNR, which may have some fluctuations within a short segment of sound file. The average error rate is about (60/11) = 5% for the 5 dB SNR case. This is very remarkable and clearly shows that GMM is suitable for low SNR situations.

6. MICROPHONE DISH DESIGN AND HARDWARE PROTOTYPE

6.1. *Array configuration*

The compound ring array is composed of 10 rings.

(i) Subarray 1 is composed of the 1 2 3 4th rings, with elements distributed as 4 4 6 8.

(ii) Subarray 2 is composed of the 2 4 5 6th rings, with elements distributed as (4) (8) 6 8.

(iii) Subarray 3 is composed of the 4 6 7 8th rings, with elements distributed as (8) (8) 6 8.

(iv) Subarray 4 is composed of the 6 8 9 10th rings, with elements distributed as (8) (8) 6 8.

The numbers in brackets are elements reused from a previous subarray. The total number of elements is $4+4+(6+8)*4 = 64$.

The top view of the ring array is shown in Figure 28. Table 5 summarizes the ring radii. The largest radius is

$$R_{10} = \frac{\delta_4 \lambda_{500\,\mathrm{Hz}}}{4\pi} = 0.6436178\,\mathrm{m}, \tag{20}$$

where δ_4 is the 4th root of the 0th-order Bessel function of

Table 5: Summary of ring radii.

Ring	1	2	3	4	5	6	7	8	9	10
Radius	R_1	$2R_1$	$3R_1$	$4R_1$	$6R_1$	$8R_1$	$12R_1$	$16R_1$	$24R_1$	$32R_1$
Number of elements	4	4	6	8	6	8	6	8	6	8

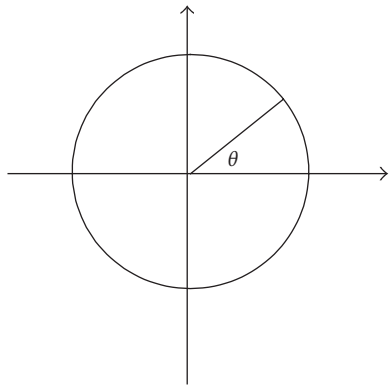


Figure 29: Definition of the angle.

the 1st kind and $\delta_4 = 11.79$. The smallest radius is

$$R_1 = \frac{1}{32} R_{10} = 0.02011306\,\text{m}. \tag{21}$$

Array elements on each ring are equally distributed. Instead of starting all the 1st element of each ring at $\theta = 0^\circ$ (see Figure 29), we use the following scheme to form a more symmetric structure for the ring array. This will reduce the fluctuation in beam pattern when signal's DOA is rotated in the plane where the ring array lies.

The 1st element of the 1st ring is placed at $\theta = 0^\circ$ and the remaining elements on that ring are anticlockwise, equally placed along the circle. For the 2nd ring, its 1st element is placed at $\theta = \pi/4$ and the remaining elements on that ring are also anticlockwise, equally placed along the circle, and so on for the other rings. Table 6 below lists the locations (expressed by degree value of θ) of each element on the rings.

6.2. Construction of the circular array

Based on the circular array configuration described in Section 5.1, we constructed a microphone array as shown in Figure 30. The array board was spliced by two pieces of wood. Holes were drilled and microphones were put into them.

6.3. Data acquisition electronics associated with the circular array

It is emphasized that it is a very challenging task to simultaneously digitize 64 channels of audio signals at 22 kHz sampling rate. With the help of NIST engineers, we built data acquisition hardware for the microphone array. Figure 31 shows how we implement the data acquisition system. Basically, there are two main parts in this system. One is the microphone data acquisition board which is analog and performs A/D, and the other is the motherboard which is digital and sends digital data to computer through Ethernet.

Microphone data acquisition board

Each microphone array board has 8 microphones. We need to have at least 64 microphones in our application, so there will be 8 analog boards in this system. There are three circuit parts on each board:

(i) the microphone amplification stage,
(ii) the digitalization stage,
(iii) the motherboard connection stage.

Motherboard

The field-programmable gate array (FPGA) motherboard is the gathering part of the data acquisition system. The cables at the top of the picture are the data collection cables connected to the 8 analog boards. The red DIP is here to fix the MAC address of the microphone array. The LEDs are to give the status of the microphone array.

