U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

FCED

Bird Call Feature Extraction and Analysis Using Signal Processing

Master in Data Science and Engineering
1º Semester - 2024/2025

Tiago Beaulieu – up202400228

João Viterbo Vieira – up202107689

João Soares – up202105364

Index

Index 2
1. Introduction 3
1.1 Goals 3
2. Data acquisition 3
3. Signal Preprocessing 4
3.1 Noise Reduction 4
3.2 Audio Segmentation: Temporal Centroid 5
3.3 Bandpass Butterworth Filter 5
4. Feature Extraction 6
4.1 Fourier Transform 6
4.2 Time Domain 7
4.2.1 Zero Crossing Rate (ZCR) 7
4.2.2 RMS Energy (Root Mean Square) 8
4.3 Frequency Domain 8
4.3.1 Spectral Centroid 8
4.3.2 Spectral Bandwidth 9
4.3.3 Spectral Rolloff 9
4.3.4 Spectral Contrast 10
4.3.5 Spectral Flatness 10
4.4 Time and Frequency Domain 11
4.4.1 Mel-Frequency Cepstral Coefficients (MFCCs) 11
4.4.2 Manual Feature Extraction 12
5. Unsupervised Learning 13
5.1 K-means 13
5.2 Results of Clustering Evaluation 14
5.3 Evaluation of Feature Effectiveness 15
6. Conclusion 17
6.1 Future Perspectives 17

1. Introduction

The analysis of birds calls bridges bioacoustics, ecology, and computational science. Bird vocalizations provide critical insights into species behavior, communication, and ecological interactions. This project explores the use of signal processing techniques to analyze and categorize bird calls, offering scalable methods for biodiversity monitoring and ecological research.

The methodology includes data acquisition via the Xeno-canto API, preprocessing to enhance audio quality, and feature extraction to capture temporal and spectral characteristics. Using unsupervised learning algorithms such as K-Means clustering, the project identifies patterns in bird calls, enabling the grouping of species based on shared acoustic traits. These findings have applications in species identification, conservation efforts, and automated monitoring systems.

1.1 Goals

The primary objective of this project is to establish a robust pipeline for extracting meaningful and representative features from bird call recordings. By applying advanced preprocessing techniques—such as noise reduction—and sophisticated signal transformations—including Fourier-based methods—we aim to isolate key acoustic attributes that capture the essence of each vocalization. Through careful design and evaluation of the feature extraction process, we seek to create a reliable foundation upon which further analytical approaches, such as clustering algorithms, can operate. By ensuring the extracted features fully encapsulate the distinguishing characteristics of bird calls, this work sets the stage for improved downstream analyses. These enhanced features can be used to better understand species distribution, inform ecological monitoring, and contribute to data-driven conservation strategies. In doing so, our methodology supports efficient, scalable approaches to ecological research that depend on the accurate and nuanced representation of avian acoustic diversity.

2. Data acquisition

The data acquisition phase is critical to establishing a reliable foundation for bird call analysis. This project retrieves audio recordings of bird vocalizations from the Xeno-canto API, a well-regarded repository for avian sound data. Xeno-canto offers an extensive and curated database, including recordings from diverse species, environments, and conditions, ensuring a robust dataset for our analysis. The API enables systematic querying and downloading of recordings, facilitating the collection of high-quality data tailored to the study's objectives.

To ensure consistency and usability, each audio file is reviewed for quality, duration, and relevance to the selected species. This step minimizes potential noise in the dataset and enhances the reliability of subsequent analyses.

Selected Bird Species

The project focuses on a selection of species known for their distinctive vocalizations and ecological significance. The chosen species are:

- **King Penguin**: Known for their unique calls, penguins present an opportunity to study vocal patterns in non-flying birds.
- **Red-tailed Hawk**: Their sharp, iconic calls provide valuable data for high-frequency vocalization analysis.
- **Greater Prairie Chicken**: With their distinctive cooing calls, these birds contribute rich low-frequency data.
- **Magpie Goose**: Their honking vocalizations add diversity to the dataset, complementing the other species.

This selection ensures a diverse dataset that challenges the clustering and signal processing models, enabling a more comprehensive evaluation of the methodologies applied.

Dataset Scope and Considerations

The dataset aims to capture a range of audio features that highlight variations in frequency, rhythm, and spectral characteristics among the selected species. Recordings were filtered to exclude excessive background noise and interruptions to ensure data integrity. The inclusion of both high- and low-frequency vocalizations ensures that the analysis encompasses a wide acoustic spectrum.

