

鳥類音声録音装置の試作と 自動種数数え上げシステム構築の試み

窪田 伊織¹ 丸山 和暉²

中嶋 響² 山崎 皓平³ 村井 亮介^{4*}

(受領日：2021年5月31日)

¹ 高知工科大学大学院工学研究科基盤工学専攻情報学コース

〒782-8502 高知県香美市土佐山田町宮ノ口185

² 高知工科大学大学院工学研究科基盤工学専攻社会システム工学コース

〒782-8502 高知県香美市土佐山田町宮ノ口185

³ 高知工科大学大学院工学研究科基盤工学専攻航空宇宙工学コース

〒782-8502 高知県香美市土佐山田町宮ノ口185

⁴ 高知工科大学地域連携機構

〒782-8502 高知県香美市土佐山田町宮ノ口185

* E-mail: murai.ryosuke@kochi-tech.ac.jp

要約：本研究では生態系調査のため、生態系の上位にあたる鳥類に注目し、鳥の鳴き声を録音する装置を作成した。また、録音した鳥の鳴き声を基に鳴いている鳥の種類を判別するシステムについて検討した。録音装置は野外での録音を可能にするため、機材を守るためハウジングと内部ケースの2重構造とした。また、鳥の種類を判別するシステムの精度検証では、録音装置で実際に録音した音声から、鳥が鳴いた箇所をトリミングし、これを学習用と評価用のデータに分けて、分類実験を行った。今回分類の対象とする鳥は比較的鳴き声わかりやすいキジバトとヒヨドリとした。メルケプストラム係数を特徴量とした Support Vector Machine を用いて鳥の分類を行った結果、正答率は全体として50%前後という結果になった。

1. はじめに

高知工科大学では心豊かな里山の再生とコミュニティの拡大を目的とし、科学技術を社会に実装する「里山基盤科学技術の社会実装モデルプロジェクト（通称：佐岡プロジェクト）」を行っている。

研究の前段階として、佐岡地区に新たなコミュニティを確立させるために、佐岡の生態系に特化した図鑑の作成を考えた。その図鑑が活用されることで、佐岡を訪れる人が増え、人々の交流が予想され

る。一般に図鑑作成には生態系調査が必要であるが、調査には専門知識や経験が必要とされており、その多くは専門家によって行われている。“松岡茂，ラインセンサスと音声録音による鳥類記録種の比較，2004”¹⁾では、鳥類の専門家である調査者が直接現地に赴いてセンサスを行うことが多い現行の体制で、信頼性や効率性を保持するのが困難な場合もあると想定し、鳥類の目録作りにおけるこれら2つの要件を高めるために録音データが鳥類調査やモニタリングに利用可能かどうかの検討を行って

録音装置の構成

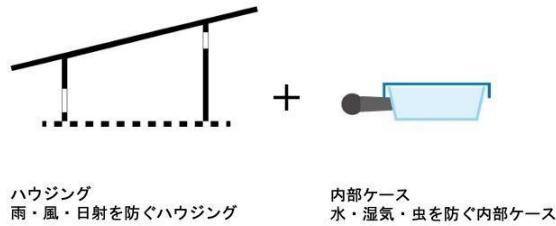


図1. 録音装置の全体構成

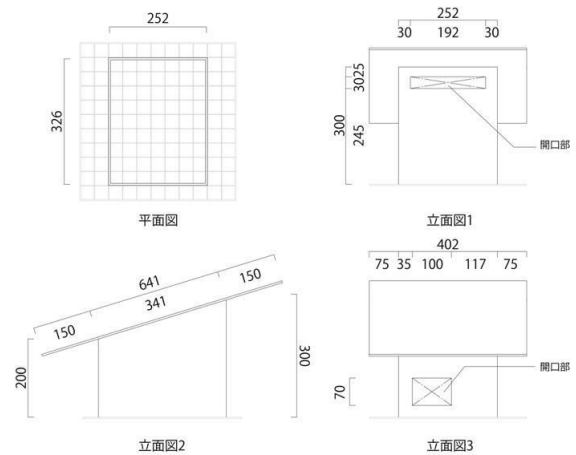


図2. ハウジングの寸法 (単位:mm)

