Key feature identification of swarming in beehives, using audio fingerprint

Bányai Dóra$^{[W5H8NT]}$ and Dániel Várkonyi as Supervisor

Faculty of Informatics, Eötvös Loránd Tudományegyetem

Abstract. This paper presents a method for processing audio recordings taken from beehives in order to analyze bee behavior. The approach involves applying feature extraction techniques on the audio signals such as zero crossing rate, spectral features, etc. The extracted features are then used as inputs for a genetic algorithm to identify patterns related to swarming.The goal is to gain a deeper understanding of bee behavior and potentially identify early warning signs of issues in the hive using the genetic algorithm optimization.

Keywords: Audio Signal Processing · Bee · Swarming · Feature Extraction · Genetic Algorithm

1 Aim of the project

The aim of this project is to develop a method for identifying swarming times in beehives using audio recordings. This project is being conducted as part of the Data Science Lab II at Eötvös Loránd University Budapest supervised by Dániel Várkonyi. The current stage of the project is a beta version, which will be further developed and expanded in the upcoming semester as part of a thesis project.

2 Literature search

2.1 Bee swarming

Bee swarming is a natural process in which a colony of bees splits into two or more parts, with one part staying in the original hive and the other part leaving to form a new colony. This process is important for the reproduction and survival of honeybee populations. According to studies, swarming is the primary means of reproduction for honeybee colonies and plays a significant role in the genetic diversity and adaptability of honeybee populations [1]

Swarming can also have negative effects on honey production, as the loss of a significant portion of the bees in a colony can reduce the number of foragers and the overall productivity of the hive [2] Therefore, identifying and monitoring the swarming times in a beehive can help beekeepers to better manage their colonies and improve honey production.

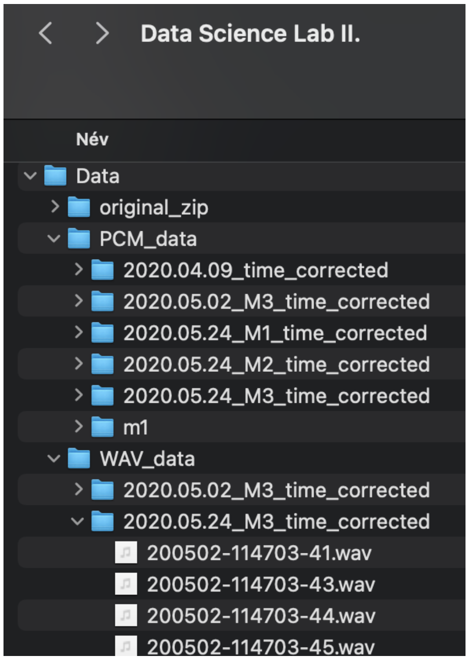
2.2 Audio Signal Processing

One approach to identifying swarming times is to observe the behavior of the bees directly, but this can be difficult and time-consuming. An alternative approach is to use audio recordings from the beehive. The sounds produced by bees during swarming, such as the buzz of the bees and the queen's mating call, are unique and can be used to identify swarming times. However, the raw audio recordings from beehives are often contaminated by background noise and other sounds, making it difficult to extract the relevant information.

This is where audio signal processing comes in. Audio signal processing is a field of study that deals with the analysis, synthesis, and manipulation of audio signals. [3] It provides a range of techniques and tools that can be used to extract relevant information from audio recordings, such as filtering to remove background noise, feature extraction to identify relevant patterns, and machine learning or genetic algorithm optimization to analyze the data and identify swarming times.

3 Architecture

3.1 Folder structure



```
Data
  original_zip
  PCM_data
    2020.04.09_time_corrected
    2020.05.02_M3_time_corrected
    2020.05.24_M1_time_corrected
    2020.05.24_M2_time_corrected
    2020.05.24_M3_time_corrected
    m1
  WAV_data
    2020.05.02_M3_time_corrected
    2020.05.24_M3_time_corrected
      200502-114703-41.wav
      200502-114703-43.wav
      200502-114703-44.wav
      200502-114703-45.wav
```

Fig. 1. Folder structure used in the project

The Data folder of the project is organized to contain audio recordings in both PCM and WAV formats. The subfolders within the data folder are clearly labeled to indicate the format of the audio files contained within them. Specifically, the `PCM_data` subfolder contains raw PCM audio recordings, which are organized into separate folders. These raw recordings are then processed using a code to convert them into WAV format and are stored in the `WAV_data` subfolder.

