

スマート畜産に向けた牛の鳴き声と乗駕行動による発情状態推定

北村 知寛† 児嶋 朋貴‡ 佐藤 精‡ 森下 忠‡ 入部 百合絵†

愛知県立大学† 愛知県農業総合試験場‡

1. はじめに

近年、畜産経営の大規模化が進み、少人数で多頭数を管理する技術が求められている。そのような背景の下、ICT を利用した家畜モニタリング技術が提案されており、特に発情状態の検知に用いられている[1]。検知には、機器の装着による家畜への負担や飼養頭数の増加による飼養者の負担から非接触な技術が求められている。

非接触による牛の発情状態推定に関する先行研究には、鳴き声情報や画像情報を用いた研究が挙げられる。Chung ら[2]は牛の鳴き声からMFCC(Mel-Frequency Cepstrum Coefficients)とその微分値である Δ MFCC, $\Delta\Delta$ MFCCの合計360次元抽出し、発情状態推定を行っていた。一方、川野ら[3]は牛の発情兆候の一つである、乗駕行動を検出するための画像データセットを構築し、適合率80%、再現率76%の精度を得た。これら先行研究は鳴き声情報あるいは画像情報により、発情状態を検出しているが、牛の個体によって発情状態・行動の表出は異なる。そのため、発情期に乗駕行動が観察されない、あるいは鳴き声を上げないことも考えられ、生産現場では発情行動による違いが見落としの要因となり得る。

本研究では牛の鳴き声情報と画像情報を用いた発情状態の推定手法を提案する。鳴き声情報からは分析窓を適応させた韻律や声道パラメータによる変化量を基にした発情推定モデルを、画像情報からは乗駕行動を深層学習によるモデルを構築した。

2. 牛の鳴き声と乗駕行動の収集

本研究では、愛知県長久手市にある愛知県農業総合試験場で飼育されているホルスタイン種・黒毛和種の合計23頭を対象に音声と動画画像を収集した。表1と表2に収集した音声と画像のデータ数をそれぞれ示す。牛は状態毎に鳴き声特性が異なる可能性があるため、鳴き声を空腹状態、発情状態、帰舎時の3つに分けて分析した。牛が単独で鳴いている音声を中心に牛毎に音声を切り出すとともに、改良スペクトル引き算法を用いてノイズを除去した。また、画像につい

表1 推定に使用した鳴き声の内訳

牛の状態	空腹	発情	帰舎時	合計
発声数	628	327	469	1424

表2 識別に使用した画像の内訳

項目	乗駕行動	1頭	複数頭
画像枚数	426	426	426

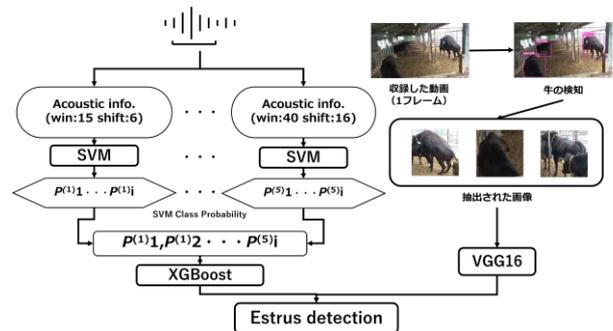


図1 本研究における発情状態推定の概要図

ては、収録した動画から1枚/秒毎に抽出した。なお、動画から切り出した画像サイズは均一でないため224×224サイズにリサイズした。また、識別に使用する画像枚数が等しくなるように最も少ないクラスの画像枚数に合わせてランダムに抽出した。次に動画を目視することで、各音声と画像に対して発情/非発情の情報をアノテーションした。本研究では、発情検知補助機器と直腸検査による判定結果に基づいて発情/非発情をラベル付けした。

3. 鳴き声と乗駕行動による発情状態推定

3.1 推定手法の概要

本研究では鳴き声識別モデルと乗駕行動識別モデルの識別結果に基づいて最終的に発情状態を推定する。なお、生産現場では発情状態の推定結果を一定の間隔で通知する機能が求められている。そのため、1時間毎に発情状態を検知させる手法を試みた。ただし、本研究では1発声あるいは1枚の画像毎に発情か否かを識別させたため、単位時間あたりに複数の識別結果が算出されることになる。そのため、予備実験より、単位時間あたりに発情と判定された発声/画像が60%以上集まった場合に最終的に発情と判断する。

図1に本研究における発情状態推定の概要図を示す。次節以降に各モデルについて詳述する。

3.2 牛の鳴き声による発情状態推定

発情はホルモンの影響による感情変化が予想されるため、感情との相関が高い音響情報を選定した。10次元の韻律情報、音源由来のパラメータとして、HNR (Harmonic to Noise Ratio) (平均, 最小, 最大), jitter, shimmer などの14次元, 声道由来のパラメータとして MFCC とその動的特徴量である Δ MFCC, $\Delta\Delta$ MFCC の各平均/標準偏差の72次元を抽出した。さらに、高次のフォルマントは感情情報を含むと考えられるため、第1~第8フォルマント周波数を抽出した。