る。ここではラインセンサスと録音聴取によって記録された鳥の種類を比較することで録音が有効な手段となりうると考えられているが、現在まで一般の条件下での推定方法はまだ確立されていない。

そこで、本研究では調査対象を鳥類に設定し、鳥類の音声録音装置の試作と録音データから鳥の種の特性を自動で行うシステムの構築を目的として録音装置の開発、録音実験、データセットの作成、鳥類の分類を行った。本稿では、その手法と結果、考察を示す。

2. 録音装置の試作

2.1 全体構成

録音装置の全体構成は、図1のように、入れ子状に2つに分かれており、雨・風・日射を防ぐハウジングと、さらに水・湿気・虫を防ぐ内部ケースから構成されている。

2.2 ハウジング

ハウジングの寸法は、図2の通りである。ハウジングの主な材料は、黒色のプラスチックダンボールを用いた。全体を支える底部は、金属網を用いており、ハウジングの壁面上部に穴を設けることで通気性を確保した。また底部と壁は結束バンドを用い固定した。

2.3 内部ケース

内部ケースにはタッパーを用い、図3のように内部に各機材を収めた。機材は、マイク・GPSモジュールアンテナ・除湿剤・SSD 240GB・Raspberry Pi3 (ラズパイ) を用いており、電源は電工ドラムを用い給電した。

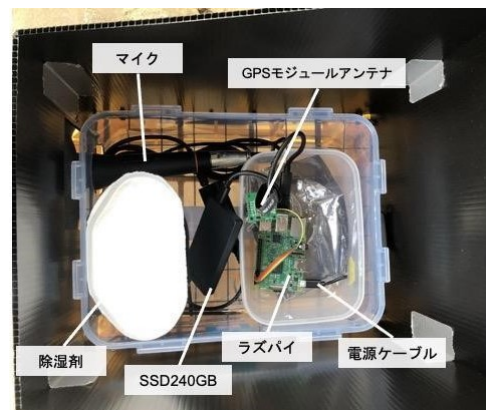


図3. 内部ケースの機材構成

2.4 細部

マイクや電源確保のため、ハウジングや内部ケースには穴を開けなければならない。その穴を塞ぐため、図4のようにハウジングにはスポンジ・ダクトテープ、内部ケースにはダクトテープを用いた。これにより雨・虫の侵入を防ぐことが出来た。

2.5 録音処理

Raspberry Pi3 を用いた録音では、次の設定で録音を行った。

- 形式: wave
- サンプルレート: 4800
- 量子化ビット数: 16bit
- チャンネル: ステレオ

今回は Sound eXchange (以下 SoX と呼ぶ) というコマンドライン上で動作するオーディオ編集ソフトを使用した。この SoX における sox コマンドと Unix においてスケジュールされた操作を行うための crontab コマンドを用いて、コマンドライン上で

詳細部分



図 4. 細部の収め

1 分ごとに 30 秒間の録音を行うシステムを構築した。これには、30 秒間録音を行う bash ファイルを、crontab を用いて 1 分ごとに実行させることによって実現している。実際に使用した bash ファイルを次に示す。

```
#!/user/bin/bash/  
NOW='date +%Y-%m-%d_%H-%M-%S'  
echo ${NOW}  
sox -c 2 -t alsa plughw:0,0  
  /mnt/usb/${now}.wav trim 0 30  
exit 0
```

このコマンドで指定している保存先 /mnt/usb/ は SSD に紐づいているディレクトリである。ここにデータを記録することで、SSD に録音データを保存した。また、今回はネット環境のない屋外での実験となるため、Raspberry Pi3 は起動後に正確な時刻を取得できない問題があった。そこで今回は Raspberry Pi 専用の GPS モジュールを用いて人工衛星から時刻を取得し、時刻同期を行なった。