This comprehensive approach to data acquisition lays the groundwork for robust signal processing and clustering analyses, ensuring the project's outcomes are both meaningful and applicable to real-world bioacoustic challenges.

3. Signal Preprocessing

This chapter presents the preprocessing procedures applied to the audio recordings before feature extraction and subsequent analysis. The steps undertaken ensure that the data is of high quality and that the resulting features accurately represent the bird calls. By systematically reducing noise, isolating the most informative temporal segments, and focusing on the relevant frequency bands, these methods collectively enhance the reliability of the dataset for clustering and classification tasks.

3.1 Noise Reduction

Overview: Bird call recordings often contain background noise and various disturbances that can obscure critical acoustic features. By applying noise reduction techniques, it is possible to improve the clarity and integrity of the recorded calls.

Methodology: Standard noise reduction algorithms were applied to each recording, identifying and attenuating non-informative background components while preserving the essential properties of the bird calls.

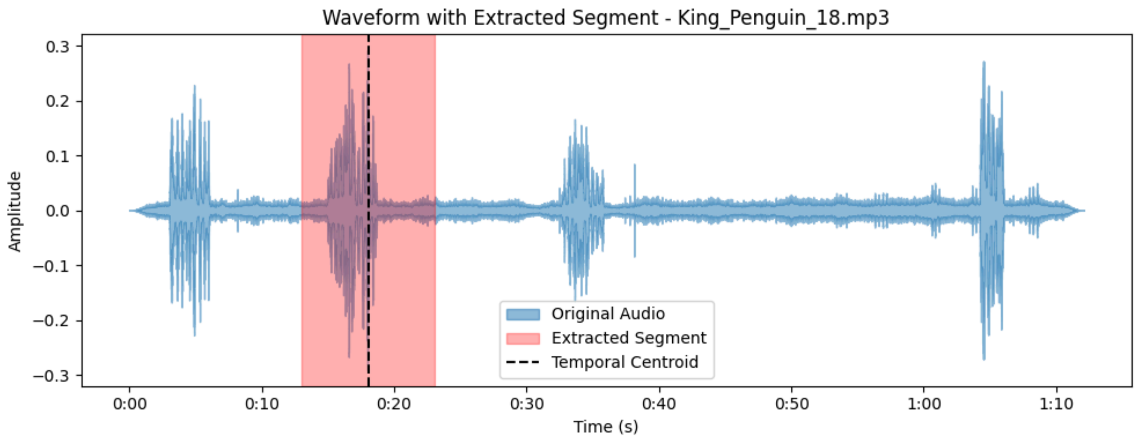
Importance and Outcome: This step ensured that the extracted features more accurately reflected the intrinsic characteristics of the bird vocalizations. The cleaner dataset resulting from noise reduction improved the reliability and precision of subsequent clustering or classification tasks.

3.2 Audio Segmentation: Temporal Centroid

Overview: Audio segmentation centered on the temporal centroid focuses analysis on the portion of the call with the most significant energy distribution, often capturing its most defining acoustic traits.

Methodology: The temporal centroid was calculated by examining the signal's energy envelope, identifying the time point at which cumulative energy reached approximately half of the total. A predefined segment around this centroid was then isolated for further analysis.

Importance and Outcome: Concentrating on the temporal centroid minimized the inclusion of noise, silence, and other irrelevant signal portions. Consequently, the extracted features represent the most informative segment of the bird call, enhancing the accuracy and effectiveness of downstream processing tasks.

