Field Instrumentation for Vocalizing Avian Survey

by

Grant Elliott

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2007

Author .....
Department of Electrical Engineering and Computer Science
May 18, 2007

Certified by .....
Dale Joachim
Professor
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

Field Instrumentation for Vocalizing Avian Survey

by

Grant Elliott

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 2007, in partial fulfillment of the
requirements for the degree of
Master of Electrical Engineering and Computer Science

Abstract

We present automated instruments to facilitate the monitoring of vocalizing species in their environment with minimal disruption. These devices offer recording and acoustic localization of bird calls and relay data via the GSM cellular network to remote servers. Further, the system may be augmented with amplifiers and speakers in order to interact dynamically with the environment as to elicit vocalized responses, for example. The complete system is designed to meet the needs the Maine Audubon Society's annual owl survey. We further propose a less computationally intensive alternative to the standard time delay of arrival algorithm and extend it to arrays of arbitrarily many microphones.

Thesis Supervisor: Dale Joachim
Title: Professor

Acknowledgments

The author would like to thank Dr. Dale Joachim for supervising this work. Special thanks also go to Eben Goodale, who conducted field work, and Jaewoo Chung and Jongu Shin, who developed client and server software.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A common task in field biology is the recording of animals in their natural environment. Audio recording is of particular use when monitoring concealed or nocturnal creatures, such as owls on which we focus, which nonetheless vocalize distinctly. Indeed, once a call is heard, it is often possible to locate the otherwise difficult to find creature.

We seek to augment traditional audio recording techniques by automating and consolidating the data collection process and supplementing it with atmospheric data. We develop autonomous nodes, to be deployed throughout a large area, such as the state of Maine, to provide audio playback and recording capabilities over cellular networks. This data is consolidated on a single display, allowing field biologists to monitor and interact with a large environment. Further, every audio recording is tagged with an estimated localization of its source, facilitating tracking of individual owls. Finally, by also presenting the biologist end user with atmospheric data, we enable the possibility of associating the behaviors of owl populations with current weather conditions.

In addition to eliminating the need for volunteers, this network offers the following improved functionality:

- Centralization: All data is accessible in real time from a centralized server. As a result, field biologists located remotely may instantaneously monitor and

interact with creatures across the entire instrumented area. In addition to easing temporal and spatial constraints, such a system facilitates cooperation between researchers.

- Localization: Unlike traditional recording equipment, the listening stations are equipped with multiple microphones and capable of establishing a direction to the source of each vocalization. This functionality may be extended to monitor the movement behaviors of vocalizing creatures and even into full tracking.

- Atmospheric Monitoring: The new listening stations will also report pressure, temperature, humidity, and light levels in real time. Combined with vocalization and acoustic tracking information, this atmospheric data enables research into the effects of weather conditions on the behavior of reclusive creatures.

Ultimately, our instrumentation will permit future studies of the hearing range of owls, the response rate of owls due to current weather or human presence, and comparison between trigger-based and naturally occurring responses in surveys. Further, the distributed data collection provided by the wide-spread deployment of these instruments will foster collaboration in analyzing and annotating the resultant database.

Chapter 2

Motivation

The study of population dynamics in owls is rendered difficult by the reclusiveness of owls, being solitary nocturnal creatures. To combat this problem, owls are typically located acoustically. Fortuitously, it is fairly simple to entice an owl to vocalize; most will respond to calls of others of their species or to calls of predator species; this technique has been commonly used in field biology for many years and techniques have been refined[18]. It further forms the basis of the Maine Audubon Society's annual census.

2.1 Maine Owl Monitoring Program

The Maine Audubon Society conducts an annual census of owls throughout the state[15]. As part of the Maine Owl Monitoring Program (MOMP), volunteers are deployed nightly through the winter months to drive predefined routes. Each volunteer carries audio recording and playback equipment. The protocol calls for the playback at each stop of three owl calls (Northern Saw-Whet Owl, Barred Owl, and Great Horned Owl) separated by periods of silent recording. The resulting recordings are subsequently analyzed by expert ornithologists to yield estimates of the number of owls of various species within listening range of each stop. A graphic depiction of such data is presented in Figure 2-1

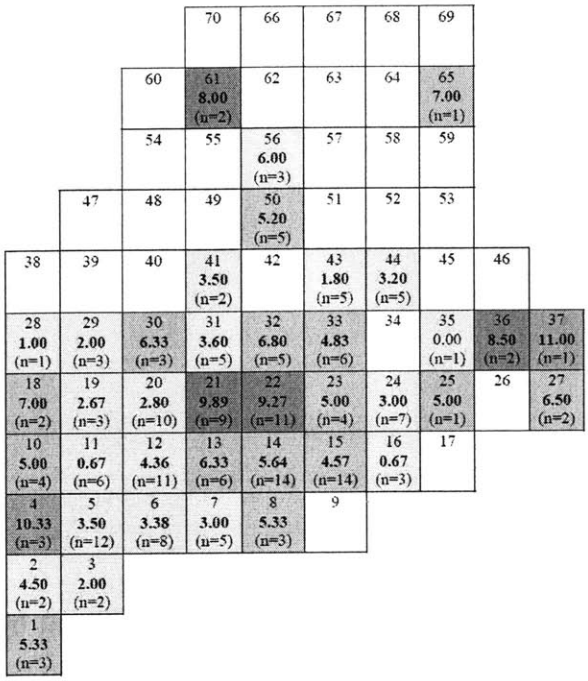


Figure 2-1: Census data for 2004 to 2006 showing average number of owls per route and number of routes, n, in each block of Maine. White blocks indicate an absence of data. Graphic taken from [15].

2.2 Cellular Phone Studies

Before using cellular technology for surveying owls, it is first necessary to establish that owls of interest respond to transmissions over the cellular network. While the frequency range of cellular telephones, with a cutoff frequency of about 4kHz, precludes use with high frequency vocalizing owls, it is sufficient for many species, including the Barred Owl and the Eastern Screech Owl. Unfortunately, cellular phone compression introduces audio artifacts, unnoticed by human perception but potentially devastating to sound integrity as perceived by an owl. Fortunately, the species of interest do not appear to perceive these artifacts, as demonstrated in Connecticut in the winter of 2006[10].

The standard playback protocol was followed, with playback and recording taking place every 0.8km along a prescribed route. Technology alternated between conventional playback and cellular phone playback, automated by a web-controlled voice-over-IP interface. The frequency of owl response to playback was comparable for both

technologies.

Based on the Connecticut experiments, we conclude that use of cellular technology in such field biology instruments is feasible.

2.3 Audio Processing

The use of acoustic localization is not new to ornithology, though most previous work has been conducted using systematic coverage and large microphone distances[17]. In such studies, localization is typically performed using the cross-correlation method (see Section 3.1.1) run on centralized computers. We introduce an approximation to this system (see Section 3.2) which is fast enough to run on an embedded microcontroller and which will produce viable localization results using tightly clustered microphones.

Current work at UCLA's Center for Embedded Networked Sensing follows a parallel path to this research, developing microphone arrays with extensive onboard processing to perform species identification and three dimensional localization. Like the system described here, the CENS system is intended to measure biodiversity with a particular interest in species which are difficult to locate visually (art birds and wrens)[22].

Though not yet integrated into this work, autonomous identification is a future goal of our research. Work in this field has progressed swiftly, as continuous field audio recordings have been popular for many years. Several computationally intensive solutions have been developed[13].

Chapter 3

Acoustic Localization

As our instrument is intended to localize sound sources in space, we examine several schemes for such localization, first examining theory and existing implementations and then proposing a far less computationally intensive approximation.

3.1 Theory

Humans and other animals identify the direction to a sound source by combining the auditory cues of interaural time difference (ITD) and interaural intensity difference (IID, commonly cited as interaural level difference or ILD)[20]. In practice these are not taken in isolation but are rather combined with visual cues. Interaural time difference refers to the time difference of arrival (TDOA) of a sound between two ears owing to the finite speed of sound. Interaural intensity difference refers to the varied energy level of sounds heard in both ears due to inverse-square falloff, presence of occluders, and the directionality of the ear. We begin by considering both methods as they apply to a two microphone (or ear) system.

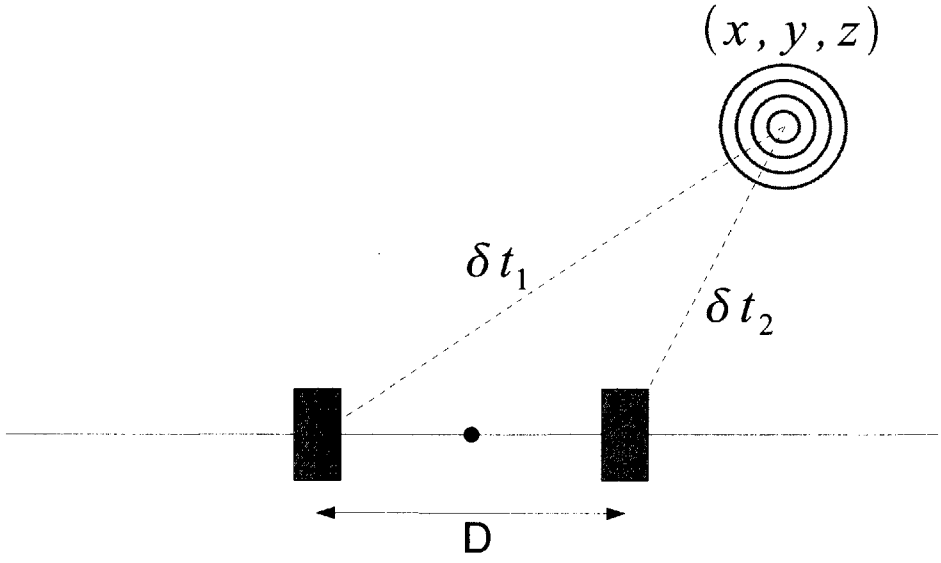


Figure 3-1: A sound source is recorded by two microphones, producing a time delay of arrival which may be used to constrain the position of the source.

3.1.1 Interaural Time Difference

Using Time Delay of Arrival

For sound sources closer to one microphone than the other (that is, those not lying on the plane that bifurcates the device) a time delay exists as sounds reach one before reaching the other. Consider the microphones and sound source depicted in Figure 3-1. Since the speed of sound, v, is constant, we have that the time delay, T, between arrivals at the microphones is given by

$$T = \frac{\sqrt{\left(x + \frac{D}{2}\right)^2 + y^2 + z^2}}{v} - \frac{\sqrt{\left(x - \frac{D}{2}\right)^2 + y^2 + z^2}}{v} \tag{3.1}$$

In the far field limit, we may treat the source as a plane wave. The difference between travel lengths is then simply the distance between microphones alone the direction to the source. Defining $\phi = 0$ to be the microphone axis we find

$$T \approx \frac{D}{v} \cos\phi \tag{3.2}$$

Alternatively, we may express this in vector notation where $\vec{l_{ij}}$ points from microphone i to microphone j and unit vector $\vec{s}$ points to the source.

$$vT_{ij} = \vec{s} \cdot \vec{l_{ij}} \tag{3.3}$$

Thus, if we can measure the time delay between signals at microphones distance D apart, we may constrain a far field sound source as lying on the cone defined by ϕ.

In the general case of a large number of microphones, a system of equations of the form of Equation 3.3 must be solved to yield the two degrees of freedom, ϕ and θ that represent the direction to the source. Least squares analysis on these two degrees of freedom is difficult in general and does not yield a linear system. A least squares analysis on three degrees of freedom (the position of the source in space, with a large error on distance to source) does yield a linear system, following substantial offline calculation dependent on geometry[11]. More often, an online search is performed.

Calculating Time Delay of Arrival

Unfortunately, calculating time delay of arrival if more difficult than it first appears. The most obvious technique is feature selection – measuring the time delay between two energy peaks, for example. This method is inherently limited by sampling rate, however. For microphones separated by ten centimeters, the largest observed delays will be only $300\mu s$, corresponding to only 13 samples at 44kHz. At slower sampling rates, as would be desirable when processing power is limited in an embedded system, the resolving capability of a feature selection method becomes impractically poor. Further, feature selection algorithms must be designed for the specific signals of interest and therefore lack generality.

A more robust technique evolves from a least squares approach. Suppose we attempt to shift one signal by the amount which minimizes the total error, defined as an integral over squares of differences. For continuous time domain signals, this

yields

$$\min_{\tau} \quad \int_{-\infty}^{\infty} (s_1(t) - s_2(t-\tau))^2 \, dt \tag{3.4}$$

$$\min_{\tau} \quad \int_{-\infty}^{\infty} s_1(t)^2 \, dt - 2\int_{-\infty}^{\infty} s_1(t)s_2(t-\tau) \, dt + \int_{-\infty}^{\infty} s_2(t-\tau)^2 \, dt \tag{3.5}$$

$$\max_{\tau} \quad \int_{-\infty}^{\infty} s_1(t)s_2(t-\tau) \, dt \tag{3.6}$$

Typically, this cross-correlation is computed equivalently in Fourier space

$$\max_{\tau} \int_{-\infty}^{\infty} X_1(\omega)\bar{X}_2(\omega)e^{j\omega\tau} \, d\omega \tag{3.7}$$

where X_1 and X_2 are Fourier representations of the signals received in the two microphones. In some cases, a frequency weighting function is also introduced[12] in order to favor sounds of interest or suppress echoes. As this method incorporates information from all frequencies, it is possible to extract time delays at a resolution higher than the sampling rate. It is also notably more robust than the feature selection technique, which is inherently restricted to those sounds which produce appropriate features.