3. 録音実験

3.1 設置場所

設置場所は、高知県香美市土佐山田町中後入にある古民家の門の軒下とした。図 5 左は、航空写真であるが、古民家は山の中の、南向き斜面に存在しており、南側が大きく開けている。また図 5 右のようにボトルクレートに録音装置を設置して録音した。

3.2 録音期間

録音は、2020 年 9 月 19 日 6 時 00 分～19 時 00 分、2020 年 12 月 2 日 14 時 47 分～2020 年 12 月 9 日 11 時 08 分、2021 年 1 月 12 日 16 時 43 分～2021 年 1 月 18 日 15 時 31 分の期間行った。この期間内には雨天の日も含まれている。



図 5. 録音装置の設置位置

(図左は、国土地理院地図より引用)

4. データセットの作成

分類を行うために録音したデータを用いて学習や評価に用いるデータセットの作成を行った。データセットはプログラムの学習に使用するもの（以下、学習用データセット）と分類結果の評価に使用するもの（以下、評価用データセット）の 2 種類に分けて作成した。学習用データセットと評価用データセットにはデータごとにどの鳥が鳴いているかラベル付けが必要なため、データの切り取り時に作業者がそれを音声から判断する必要がある。しかし、作業を行う者には専門知識や経験がないため鳥の種類判別が厳しい。そこで今回は、録音装置の設置場所で鳴く頻度の多いヒヨドリと、鳴き声の特徴的なキジバトに限定して切り取りを行った。

4.1 データセット作成手順

本研究では、データ確認作業を日付毎に二人で分担し、切り取り作業を行った。鳴き声の確認には、音声ファイル（WAV 形式）を図 6 に示した音声編集ソフト Audacity を用いた。今回は 1 種類の鳥が一羽で鳴いたときのデータをもとに分類することを目標とし、鳴き声の確認できた場合は、手動で切り取り作業を行った。図 7 は切り取り作業の手順を示している。赤い枠線が示す箇所を鳥の鳴き声として保存する場合、選択ツールを用いて不要範囲の切り取りを手動で行った。前後間隔はできるだけ短くするものとし、0.5 秒から 1 秒程に切り取ったものを WAV 形式で書き出した。

4.2 鳴き声の区切り方

今回対象としたヒヨドリとキジバトは鳴き方のそれぞれ特徴がある。データの区切り方をそれぞれ示す。まず図 8 にヒヨドリの場合の区切り方を示す。ヒヨドリは短い鳴き声を連続させるため、1 鳴