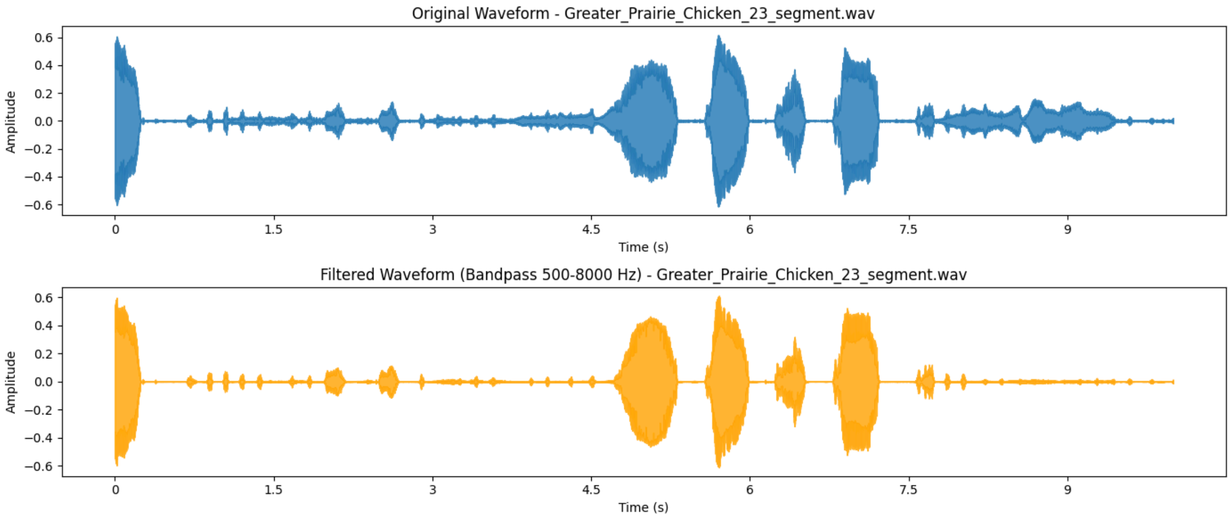


3.3 Bandpass Butterworth Filter

Overview: A bandpass Butterworth filter was employed to emphasize frequencies typically associated with bird vocalizations, while attenuating both lower-frequency environmental noise and higher-frequency interference.

Methodology: The filter was configured to retain frequencies between 500 Hz and 8000 Hz, ensuring that the relevant components of the bird calls were preserved. Frequencies outside this range were reduced, improving the clarity of the signals.

Importance and Outcome: Focusing on the frequency band where bird calls predominantly occur enhanced the quality of the extracted spectral features. This improved signal quality facilitates more reliable clustering and classification in subsequent analytical stages.



4. Feature Extraction

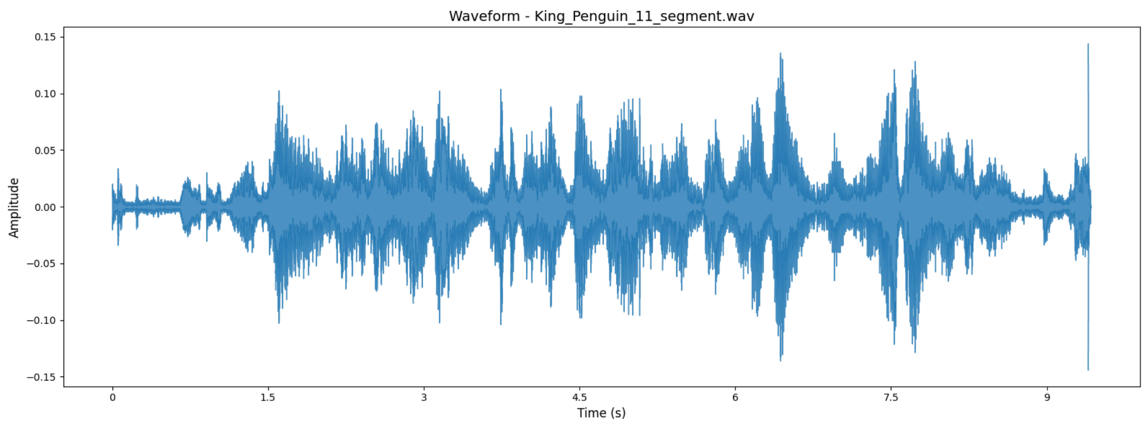
Feature extraction is a critical step in audio signal analysis, particularly when studying complex sounds like bird calls. By extracting meaningful features, we can quantitatively represent the audio data, enabling better classification, recognition, and analysis. Each feature provides unique insights into different aspects of the sound, such as pitch, energy, or spectral characteristics, which are essential for understanding and distinguishing various audio patterns. Below is an overview of key features commonly used in audio analysis and their relevance to the study of bird calls.