Following this formulation of time delay of arrival, we may extract localization from multiple microphones using the more robust method of spatial likelihood functions[1]. In effect, each coordinate in space is assigned a probability of generating the signals observed by an array and these likelihood functions simply multiply. The spatial likelihood function is given by

$$\sum_k \sum_l \int_{-\infty}^{\infty} \frac{X_1(\omega)\bar{X}_2(\omega)e^{j\omega\tau_{kl}(\vec{x})}}{|X_1(\omega)| \, |X_1(\omega)|} \, d\omega \tag{3.8}$$

where $\tau_{kl}(\vec{x})$ is the time difference associated with with the pair of microphones k and l for a theoretical sound at position $\vec{x}$ and is referred to as the array steering delay.

This technique is particularly powerful because it can be augmented with spatial observability functions, which account for occlusion effects, so that these do not improperly impact a location's likelihood score. Unfortunately, because the calculation

requires full knowledge of the recorded waveforms, it may only be implemented if fast communication to a central processing system is available. This method also requires considerable computational and storage resources. Further, a complete analysis of the spatial likelihood requires application of techniques originally developed for machine vision.

Ultimately, interaural time difference is most effective for sounds with wavelength larger than microphone spacing. High frequency waves will encounter several full periods between the microphones, making phase difference measurements misleading. For a human, the usable range corresponds to sounds below about 10kHz, encompassing most of the human auditory range. Indeed, it has been suggested that interaural time difference is the dominant mechanism for localization in humans[23]. For a listening device four inches on edge, this corresponds to sounds below approximately 8kHz, a reasonable range for observing the target owls.

Our instrument uses a proposed less computationally intensive variant of the time delay of arrival algorithm, presented in Section 3.2.

3.1.2 Interaural Intensity Difference

In addition to time delay, acoustic signals degrade in intensity with distance according to an inverse square law. Thus, we ideally have

$$\frac{I_1}{I_2} = \frac{r_2^2}{r_1^2} \tag{3.9}$$

In the far field this ratio approaches unity and the technique appears to be rendered useless. However, for strongly directional microphones or microphones separated by an acoustic occluder (both being true of ears), a considerable intensity difference will exist owing to the asymmetry of the receiver. Unfortunately, this effect is notably more difficult to characterize, as it depends on the angular response of the microphone and the nature of the occluder.

Additionally, interaural intensity differences are useful primarily for higher frequency sounds, as wavelengths larger than the occluder are significantly less affected.

It is worth noting that most artificial methods of producing localized sound, such as stereo and surround speaker systems, function by producing intensity disparities. These systems lack directional bass because low frequency cannot be localized using this method.

As a result of these considerations, interaural intensity difference is less frequently used for sound localization than is interaural time difference. Nonetheless, it has been used successfully in systems for which online calibration or learning are possible[8].

3.1.3 Elevation

As described for a two microphone system, interaural time and intensity differences only constrain a sound source to within a cone in space. Extension to full spacial direction finding (that is, constraining the source to lie on a line) is relatively straightforward, as one may simply add additional microphone pairs. Four microphones arranged in a tetrahedron (or any non-coplanar configuration) is the minimal case[8]. From an engineering perspective, the incremental cost (in terms of monetary cost, development time, and system complexity) of additional microphones is very small, making this solution attractive.

However, nature has almost exclusively produced two-eared solutions, so it is worth discussing the biological mechanism for measuring elevation. Two methods dominate the resolution of this ambiguity. The first is motion; by repositioning the pair of ears in a known way one may further constrain the direction of the source. However, full three space localization occurs even if the head is stationary. This process is dominated by the frequency "coloring" caused by the pinna of the outer ear. This weighting of the sounds' Fourier representation is elevation dependent, so elevation can be determined if assumptions are first made about the uncolored Fourier transform. While effective in natural environments, which typically exhibit a consistent frequency distribution, this mechanism can fail with artificial sounds as simple as pure tones. This mechanism also requires significant online learning[20][19].

3.1.4 Advanced Techniques

Several more complex combinations of these techniques have also developed in nature. Of particular note is the fly *Ormia ochracea*, which must precisely localize the 5kHz pitch generated by crickets in order to lay eggs. With ears only 0.5mm apart, no interaural intensity delays can be measured and interaural time delays are on the order of 1.5μs. Nonetheless, experiments show the fly can localize cricket chirps to within 2^o azimuth.

In fact, the fly's eardrums are coupled. First, a mechanical resonance increases effective interaural time differences up to a maximal 55μs. Moreover, the tympanal mechanics are such that sources positioned strongly to either side produce considerably larger vibrations. Thus very small interaural time delays are converted into measurable interaural intensity differences[14].

This system introduces the possibility of mechanical computation in the calculation of time delay of arrival and cannot be emulated exclusively in software. It does, however, inspire a possible nanoscale directional microphone design to be incorporated into future work of this nature.

3.2 Proposed Algorithm

We now propose a computationally simple approximation to the traditional cross-correlation method of determining time delay of arrival and subsequently extend it to the case of more than two microphones. We also demonstrate its applicability to localizing sources in space using the tetrahedral configuration of the instrument designed here.

3.2.1 Measuring Time Delay of Arrival

We demonstrate the time delay of arrival algorithm on the three signals shown in Figure 3-2. These simulated signals represent time delayed sounds in the presence of broadband noise. We use the three microphone case for simplicity of representation

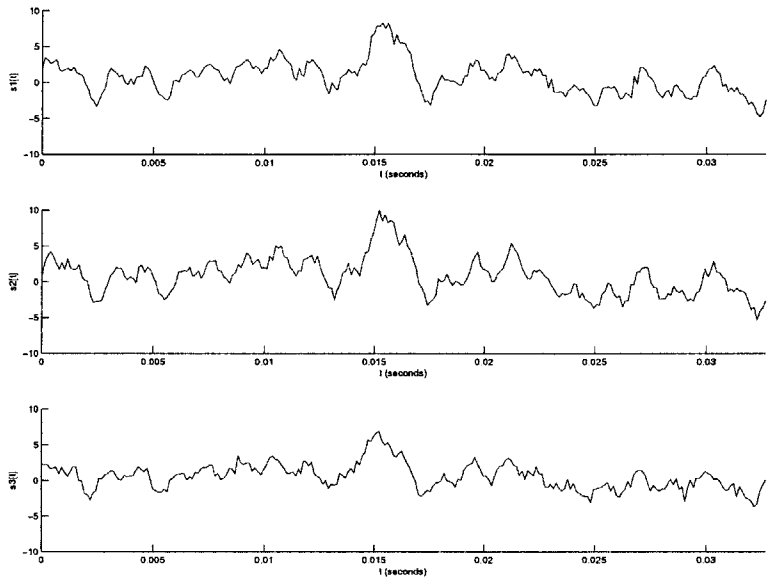


Figure 3-2: Simulated microphone signals used to demonstrate the fast TDOA algorithm. Signal 1 leads signal 2 by 0.6 samples and trails signal 3 by 1.3 samples.

ij	T	τ_0	k	$\tau_0 \pm \sigma$	$\tau_0 + \Delta \pm \sigma$
12	0.60	0.54	29.4	0.54 ± 0.08	0.63 ± 0.08
13	-1.30	-1.27	18.4	-1.27 ± 0.10	-1.41 ± 0.10
23	-1.90	-2.20	15.9	-2.20 ± 0.11	-2.04 ± 0.11

Table 3.1: Actual (T), best fit (τ_0), and constrained ($\tau_0 + \Delta$) time delays of arrival for the three sample signals given in Figure 3-2. Notice that the Δ correction forces the calculated TDOAs to obey the physical constraint imposed by Equation 3.18.

only; analysis of four or more microphones proceeds similarly and equations for the general case are presented.

Two Microphones

In order to reduce onboard processing requirements, it is desirable to eliminate the need for the intensive tasks of Fourier transform and search in the time delay of arrival algorithm presented in Equation 3.7. The former is easily removed if frequency weighting is not required, as the cross-correlation may be evaluated in the time domain. Define $x_i[t]$ to represent the t^{th} sample of the i^{th} microphone. The calculated

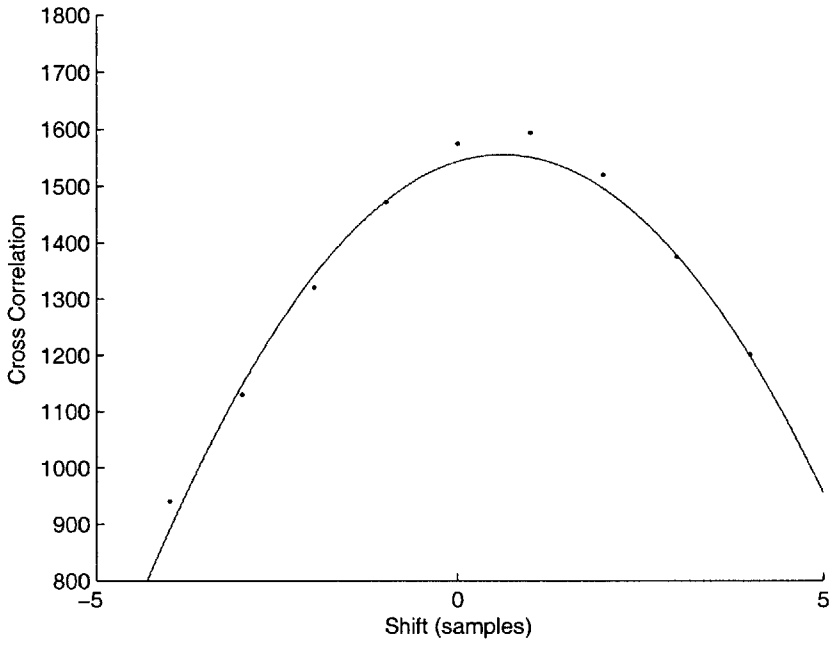


Figure 3-3: A parabola fit suffices to calculate the most likely TDOA between two microphones using the cross-correlation in the time domain. Here, the true TDOA between the simulated signals is 0.60 samples and the fit calculates a most likely TDOA of 0.54 ± 0.08 samples

TDOA between microphones i and j is then simply the value of τ_{ij} which maximizes

$$\chi_{ij}(t_{ij}) = \sum x_i[t] x_j[t - \tau_{ij}] \tag{3.10}$$

Two problems now arise. First, a search for the optimal value of τ is still necessary and, even more significant, we are now restricted to choosing τ as an integer representing a number of samples. In a small device sampling at a relatively low frequency, the range of possible TDOAs is small. For instance, the instrument developed here measures four inches between microphones. The maximum TDOA is about $300\mu s$ or only 2.3 samples at 7.813kHz. As such, subsample resolution is necessary. We alleviate both concerns by using a parabolic approximation to the cross-correlation as follows.

For small delays, the cross-correlation is well approximated by the autocorrelation shifted by the time delay of arrival. Further, the central peak of the autocorrelation

is even and may consequently be Taylor expanded as a parabola. We estimate the parameters of this parabola (notably the location of the central peak) by evaluating the cross-correlation at several values of τ symmetrically distributed about 0 and extending past the largest possible delay (based on the geometry of the microphones) and fitting using a least squares approach. A closed form for this polynomial fit exists and is given by

$$\begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \Sigma 1 & \Sigma\tau_i & \Sigma\tau_i^2 \\ \Sigma\tau_i & \Sigma\tau_i^2 & \Sigma\tau_i^3 \\ \Sigma\tau_i^2 & \Sigma\tau_i^3 & \Sigma\tau_i^4 \end{pmatrix}^{-1} \begin{pmatrix} \Sigma C_i \\ \Sigma\tau_i C_i \\ \Sigma\tau_i^2 C_i \end{pmatrix} \tag{3.11}$$

where the resultant parabolic dependence of the cross-correlation χ is described by

$$\chi(\tau) = c_2\tau^2 + c_1\tau + c_0 \tag{3.12}$$

$$= \chi_{max} - k\left(\tau - \tau_0\right)^2 \tag{3.13}$$

so that the vertex and width parameter are given by

$$\tau_0 = -\frac{c_1}{2c_2} \tag{3.14}$$

$$k = c_2 \tag{3.15}$$

τ_0 is thus the most likely TDOA. It may further be shown that the error margin on this calculation is given by

$$\sigma = \frac{1}{\sqrt{5k}} \tag{3.16}$$

Figure 3-3 shows data points calculated in this way for simulated signals 1 and 2, separated by 0.6 samples (and in the presence of broadband noise). The parabolic fit calculates $t_{012} = 0.54$ and $k_{12} = 29.4$. More conventionally, it computes a TDOA of 0.54 ± 0.08 samples, roughly one standard deviation from the true result. Table 3.1 summarizes the results for all three pairs of signals (for four microphones, 6 pairs would exist).