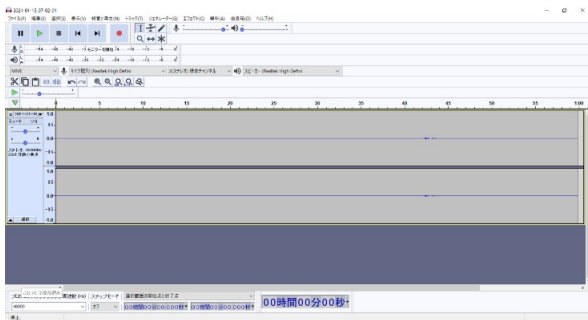


図 6. 音声編集ソフト Audacity の再生画面

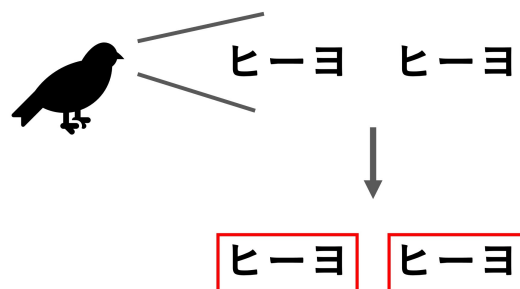


図 8. ヒヨドリの区切り方

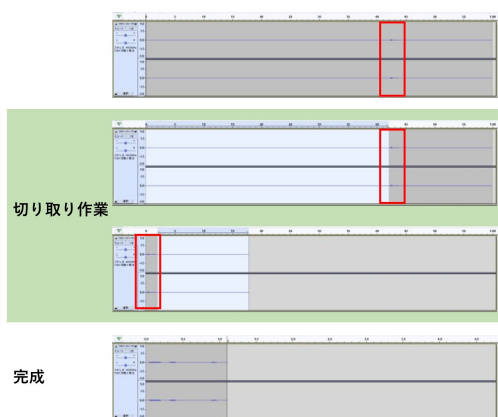


図 7. 切り取り作業の手順

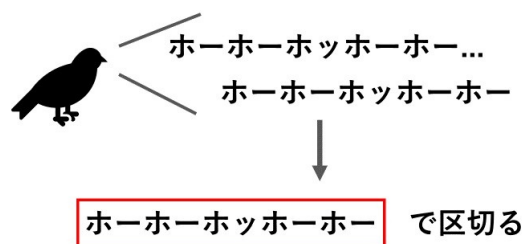


図 9. キジバトの区切り方

きごとにトリミングした。次に図9にキジバトの場合の区切り方を示す。キジバトは1フレーズを繰り返しながら鳴き続けるため、作業するもの同士で1フレーズの長さを共有し、作業を行った。

4.3 学習用データセット

学習用データセットは2020年9月19日の録音データと2020年12月2日から2020年12月9日までの録音データを用いて作成した。録音データを聞いて、時間を問わずヒヨドリとキジバトが鳴いている箇所の切り取りを行った。データ数を表1に示す。

切り取りの対象は確実に聞こえるものとし、鳴き声が音声編集ソフト Audacity 上で波形に表れているもの、波形の有無は関係なく音がはっきりと聞こえるものを収集した。

4.4 評価用データセット

評価用データセットの作成には2021年1月13日から2021年1月15日までの録音データを使用した。

野鳥の頻出時間帯は先行研究^{1,2)}から日の出前後にであることが明らかにされているため、日の出10分前から日の出10分後の20分間に限定して切

り取りを行った。データ数は表2に示す。切り取りの対象は、音がはっきり聞こえるものに加え、かすかに聞こえたものを収集した。

5. 鳥類の音声分類実験

作成したデータセットを使用し、Support Vector Machine やメルケプストラム係数を用いて2種類の鳥を判別した。本章ではSupport Vector Machine、メルケプストラム係数について解説した上で、実験環境、実験方法、結果、考察について述べる。

5.1 Support Vector Machine

Support Vector Machine (以下SVMと呼ぶ)とは、教師有りの機械学習アルゴリズムの一つである。この手法は、学習データの特徴量空間に対して識別境界と呼ばれる境界を引くことで分類を行うものである。この識別境界は式(1)で表される。

$$y = \text{sign} \left(\sum_{i=1}^N y_i w_i K(x, x_i) + b \right) \quad (1)$$

x は評価したいデータの特徴量、 x_i は N 個ある学習データの特徴量、 y_i は特徴量のクラスラベル、 w_i を x_i に対する重み、 b はバイアス項、 $K(x, x_i)$ はカーネル関数を示している。この識別境界はマージ

表 1. 学習用データセットのデータ数

日付	ヒヨドリ	キジバト
2020/9/19	0	22
2020/12/2	0	0
2020/12/3	0	0
2020/12/4	0	0
2020/12/5	0	0
2020/12/6	85	13
2020/12/7	92	0
2020/12/8	52	7
2020/12/9	50	21
計	279	63