3.2 Filename exploration

To facilitate the analysis of the input PCM audio recordings, a notebook named `filename_exploration` was created. This notebook allows for the examination of the filenames of all possible PCM files in a single system. Initially, the focus of this examination was on identifying the specific times of swarming as provided in `labels.xlsx`. However, it was later discovered that this information was not included among the audio files themselves. As a result, the identification of swarming times relies on identifying the audio file closest in time to the reported swarming event.

3.3 Bee Data Class

This class provides the ability to process and convert sound files, as well as perform feature extraction. The class utilizes two methods of initialization: the option to load pre-processed data from `pickle` files or to specify a folder for the conversion of audio files and feature extraction. The latter option allows for the subsequent storage of the processed data in `pickle` files. To fully understand the capabilities of the class, will be explained for each key function in detail.

3.4 convert files from folder to wav

The proposed class utilizes a specified folder as a parameter for processing. Within this folder, PCM files are systematically converted to WAV format and subsequently stored within a designated `WAV_data` folder for further use. The execution of this function is contingent upon the specification within the class' initialization function. This is due to the possibility of pre-existing WAV files and the desire to avoid unnecessary duplication of file creation.

3.5 feature extraction

It performs feature extraction on each audio file present in the designated folder. It calculates the mean and variance of all extracted features.

rms Root Mean Square (RMS) is a measure of the amplitude of a sound signal.

stft Short-Time Fourier Transform (STFT) is a time-frequency analysis method that is used to analyze the frequency content of a signal over time.

spectral centroid The spectral centroid is a measure of the center of mass of a signal's frequency spectrum. It represents the frequency at which the energy of a sound is centered and is calculated by taking the weighted mean of the frequencies present in a signal.

spectral bandwidth Bandwidth is a measure of the range of frequencies in a continuous band of frequencies. It is defined as the difference between the upper and lower frequencies. For a signal that oscillates around a centroid, the bandwidth can be calculated as the sum of the maximum deviations of the signal on both sides of the centroid at a specific time frame.

spectral rolloff Roll-off is a term used to describe the action of a particular type of filter that gradually reduces the amplitude of frequencies outside a specified range. This process is known as roll-off because it is not abrupt but rather a gradual decrease in amplitude.

spectral contrast In audio signals, the spectral contrast is a measure of the variation in energy across different frequencies at each timestamp. As many audio files contain frequencies whose energy changes over time, it can be challenging to accurately measure this energy level. Spectral contrast provides a method for measuring this variation in energy, which is important for understanding the characteristics of an audio signal.

spectral flatness Spectral flatness is a measure of the "peakiness" of a sound's power spectrum.

zero crossing rate The zero crossing rate is a measure of the number of times the signal changes from positive to negative or negative to positive within a given time frame.

harmonic Harmonic refers to the relationship between different pitches in music, specifically when they are integer multiples of a common fundamental frequency. Harmonics are also known as overtones and they are heard as a pleasant and rich sound.

mfcc Mel-Frequency Cepstral Coefficients (MFCCs) are a set of coefficients that represent the spectral envelope of a sound in a more perceptually meaningful way. MFCCs are derived from the logarithmically scaled magnitude of the short-term Fourier transform of a sound.

tonnetz Tonnetz is a geometrical representation of the harmonic relations between different pitches in music.

3.6 give swarming score

This function was designed for use in the genetic algorithm's optimization process. Upon examination of the filenames, it was determined that no audio materials pertained specifically to the exact swarming times. As such, a scoring system was implemented to determine the proximity of each audio material to a swarming time point. The function calculates the closest swarming time for each audio's time stamp, and assigns a score accordingly. If the audio was recorded prior to the swarming, points are awarded based on the minutes between the audio time and the swarming time. Conversely, if the audio was recorded after the swarming, points are multiplied by a penalty factor. This approach is necessary to allow the algorithm to learn and optimize for the period leading up to the bee swarming.

3.7 create df from features

This function creates a dataframe by combining the normalized features, the swarming score, and the date and time information obtained from the file name.