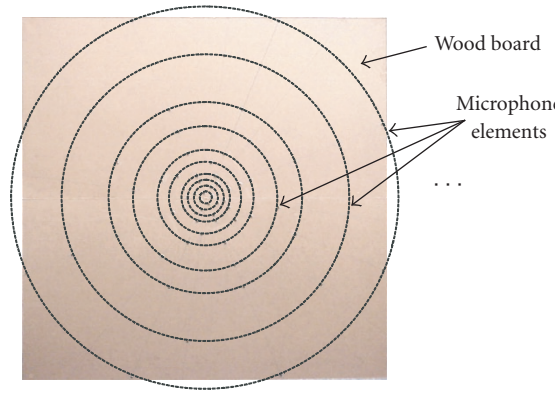
7. EXPERIMENTAL RESULTS

7.1. DOA estimation results

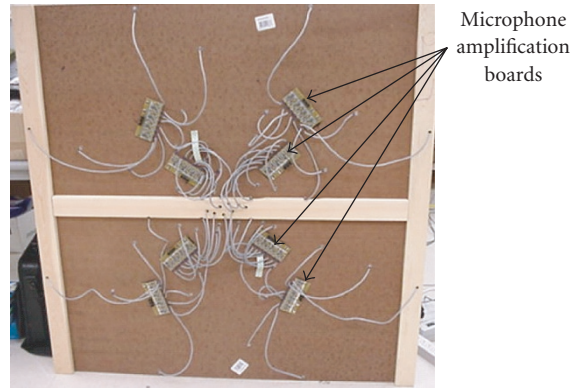
The microphone dish system was fully tested by taking many measurements in our laboratory as well as in an open field (parking lot). Table 7 summarizes one experiment in the open field. The experimental setup involved microphone dish and the data acquisition system to collect the data, two laptops displaying the sound sources. One source emulated the bird and the other emulated an aircraft. The distance between the sources and the microphone dish was about 40 ft. It can be seen that there are about 10% difference between the estimated and expected angles in the θ direction and 2% difference in the ϕ direction. The expected angles should not be treated as true DOA values as there are three major sources of errors. One is that the distance between the sources and the microphone dish is still not long enough. Hence, the microphone dish cannot be treated as a point. The second reason is that distance measurements were done manually and may have some inherent errors. The third reason is that the number of elements in the dish is only 64. The angular resolution will be much better if more than 100 elements are used in the array.

Table 6: Angular distribution of array elements.

	1	2	3	4	5	6	7	8
Ring 1	0	90	180	270	—	—	—	—
Ring 2	45	135	225	315	—	—	—	—
Ring 3	90	150	210	270	330	30	—	—
Ring 4	135	180	225	270	315	0	45	90
Ring 5	180	240	300	0	60	120	—	—
Ring 6	225	270	315	0	45	90	135	180
Ring 7	270	330	30	90	150	210	—	—
Ring 8	315	0	45	90	135	180	225	270
Ring 9	0	60	120	180	240	300	—	—
Ring 10	45	90	135	180	225	270	315	0



(a) Front view of the circular array

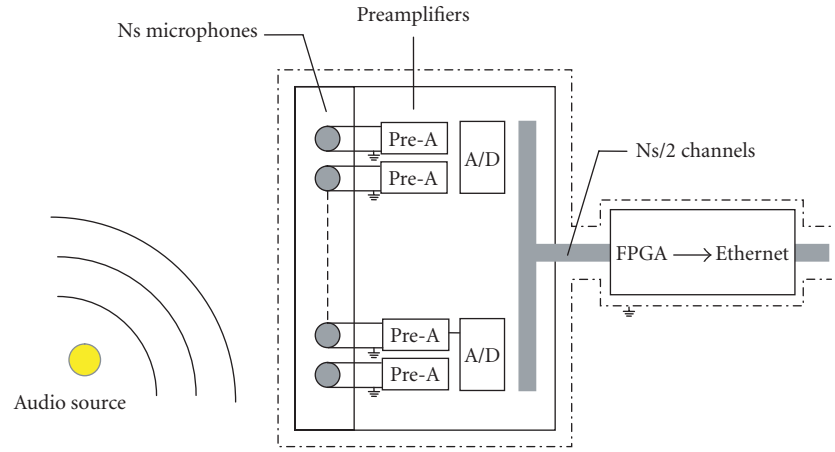


(b) Rear view of the array

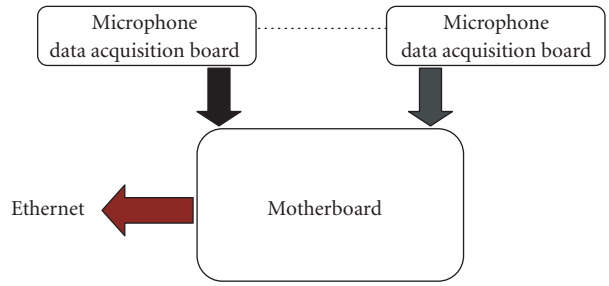
Figure 30: Finished circular microphone array. It has a dimension of 1.2 meter in diameter.

