2aAB4. Bat species identification from zero crossing and full spectrum echolocation calls using Hidden Markov Models, Fisher scores, unsupervised clustering and balanced winnow pairwise classifiers

Ian Agranat*

*Corresponding author's address: Wildlife Acoustics, Inc., Concord, MA 01742-4939, ian@wildlifeacoustics.com

A new classification technique for the identification of bats to species from their echolocation calls is presented. Data is compiled and split in half for training and testing classifiers including 8,782 recording files (bat passes) with 223,123 candidate calls (pulses or extraneous noise) representing 17 species of bats found in North America including 10 species of Myotis. Some files are of high quality consisting of hand-selected search phase calls of tagged free flying bats while others are from a variety of field conditions including both active (attended) and passive (unattended) recordings made with a variety of zero crossing and full spectrum recording equipment from multiple vendors. Average correct classification rates are 88.8% with an average of 39.9% of all files identified to species. Most importantly, classifiers for two species of U.S. endangered bats, Myotis sodalis and Myotis grisescens have a positive predictive value of 100% and 98.6% respectively and a true positive rate of 67.4% and 93.8% respectively suggesting that the classifiers may be well suited to the accurate detection of these endangered bats.
INTRODUCTION

Populations of bats are commonly monitored acoustically because many species echolocate while foraging at night. Echolocation calls are typically recorded using zero crossing detectors or full spectrum recorders at high sample rates, and the resulting recordings are then analyzed with software to display the frequency modulated sweeps common to many bat species. Identifying bat species from their echolocation calls is desirable for management of biodiversity and compliance with environmental regulations. Human experts have proven that it is possible in many cases to identify bats by analysis of their echolocation calls. However, variation at several levels makes some species indistinguishable (Barclay, 1999). Many bats produce a wide range of calls to adapt to their physical surroundings, including the presence of other bats, and these calls often converge across species to very similar call types making identification difficult. Search phase calls are best suited for the acoustical identification of bats because they are the most commonly encountered in the field and have been shown to have species-specific characteristics (Allen, Burt, and Miller, 2007). However, field recordings collected from unattended passive monitoring sites will likely have a wide variety of call types present (e.g. clutter, feeding buzz, etc.) and quality (e.g. near and far, insect noise, and echoes).

There have been several efforts to develop algorithms for the automatic classification of bat calls including (Parsons and Jones, 2000), (Fukui, Agetsuma, and Hill, 2004), (Skowronski and Harris, 2006), (Corcoran, 2007), (Redgwell, Szewczak, Jones, and Parsons, 2009), (Britzke, E.R., J. Duchamp, R.S. Swihart, K.M. Murray, and L.W. Robbins, 2011) et. al. Prior published efforts are generally limited in scope attempting to classify from among only a dozen or so individual species using a relatively small number of hand-labeled search phase calls. Their performance when presented with the likely variety of call types and quality typical of unattended field recordings is unknown. Prior methods generally involve the extraction of discrete parameters (e.g. mean frequency, etc.) from each call pulse to form a feature vector, and then use any number of well-known techniques to classify to species including Discriminant Function Analysis, Artificial Neural Networks, Random Forest, and Support Vector Machines. While these parameters have demonstrated good classification rates on high quality search phase calls, consistent determination of parameters in noisy environments can be challenging and these parameters may not contain sufficient discriminant information to separate a larger variety of call types into classes. Finally each of these methods was developed specifically for analysis of either zero crossing or full spectrum recordings, not both, limiting the choices of recording systems that can be utilized by prior methods (Allen, C.R., S.E. Romeling, and L.W. Robbins, 2011).

The objective of this research is to develop new techniques for accurately classifying the echolocation calls of bats on a large scale in the presence of a wide variety of call types, field conditions and recording technologies. The author’s prior research in animal vocalization classification (Agranat, 2009) provides an imperfect starting point and motivation to find improvements with the hope that these new techniques may also find applications in other acoustic classification problems including birds, frogs, and cetaceans.