3.8 save df to pickle

The processed dataframe can be saved for later use, allowing for the avoidance of time-consuming feature extraction processes in the future.

3.9 visualize audio

The proposed method utilizes a sound file, specified as an input parameter, to generate a set of 12 visual representations. These visualizations include the original audio plot, as well as spectrograms and chromagrams of the audio's features, and waveplot representations. These diagrams provide a comprehensive visual representation of the audio file. An example of the visualizations can be found in the `visualization.ipynb` file.

3.10 Genetic Algorithm Class

The presented class includes functionalities for executing a genetic algorithm, including selection, crossover, and mutation operations. It implements the standard procedures of a genetic algorithm, enabling its application to a variety of optimization problems. The only exception, which is a bit more specific to the `Bee Data Class`, is the fitness function.

Initialization of Population Population initialization refers to the process of creating the initial set of individuals, also known as the population, that will be used in a genetic algorithm. It is an essential step in the algorithm as it sets the starting point for the evolutionary process.

Fitness Function The fitness function is a crucial component of a genetic algorithm as it evaluates the quality of the solutions (individuals) in the population. It assigns a value to each individual based on how well it solves the problem at hand. The fitness function plays a key role in the selection process, as it allows the genetic algorithm to identify the best individuals to use as parents for the next generation. In this study, the fitness function is based on the calculation of the mean squared error (MSE) between the individual value and all the feature values of the sound files. The function returns the swarming score of the sound file with the smallest MSE and multiply it with the MSE score. The smaller this score, the better the resemblance of the sound file to the individual values, particularly in relation to the time of the swarming.

Selection In this implementation of the genetic algorithm, roulette wheel selection is employed as the strategy for selecting individuals from the population for reproduction. This method assigns a probability of selection to each individual based on their fitness, where individuals with higher fitness have a higher probability of being chosen to propagate their genetic information to the next generation.

Crossover Crossover combines the genetic information of two individuals to create offspring with novel characteristics. In this implementation, a random point is selected in the genetic sequence of the individuals for the crossover process. At this point, the genetic information between the two individuals is swapped to create new combinations of genetic information.

Mutation Mutation, as a genetic operator, is applied in order to introduce random variations into the genetic information of individuals in the population. With a probability parameter set, mutation is applied on individuals, where two positions in the individual's genetic sequence are randomly selected and changed, in order to introduce diversity and explore new solutions in the population.

When to stop Three types of stop conditions have been implemented in the genetic algorithm to ensure its efficient and timely termination. The first kind of stop condition, triggered when the fitness limit set in the parameters is reached, the second stop condition, triggered when the maximum number of iterations is reached, and the third stop condition, which is optional, is triggered when a specified number of generations in which the best score has not changed, as defined in the parameters, is reached.

4 Result

The genetic algorithm was utilized to evolve a population of individuals towards a solution that minimizes a predefined fitness function. The result of this process is an individual that was found to be the most optimal among the population during the evolution process. The resulting individual can be used to predict swarming behavior and to find individuals that closely approximate the results of a swarming. However, it should be noted that the fitness function used in the algorithm may have played a role in several early stoppings of the algorithm. To mitigate this issue, it may be necessary to re-evaluate and potentially modify the function. Since, in testing, it was found that increasing the probability of mutation and / or expanding the population size did not lead to better results.

5 Future work

In future work, the feature parameters can be further developed and fine-tuned to improve performance. Furthermore, exploring the extraction of additional features could enhance the algorithm's capability. The visualization aspect of the algorithm can also be improved by developing specific diagrams for a given period of time, which can provide greater insight into the results. Additionally, the fitness function can be reevaluated to explore new ideas that could potentially bring more efficient and effective results.

5.1 Git repository

The source code of this project can be found here: git repository

References

1. Rinderer, T.E., de Guzman, L.I., and Stelzer, J.A. (1990) Reproduction in honey bee colonies. Apidologie 21: 191-204
2. Otis, G.W., Danka, R.G., and Harris, J.W. (1996) Honey bee colony strength and swarming. Journal of Apicultural Research 35: 11-19.
3. Audio Signal Processing- Understanding Digital Analog Audio Signal Processing, https://www.pathpartnertech.com/