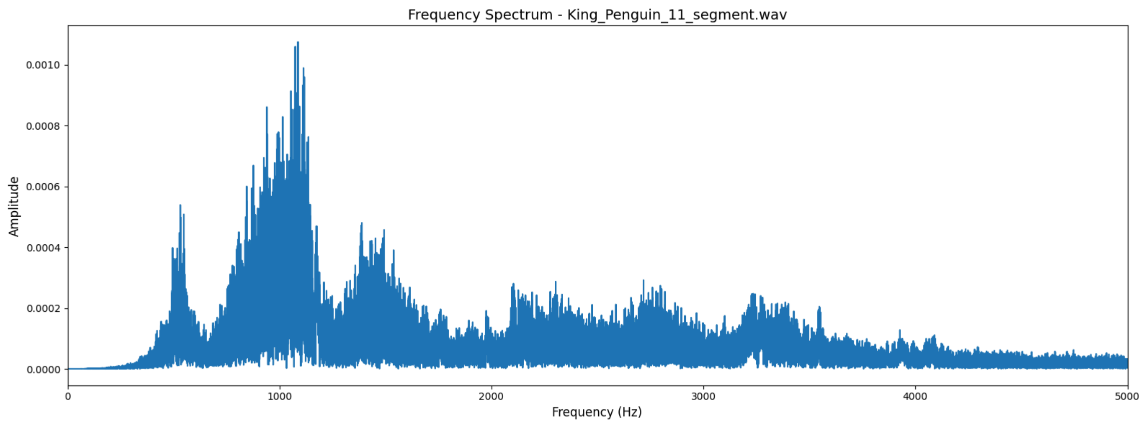
4.1 Fourier Transform

The Fourier Transform (FT) is a mathematical technique that converts a signal from its time-domain representation to its frequency-domain representation, decomposing complex signals into sinusoidal components to analyze frequency-specific characteristics, including amplitude and phase. In the context of bird call analysis, the FT plays a crucial role by identifying dominant frequency patterns, which are essential for classification. It also facilitates noise analysis by revealing unwanted components, such as low-frequency hums or high-frequency interference, thereby supporting preprocessing. Additionally, frequency-domain features derived from the FT simplify data representation while emphasizing relevant characteristics.

The implementation of the FT involves the use of the Fast Fourier Transform (FFT), an efficient algorithm for computing frequency components from signal segments. The FFT output consists of complex numbers whose magnitudes form the amplitude spectrum, representing the energy distribution across frequencies. This spectrum can be visualized to highlight energy concentrations, aiding both analysis and interpretation.

The FT is particularly valuable for signal analysis as it enables the identification of species-specific frequency signatures, contributing to the classification and clustering of bird calls. Furthermore, it validates preprocessing steps by confirming the removal of noise and the retention of relevant frequency bands. By incorporating Fourier analysis, this thesis integrates both time- and frequency-domain insights, enhancing the feature extraction process and improving the accuracy of bird call classification.





4.2 Time Domain

4.2.1 Zero Crossing Rate (ZCR)

The **Zero Crossing Rate (ZCR)** is a measure of how frequently the amplitude of an audio signal crosses the zero axis within a given time frame. This feature captures the rate of sign changes in the signal, reflecting the signal's temporal dynamics. It is commonly associated with the noisiness or abruptness of a sound, as rapid changes in direction result in higher ZCR values.

Relevance to Bird Call Analysis:

Characterizing Call Quality:

- Smooth, tonal bird calls typically exhibit low ZCR values, reflecting their stable amplitude and frequency patterns.
- Noisier or more percussive calls, such as sharp chirps or squawks, have higher ZCR values due to frequent amplitude changes.

Species and Behavior Differentiation:

- ZCR can help distinguish between calls with tonal characteristics, such as mating songs, and abrupt, noisy calls, such as alarm signals or distress calls.

In bird call classification, ZCR is particularly effective for identifying alarm calls, which are characterized by abrupt amplitude changes and high energy transitions. This feature provides a quantitative measure to differentiate behavioral contexts and sound types, contributing to more accurate classification.

4.2.2 RMS Energy (Root Mean Square)

Root Mean Square (RMS) Energy quantifies the average power of an audio signal over time, providing a measure of its perceived loudness. It is calculated as the square root of the mean of the squared amplitude values within a segment. This feature effectively captures variations in signal intensity, making it a critical parameter in audio analysis.

Relevance to Bird Call Analysis

- **Loudness Differentiation**:
    - RMS Energy distinguishes between soft, distant calls and loud, close-range vocalizations, providing insights into the bird's proximity and environmental context.
- **Behavioral and Contextual Indicators**:
    - Variations in intensity can reflect different behavioral states, such as urgency or aggression. Louder calls may indicate dominance or alarm, while softer calls might signify non-threatening communication or distant vocalizations.