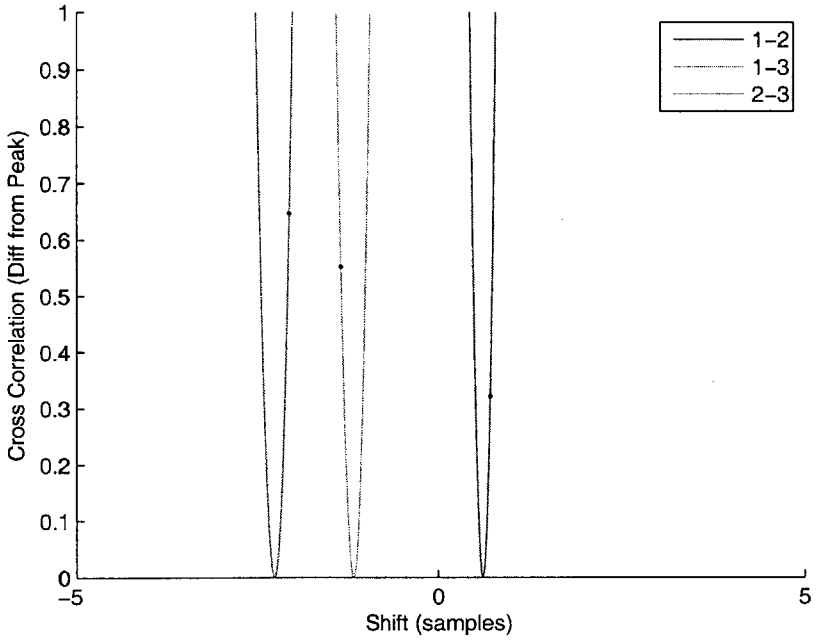


Figure 3-4: Using the best fit method in Section 3.2.1 to adjust τ_{0ij} from the parabola vertices to the points indicated by blue dots so as to meet the constraint given in Equation 3.18. The total error over all microphone pairs decreases as a result.

Many Microphones

For n microphones, the number of pairs (and therefore the number of calculable τ_0) is given by

$$\binom{n}{2} = \frac{n(n-1)}{2} \tag{3.17}$$

but only $n-1$ degrees of freedom are actually present (as is obvious if one considers the basis of time delays between a particular microphone and all others). Being differences between unmeasurable absolute travel times, the true TDOAs, notated T, must satisfy

$$T_{ij} + T_{jk} = T_{ik} \tag{3.18}$$

which allows us to define TDOAs for any pair in terms of a basis of $n-1$ independent pairs.

Notice in Table 3.1 that the τ_0 calculated using the parabola fit method do not

meet this constraint. We could choose only to calculate the TDOA for a set of $n-1$ independent pairs, but this introduces an asymmetry and does not fully utilize the available information. Instead we demonstrate a method for combining the redundant TDOA calculations to yield a consistent set.

We seek to adjust each of our calculated τ_{0ij} by some amount Δ_{ij} so that the constraints given by equation 3.18 are satisfied:

$$\tau_{0ij} + \Delta_{ij} + \tau_{0jk} + \Delta_{jk} = \tau_{0ik} + \Delta_{ik} \tag{3.19}$$

In particular, we choose the set of Δ_{ij} which satisfies these constraints with the minimum depreciation in the sum of the cross correlations, χ_{ij}. The total depreciation of all χ_{ij} is then given by

$$\sum k_{ij}\Delta_{ij}^2 \tag{3.20}$$

For notational clarity, we consider only τ_{0ij}, Δ_{ij}, and k_{ij} where $j > i$ since

$$\tau_{0ij} = -\tau_{0ji} \tag{3.21}$$

$$\tau_{0ii} = 0 \tag{3.22}$$

$$k_{ij} = k_{j,i} \tag{3.23}$$

We further consider only the $n-1$ choices of Δ of the form Δ_{in} since all other Δ_{ij} may be written using Equation 3.19

$$\Delta_{ij} = \left(\tau_{0in} - \tau_{0jn} - \tau_{0ij}\right) + \Delta_{in} - \Delta_{jn} \tag{3.24}$$

where the parenthesized term is simply a sum of previously computed constants.

We now rewrite Equation 3.20 in terms of the $n-1$ parameters Δ_{in}

$$\sum_{i \neq n} k_{in}\Delta_{in}^2 + \sum_{i \neq n, j \neq n, i<j} k_{ij}\left(\left(\tau_{0in} - \tau_{0jn} - \tau_{0ij}\right) + \Delta_{in} - \Delta_{jn}\right)^2 \tag{3.25}$$

Since we seek to minimize Equation 3.25, we set partial derivatives with respect to each parameter to zero, yielding a set of $n-1$ linear equations, the i^{th} of which is

given by

$$\begin{aligned} k_{in}\Delta_{in} \quad &+ \quad \sum_{i<j<n} k_{ij}\left(\left(\tau_{0in}-\tau_{0jn}-\tau_{0ij}\right)+\Delta_{in}-\Delta_{jn}\right) \\ &- \quad \sum_{0<j<i} k_{ji}\left(\left(\tau_{0jn}-\tau_{0in}-\tau_{0ji}\right)+\Delta_{jn}-\Delta_{in}\right) \end{aligned} \tag{3.26}$$

Finally, we write these simultaneous linear equations in matrix form

$$\mathcal{M}\vec{\nu} = \vec{N} \tag{3.27}$$

where

$$\mathcal{M}_{ij} = \begin{cases} -k_{ji} & i > j \\ k_{in} + \sum_{0<m<i} k_{mi} + \sum_{i<m<n} k_{im} & i = j \\ -k_{ij} & i < j \end{cases} \tag{3.28}$$

$$= \begin{cases} \sum_{m\neq i} k_{im} & i = j \\ -k_{ij} & i \neq j \end{cases} \tag{3.29}$$

$$\nu_i = \Delta_{in} \tag{3.30}$$

$$\begin{aligned} N_i &= \sum_{i<j<n} k_{ij}\left(\tau_{0in}-\tau_{0jn}-\tau_{0ij}\right) \\ &\quad - \sum_{0<j<i} k_{ji}\left(\tau_{0jn}-\tau_{0in}-\tau_{0ji}\right) \end{aligned} \tag{3.31}$$

$$= \sum_{j\neq n} k_{ij}\left(\tau_{0in}-\tau_{0jn}-\tau_{0ij}\right) \tag{3.32}$$

in which we have relaxed the notation convention of considering only τ_{0ij} and k_{ij} with $i < j$ and used the relationships given in Equation 3.21. Solving the matrix equation may now be done offline or online.

The three microphone case is particularly simple. The equation is

$$\begin{pmatrix} k_{12}+k_{13} & k_{12} \\ k_{12} & k_{12}+k_{23} \end{pmatrix} \begin{pmatrix} \Delta_{13} \\ \Delta_{23} \end{pmatrix} = \begin{pmatrix} k_{12}\left(\tau_{013}-\tau_{023}-\tau_{012}\right) \\ -k_{12}\left(\tau_{013}-\tau_{023}-\tau_{012}\right) \end{pmatrix} \tag{3.33}$$

with solution

$$\Delta_{12} = \frac{k_{13}k_{23}}{k_{12}k_{13}+k_{12}k_{23}+k_{13}k_{23}}(\tau_{013}-\tau_{012}-\tau_{023}) \quad (3.34)$$

$$\Delta_{13} = \frac{k_{12}k_{23}}{k_{12}k_{13}+k_{12}k_{23}+k_{13}k_{23}}(\tau_{012}-\tau_{013}+\tau_{023}) \quad (3.35)$$

The results of applying these corrections to the simulated sounds are also presented in Table 3.1. Notice that the results now meet the difference constraint and the total error over all microphone pairs has been decreased.

For the instrument designed here we use the four microphone case, the solution of which is notably more complicated. Just as in the three microphone case, the linear system defined here yields three time delay measurements, representing delays between each of three microphones and a reference microphone.

3.2.2 Source Localization

We now present a simple method for calculating source location in the special case of microphones located at the vertices of a tetrahedron, as they are in our instrument. Following the far field time delay formulation in Equation 3.3, we derive equations for each of the three TDOA values produced by the algorithm presented in Section 3.2.1

$$\frac{v}{a}\tau_{41} = \vec{s}\cdot\left(0,\frac{1}{\sqrt{3}},-\frac{\sqrt{2}}{\sqrt{3}}\right) \quad (3.36)$$

$$\frac{v}{a}\tau_{42} = \vec{s}\cdot\left(\frac{1}{2},-\frac{1}{2\sqrt{3}},-\frac{\sqrt{2}}{\sqrt{3}}\right) \quad (3.37)$$

$$\frac{v}{a}\tau_{43} = \vec{s}\cdot\left(-\frac{1}{2},-\frac{1}{2\sqrt{3}},-\frac{\sqrt{2}}{\sqrt{3}}\right) \quad (3.38)$$

in which a is the edge length of the tetrahedron, $\vec{s}$ is a unit vector pointing to the source, and the enumerated vectors are derived from tetrahedron geometry. The microphone pair constraint of Section 3.2.1 effectively integrates information from the remaining three time delays into these three, ensuring that they are consistent with the definition of time delay of arrival, though geometric inconsistencies may

remain. With only these three equations, we may treat the system as linear and solve for $\vec{s}$, ignoring for the moment that $\vec{s}$ must be a unit vector.

$$\vec{s} = \frac{v}{a} \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \\ -\frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}^{-1} \begin{pmatrix} \tau_{41} \\ \tau_{42} \\ \tau_{43} \end{pmatrix} \tag{3.39}$$

If the calculated time delays are, in fact, consistent with the geometry of microphone placement, this system will yield a unit source vector. If $\vec{s}$ is does not have unit magnitude, we may nonetheless assume that it points in the proper direction and take its magnitude as an indicator of error. In fact, this approximation is used by general least squares solutions which minimize over three degrees of freedom, similarly assuming the result has unit magnitude[11].

Chapter 4

Instrumentation

The instrument consists of three separate devices: an audio and data collection and processing instrument referred to as the Hanger (as its preferred deployment is suspended beneath a tree), an audio amplifier for playback of prerecorded sounds, and a standard cellular telephone which relays audio and data to a base station.

4.1 The Hanger

The Hanger, shown in Figure 4-1 forms the core of the instrument, interacting with the cellular phone while collecting and processing audio and atmospheric data. In particular, the Hanger records from four microphones and uses the four signals both to calculate the relative position of the source and to synthesize a combined omnidirectional signal to be transmitted via the cellular network. In addition, pressure, temperature, humidity, and light levels are monitored so that they may later be correlated with owl vocalizations and behavior. We provide an overview of the major systems implemented in the Hanger, including central control, data storage, audio capture, atmospheric sensing, communications, and power management.

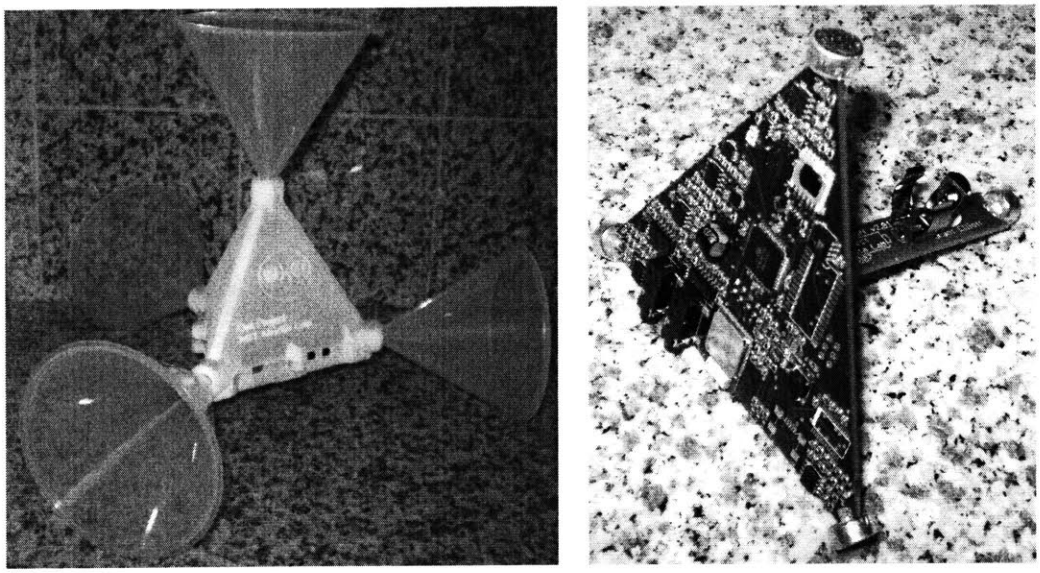


Figure 4-1: Photograph of an assembled hanger in and out of its enclosure.