7.2. *Beamformer outputs*

Once the DOAs are estimated, the beamforming algorithm eliminated the effects of interference and collected a clean bird signal. Figure 32 shows the signals before and after the beamforming. It can be seen that before the beamforming, the signals were very noisy. However, after the beamforming, the signals were very clean. We could not hear any interference and background noise after the beamforming.



(a) Data acquisition schematics



(b) Hardware relationships

Figure 31: Data acquisition system for the circular microphone array.

Table 7: Summary of DOA estimation and the expected DOA values.

	DOA estimation		Expected DOA	
	θ	ϕ	θ	ϕ
Interference 1	171	30	180	28.52
Bird 1	329	40	360	40.06

7.3. Bird classification performance

We set up our microphone array in the parking lot, with one PC speaker playing bird sounds, and another PC speaker playing aircraft noise. Table 8 is the result of bird classification. It can be seen that the GMM algorithm correctly identified the bird species. The real error rate for real open-field experiments is part of future work.

8. CONCLUSIONS

In this paper, we described a new acoustic system for bird monitoring. The system has been designed, implemented, tested, and evaluated. The array consists of a number of concentric rings of microphones. The proposed circular microphone array has only two DOA ambiguities whereas a typical conventional linear array has a cone of ambiguities. Moreover, the circular array is directional and has a long range. Real-time data acquisition electronics have been designed and implemented for this array. Our current system can digitize 64 microphones simultaneously at a sampling rate of 22 kHz. A wideband DOA estimation algorithm for the circular array has been developed and evaluated as well. The performance of the DOA estimation is much better than that of linear arrays. A novel algorithm for wideband beamforming was also developed in this research. The ring and element spacing is carefully designed to yield optimal performance in beamforming and costs. As a result, the circular array has much better directional interference rejection capability than linear arrays. A bird classification algorithm based on Gaussian mixture model (GMM) has been designed and evaluated. GMM can achieve an average of 90% classification rate for 5 dB SNR. Finally, extensive laboratory tests and field experiments have been successfully performed.

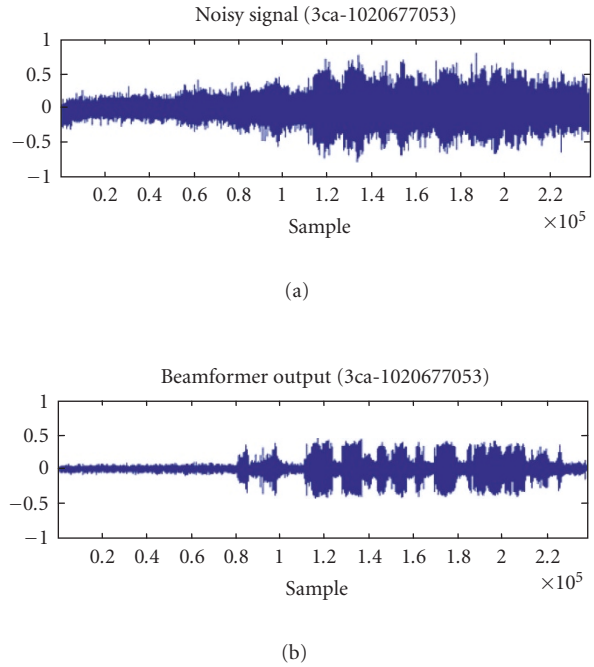


Figure 32: Signals before beamforming and after beamforming.

Table 8: Experimental results of bird classification with beamforming.

	Prob. as Canada Goose	Prob. as Chip Sparrow	Decision
Canada Goose test data without aircraft noise	**−89.1120**	−103.2840	Canada Goose
Canada Goose test data with aircraft noise	**−100.9485**	−149.3710	Canada Goose
Chip Sparrow test data without aircraft noise	−131.2033	**−92.0052**	Chip Sparrow
Chip Sparrow test data with aircraft noise	−110.9642	**−92.4776**	Chip Sparrow

ACKNOWLEDGMENTS

This research was supported by the Air Force Office of Scientific Research under Contract F49620-02-C-0044. Prior research supported by NIST under Contract SB1341-02-W-1140 laid a solid foundation for this work. The encouragements from Dr. Willard Larkin are deeply appreciated.