RMS Energy can be utilized to analyze the intensity patterns in bird calls to infer behavioral traits, such as dominance or aggression, in communication. For example, a series of loud, high-energy calls might signal territorial disputes or alert other birds to a threat.

4.3 Frequency Domain

4.3.1 Spectral Centroid

The **Spectral Centroid** is a measure that indicates the "center of mass" of a signal's frequency spectrum. It is calculated as the weighted average of all frequencies, with their magnitudes serving as weights. Essentially, it reflects where the majority of the signal's energy is concentrated in the frequency domain.

Relevance to Bird Call Analysis

- **Pitch Characterization**:
    - High spectral centroid values correspond to higher-pitched calls, while lower values indicate deeper or lower-pitched tones. This makes it a useful feature for characterizing the pitch of bird calls.
- **Species Differentiation**:

    - Many bird species exhibit distinct frequency ranges in their vocalizations. The spectral centroid can help distinguish species with higher vocal ranges, such as songbirds, from those with lower ranges, such as larger birds like owls.

The spectral centroid can be used to compare the calls of different bird species. For instance, songbirds typically have higher spectral centroids due to their high-pitched vocalizations, while birds like owls exhibit lower centroids because of their deeper, low-frequency calls..

4.3.2 Spectral Bandwidth

Spectral Bandwidth measures the range of frequencies surrounding the spectral centroid, indicating the spread or distribution of energy in the frequency spectrum. It is calculated as the standard deviation of the frequencies weighted by their magnitudes. This feature provides insights into the complexity and texture of a signal.

Relevance to Bird Call Analysis

- **Complexity of Calls**:
    - Calls with a narrow spectral bandwidth are often simple and tonal, while wider bandwidths suggest dynamic calls with a mixture of tonal and noisy elements.
- **Behavioral and Species Differentiation**:
    - Spectral bandwidth aids in distinguishing bird species and behaviors by analyzing the complexity of their vocalizations. For example, elaborate calls may indicate mimicry or advanced communication patterns.

Spectral bandwidth is particularly useful in studying mimic bird, which produce calls with wide bandwidths due to their ability to mimic diverse sounds. In contrast, simpler tonal calls with narrow bandwidths are often associated with songbirds performing mating calls.

4.3.3 Spectral Rolloff

The **Spectral Rolloff** is a measure that identifies the frequency below which a specified percentage (commonly 85%) of the total spectral energy is concentrated. It provides insights into the distribution of energy across the frequency spectrum, reflecting whether the signal's energy is dominated by lower or higher frequencies.

Relevance to Bird Call Analysis

- **Energy Distribution Analysis**:
    - Higher rolloff values are associated with high-frequency calls, such as sharp chirps, while lower rolloff values indicate low-frequency calls, like deep coos.
- **Signal Categorization**:

- Spectral rolloff can differentiate between tonal calls with concentrated energy in specific frequency bands and broadband calls that spread energy across a wider range of frequencies.

In noisy environments, spectral rolloff is particularly useful for distinguishing bird calls from ambient noise. For instance, it can classify high-frequency bird calls in habitats with significant low-frequency background noise, ensuring accurate call analysis.

4.3.4 Spectral Contrast

Spectral Contrast measures the difference in intensity between peaks (high-energy frequencies) and valleys (low-energy frequencies) in the frequency spectrum. This feature provides insights into the dynamic range and harmonic content of a signal, capturing the distribution of energy variations across different frequency bands.

Relevance to Bird Call Analysis

- **Dynamic Range and Harmonic Content**:
  - Spectral contrast highlights the richness and complexity of bird calls. Calls with high spectral contrast often have pronounced harmonic structures, while monotone or less dynamic calls exhibit lower contrast.
- **Call Characterization**:
  - By analyzing variations in intensity, spectral contrast differentiates between simple calls with uniform energy and complex calls with diverse intensity patterns.

Spectral contrast is particularly useful for studying territorial songs, which often feature rich harmonic structures and dynamic intensity variations. For example, birds employing elaborate calls to assert dominance or attract mates can be identified and analyzed using this feature.

4.3.5 Spectral Flatness

Spectral Flatness is a measure that quantifies how uniform or "flat" the energy distribution is across the frequency spectrum. It is calculated as the ratio of the geometric mean to the arithmetic mean of the power spectrum. High spectral flatness indicates a noise-like signal with energy spread evenly across frequencies, whereas low flatness corresponds to tonal signals with concentrated energy in specific frequencies.