4.1.1 Control

The Hanger is controlled by an Atmel Atmega324P, an 8-bit RISC microcontroller. The firmware is responsible for providing an interface to the microphones, sensors, and communication channels while simultaneously implementing the time delay of arrival scheme outlined in Section 3.2. This choice of microcontroller allows us to avoid several external circuits by providing onboard UARTs, used in implementing the Bluetooth and USB connections (see Section 4.1.6), an SPI interface, used to communicate with the pressure and temperature sensors (see Section 4.1.5), EEPROM, and real time clock (see Section 4.1.2), ADCs, used for audio monitoring (see Section 4.1.3) as well as compass and light level readout, PWM generators, used for audio synthesis (see Section 4.1.4), and timers, used for humidity readout.

The onboard timers are also used to trigger interrupts which dictate the instrument's behavior. In particular, the microphone channels must be read at 7.8125 kHz (a rate chosen to correspond with the cellular phone's 4kHz limit and for the convenience of generating such a clock from the onboard 8MHz oscillator). This timing is critical for ensuring the success of the time delay of arrival algorithm. A timer interrupt reads the ADCs, updates the synthesizers, and may set any of sev-

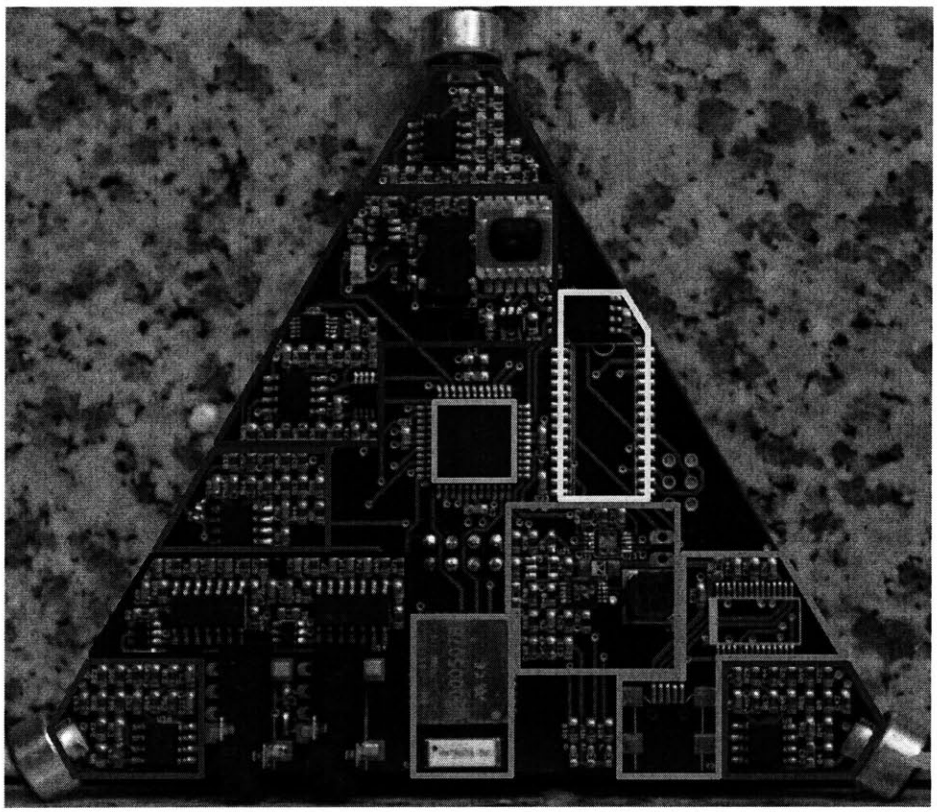


Figure 4-2: Photograph of the bottom of a Hanger, highlighting the major subsystems. Cyan: Control (Section 4.1.1), Yellow: Data Storage (Section 4.1.2), Red: Audio Capture (Section 4.1.3), Blue: Audio Synthesizers (Section 4.1.4), Purple: Atmospheric Sensors (Section 4.1.5), Green: Communications (Section 4.1.6), Orange: Power (Section 4.1.7).

eral flags indicating additional computations to take place during the period between interrupts. These flags include reading of sensors, processing audio buffers, and transmitting data. (In practice, the lack of reentrant interrupts makes this process notably more complicated than presented here.)

4.1.2 Data Storage

The Hanger features an optional EEPROM, interfaced over SPI, for the storage of collected data. This feature is of particular use if the Hanger is to be deployed in the absence of a cellular phone, to function as a data logger. Further a real time clock permits accurate timestamping of such recorded data.

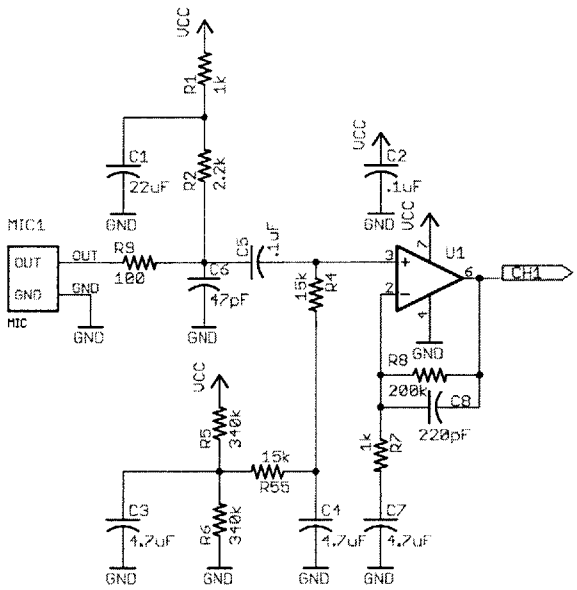


Figure 4-3: Schematic for one of the Hanger's four microphone amplifiers. The circuit consists of a non-inverting op amp, coupled with low and high pass filters. See Section 4.1.3

Additionally, the EEPROM may be used to store pre-recorded audio, which may be synthesized (see Section 4.1.4) and played back through the amplifier (see Section 4.2). In this way, an isolated Hanger may nonetheless implement the full Audubon protocol.

4.1.3 Audio Capture

Audio is captured by four electret microphones positioned at the vertices of a tetrahedron four inches on edge. Each microphone is biased and amplified using the circuit in Figure 4-3 based on an ST Microelectronics application note[5]. Microphone bias voltage is filtered by R_1 and C_1 while R_3, R_5, R_6, C_3, and C_4 form a π bridge to provide the half voltage offset rail for the operational amplifier. The gain of the amplifier is given by

$$G = \left(1 + \frac{R_8}{R_7}\right)\left(\frac{R_4}{R_4 + R_2}\right) = 175 \tag{4.1}$$

Signals from the microphone amplifiers are read by the microcontroller's onboard ADCs sampling at $7.8125kHz$, allowing capture of sounds up to $3.9kHz$ (the Nyquist

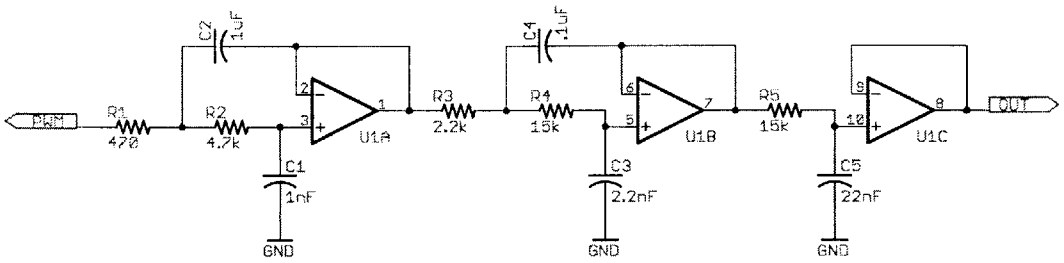


Figure 4-4: Schematic for one of the Hanger's two audio synthesizers, consisting of a pulse width modulator with the carrier removed by a fifth order Chebyshev filter. See Section 4.1.4

limit). As such, we should filter the input signal to reduce components at higher frequencies. Such a lowpass filter is formed by R_8 and C_8 while two highpass filters are formed by R_2, R_4, and C_5 and by R_7 and C_6. These have cutoff frequencies given by

$$f_{ch} = \frac{1}{2\pi R_8 C_8} = 3.6kHz \tag{4.2}$$

$$f_{cl_1} = \frac{1}{2\pi(R_2+R_4)C_5} = 93Hz \tag{4.3}$$

$$f_{cl_2} = \frac{1}{2\pi R_7 C_7} = 34Hz \tag{4.4}$$

R_9 and C_6 form an additional low pass filter to reduce EMI.

The microcontroller maintains audio buffers for 256 samples (33.2ms) of each microphone. When these buffers are full, they are processed by the TDOA algorithm, although the audio capture interrupt continues to occur so that synthesizers and state may be updated. Following the termination of an interesting sound (as defined by the firmware's triggering function), TDOA measurements for that sound are merged and constrained (see Section 3.2.1) and the resulting delays are transmitted via Bluetooth to the cellular phone, which in turn relays them to the central server.

4.1.4 Synthesizers

The audio synthesizers consist of pulse width modulation generators connected to low pass filters. The PWM generators are integrated into the microcontroller and operate

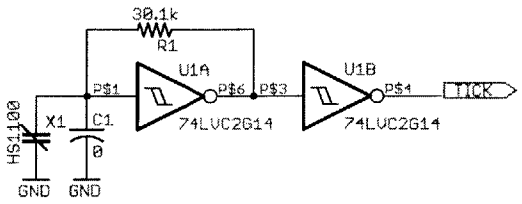


Figure 4-5: Schematic for the relaxation oscillator used to measure the capacitance of the humidity sensor. See Section 4.1.5.

with a carrier frequency of $15.625kHz$ (chosen so that 8-bit resolution is achievable with an 8MHz system clock). The relatively low ratio of carrier to signal frequency necessitates a high order low pass filter; we choose a fifth order Chebyshev filter[16] as shown in Figure 4-4. The first two stages are staggered second order filters while the last is a first order output RC network, with transfer functions

$$TF_1 = \frac{1}{s^2 R_1 R_2 C_1 C_2 + s\left(R_1 + R_2\right) C_2 + 1} \tag{4.5}$$

$$TF_2 = \frac{1}{s^2 R_3 R_4 C_3 C_4 + s\left(R_3 + R_4\right) C_4 + 1} \tag{4.6}$$

$$TF_3 = \frac{1}{1 + sR_5 C_5} \tag{4.7}$$

4.1.5 Atmospheric Sensors

Pressure, temperature, humidity, light, and compass sensors are read periodically (typically every 10 minutes) and reported to the cellular phone via Bluetooth. The cellular phone then immediately relays this information to a central server.

Pressure is measured by piezoresistance using Intersema's MS5534, interfaced using SPI. The achievable range is 10 to 1100mbar with a precision of ± 0.5mbar and absolute accuracy of ± 1.5mbar. Since this pressure measurement exhibits a second order dependence on temperature, the package also contains a temperature sensor on the range $-40°$C to $125°$C with a precision of $\pm 0.2°$C and absolute accuracy of $\pm 0.8°$C. Temperature compensation is performed on the microcontroller.

Humidity is measured using Humirel's HS1100, the capacitance of which varies with relative humidity according to a third order polynomial. Capacitance is mea-

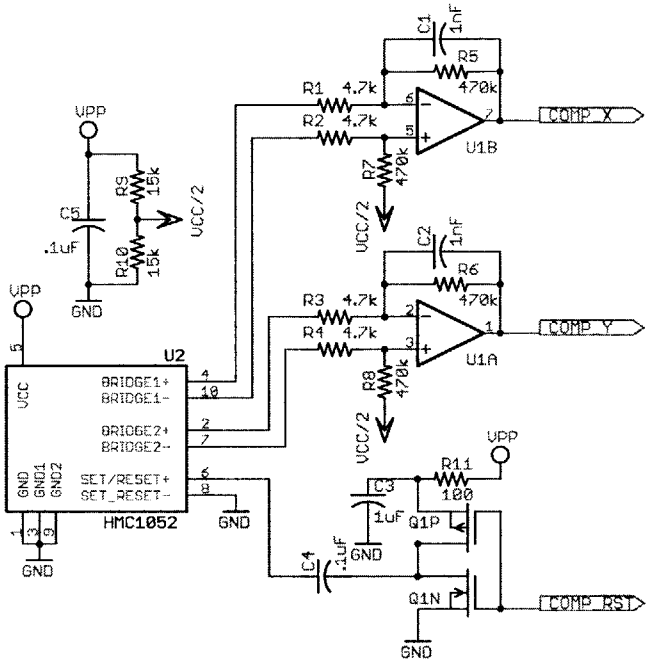


Figure 4-6: Schematic for the differential amplifier which reads the two axis magnetometer. Also shown is the reset strap circuit, which produces brief pulses of current to align magnetic domains. See section 4.1.5.

sured using the relaxation oscillator shown in Figure 4-5, which generates a square wave with frequency

$$f = \frac{2}{RC_{sense}} \tag{4.8}$$

This frequency is measured using one of the microcontroller's onboard counters, converted to a capacitance using the known resistance R, and finally converted to a relative humidity using the third order transfer function.