REFERENCES

[1] Federal Aviation Agency, "Wildlife strikes to civil aircraft in the United States," 2001.

[2] C. Kwan, K. C. Ho, G. Mei, et al., "An automated acoustic system to monitor and classify birds," in *Proceedings of 5th Annual Meeting of Bird Strike Committee-USA/Canada (Bird Strike '03)*, Toronto, Canada, August 2003.

[3] K. M. Buckley and L. Griffiths, "Broad-band signal-subspace spatial-spectrum (BASS-ALE) estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 953–964, 1988.

[4] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[5] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 823–831, 1985.

[6] C. Kwan, K. C. Ho, G. Mei, et al., "Phase 1 Progress Report 3 to the Air Force," Internal Research Report, Intelligent Automation, Poway, Calif, USA, 2003.

[7] Y. Li, K. C. Ho, and C. Kwan, "3-D array pattern synthesis with frequency invariant property for concentric ring array," to appear in *IEEE Transactions on Signal Processing*.

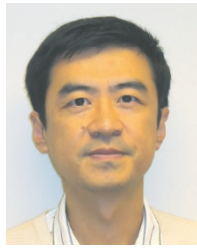
[8] E. Kreyszig, *Advanced Engineering Mathematics*, John Wiley & Sons, New York, NY, USA, 7th edition, 1993.

[9] C. O. Stearns and A. C. Stewart, "An investigation of concentric ring antennas with low sidelobes," *IEEE Transactions on Antennas and Propagation*, vol. 13, no. 6, pp. 856–863, 1965.

[10] NIST web site, http://www.nist.gov/speech/tests/spk.

[11] NIST web site, http://www.nist.gov/smartspace/.

[12] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 1, pp. 72–83, 1995.



C. Kwan received his B.S. degree in electronics with honors from the Chinese University of Hong Kong in 1988 and his M.S. and Ph.D. degrees in electrical engineering from the University of Texas at Arlington in 1989 and 1993, respectively. From April 1991 to February 1994, he worked in the Beam Instrumentation Department of the Superconducting Super Collider (SSC) Laboratory in Dallas, Texas, where he was heavily involved in the modeling, simulation, and design of modern digital controllers and signal processing algorithms for the beam control and synchronization system. He received an invention award for his work at SSC. Between March 1994 and June 1995, he joined the Automation and Robotics Research Institute in Fort Worth, where he applied neural networks and fuzzy logic to the control of power systems, robots, and motors. Since July 1995, he has been with Intelligent Automation, Inc. in Rockville, Maryland. He has served as the Principal Investigator/Program Manager for more than 65 different projects, with total funding exceeding 20 million dollars. Currently, he is the Vice President, leading research and development efforts in signal/image processing and controls. He has published more than 39 papers in archival journals and has had 90 additional refereed conference papers.

K. C. Ho received the B.S. degree with first class honors in electronics and the Ph.D. degree in electronic engineering from the Chinese University of Hong Kong, Hong Kong, in 1988 and 1991, respectively. He was a Research Associate in the Royal Military College of Canada from 1991 to 1994. He joined the Bell-Northern Research, Montreal, Canada, in 1995 as a member of the scientific staff. He was a faculty member in the Department of Electrical Engineering at the University of Saskatchewan, Saskatoon, Canada, from September 1996 to August 1997. Since September 1997, he has been with the University of Missouri-Columbia, where he is currently an Associate Professor in the Electrical and Computer Engineering Department. His research interests are in statistical signal processing, array processing, source localization, subsurface object detection, wavelet transform, wireless communications, and the development of efficient adaptive signal processing algorithms for various applications including landmine detection, echo cancellation, and time-delay estimation. He is currently serving as an Associate Editor of the IEEE Transactions on Signal Processing and the IEEE Signal Processing Letters. He received the Junior Faculty Research Award from the College of Engineering of the University of Missouri-Columbia in 2003. He has three United States patents, three Canadian patents, and three European patents in the area of mobile communications.

G. Mei received his B.E. degree in electrical engineering and his M.E. degree in communications and electrical systems from the University of Science and Technology of China in 1995 and 1998, and his M.S. degree in communications from the University of Maryland at College Park in 2000. Since June 2002, he has worked as a Research Engineer in Intelligent Automation, Inc. in Rockville, Maryland. He has published one paper in archival journals and has had four additional refereed conference papers.

Y. Li was born in China. He received the B.S. and M.S. degrees in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1997 and 2000, respectively, and the Ph.D. degree in signal processing from the University of Missouri-Columbia, USA, in 2005. He is currently a System Engineer with National Semiconductor Corp., Santa Clara, Calif. His research interests include array processing, audio and speech processing.