PARAMETERIZATION OF THE ECHOLOCATION CALLS OF BATS

Prior work has generally relied on a series of discrete parameters to describe the frequency modulated sweep of a bat’s echolocation call. This approach is limiting in that call parameters cannot always be extracted consistently from noisy environments, and may miss important subtleties present in the call needed for discrimination. Some pairs of species such as Myotis lucifugus and Myotis sodalis have significant overlap in discrete parameters such as call duration, characteristic frequency, start slope, slope at characteristic frequency, and cumulative normalized slope such that these species cannot be differentiated (Szewczak, 2011).

In many other disciplines including speech recognition, handwriting recognition, face recognition, electrocardiogram analysis, DNA sequencing, etc., the Hidden Markov Model (HMM) has proven to be a powerful generative model for representing variable length observation sequences. HMMs are the basis of the author’s prior work on bird song classification as well. In simple terms, an HMM is a generative model $\varphi$ that has a probability of generating a sequence of observations. The model is represented as a collection of states $X_t$ with each state having a probability $b_{ij}$ of emitting an observation $y_j$, and transition probabilities $a_{ij}$ of moving from one state $X_t$ to the next state $X_{t+1}$ at each step $t$. The observations can model discrete symbols or continuous functions, the latter commonly
represented as a \( d \)-dimensional Gaussian Mixture Model (GMM) with mixing coefficients \( \mathbf{m}_{ik} \), mean vectors \( \mu_{ik} \) and covariance matrices \( \Sigma_{ik} \).

In the GMM case, the emission probability is given by the Gaussian probability density function:

\[
b_{ik}(y) = \mathbf{m}_{ik} \arg\max_{x} \left[ (2\pi)^{-d/2} |\Sigma_{ik}|^{-1/2} e^{-\frac{1}{2}(y - \mu_{ik})'\Sigma_{ik}^{-1}(y - \mu_{ik})} \right]
\]  

(1)

Given an HMM \( \phi \) and an observation sequence \( y \), the prior probability \( P(y|\phi) \) can be calculated by using the iterative Viterbi algorithm:

\[
V_{0,k} = b(y_0|k) \pi_k \\
V_{t,k} = b(y_t|k) \arg\max_{x \in \mathcal{X}} (a_{x,k} V_{t-1,x}) \\
P(y|\phi) = \arg\max_{x \in \mathcal{X}} (V_{T,x})
\]

(2)  

(3)  

(4)

In zero-crossing recordings, the frequency modulated sweep of the bat’s echolocation call through time is represented as a series of points, or “dots”, one dot for every \( N \) zero crossings in the signal. The time between dots is stored in the recording. For analysis, the frequency and time position of each dot can be easily calculated, and this is how a plot of a bat’s frequency sweep through time can be generated. It is natural to think of the sequence of dots from a zero-crossing recording as an observation sequence, with each dot corresponding to an observation.

The approach described herein uses HMMs to model the variable length sequence of dots normalized to the same division ratio to represent the echolocation calls of bats. To help with model alignment and because the slope of the frequency sweep is known to be one of the important parameters used in prior efforts, a 2-dimensional feature vector modeled as a GMM for each state is used. One dimension represents the frequency of each dot, and the second dimension represents the change of frequency from the previous dot. These values are normalized to unit variance across the training data, and a diagonal covariance matrix is used for convenience.

A Hidden Markov Model can then be used to model bat echolocation calls. A reasonable estimate for an initial model would be one state for each dot with states having left-to-right transitions probabilities through the duration of the call. In practice, clustering observations into a smaller number of states using K-means also works well and is computationally more efficient.

By using HMMs to model echolocation calls, every detail of the call’s frequency sweep through time is captured without relying on discrete parameters to approximate isolated portions of the call. Hidden Markov Models have also been proven in other applications to be robust against noise and variation.