Light levels are measured simply by using a phototransistor and resistor, so that the voltage across the resistor grows linearly with the current in the phototransistor, and therefore approximately linearly with photon density. The phototransistor chosen is Microsemi's LX1972IBC, designed with a wide frequency response meant to mimic that of the human eye.

Finally, compassing is performed using Honeywell's HMC1052 dual magnetometer and a differential amplifier, as shown in Figure 4-6. The magnetic domains of the

sensor must be periodically reset by pulsing 600mA through the set/reset strap. The charge for this pulse is stored in C_5 when Q_1 is enabled and sunk through the 6Ω strap when Q_2 is enabled.

4.1.6 Communications

Two externally accessible communication systems are available. A Bluetooth interface provides a standardized interface to a nearby cell-phone while a USB slave interface offers a computer connection for reading out data following a data logging session and for debugging. Both interfaces are implemented using the microcontroller's onboard UARTs and commercial modules which implement their respective protocols. Bluetooth is implemented with Blue Radio's BRC46AR and USB is implemented with FTDI's FT232R.

It is worth noting that Bluetooth is an expensive and power consuming choice, particularly for interfacing to a directly connected cell phone. Unfortunately, no wired data standard exists within the cell-phone industry. (Though USB is fairly common, implementing a USB host and porting proprietary drivers is a solution deemed too complex.) The Bluetooth interface may be omitted if a cell phone is available with a serial data interface (such as Nokia's FBus) or USB On-the-Go (which allows two USB slaves to connect).

4.1.7 Power

The device is powered by a single primary lithium-ion cell with a capacity of 2.6 amp-hours and a nominal voltage of $3.6V$. A four switch buck-boost converter (Linear Semiconductor's LTC3440), operating at 2MHz, regulates the supply voltage to 3.3V over the complete voltage range of the battery. To conserve power, onboard systems not in use are disconnected. In particular, synthesizers and atmospheric sensors are powered down when unneeded and the Bluetooth interface operates at a minimum broadcast power.

An optional lithium-ion battery charger based on Linear Semiconductor's LTC4053

System	Max	Avg	Usage	Notes
Microphones	4×1.5mA	6mA	6mA	
Synthesizers	2×3mA	3mA	0mA	Disabled when unused
Pressure	1mA	1μA	0mA	Disabled when unused
Humidity	0.1mA	0.1mA	0mA	Disabled when unused
Light	2mA	2mA	0mA	Disabled when unused
Compass	4A	20mA	0mA	Disabled when unused
Microcontroller	4.5mA	4.5mA	4.5mA	
Bluetooth	39mA	1mA	1mA	Depends heavily on usage and mode

Table 4.1: Current draw of various Hanger systems. Maximum current draw may occur only in pathological cases. Average draw refers to conditions while operating while usage draws include typical usage rates.

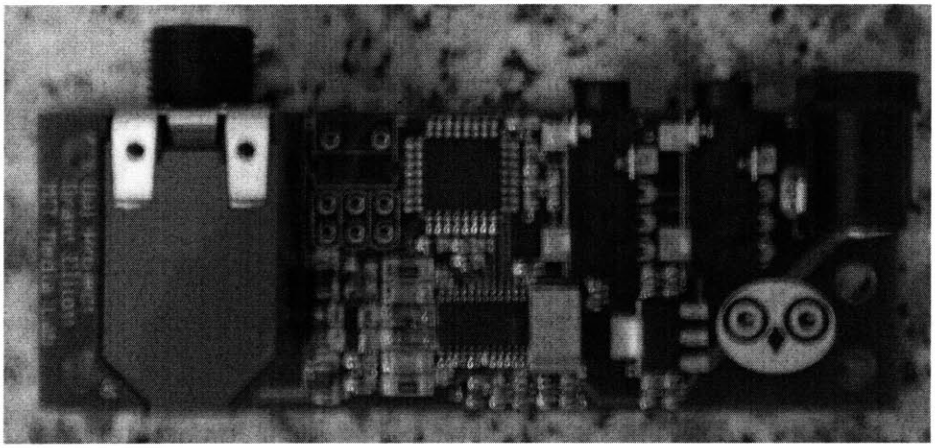


Figure 4-7: Photograph of the external speaker amplifier. See Section 4.2.

is also provided. The charging circuit is powered from the USB supply and features a disconnection jumper so as not to inject current into nonrechargable batteries (such as primary lithium cells).

Estimated current consumption of the Hanger's onboard systems are presented in Table 4.1. Note that load may change with time, as many systems are disabled when unused.

4.2 Amplifier

An external amplifier drives an 8Ω horn speaker, for broadcasting prerecorded sounds over large areas. Since this amplifier is used infrequently and at high power (up to

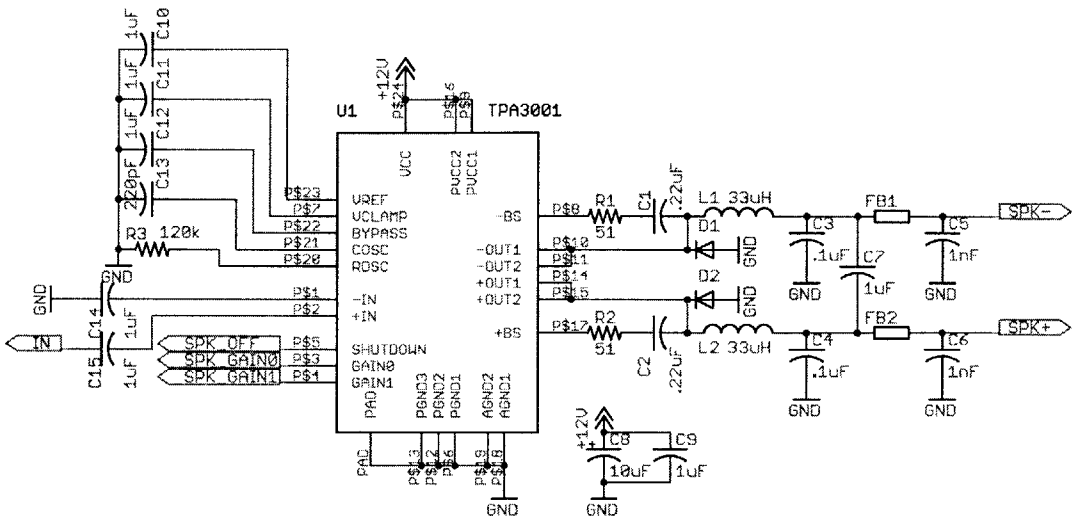


Figure 4-8: Schematic for the external class-D amplifier, based on Texas Instrument's TPA3001D. Note the filtering of the differential output stage. See Section 4.2.

20W), we choose a class-D amplifier. For simplicity, we build the amplifier around TI's TPA3001D, a monolithic amplifier on a chip. All that remains is to provide filtering for the differential output stage (a differential output being chosen due to the length of cable connecting the speaker to the amplifier). An Atmega48 microcontroller monitors the input signal and controls the amplifier's gain and muting. In particular, the speaker is muted when presented with signals presumed to be static noise. The amplifier is intended to be powered from an external battery, with voltage between 8V and 18V. The supply voltage determines the maximum allowed gain before saturation. In many cases, a 9V battery proves sufficient. The schematic is depicted in Figure 4-8.

Chapter 5

Software

Software to coordinate communication and interaction with remote Hangers was developed by students at the MIT Media Lab.

Communication between the Hanger, cellular phone, and central server is handled by a Java applet which acts as a Bluetooth master when exchanging data with the Hanger and as a client when exchanging data with the server.

The end user is presented with two web-based interfaces. The first is a scheduler, which coordinates voice-over-IP communication with the deployed cellular phones and allows the user to plan recording and playback events in advance. This software is an evolution of that used in the Connecticut study (see Section 2.2) . The second web interface is a mapping application which depicts audio and sensor data events in real time. Using the Google Maps API, this interface provides a graphical depiction of the atmospheric data over the instrumented region while plotting the locations of audio events as they occur.

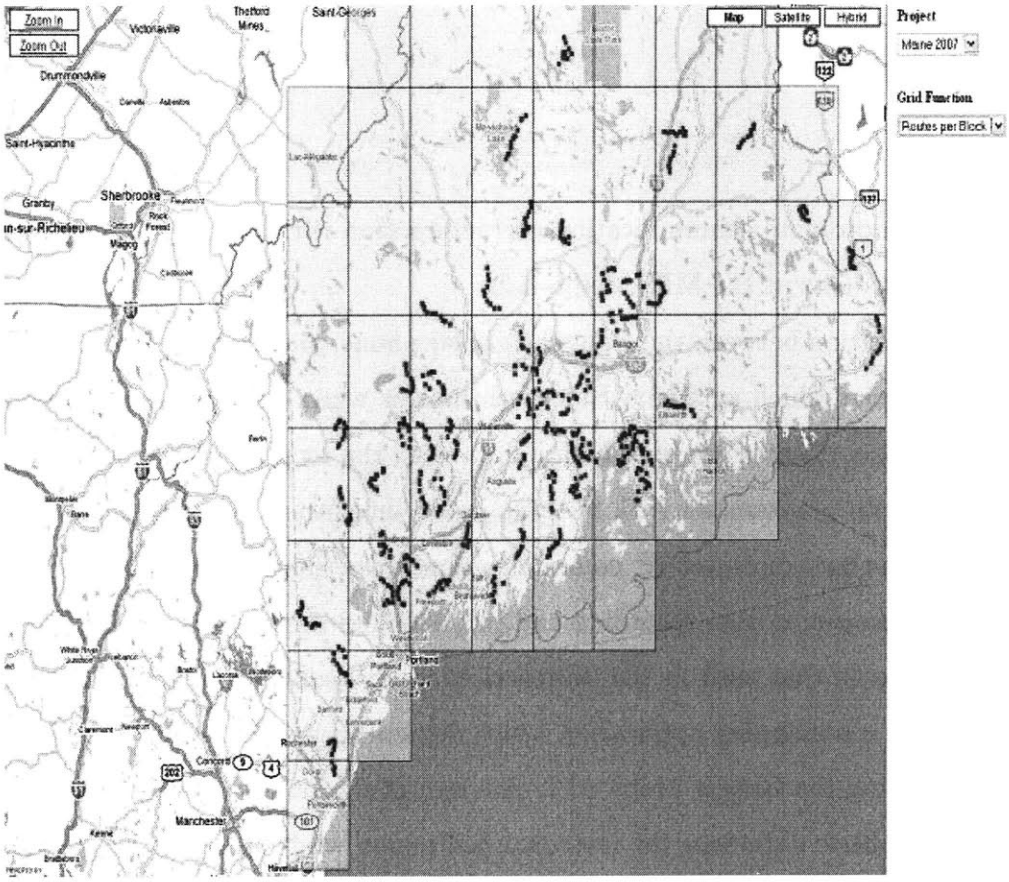


Figure 5-1: Screenshot of the mapping interface, showing owl events graphically.

Chapter 6

Conclusion

The instrument described here has been prototyped, tested, and put into low quantity production, with 73 Hangers and amplifiers constructed. In the immediate future, they will be tested by Eban Goodale, a biologist studying local owl populations, and later will be used in the 2007-2008 Maine Owl Monitoring Survey. Interest has also been expressed by groups at Tulane University, with whom we have associated in this work.

Several minor issues associated with the cellular phone have thus far arisen. First, the proprietary connector used on Nokia phones requires the production of custom cables and the location of suitable mating connectors. A more interesting problem arises in handling the cellular phone's internal feedback. When operating with a headset, the Nokia N80 feeds back a small portion of the microphone signal to the earpiece. In our case, the microphones can easily hear the high power horn speaker and a feedback loop is generated, destroying audio integrity. It should be noted that this problem varies between phone models. Temporarily, the amplifiers and Hangers are operated from different phones. In the long term, a firmware solution will be implemented so that the microphones will be disabled during broadcast and the speaker will be disabled all other times.

All atmospheric sensing and communication to the centralized server function as expected. The latest mapping software graphically depicts this data and demonstrates the power of this monitoring.

The proposed time delay of arrival algorithm has been tested informally and found to yielding consistent results when localizing broadband sounds. A more formal testing procedure should be conducted, to establish the capabilities of this system. Further, the conversion from TDOAs to location in space has not yet been implemented.

Ultimately, the device is a success, meeting the design criteria. The future months will reveal the benefits of such autonomous audio monitoring over large areas.

Bibliography

[1] P. Aarabi, *The fusion of distributed microphone arrays for sound localization*, EURASIP Journal on Applied Signal Processing Special Issue on Sensor Networks 4 pp. 338347, 2003.

[2] P. Aarabi, *Self localizing dynamic microphone arrays*, IEEE Transactions on Systems, Man, and Cybernetics, Part C 32 (4)pp. 474484, 2002.

[3] U. Bub, M. Hunke, and A. Waibel, *Knowing who to listen to in speech recognition: Visually guided beamforming*, Proc. Int. Conf. Acoustics, Speech, and Signal Processing, 1995, pp. 848-851.