表 2. 評価用データセットのデータ数

日付	ヒヨドリ	キジバト
2021/1/13	10	11
2021/1/14	16	14
2021/1/15	8	10
計	34	35

ンの最大化と呼ばれる識別境界に最も近いデータまでの距離が最大になるように識別境界を決める手法によって導き出される。これは式 (2) で表される目的関数を最小化することで求められる。

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \zeta_i \quad (2)$$

式 (2) における w は w_i から構成される重みベクトルで C は分類誤りを考慮するかの指標であるコストパラメータ、 ζ_i は分類誤りを起こしたデータが識別境界を超過した距離であるスラック変数である。

また、式 (1) におけるカーネル関数は非線形の識別境界を決定するために使用されるものである。代表的なカーネル関数として式 (3) の RBF カーネルがあげられる。

$$K(x, x_i) = \exp(-\gamma \|x - x_i\|^2) \quad (3)$$

式 (3) における γ は識別境界の複雑さを決めるパラメータである。

RBFカーネルにおいては、式 (2) のコストパラメータ C と式 (3) の γ を手で決める必要があるが、このパラメータを最適化する方法として、グリッドサーチなどの手法があり、パラメータを調節する際によく用いられている。グリッドサーチは、パラメータ

の組み合わせに対して総当たりで精度を検証することで最適なパラメータを求める方法である。

5.2 メルケプストラム係数

メルケプストラム係数とは、人間の聴覚特性を考慮した音の特徴量である。人の聴覚において重要な低周波成分を引き延ばして特徴量としており、音声認識などで使用されることが多い。メルケプストラム係数は以下の手順で求められる。

手順 1 フーリエ変換による周波数軸への変換

手順 2 メル周波数スペクトルグラムへの変換

手順 3 メルケプストラム係数への変換

手順 1 では、短期間フーリエ変換を用いて音の振幅を周波数軸に変換する。

手順 2 では、メル尺度³⁾と呼ばれる指標を基にしたメルフィルタバンクを利用してメル周波数スペクトルグラムに変換する。メル尺度は 1000Hz の音を 1000 メルと定義し、この音の 2 倍高く聴こえる音を 2000 メル、半分に聴こえる音を 500 メルとして定義したものである。メル尺度は以下の式 (4) で表される。

$$m = m_0 \log \log \left(\frac{f}{f_0} + 1 \right) \quad (4)$$

式 (4) における f_0 と m_0 は 1000Hz は 1000 メルという制約を基にしたパラメータであり、次の式 (5) を用いて求められる。

$$m_0 = \frac{1000}{\log \left(\frac{1000\text{Hz}}{f_0} + 1 \right)} \quad (5)$$

f_0 と m_0 は、 $f_0 = 700$ 、 $m_0 = 2595$ などが用いられている。

手順 3 では、メル周波数スペクトルグラムを離散コサイン変換し、ローパスフィルタによって低次元成分を取り出す処理を行う。これによってメルケプストラム係数を求めることができる。

5.3 実験環境

今回は SVM とメルケプストラム係数による鳥類データの学習と分類について音声処理ライブラリ `librosa` や、機械学習ライブラリ `scikit-learn` を用いて実装した。基盤となるプログラム言語は `python` を用いた。言語のバージョンなどは以下の通りである。