**DISCOVERING CALL TYPES WITH UNSUPERVISED CLUSTERING**

Given a large collection of files, each containing several echolocation calls and labeled with some imperfect degree of accuracy as belonging to a particular species, one objective is to automate the process of discovering those call types that are common to the species and therefore good candidates on which to base the classifiers. Another objective is to be robust against the inclusion of some files that are incorrectly labeled and noise incorrectly extracted as calls.

Various techniques for clustering HMMs are described by (Butler, 2003), (Chen, Man, and Nefian, 2005), et. al. used for analysis of speech, whale song, face recognition and other applications and are ideal for clustering variable-length time-series data. Training files are selected with the same species label and individual echolocation calls are extracted (or possibly other noise fragments erroneously extracted from field recordings) on which to perform agglomerative clustering. In one approach, each of these echolocation calls is represented by a single HMM. All of the HMMs are compared pairwise and the two most similar HMMs are merged into a single HMM. This process is repeated until the desired number of clusters has been realized. Unfortunately, this approach does not scale well for large data sets. Instead, a single HMM is trained with all of the calls, and then the inner product of Fisher scores is used to measure the similarity of calls. The Fisher score is defined as:
\[ U_y = \nabla \varphi \log P(y|\varphi) \]  

In the context of HMMs, the Fisher score represents how an observation sequence fits the model with respect to each of the model’s underlying parameters. The Fisher score is the gradient of the log probability of the observation sequence with respect to each of the model parameters. Two variable length observation sequences can be compared by taking the inner product of their Fisher scores. The angular distance between these two observations in hyperspace is thus:

\[ \cos^{-1}(U_x \cdot U_y) \]  

The algorithms perform agglomerative clustering by grouping the most similar calls together. This generally results in a small number of large clusters per species with a number of smaller clusters and un-clustered outliers. The larger clusters represent the most common call types present throughout a large portion of the training data (e.g. search phase calls). Smaller clusters may represent uncommon or highly variable calls, or even groups of misclassified calls that are similar to each other but don’t actually belong in the data set. The largest cluster by membership is included among the call types of interest. Additionally, any other similarly large clusters with at least 50% of the membership of the largest cluster are also included. In this way, a handful of HMMs are formed that are trained on the most common call types of a given species, and outliers are explicitly discarded.

To illustrate the power of this approach, consider the *Myotis lucifugus* (MYLU) and *Myotis sodalis* (MYSO) classification problem described previously that could not be easily solved using simple parametric measurements. This new clustering technique applied to a library of 93 MYSO recordings and 68 MYLU recordings discovers one MYLU cluster (with 332 member calls) and one MYSO cluster (with 456 member calls). HMMs are trained with left-to-right topology with a number of states equal to the average sequence length. A plot of the mean frequency with error bars of each state is shown for these clusters.

**FIGURE 1.** HMMs separating *Myotis lucifugus* (MYLU) and *Myotis sodalis* (MYSO).

Note that the MYLU cluster and the MYSO cluster overlap significantly in several places, but diverge slightly in states 5-10. Traditional parameters like the characteristic frequency (Fc), characteristic slope (Sc), initial slop (S1)
and others measure portions of the call where discrimination is not possible. On the other hand, HMMs allow us to model the entire call holistically and can tease apart subtle differences to greatly improve classification performance. The HMMs automatically discover differences in the calls encoded in the model parameters. Analysis of these model parameters suggests that these two species can be differentiated from a subtle difference in the slope of the frequency trajectory at around 55kHz. This discovered feature does not contribute significantly to the traditional discrete parameters used by prior methods and solves the classification problem where others have struggled.

