Experimental assessment of distance-from-source effects on spectral classification of bird vocalizations recorded in natural environments.

Jordon E. Davis* and Kathryn E. Sieving, Wildlife Ecology & Conservation, University of Florida

Introduction

Our research addresses causes of vocal complexity in birds of Family Paridae. With Terabytes of automated recordings to sift through for parid calls, we seek to utilize spectral correlation for data retrieval. We assume that spectral detection will decline with sound degradation by distance from source (Fig. 1), with denser vegetation and via masking by a variety of natural and human background noises. We present an experimental assessment of distance effects and other factors on spectral detection of 4 common parid vocalizations (Obj. 1), and a test of one strategy to overcome these issues (Obj. 2) using Avisoft SASLab Pro.

Objectives

Objective 1: To experimentally determine loss of spectral detection accuracy with distance from source, vegetation density, and masking by background noise using 4 calls (Fig. 2) from 2 common parid spp. in their natural habitats in Florida, USA: Eastern Tufted Titmouse (ETTI; Baeolophus bicolor) and Carolina Chickadee (CACH; Poecile carolinensis).

Objective 2: To evaluate a strategy for minimizing degradation of spectral detection accuracy.

Methodology

Objective 1: Call degradation by distance, habitat and masking

- Recorded birds w/ background noise and created (a) standardized playback files (PF) & (b) spectral tags (ST) from clean recordings.
- Recorded playbacks of PF at 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, & 100 m on open pine and dense (oak) forest transects (8 transects x 11 distances).
- Ran continuous spectral correlation of ST on recorded files (Table 1).

Objective 2: Enhancing detections via tag set construction

We compared the performance of 5 different tag sets varying in # of tags and composition when correlated with files recorded across the range of distances: (1) Small tag sets (~10 tags) with (a) pure tonal (clean) or (b) purely degraded tags (recorded at distances > 30m from playback speaker) or (c) a mix of (a) and (b); and (2) Large (doubled) tag sets with (b) or (c) compositions (only CACH_D tags were used in this analysis).

Results

Table 1. Continuous Spectral Correlation Settings

<table>
<thead>
<tr>
<th>Call Type</th>
<th>High Pass</th>
<th>Low Pass</th>
<th>ID Threshold</th>
<th>Masking</th>
<th>Freq. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) CACH D</td>
<td>3050</td>
<td>6700</td>
<td>0.835</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>(b) CACH HZ</td>
<td>6630</td>
<td>9040</td>
<td>0.86</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>(c) ETTI D</td>
<td>1300</td>
<td>7300</td>
<td>0.65</td>
<td>0.15</td>
<td>0</td>
</tr>
<tr>
<td>(d) ETTI Z</td>
<td>4300</td>
<td>9150</td>
<td>0.795</td>
<td>0.1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. General Linear Model Results

<table>
<thead>
<tr>
<th>Call Degradation</th>
<th>Distance (m)</th>
<th>Note Type</th>
<th>Habitat</th>
<th>Dist^* Note</th>
<th>Noise Type</th>
<th>Error</th>
<th>Total</th>
<th>Model R^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>10.05</td>
<td>0.5</td>
<td>8.3</td>
<td>10.4</td>
<td>0.096</td>
<td>609</td>
<td>240.2</td>
<td>0.825</td>
</tr>
<tr>
<td>Habitat</td>
<td>22.2</td>
<td>22.2</td>
<td>370.5</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Note Type</td>
<td>10.05</td>
<td>11.1</td>
<td>184.2</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habitat^* Note</td>
<td>3.9</td>
<td>1.0</td>
<td>16.9</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dist^* Note</td>
<td>3.1</td>
<td>0.1</td>
<td>1.7</td>
<td>0.0094</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noise Type</td>
<td>2.3</td>
<td>0.8</td>
<td>13.0</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>609</td>
<td>36.5</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig 1: Degradation by distance of pure tonal recordings of various parid calls and notes between 1 to 100m from sound playback source. (JR Lucas, unpublished data).

Fig 2: We chose the 4 note types for contrasting bandwidths and transmission performances over distance.

Fig 3: Distance from source and habitat type (vegetative density) have substantial and immediate effects on the detection accuracy using standard settings for continuous correlation (Table 1) using Avisoft.

Fig 4: Detection accuracy is reduced by half (on average) for all note types in more dense hardwood (oak) habitat than open pine woodlands (e.g., like background photo).

Fig 5: Small plane flyovers and birds calling close to the microphone during playback transects decreased tag detection accuracy in Avisoft, but not as much as insects (e.g., cicada spp.).

Fig 6: Increasing tag set size (doubling the #) and including tags from degraded recordings made at distances > 30m from sound source increased overall detection accuracy (F, df = 12, p<0.0001).

Fig 7: A large tagset (16 tags of pure tonal exemplars + tags made from CACH_D notes recorded > 30m from source) performed best over distances, and significantly better at > 30m than a small 'clean' tagset (F, df = 2, p = 0.001).

Analysis

Detection accuracy (# true positive / # false negative+#true positive detections) was assessed under effects of distance from source, noise type, parid call type, habitat type (Obj. 1) and tag set composition (Obj. 2) with CLM (normal tag models). Detection accuracy was angular-transformed, arcsin(sqrt(prop)) to normalize the detection proportion.

Objectives

Objective 1: Call degradation by distance, habitat and masking

- Significant reductions occurred in detection accuracy for note types: in oak habitat; with increasing distance (Figs 3, 4); and when insects were calling (Fig. 5).
- Titrums (ETTI) D notes transmit poorly in open pine (but not dense oak) than other notes tested (Fig. 4).

Objective 2: Enhancing detections via tag set construction

- Using Avisoft SASLab Pro, we compared the performance of 5 different tag sets varying in # of tags and composition when correlated with files recorded across the range of distances: (1) Small tag sets (~10 tags) with (a) pure tonal (clean) or (b) purely degraded tags (recorded at distances > 30m from playback speaker) or (c) a mix of (a) and (b); and (2) Large (doubled) tag sets with (b) or (c) compositions (only CACH_D tags were used in this analysis).

Conclusions

- Detection accuracy varies with factors that affect sound degradation (distance from source; vegetation density; spectral properties of calls, and background noise).
- Increasing tag set size and using both pure and degraded tonal tags can enhance call detection accuracy in ambient recordings.

Useful References & Acknowledgements

- See hand-out provided.