Relevance to Bird Call Analysis

- **Characterizing Signal Tonality**:
  - Signals with low spectral flatness are typically tonal and harmonic, while those with high flatness resemble broadband noise or chaotic calls.
- **Noise Differentiation**:
  - Spectral flatness is effective in distinguishing actual bird calls from background noise or interference in recordings, aiding in data preprocessing.

- **Behavioral Insights**:
  - Calls with high flatness may indicate alarm or distress signals, while tonal calls with low flatness often signify communication or mating behaviors.

Spectral flatness is particularly useful for isolating harmonic bird songs in recordings from noisy environments, such as forest ecosystems with ambient wind or water sounds. It can also help classify calls based on their tonal structure, distinguishing melodic songs from noisy, broadband vocalizations.

4.4 Time and Frequency Domain

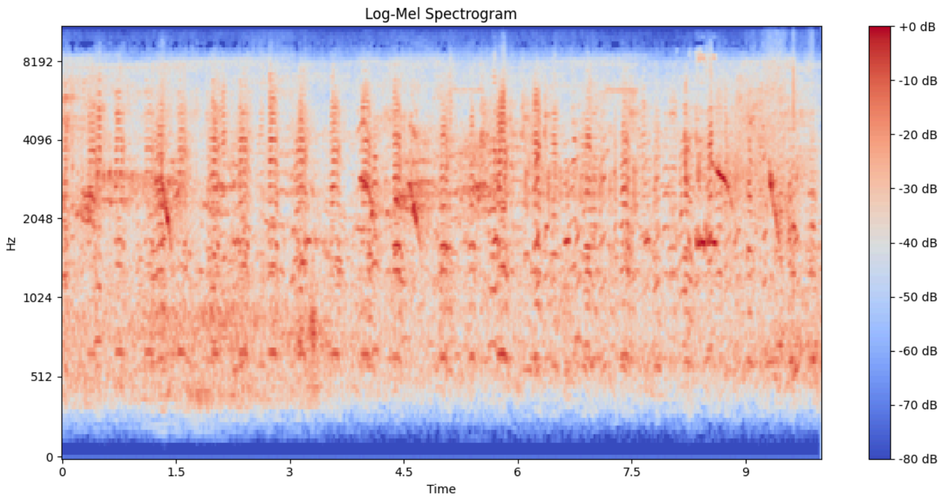
4.4.1 Mel-Frequency Cepstral Coefficients (MFCCs)

Mel-Frequency Cepstral Coefficients (MFCCs) are features derived from the short-term power spectrum of a sound, designed to model human auditory perception on the Mel scale. They are computed by applying a logarithmic transformation to the signal's power spectrum, followed by a discrete cosine transform (DCT), which compresses the information into a set of coefficients representing the overall spectral shape.

Relevance to Bird Call Analysis

- **Timbre and Structural Characteristics**:
  - MFCCs capture nuanced differences in the timbre and structure of bird calls, providing valuable features for species classification.
- **Noise Robustness**:
  - Their design makes them relatively robust to background noise and variations in recording conditions, making MFCCs ideal for real-world ecological studies.
- **Species Differentiation**:
  - Subtle acoustic differences encoded by MFCCs enable the identification of specific bird species based on unique vocal patterns.

MFCCs are widely used for identifying bird species by their distinct vocal signatures. For example, subtle differences in frequency modulation and harmonic content between species can be effectively captured using MFCCs, even in noisy environments or across varying recording setups.



4.4.2 Manual Feature Extraction

Basic Spectrograms

We developed a custom-designed feature extraction technique specifically tailored for bird call analysis, referred to as a basic spectrogram. This approach, while lightly leveraging the Short-Time Fourier Transform (STFT) functionality from the Librosa package, primarily employs an original processing pipeline. By sampling probably overlapping windows across the signal and calculating the magnitude of frequencies divided into predefined frequency bins, this method generates a frequency matrix. The resulting matrix is transposed and visualized as a heatmap, providing an intuitive graphical representation of the temporal and spectral dynamics of bird calls.