**CLASSIFICATION OF CALLS**

Given an unclassified observation sequence representing an unknown bat call and a set of trained HMMs representing different species classes, a simple method of classification is to calculate the prior probability that each HMM generated the sequence and choose the class with the highest probability. This is essentially what the author’s prior work on birdsong classification was based on. However, it turns out that this technique does not produce reliable results. Hidden Markov Models are good generative models and are optimized to maximize the prior probabilities of the training sequences, but this is not the same as maximizing the discrimination between classes. Classification can break down in a number of scenarios. In the author’s prior work, some models with loop states and broader GMM variances became “greedy” allowing unrelated sequences to produce high scores. Additionally, short duration sequences that fit closely to a portion of a model trained on longer sequences scored high as well. When two classes are similar, the optimization for maximizing the probabilities of the training data without regard to maximizing the margin between classes can result in high rates of misclassification.

Fisher scores can improve discrimination of generative models. Using Fisher scores, a variable length observation sequence applied to a Hidden Markov Model can be transformed into a fixed-length feature vector in high dimensional space with one dimension corresponding to each of the underlying model parameters (e.g. state transition probabilities, means and variances of GMMs, etc.). This technique has been used successfully in many applications including sign language recognition (Aran and Akarun, 2009).

The Fisher score of an observation sequence for an HMM corresponding to the related class reveals which model parameters are most significant in their contribution to the prior probability. The Fisher score of the same sequence for unrelated HMMs can also be revealing in exposing similarities and differences between classes.

The method described uses a Fisher score vector and prior probability for all the included HMM clusters for all species as a classification feature vector given an unknown observation sequence. The number of parameters in each HMM can be quite large. An N-state model would have N initial state probabilities, N^2 state transition probabilities, and N GMMs each containing (in our case) one mixture in two dimensions with a mean and diagonal covariance matrix for a total of 5N+N^2 parameters. A 16 state model would therefore contribute 336 parameters to the feature vector. Given dozens of species, our feature vector can quickly grow to on the order of 10,000 dimensions. Only a small portion of these dimensions are expected to be significant for discrimination, many dimensions will be noisy, and many dimensions will have no significance (e.g. representing model parameters for seldom visited states or state transitions).

Binary classifiers based on Support Vector Machines are popular tools for handling highly dimensional spaces, but very good results are achieved using the simpler balanced Winnow methods. The Winnow algorithm is capable of rapid convergence on linearly separable data even when few variables are relevant (Kivinen and Warmuth, 1995). The balanced Winnow algorithm finds a hyper-plane separating two classes in multidimensional space defined by a positive and negative weight vector as follows:

\[ \sum_{i=0}^{n} (w_i^+ x_i - w_i^- x_i) > 0, \text{ class } c = 1. \text{ Otherwise } c = -1. \] (7)

The classifier is trained with an online learning algorithm. When a mistake is made, the weights are updated exponentially as follows:

\[ w_i^{t+1} = \frac{w_i^t e^{c p x_i}}{\sum_{t=0}^{m} w_i^{t+1}} \] (8)
After training, the weights of the Winnow hyper-planes can be analyzed to reduce the dimensionality of the problem significantly in the final classifier by eliminating the dimensions that do not contribute significantly to the classification.

Identification of bats to species is a multi-class classification problem solved with binary classifiers capable of separating data sets with linear hyper-planes in highly dimensional space. One approach is a one-against-all strategy in which a binary classifier is trained for each class in an attempt to find a hyper-plane that separates members in the class from members in all other classes. Such an approach if possible has the benefit of excluding noise and other non-identifiable calls from analysis. Unfortunately, it is not likely given the number of classes and the similarity among some classes that a one-against-all approach is solvable with a linear separating hyper-plane. Instead, an exhaustive pair-wise strategy is used with a binary classifier trained to find a separating hyper-plane between each pair of classes. In addition, one extra class is defined to represent the unidentifiable sequences derived from the discarded HMM clusters. Using this method, Winnow converges on a linear separating hyper-plane with average error less than 3% for all pairwise combinations.

The outputs of the pairwise classifiers are combined using a hinge loss function to predict the winning class with a confidence factor. If the winning class turns out to be the unidentifiable sequence class, then the call is rejected as unknown.