- python: 3.7.9
- scikit-learn: 0.23.2
- librosa: 0.8.0

今回は、ヒヨドリとキジバトの 2 種類の鳥類のみを分類することを想定している。これらの 2 種類の鳥類のデータをメルケプストラム係数に変換し、時

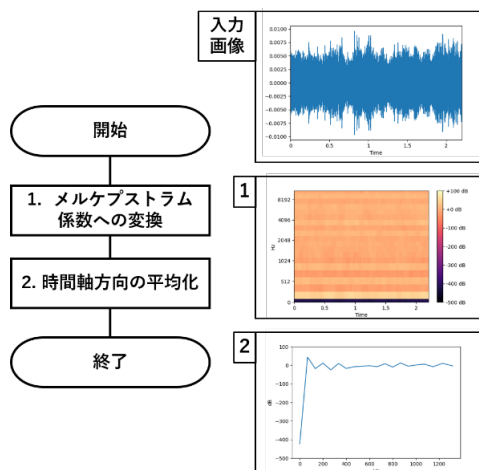


図 10. 特徴量への変換

時間軸方向に平均化した後、これを特徴量として学習と分類を行った。この特徴量の変換の流れを図 10 に示す。

特徴量への変換では、振幅で表される音声データを周波数軸のメルケプストラム係数に変換した。この変換は `librosa` の `feature.mfcc` 関数を用いた。関数には短時間フーリエにおける窓のずらし幅 512 と音声のサンプリングレートを与えた。ここでは、短時間フーリエ変換を基にメルケプストラム係数を求めるため、短時間フーリエ変換における窓のずらし幅に応じて短時間の間にいくつかのメルケプストラム係数が求められることになる。このため、時間軸方向に対して平均化を行うことで、最終的に一つの音声データに対して、20次元の特徴量を得ることになる。これを特徴量として学習と分類に用いた。

学習と分類に使用するカーネルは、非線形カーネルの一種である RBF カーネルを使用した。これは、`scikit-learn` における `svm.SVC` 関数を用いた。今回は使用するヒヨドリとキジバトのデータの数に偏りがあるため、SVM の識別境界を決定するにあたってクラス内で重みづけを行った。この重みはデータ数 / (クラス数 × クラスのデータの数) で求められる値を基に決定した。

また、SVM のハイパーパラメータについては、グリッドサーチを用いて最適化を行った。今回はグリッドサーチにおける交差検証の回数を 5 とし、正答率で評価した。指定したハイパーパラメータは C と γ で、 10^{-10} から 10^{10} までの値を 10 倍ずつ手動で与えて正答率を観察しながらパラメータを調節した。

5.4 実験方法

今回実験に使用するデータは 4.3 節におけるヒヨドリ 279、キジバト 63 の学習用データ、4.4 節における 1 月 13 日から 1 月 15 日までの日の出時間前後のヒヨドリ 34、キジバト 35 の評価用データを使用した。

事前実験では、web 上のサイトのデータを使用して 12 種類、全 626 のデータから学習モデルを作成し、正答率による精度評価を行ったが、撮影環境のばらつきや、環境音、他の鳥の鳴き声の影響が大きくなり、過学習を引き起こすことがあったため、今回は本報告において提案した録音システムによる録音データを使用して精度評価を行った。

実験でははじめに、学習用データを用いて最良パラメータの評価を行った。ここでは、4.3 節における学習用データを更に学習用 75 %、評価用 25 % のデータセットに分けて、グリッドサーチにおけるパラメータの正答率の評価を行った。その後、ここで最も良かったハイパーパラメータと学習用データを用いて学習させたモデルを使用して、4.4 節の評価用データによる分類の精度評価を行った。

5.5 結果

グリッドサーチによる最良パラメータの精度評価では、はじめに `svm.SVC` 関数におけるデフォルトパラメータ $C = 1$ 、 $\gamma = \text{'scale'}$ の値の正答率を評価した。 $\gamma = \text{'scale'}$ とは $1 / (\text{データ数} \times \text{データの分散})$ から求められる値である。4.3 節における学習用データを更に学習用 75 %、評価用 25 % のデータセットに分けて、グリッドサーチにおけるパラメータの正答率の評価を行ったところ、このパラメータにおける学習用データの正答率は 87.89 %、評価用データの正答率は 88.372 % であった。これに対して、グリッドサーチによって調節した最良パラメータ $C = 10$ 、 $\gamma = 0.01$ においては、学習用データ、評価用データともに正答率が 100 % となった。