[4] V. Cevher, J. McClellan, *Sensor array calibration via tracking with the extended Kalman filter*, Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on , vol.5, no.pp.2817-2820 vol.5, 2001.

[5] R. Cittadini and F. Poulin, *TS971 Bases Electret Condenser Microphone Amplifier*, `http://www.st.com/stonline/products/literature/an/8698.pdf`. June 17, 2002.

[6] B. Dalton and M. Bove. *Audio-based self-localization for ubiquitous sensor networks*, 118th Audio Engineering Society Convention, 2005.

[7] A. Galstyan, B. Krishnamachari, K. Lerman, and S. Pattem, *Distributed online localization in sensor networks using a moving target*, Proc. IEEE Information Processing in Sensor Networks (IPSN), April 2004, pp. 6170.

[8] K. Guentchev and J. Weng, *Learning-based three dimensional sound localization using a compact noncoplanar array of microphones*, Proc.1998 AAAI Symp. Intelligent Environments, 1998.

[9] X. Ji and H. Zha, *Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling*, Proc. IEEE INFOCOM, March 2004.

[10] D. Joachim, *Broadcasting and recording bird vocalizations through cellular telephone channels.*

[11] D. Joachim, *Notes on Microphone Array Fusion.* Unpublished notes.

[12] C. Knapp; G. Carter, *The generalized correlation method for estimation of time delay*, Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], vol.24, no.4, pp. 320- 327, Aug 1976.

[13] J. Kogan, D. Margoliash, *Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study.* Journal of the Acoustical Society of America (1998) 103:2185-2196.

[14] A. Mason, M. Oshinsky, and R. Hoy. *Hyperacute directional hearing in a microscale auditory system.* Nature, vol.410, no.6829, pp.686-690, Apr 2001.

[15] Maine Audubon, *Maine Owl Monitoring Program*, `http://www.maineaudubon.org/conserve/citsci/owl.shtml`.

[16] Maxim/Dallas Semiconductor, *Application Note 1795: Analog Filter Design Demystified.* `http://www.maxim-ic.com/appnotes.cfm/appnote_number/1795/`

[17] D. J. Mennill, J. M. Burt, K. M. Fristrup, and S. L. Vehrencamp. *Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest.* J. Acoustic Soc. America 119: 2832-2839.

[18] National Park Service, *Draft Manual for Conducting Acoustics and Soundscape Studies in National Parks.* August 20, 2005

[19] B. Razavi, W. O'Neill; G. Paige, *Both Interaural and Spectral Cues Impact Sound Localization in Azimuth*, Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on , vol., no.pp. 587- 590, March 16-19, 2005.

[20] R. Sekuler and R. Blake, *Perception.* New York, NY: McGraw-Hill, 2002. Chpt 12.

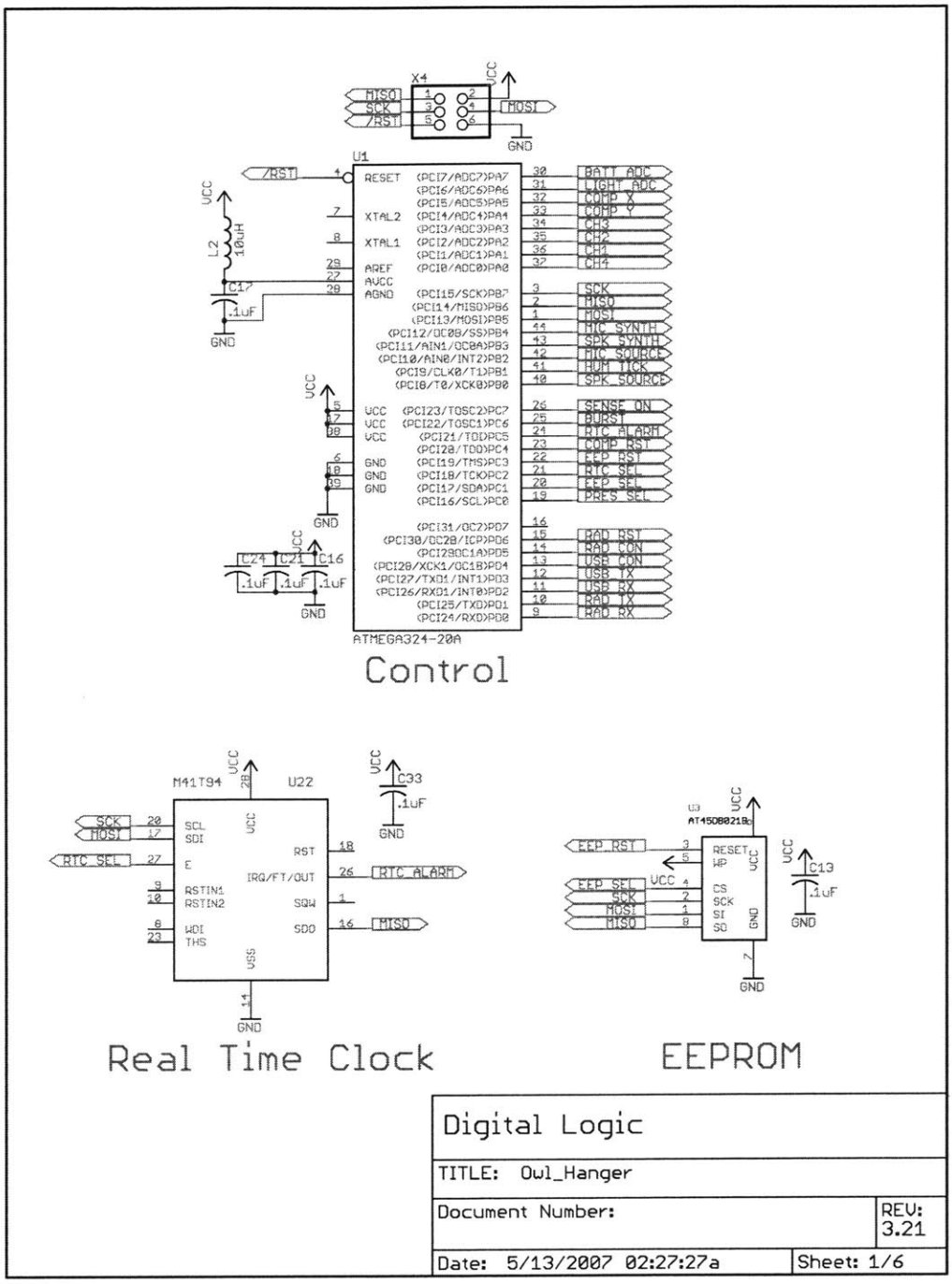
[21] C. Taylor, A. Rahimi, J. Bachrach, and H. Shrobe, *Simultaneous Localization and Tracking in an Ad Hoc Sensor Network*, IPSN 2005.

[22] C. Taylor, E. Stabler, *Adaptive Communication in Accoustic Sensor Arrays.* `http://research.cens.ucla.edu/portal/page?_pageid=56,56051,56_56052&_dad=portal&_schema=PORTAL`

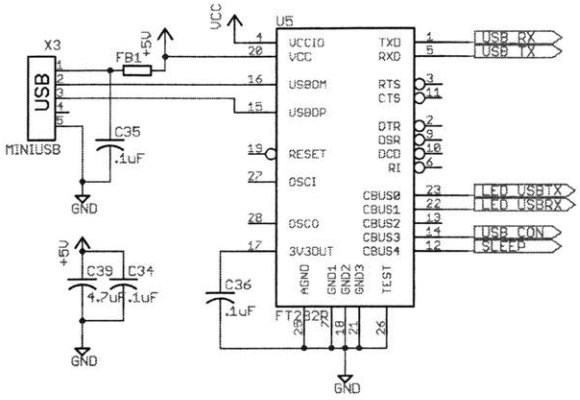
[23] F. Wightman and D. Kistler, *The dominant role of low-frequency interaural time differences in sound localization*, J. Acoust. Soc. Am. 91, pp. 16481660, 1992.

Appendix A

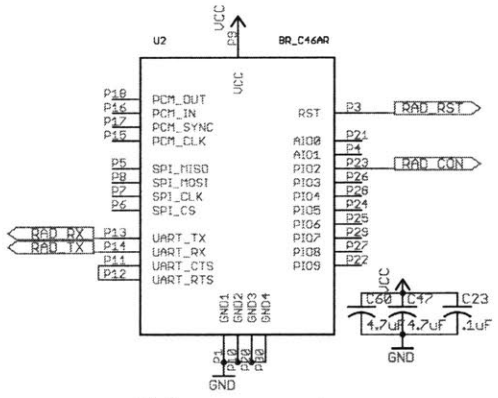
Hanger Design



Digital Logic	
TITLE: Owl_Hanger	
Document Number:	REV: 3.21
Date: 5/13/2007 02:27:27a	Sheet: 1/6