At the file level, there may be a sequence of echolocation calls recorded during a “bat pass” over several seconds. Each call is classified as above with the confidence factors accumulated by species and normalized across the file. The file is considered identified to species if the normalized confidence factors exceed some threshold.

\[
W_i, t+1 = \frac{w_i t e^{-\beta \eta_i}}{\sum_{i=0}^{\infty} w_i t+1}
\]

(9)

DISCUSSION AND RESULTS

Recordings of bats were acquired from three different sources in North America for this research. Combined, the data includes 17 species of bats found throughout North America including 10 species of *Myotis*. The dataset has 8,782 files from which 223,123 candidate calls were extracted. These recordings were made in a variety of field conditions including both active (attended) and passive (unattended) recordings made with a variety of zero crossing and full spectrum recording equipment from multiple vendors. The dataset is not evenly distributed among species with 38.0% of the files representing *Eptesicus fuscus* and only 0.6% representing *Corynorhinus townsendii*.

Zero crossing files were normalized to a division ratio of 8 and full spectrum files were converted to zero crossing with Kaleidoscope. Individual echolocation calls are extracted from each file by looking for chains of between 10-100 dots forming smooth trajectories. After a call is extracted, 50ms of the file is skipped to avoid echoes before looking for the next call. Each call is represented as a sequence of two-dimensional feature vectors, one vector for each dot, with dimensional components representing the frequency and change in frequency normalized to unit variance.

Each dataset is split in half with approximately one half of the files used for training the classifiers and the other half used for testing the classifiers against previously unconsidered recordings.

The classifier is built from the training data as follows: For each species, a Hidden Markov Model is trained on all of the corresponding calls. For each pair of calls corresponding to each species, the inner product of their Fisher scores is used to determine their similarity expressed as an angular separation in hyperspace. Agglomerative clustering is performed by merging calls from the smallest angular separation up to some maximum. This generally results in a few large clusters per species which are kept while discarding the others. A feature vector is generated for each call by combining the Fisher scores for all of the models and their prior probabilities.

For each pair of clusters across all species, a balanced Winnow classifier is trained to find an optimal hyper-plane separating each pair. Additional Winnow classifiers are trained between each cluster and all discarded clusters.
which are labeled as “no identification”. On average, the Winnow classifiers successfully separated 97.17% of call pairs.

For testing, a feature vector is extracted from each call of each test file as above. The feature vector is then classified by each of the binary Winnow classifiers and a hinge loss function is utilized to determine the overall winning class with a confidence level indicated by the combined classifier margins. If the “no identification” label wins, then the call is discarded. At the file level, the confidence of each winning call is combined to determine an overall winning species classification with an overall confidence score.

Classifier performance is shown in the tables below for both training and testing data with minimum margin thresholds optimized for accuracy (i.e. low false positive rates) preferred to sensitivity (i.e. high true positive rates). The True Positive Rate (TPR), or sensitivity, is the number of correct classifications divided by the total number of corresponding inputs for a given classifier and is inversely related to Type II errors. The Correct Classification Rate (CCR) is the number of correct classifications divided by the total number of classifications for a given input. The Positive Predictive Value (PPV) is a measure of classifier performance and is the number of correct classifications divided by the total number of classifications for a given classifier. Note that PPV is normalized to account for the differences in the distribution of species in the datasets.