一方、人間の音声は感情によって必要な時間解像度が変化すると報告されている[4]。本研究では牛のホルモン変化による感情に注目したが、牛も人と同様の傾向を有している可能性がある。本研究では、空腹、発情、帰舎時の3つの異なる状態を対象としているため、異なる分析窓長に分けて各特徴量を抽出した。設定した窓長は15ms, 20ms, 25ms, 30ms, 40ms, シフト長はそれぞれ6ms, 8ms, 10ms, 12ms, 16msである。本研究では窓長毎に特徴量を抽出するため、識別には、スタッキング手法によるアンサンブル学習を採用した。図1に示すように、窓長毎の特徴量を複数のSVM(Support Vector Machine)に入力し、その出力値をXGBoost(eXtreme Gradient Boosting)に入力し、発声毎に発情有無を出力する。

3.3 牛の画像による乗駕行動識別

乗駕行動を識別するにあたって、まず画像分類を実施した。動画から抽出した牛の画像を観察したところ、発情期行動である乗駕行動の画像、単独の画像、複数頭が密になっている画像の3種類に大別されることが分かった。そのため、画像毎に3種類のいずれかをラベリングした(表2)。これらの画像を用いて深層学習の一つであるVGG-16[5]によるクラス識別を行った。学習・テスト・バリデーションの比率は8:1:1である。

4. 識別結果と考察

本章では識別結果について述べる。評価方法はいずれも10分割交差検証である。鳴き声モデルと乗駕行動モデルを統合した結果を表3に示す。1時間毎に識別結果を出力させたところ、発情状態を正しく発情状態と推定した値 Sensitivity が97.5%となった。牛は個体によって発情状態・行動の表出が異なるため、本研究では鳴き声情報と画像情報を補完する発情推定手法を提案した。これにより発情行動の違いによる生産現場での見落としの減少が期待できる。しかし、一部誤識別が生じていたため、牛の鳴き声モデルと乗駕行動モデルの各々の精度を検証した。

表3 鳴き声モデルと乗駕行動モデルを統合した単位時間毎の発情状態の推定結果

Sensitivity	予測識別結果	
	発情	非発情
97.5%	39	1

表4 1発声毎の状態識別の結果

分析窓長	Accuracy	Sensitivity
窓長変更	88.2%	99.6%
窓長固定	82.4%	94.1%

表5 1画像毎の乗駕行動識別の結果

項目	Accuracy	Sensitivity
乗駕行動識別	71.9%	81.9%

牛の1発声当たりの識別結果を表4に示す。本研究では異なる窓長を採用したため、窓長を変更/固定した場合の比較結果を示す。窓長を変更した場合、Accuracy が88.2%、Sensitivity は99.6%であり、窓長固定よりも精度が向上した。この結果から、牛の鳴き声も感情によって必要な時間解像度が異なる可能性がある。また、発情状態の識別に有用な音響情報を確認するためSteel-Dwass 検定を実施した。有意差の認められた音響情報の一つである、HNR は音声の周期性の度合いを示しているが、他の動物と同様に牛の発情時もHNRは低く、標準偏差が小さいことが分かった。次に牛の1画像当たりの乗駕行動の識別結果を表5に示す。Accuracy が71.9%、Sensitivity が81.9%であり、鳴き声と比較すると、精度の低い結果となった。本研究は単位時間あたりの識別を目指しているため、画像毎の識別ではなく、一定の時間内における前後動作を考慮した画像の識別手法を検討していく。

5. おわりに

本研究では、牛の鳴き声を用いた状態識別による発情状態推定と牛の画像を用いた乗駕行動識別により発情状態推定を行った。その結果、鳴き声による状態識別ではAccuracy が88.2%となり、画像による乗駕行動識別では81.9%の精度を得た。発情状態推定ではSensitivity が97.5%となった。今後の課題としては、データの拡充を行った上での識別精度の向上が必要である。

参考文献

- [1] 坂口, 日本畜産学会報 81(4), pp.413-419, 2000
- [2] Y.Chung et al, Asian-Australasian journal of animal sciences 26.7, pp.1030-1037, 2013.
- [3] 川野百合子 et al, 人工知能学会全国大会論文集, 第34回(2020), pp.1N5GS1305-1N5GS1305
- [4] 藤村浩司, 電子情報通信学会論文誌 D 105.3 (2022): 154-166.
- [5] K. Simonyan et al, "Very Deep Convolutional Networks for Large-Scale Image Recognition", In Intl. Conf. on Learning Representations (ICLR), 2015