Z. Ren received her B.S. (1983) and M.S. (1989) degrees in electrical engineering from the Beijing University of Aeronautics and Astronautics, China. She also obtained an M.S. degree in electrical engineering in 2000 from the Johns Hopkins University. From 1983 to 1986 and 1989 to 1995, she worked as a Design Engineer at Beijing Institute of Radio Measurement in China. From 1995 to 1997, she worked at Glocom Inc, USA, where she was a Design Engineer and involved in several projects that were related to satellite communication terminals. Since 1997, she has been with Intelligent Automation, Inc (IAI), USA, where she is currently a Senior Electronic Engineer. Her research interests include signal processing, image processing, and fault diagnostics and applications. Over the last 8 years with IAI, she has worked on many different research projects in the above areas funded by various US government agencies such as DoD and NASA. She has also published over 10 journal and conference papers in the related areas.

R. Xu received the B.S. degree from Jiangsu University in 1982, and the M.S. degree in 1988 from Xi'an Tiaotong University, China, both in electrical engineering. From 1982 to 1985 and 1988 to 1993, he was teaching at Jiangsu University as an Assistant Professor. From 1993 to 1994, he was a Visiting Scholar at Lehrstuhl für Allgemeine und Theoretische Elektrotechnik, Universitaet Erlangen-Nuernberg, Germany. Since 1994, he has been with Intelligent Automation, Inc (IAI), USA, where he is currently a Principal Engineer. Mr. Xu's research interests include array signal processing, image processing, fault diagnostics, network security, and control theory and applications. Over the last 11 years with IAI, he has worked on many different research projects in the above areas funded by various US government agencies such as DoD and NASA. He has also published over 20 journal and conference papers in the related areas.

Y. Zhang received her B.S. and M.S. degrees in automatic control from Shanghai Jiaotong University in China in 1996 and 1999. She then obtained her M.S. and Ph.D. degrees in electrical engineering from Georgia Institute of Technology in 2001 and 2005, respectively. Since July 2005, she has been with Intelligent Automation, Inc (IAI), USA, where she is currently a Research Engineer. Her research interests include speech and image processing, intelligent systems, and controls.

D. Lao received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China, China, in 1988 and 1993, respectively. He received the Ph.D. degree from New Jersey Institute of Technology, Newark, USA, in 2003. Currently he is working on various digital signal processing projects at Intelligent Automation, Inc.

M. Stevenson received his Bachelor's degree from the University of Delaware in 1995. During that time, he was a member of a number of professional organizations, including the American Society of Mechanical Engineers (ASME) and the National Society for Professional Engineers (NSPE)—the latter of which he served as the President of the student chapter from 1993 to 1995. He joined Intelligent Automation, Inc. (IAI) in June 1995 and has been involved in a number of projects for various government agencies as well as for private-sector companies, including Ford Motor Company, Computer Sciences Corporation, and Raytheon. He currently holds the position of a Senior Mechanical Engineer at IAI.

V. Stanford works at the NIST Information Technology Laboratory and is currently leading the NIST Smart Space project in its Information Access Division. He received a B.A. degree in mathematics in 1974, and has spent thirty years in systems engineering, software, and applied statistical signal processing. His project developed the NIST smart data flow system, an abstract data transport mechanism for multimodal interface sensor data to capture terascale databases, and been used to integrate technologies for speaker, face, and speech recognition, using multiple sensors. He managed the development of Mk-III microphone array. He was a Lead Engineer for IBM Bethesda Spoken Language Systems Group, and developed the IBM recognition products, which won a Comdex Show Prize and other industry awards. He holds patents on speech recognition techniques including word spotting, signal processing, and spoken interfaces. He was an Invited Speaker at the opening ceremonies of the US Patent and Trademarks Office Patent database because these patents were widely referenced. He has worked on many complex systems engineering projects involving statistical pattern classification, and signal processing in the areas including electrocardiograms, phonocardiograms, acoustics, radar, sonar, and seismic signature analysis.

C. Rochet is currently a Research Engineer at the University of Karlsruhe on the Computers In the Human Interaction Loop (CHIL) project. He holds a Masters degree in computer science from the Ecole Internationale des Sciences du Traitement de l'Information in France. While at the National Institute for Standards & Technology (NIST), he developed the Mark III microphone array for the Smart Space Lab and Meeting Room projects. For this hardware sensor, he coded several clients including some integrated in a distributed processing for real-time beamforming and source localization.