**TABLE 1.** Classifier results on training data

<table>
<thead>
<tr>
<th>Species</th>
<th>Files</th>
<th>Calls</th>
<th>TPR</th>
<th>CCR</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corynorhinus townsendii</td>
<td>29</td>
<td>451</td>
<td>72.4%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Eptesicus fuscus</td>
<td>1,672</td>
<td>38,036</td>
<td>32.1%</td>
<td>98.2%</td>
<td>95.8%</td>
</tr>
<tr>
<td>Lasionycteris noctivagans</td>
<td>72</td>
<td>1,521</td>
<td>56.9%</td>
<td>100.0%</td>
<td>99.9%</td>
</tr>
<tr>
<td>Lasiurus borealis</td>
<td>778</td>
<td>9,334</td>
<td>68.5%</td>
<td>91.3%</td>
<td>98.8%</td>
</tr>
<tr>
<td>Lasiurus cinereus</td>
<td>265</td>
<td>3,940</td>
<td>57.4%</td>
<td>98.1%</td>
<td>99.4%</td>
</tr>
<tr>
<td>Myotis septentrionalis</td>
<td>71</td>
<td>2,139</td>
<td>71.8%</td>
<td>98.1%</td>
<td>91.4%</td>
</tr>
<tr>
<td>Myotis sodolii</td>
<td>47</td>
<td>1,577</td>
<td>97.9%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Myotis californicus</td>
<td>81</td>
<td>2,119</td>
<td>60.5%</td>
<td>92.5%</td>
<td>99.6%</td>
</tr>
<tr>
<td>Myotis ciliolabrum</td>
<td>57</td>
<td>1,906</td>
<td>63.2%</td>
<td>97.3%</td>
<td>97.4%</td>
</tr>
<tr>
<td>Myotis grisescens</td>
<td>33</td>
<td>1,453</td>
<td>93.9%</td>
<td>100.0%</td>
<td>98.8%</td>
</tr>
<tr>
<td>Myotis keenii/evotis</td>
<td>363</td>
<td>7,335</td>
<td>36.4%</td>
<td>98.5%</td>
<td>99.3%</td>
</tr>
<tr>
<td>Myotis lucifugus</td>
<td>366</td>
<td>13,080</td>
<td>56.3%</td>
<td>94.5%</td>
<td>94.4%</td>
</tr>
<tr>
<td>Myotis thysanodes</td>
<td>36</td>
<td>848</td>
<td>91.7%</td>
<td>100.0%</td>
<td>99.3%</td>
</tr>
<tr>
<td>Myotis volans</td>
<td>45</td>
<td>1,258</td>
<td>86.7%</td>
<td>100.0%</td>
<td>99.7%</td>
</tr>
<tr>
<td>Myotis yumanensis</td>
<td>370</td>
<td>23,010</td>
<td>72.2%</td>
<td>98.9%</td>
<td>99.8%</td>
</tr>
<tr>
<td>Nycticeius humeralis</td>
<td>51</td>
<td>2,251</td>
<td>86.3%</td>
<td>100.0%</td>
<td>97.7%</td>
</tr>
<tr>
<td>Perimyotis subflavus</td>
<td>60</td>
<td>1,504</td>
<td>96.7%</td>
<td>100.0%</td>
<td>98.2%</td>
</tr>
<tr>
<td><strong>Totals/Average</strong></td>
<td>4,396</td>
<td>111,762</td>
<td>70.6%</td>
<td>98.1%</td>
<td>98.2%</td>
</tr>
</tbody>
</table>