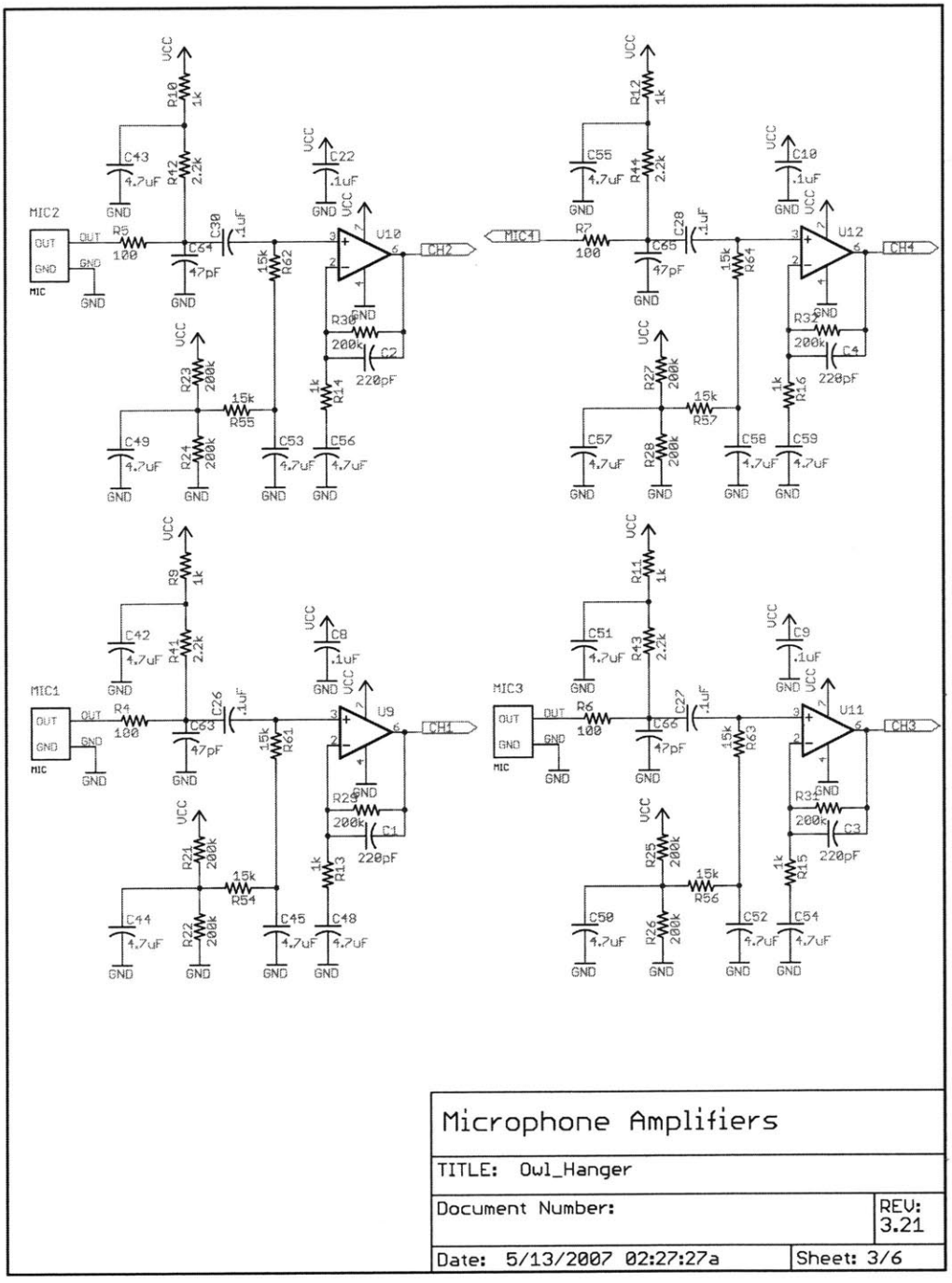


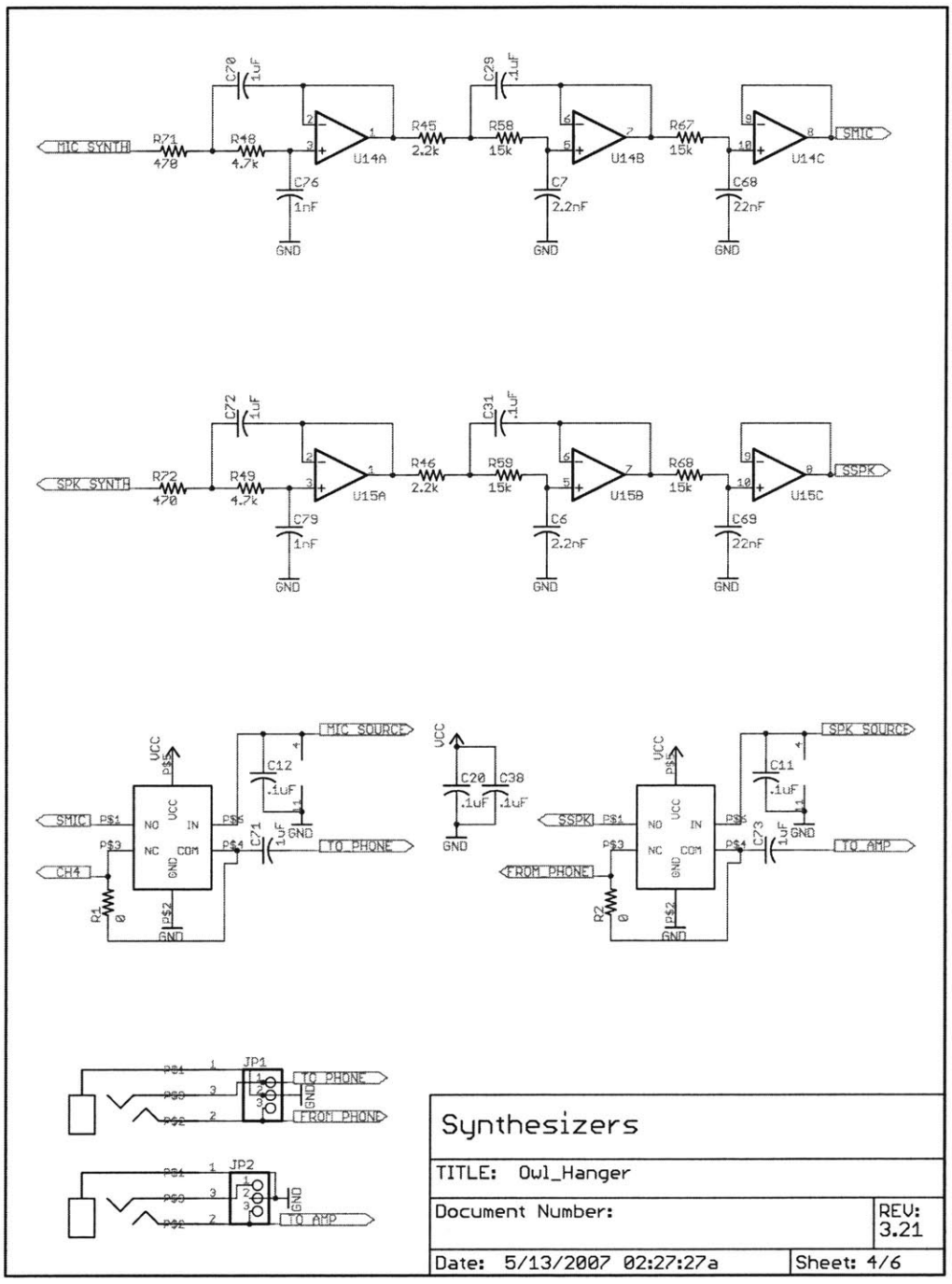
USB

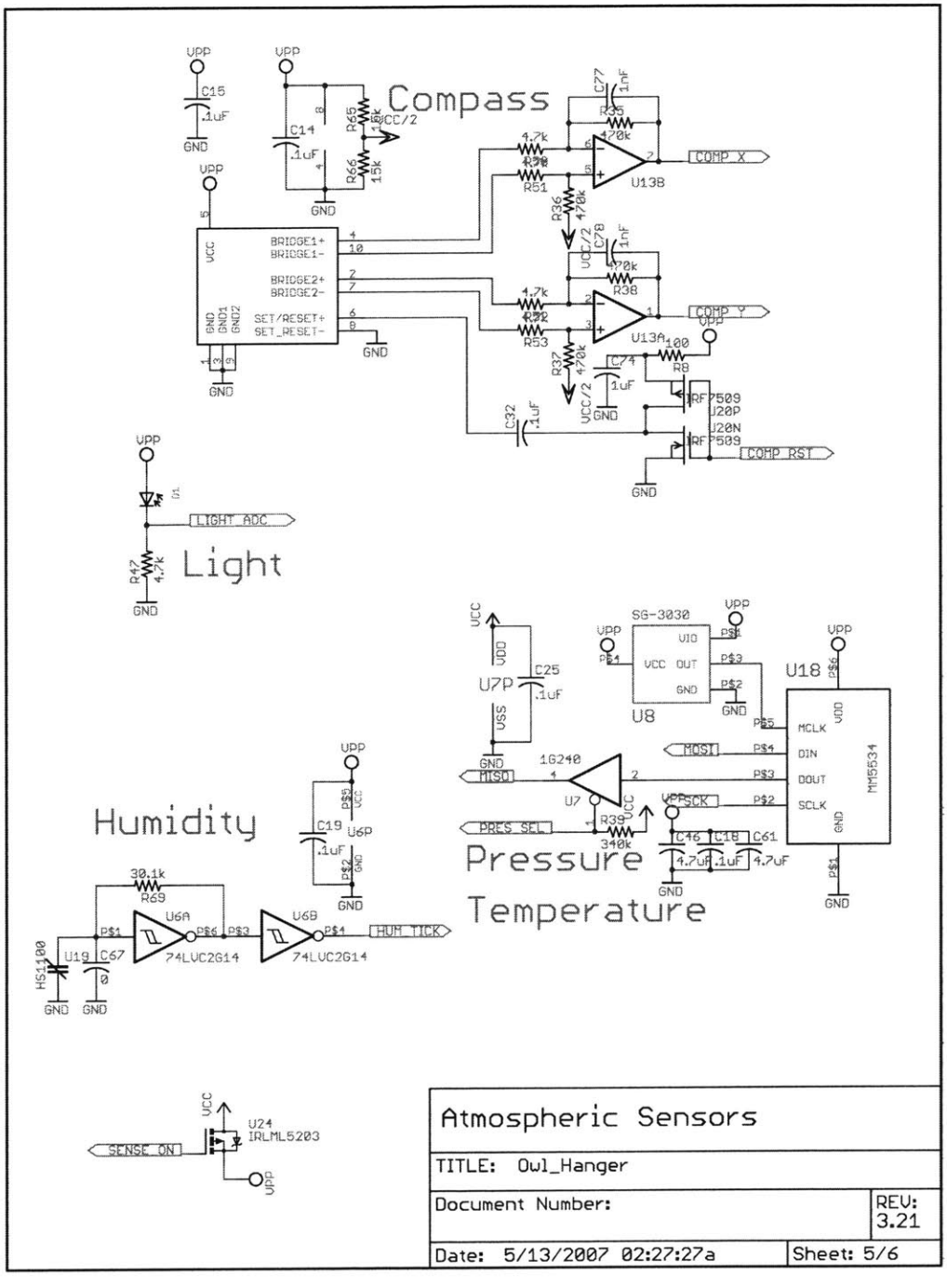


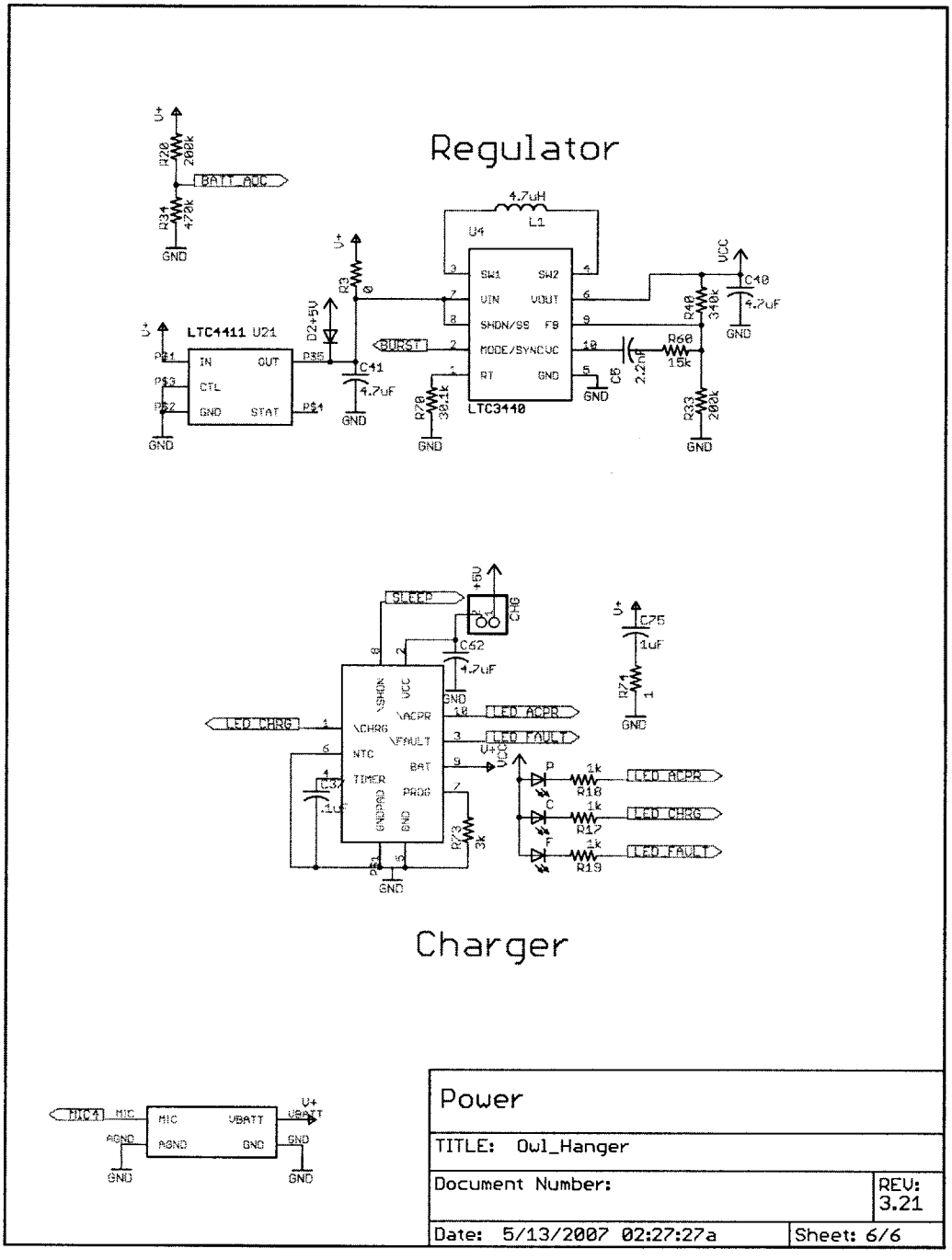
Bluetooth

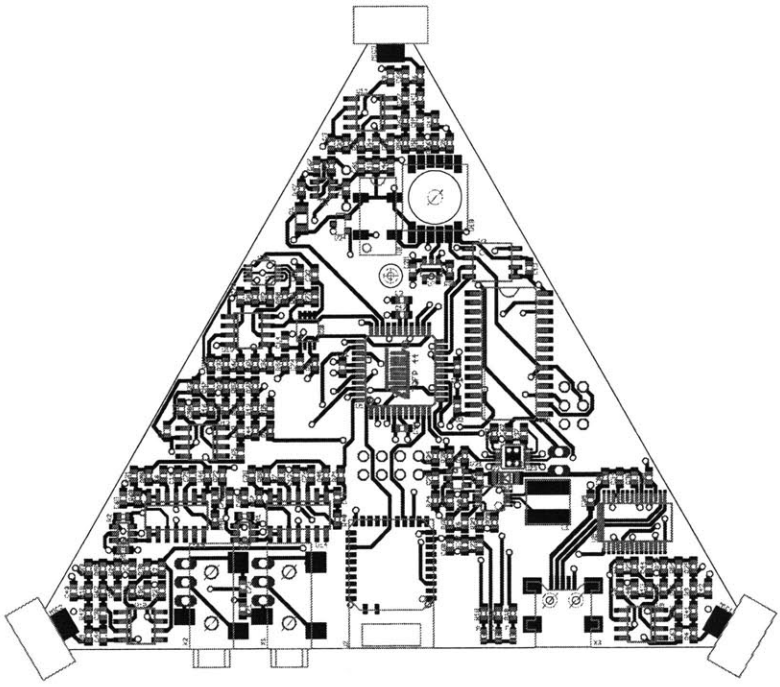
Communication	
TITLE: Owl_Hanger	
Document Number:	REV: 3.21
Date: 5/13/2007 02:27:27a	Sheet: 2/6