この最良のパラメータ $C = 10$ 、 $\gamma = 0.01$ と 4.4 節の学習用データを用いて学習後、4.4 節の評価用データでヒヨドリとキジバトの二値分類を行った。二値分類におけるクラスは正例と負例と呼ばれるクラスに分けられるが、今回は正例をヒヨドリ、負例をキジバトとして分類している。この分類結果を表 3 に示す。

ここでいう正答率とは、評価データにおける予測全てにおいて、実際に正解であったものの確率を指している。適合率は評価データにおける予測にお

表3. 評価用データにおける二値分類結果

日付	正答率	適合率	再現率	F 値
2021/1/13	47.619 %	47.619 %	100 %	64.516 %
2021/1/14	53.333 %	53.333 %	100 %	69.565 %
2021/1/15	44.444 %	44.444 %	100 %	61.538 %
全体	49.275 %	49.275 %	100 %	66.019 %

いて正例と予測したもののうち、実際に正例であったものを示している。今回は正例をヒヨドリとしているため、ヒヨドリと予測したもののうち、実際にヒヨドリであった確率を示している。再現率とは、実際に正例であったもののうち、正例と予測されたものの割合である。F 値は適合率と再現率両方の値を考慮した指標で、適合率と再現率の調和平均をとることで求められる。

全体として再現率は 100 %なのに対して、正答率は 50 %前後という結果となった。また、正答率と再現率の値が同じという結果になった。

5.6 考察

最良のパラメータ $C = 10$ 、 $\gamma = 0.01$ と 4.3 節の学習用データを用いて学習後、4.4 節の評価用データでヒヨドリとキジバトの二値分類を行った結果、再現率は 100 %に対して正答率が 50 %前後になった。また、正答率と再現率の値が同じという結果になった。これは、評価データにおける分類結果がヒヨドリとキジバトにかかわらず、全て正例であるヒヨドリと分類されているためであると考えられる。

キジバトの分類が全て失敗している原因としてはデータの質の問題があげられる。今回使用したキジバトのデータについてはキジバトが遠くで鳴いているものが多く、キジバトの鳴き声が小さく分類ができなかったと考えている。

分類の失敗については、その他に考えられる原因として特徴量にメルケプストラム係数を時間軸で平均化している点があげられる。今回取得したメルケプストラム係数は短時間フーリエ変換により時間軸方向に複数の値がでてくるため、これを平均化して特徴量としているが、鳥はデータ上のどのタイミングで鳴いているかわからないため、時間軸で平均化した処理は妥当ではない可能性がある。

6. 現在把握している改善点

6.1 録音装置の開発

周りに雨などを防ぐ建物がない場合について、雨

への耐性や周りの環境に配慮した、ハウジングの形状や素材を考える必要がある。また、マイクについて、今回のものと指向性の異なるものや性能の良いものを用いることで、ノイズの低減につながる可能性がある。

6.2 録音実験

今回の録音実験では、ボトルクレートに固定した設置方法であったが、電源供給が困難な場所の場合、木などその場に存在する物への固定方法など別途考えなければならない。

6.3 データセットの作成

今後の課題点としてデータの質を均一にすることが挙げられる。これを改善するためには切り取り方法の見直しが必要である。今回は複数人の手作業でデータの切り取りを行い、前後間隔についても鳴き始めまでの時間をできるだけ短くしていた。

今後、鳴き声が波形に現れている場合は切り取りの際に鳴き始めの定義を細かく決めることでデータを均一にすることができる。しかし、今回は波形に現れていないものも切り取りの対象としているため、再度方法を検討する必要がある。

6.4 鳥類の分類実験

今回の鳥類の分類における実験の課題点として、メルケプストラム係数を時間軸で平均化している点が挙げられる。本報告では、メルケプストラム係数における鳥の鳴き声の分類の妥当性についても検証が不十分である。このため、今後はメルケプストラム係数における特徴量の抽出方法の見直しと、その特徴量における鳥の鳴き声の分類の妥当性の検証を行う必要がある。また、メルケプストラム係数だけではなく、短時間フーリエ変換などの他の指標を使用して学習と分類を行いたいと考えている。