**TABLE 2.** Classifier results on testing data

<table>
<thead>
<tr>
<th>Species</th>
<th>Files</th>
<th>Calls</th>
<th>TPR</th>
<th>CCR</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corynorhinus townsendii</td>
<td>29</td>
<td>515</td>
<td>10.3%</td>
<td>75.0%</td>
<td>97.4%</td>
</tr>
<tr>
<td>Eptesicus fuscus</td>
<td>1,671</td>
<td>38,036</td>
<td>22.3%</td>
<td>94.7%</td>
<td>89.3%</td>
</tr>
<tr>
<td>Lasionycteris noctivagans</td>
<td>71</td>
<td>1,482</td>
<td>15.5%</td>
<td>91.7%</td>
<td>99.6%</td>
</tr>
<tr>
<td>Lasiurus borealis</td>
<td>776</td>
<td>9,058</td>
<td>49.0%</td>
<td>81.7%</td>
<td>72.0%</td>
</tr>
<tr>
<td>Lasiurus cinereus</td>
<td>265</td>
<td>3,727</td>
<td>45.3%</td>
<td>97.6%</td>
<td>98.1%</td>
</tr>
<tr>
<td>Myotis septentrionalis</td>
<td>70</td>
<td>1,856</td>
<td>18.6%</td>
<td>86.7%</td>
<td>72.3%</td>
</tr>
<tr>
<td>Myotis sodolii</td>
<td>46</td>
<td>1,559</td>
<td>67.4%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Myotis californicus</td>
<td>82</td>
<td>1,915</td>
<td>17.1%</td>
<td>56.0%</td>
<td>91.2%</td>
</tr>
<tr>
<td>Myotis ciliolabrum</td>
<td>56</td>
<td>1,711</td>
<td>21.4%</td>
<td>80.0%</td>
<td>73.1%</td>
</tr>
<tr>
<td>Myotis grisescens</td>
<td>32</td>
<td>1,382</td>
<td>93.8%</td>
<td>100.0%</td>
<td>98.6%</td>
</tr>
<tr>
<td>Myotis keenii/evotis</td>
<td>363</td>
<td>7,725</td>
<td>34.7%</td>
<td>95.5%</td>
<td>92.3%</td>
</tr>
<tr>
<td>Myotis lucifugus</td>
<td>363</td>
<td>13,482</td>
<td>32.8%</td>
<td>86.9%</td>
<td>70.8%</td>
</tr>
<tr>
<td>Myotis thysanodes</td>
<td>38</td>
<td>907</td>
<td>36.8%</td>
<td>87.5%</td>
<td>93.3%</td>
</tr>
<tr>
<td>Myotis volans</td>
<td>45</td>
<td>1,138</td>
<td>17.8%</td>
<td>100.0%</td>
<td>87.9%</td>
</tr>
</tbody>
</table>
Average correct classification rate on test data was 88.8% with an average of 39.9% of all files identified to species. Most classifiers did very well, but some had higher classification errors, most notably Corynorhinus townsendii (COTO). As mentioned earlier, COTO is the most underrepresented in the training data with only 29 training files. The performance of classifiers on the training data is significantly better compared to the test data. For training data, the average correct classification rate was 98.1% with 70.6% of all files identified to species. This suggests that additional training data may improve classification performance.

The most important result is that two species of U.S. endangered bats including *Myotis sodalis* (MYSO) and *Myotis grisescens* (MYGR) are accurately detected and identified. The MYSO classifier has a Correct Classification Rate of 100%, a False Positive Rate of 0.0%, and a True Positive Rate of 67.4% meaning that approximately two in every three bat passes will be detected, and once detected, will be accurately identified. The MYGR classifier has a Correct Classification Rate of 100%, a False Positive Rate of 0.1%, and a True Positive Rate of 93.8% meaning nearly all bat passes will be detected and identified correctly. These results suggest that the methods described are well suited to the accurate detection of these protected species. That being said, the size of the data is limited and more testing is needed.

**FUTURE WORK**

Up to this point, the author has avoided review of the underlying recordings in any detail. In fact, the strength of these methods is the ability to build classifiers from raw data without further human analysis. One obvious next step is to study in greater detail common misclassifications to better understand how the methods can be improved further. Additionally, analysis of clustering and HMM parameters may offer insights into the structure of the echolocation calls of bats. We are also interested in applying these classification methods to other animal vocalizations such as birds, frogs and cetaceans.

**ACKNOWLEDGEMENTS**

I would like to thank the following people and organizations for making bat call recordings available for our research:

Dr. Cori Lausen, Birchdale Ecological Ltd., B.C. Canada

Lynn Robbins, Professor, Dept. of Biology, Missouri State University

U.S. National Park Service

**BIBLIOGRAPHY**


Parsons and Jones. (2000). Acoustic Identification of Twelve Species of Echolocating Bat by Discriminant Function Analysis and Artificial Neural Networks. Bristol: School of Biological Sciences, University of Bristol.