MIC2
MIC1
MIC3
MIC
OUT
GND
VCC
R10 1k
R12 2.2k
C43 4.7uF
C30 .1uF
C54 47pF
R5 100
R62 15k
U10
CH2
C22 .1uF
R30 200k
C2 220pF
R14 1k
R23 200k
R55 15k
R24 200k
C49 4.7uF
C53 4.7uF
C56 4.7uF
R12 1k
R44 2.2k
C55 4.7uF
C28 .1uF
C65 47pF
R7 100
R64 15k
U12
CH4
C10 .1uF
R32 200k
C4 220pF
R27 200k
R57 15k
R28 200k
C57 4.7uF
C58 4.7uF
C59 4.7uF
R16 1k
R9 1k
R41 2.2k
C42 4.7uF
C26 .1uF
C63 47pF
R4 100
R61 15k
U9
CH1
C8 .1uF
R29 200k
C1 220pF
R21 200k
R54 15k
R22 200k
C44 4.7uF
C45 4.7uF
C48 4.7uF
R13 1k
R41 1k
R43 2.2k
C51 4.7uF
C27 .1uF
C66 47pF
R6 100
R63 15k
U11
CH3
C9 .1uF
R31 200k
C3 220pF
R25 200k
R56 15k
R26 200k
C50 4.7uF
C52 4.7uF
C54 4.7uF
R15 1k
Microphone Amplifiers
TITLE: Owl_Hanger
Document Number:
REV: 3.21
Date: 5/13/2007 02:27:27a
Sheet: 3/6

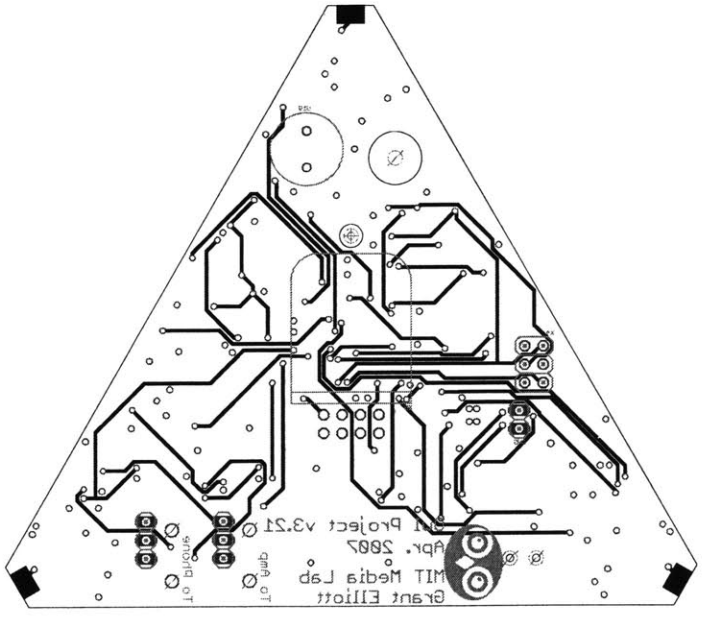

MIC SYNTH
R71
470
R48
4.7k
C70
1uF
C76
1nF
GND
U14A
R45
2.2k
C29
.1uF
R58
15k
C7
2.2nF
U14B
R67
15k
C68
22nF
U14C
SMIC
SPK SYNTH
R72
470
R49
4.7k
C72
1uF
C79
1nF
U15A
R46
2.2k
C31
.1uF
R59
15k
C6
2.2nF
U15B
R68
15k
C69
22nF
U15C
SSPK
MIC SOURCE
SMIC
CH4
R1
0
C12
.1uF
C71
1uF
TO PHONE
VCC
C20
.1uF
C38
.1uF
SPK SOURCE
SSPK
FROM PHONE
R2
0
C11
.1uF
C73
1uF
TO AMP
JP1
JP2
Synthesizers
TITLE: Owl_Hanger
Document Number:
REV: 3.21
Date: 5/13/2007 02:27:27a
Sheet: 4/6


Compass
Light
Humidity
Pressure
Temperature
Atmospheric Sensors
TITLE: Owl_Hanger
Document Number:
REV: 3.21
Date: 5/13/2007 02:27:27a
Sheet: 5/6

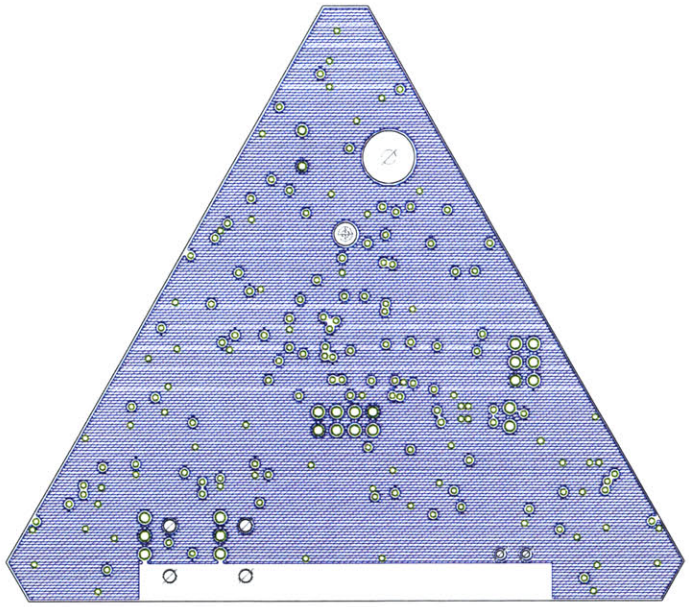

Regulator
BATT_ADC
R20 200k
R34 470k
GND
4.7uH
L1
U4
LTC3440
SW1
SW2
VIN
VOUT
SHDN/SS
FB
MODE/SYNC
RT
GND
BURST
D2+5V
R3 0
C41 4.7uF
R70 38.1k
C5 2.2nF
R60 15k
R40 340k
R33 200k
C10 4.7uF
VCC
LTC4411 U21
IN
OUT
CTL
GND
STAT
Charger
SLEEP
+5V
CHG
C62 4.7uF
C75 1uF
R74 1
LED_CHRG
LED_ACPR
LED_FAULT
\SHDN
VCC
\ACPR
\CHRG
\FAULT
NTC
BAT
TIMER
PROG
GND
C3 .1uF
R73 3k
R18 1k
R17 1k
R19 1k
MIC4
MIC
AGND
VBATT
Power
TITLE: Owl_Hanger
Document Number:
REV: 3.21
Date: 5/13/2007 02:27:27a
Sheet: 6/6



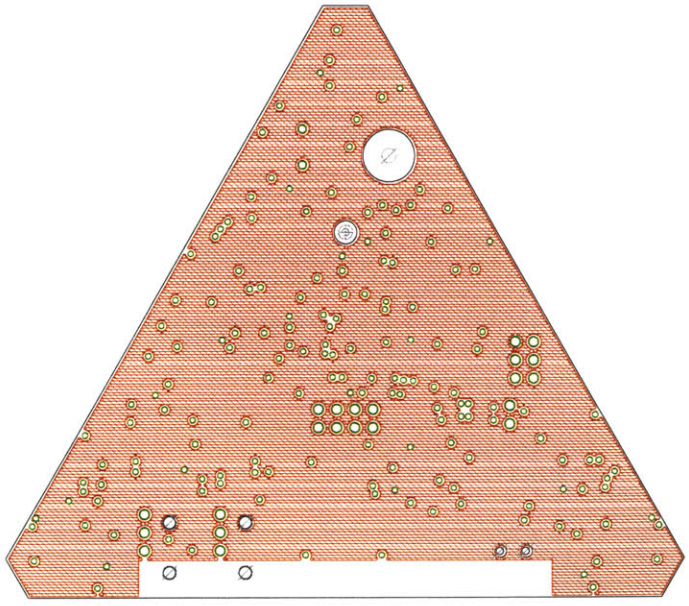
Top Layer



Bottom Layer



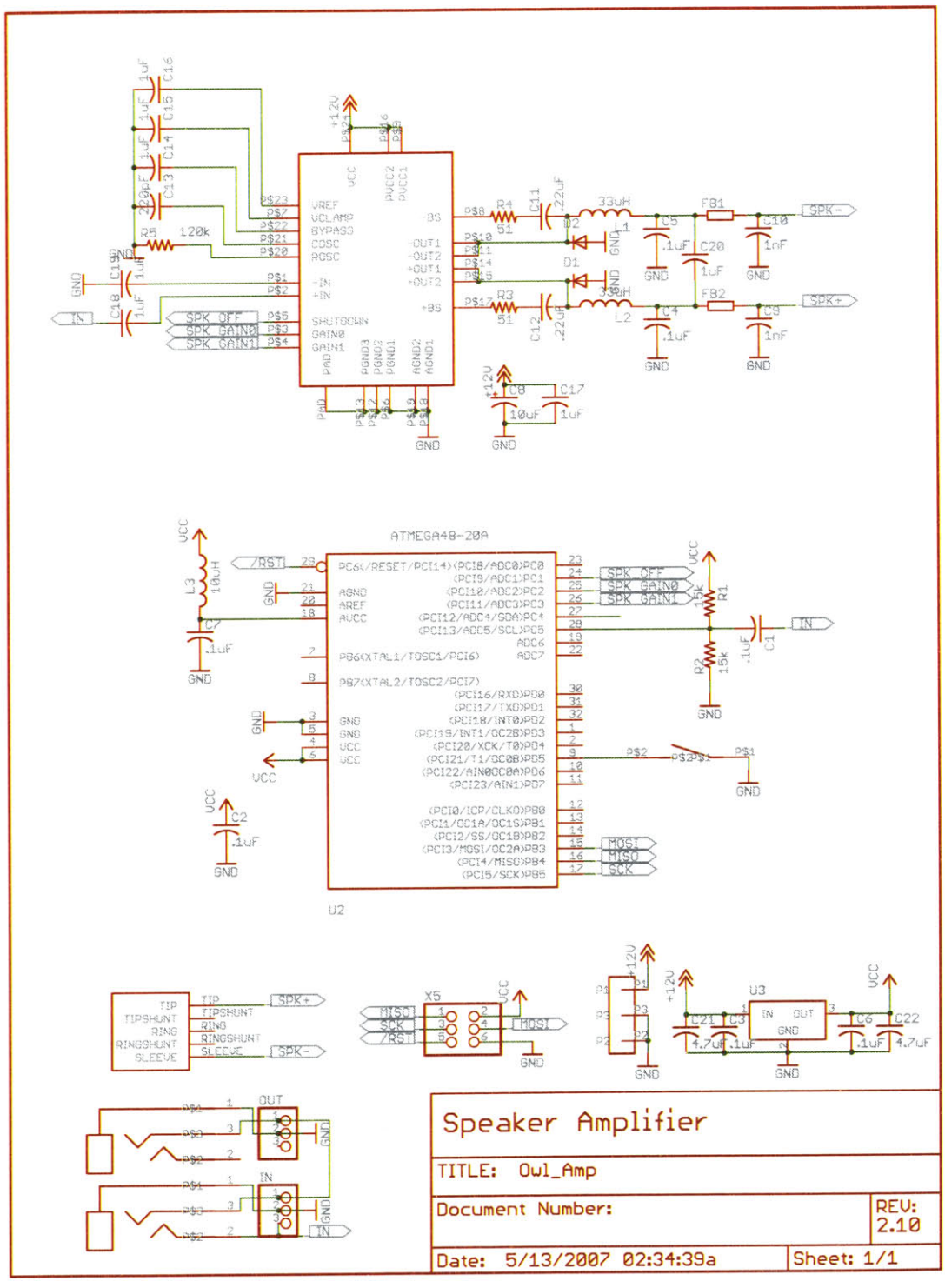
Inner Layer 2 (+3.3V Plane)

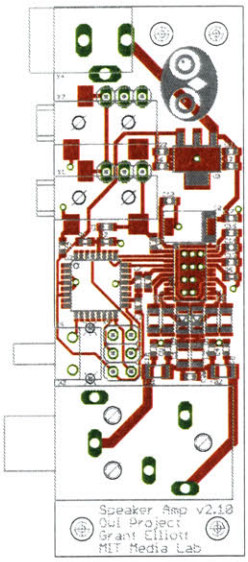


Inner Layer 3 (GND Plane)

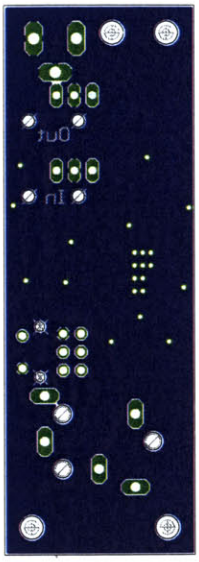
Appendix B

Amplifier Design


Speaker Amplifier
TITLE: Owl_Amp
Document Number:
REV: 2.10
Date: 5/13/2007 02:34:39a
Sheet: 1/1
ATMEGA48-20A
U2
SPK-
SPK+
IN
MOSI
MISO
SCK
/RST
VCC
GND



Top Layer



Bottom Layer