7. まとめ

本研究では生態系調査のため、生態系の上位にある鳥類に注目し、鳥の鳴き声を録音し、その鳴き声から分類を行うシステムの作成を目的とした。そして、録音装置の開発、録音実験、データセットの作成、鳥類の分類を行った。

録音装置は、ハウジングと内枠ケースの2つの構造で内部の機材を守ることによって野外でも録音可能な装置を作成した。しかし古民家の軒下に設置するなど装置以外に雨を防ぐ建物がある前提で作成したため、そのような建物がなくても設置できる

ようハウジングの素材や形状を再度検討する必要がある。また、作成した録音装置を利用して、2020年9月から2021年1月までの特定の期間において録音実験を行い、得られたデータから学習用データとしてヒヨドリ279個、キジバト63個を作成した。また、評価用データとしてヒヨドリ34個、キジバト35個のデータを作成した。これらのデータは、録音実験で得られたデータを基に複数人が手作業で鳥の鳴き声を切り取って作成したものであるが、切り取りのタイミングを明確に決めていないなど、データの均一性に問題があった。このため、データセット作成の際に切り取りの定義を明確にすることによりデータの均一化を図る必要がある。

鳥類の分類実験では、作成した学習用データセットと評価用データセットから Support Vector Machine を用いて分類を行ったところ、結果として50%前後の正答率が得られた。今回の分類実験では、キジバトとヒヨドリの2クラス分類となっているため、これらの鳥以外の多クラス分類を行う方法について検討する必要がある。また、今回はデータセットを特徴量に変換する際に、メルケプストラム係数を時間軸で平均化したものを使用しているが、鳥はデータ上のどのタイミングで鳴いているかわからないため、この方法は適さない可能性がある。このため、特徴量の抽出方法の見直しと妥当性の検証を行う必要があると考える。

文献

- 1) 由井正敏, “森林原野生鳥類のラインセンサス法の研究:I 記録率の日周変化”, 山階鳥研報 vol. 10, no. 1-2, pp. 70-81, 1978.
- 2) 植田睦之, 平野敏明, 黒沢令子, “長時間の録音データから鳥のさえずり状況を知るための聞き取り時間帯の検討”, Bird Reserch, vol. 8, pp. T1-T6, 2012.
- 3) Stevens, Stanley Smith, John Volkman, and Edwin B. Newman. “A scale for the measurement of the psychological magnitude pitch.”, The Journal of the Acoustical Society of America, vol. 8, no. 3, pp. 185-190, 1937.

Prototype of a Bird Voice Recording Device and Automatic Species Counting System

Iori Kubota¹ Kazuki Maruyama²

Hibiki Nakashima² Kohei Yamasaki³ Ryosuke Murai^{4*}

(Received: May 31th, 2021)

¹Informatics Course, Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami City, Kochi 782–8502, JAPAN

²Infrastructure Systems Engineering Course, Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami City, Kochi 782–8502, JAPAN

³Aerospace Engineering Course, Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami City, Kochi 782–8502, JAPAN

⁴Research Organization for Regional Alliances, Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami City, Kochi 782–8502, JAPAN

* E-mail: murai.ryosuke@kochi-tech.ac.jp

Abstract: This study focused on birds at the apex predators of the ecosystem for the purpose of investigating the ecosystem. A prototyping of a bird voice recording device and an automatic species counting system were created. The recording device has a dual structure for outdoor use with a housing and inner case. To verify the accuracy of bird species discrimination, a classification experiment was conducted by creating learning data and evaluation data from the voice recorded by the recording device. The classification targets were Eastern turtle dove and Brown-eared bulbul, whose voice are relatively easy to understand. The classification was performed using the Support Vector Machine with the Mel-frequency cepstral coefficient (MFCC) as the feature quantity. The result was a correct answer rate of around 50%.