Key Components

- **Signal Windowing:** The audio signal is segmented into overlapping time windows (e.g., 30 ms), enabling localized frequency analysis. This segmentation facilitates a detailed examination of temporal variations in the audio signal.
- **Frequency Binning:** The frequency spectrum is divided into bins (e.g., 200 bins) corresponding to the typical range of bird vocalizations (500 Hz to 8000 Hz). The bin ranges are dynamically calculated using the formula: This customization aligns the frequency bins with the specific characteristics of bird calls, ensuring accurate representation of the vocalizations.
- **Magnitude Calculation:** The STFT output is processed to compute the relative magnitude of frequencies within each bin. This step captures the energy distribution over time and frequency, providing a detailed spectral profile of the signal.
- **Matrix Representation:** The final output is a time × frequency_bins matrix that visualizes the evolution of frequency content over time. This matrix forms the foundation for creating heatmap representations, offering clear insights into the spectral dynamics of the analyzed signals.

Utility in Bird Call Analysis

- **Manual Feature Design:** This custom spectrogram implementation allows for the creation of tailored frequency bins and ranges, specifically designed to align with the unique characteristics of bird vocalizations.
- **Temporal and Spectral Dynamics:** The spectrogram effectively captures unique time-frequency patterns, such as chirps, trills, and modulations. These intricate structures are essential for identifying and differentiating bird calls.
- **Compact and Interpretable Representation:** By converting raw audio data into a concise matrix format, the method reduces computational complexity while preserving critical spectral and temporal features. This matrix can be directly visualized as a heatmap, facilitating intuitive analysis.
- **Visual Insights:** Heatmap-ready visualizations provide an intuitive way to compare bird calls. They highlight harmonic patterns and spectral shifts, enhancing the interpretability of the data.

Applications

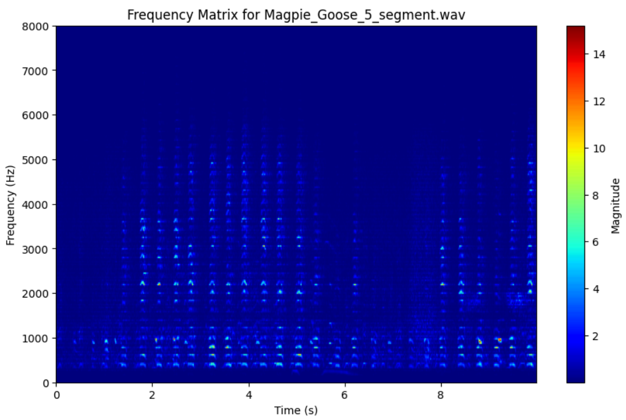
1. **Species Identification**: Differentiates bird species by analyzing their unique vocal patterns.

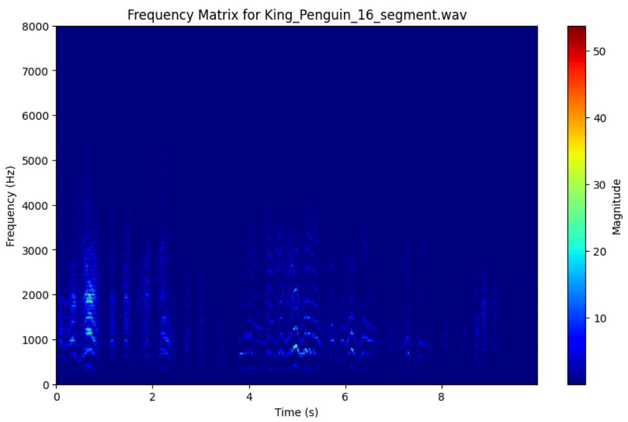
2. **Behavioral Analysis**: Investigates call sequences, interactions, and temporal variations to infer behavioral contexts.
3. **Noise Resilience**: By focusing on frequency magnitudes, the method minimizes the impact of background noise, enhancing the clarity of extracted features.

Highlights

The custom spectrogram demonstrates the value of domain-specific, manually designed feature extraction:

- **Customizable Frequency Bins and Ranges**: Offers flexibility for detailed studies of bird calls.
- **Visualization Techniques**: Enhances the interpretability of spectral features, including harmonics and frequency shifts.
- **Robust Alternative to Generic Methods**: Emphasizes flexibility and precision, addressing the unique challenges of bioacoustic feature extraction.





5. Unsupervised Learning

5.1 K-means

K-Means is an **unsupervised machine learning algorithm** that partitions data into **k distinct clusters** based on feature similarity. It works by minimizing the **inertia**, which is the sum of squared distances between each data point and the centroid of its assigned cluster. The algorithm follows these steps:

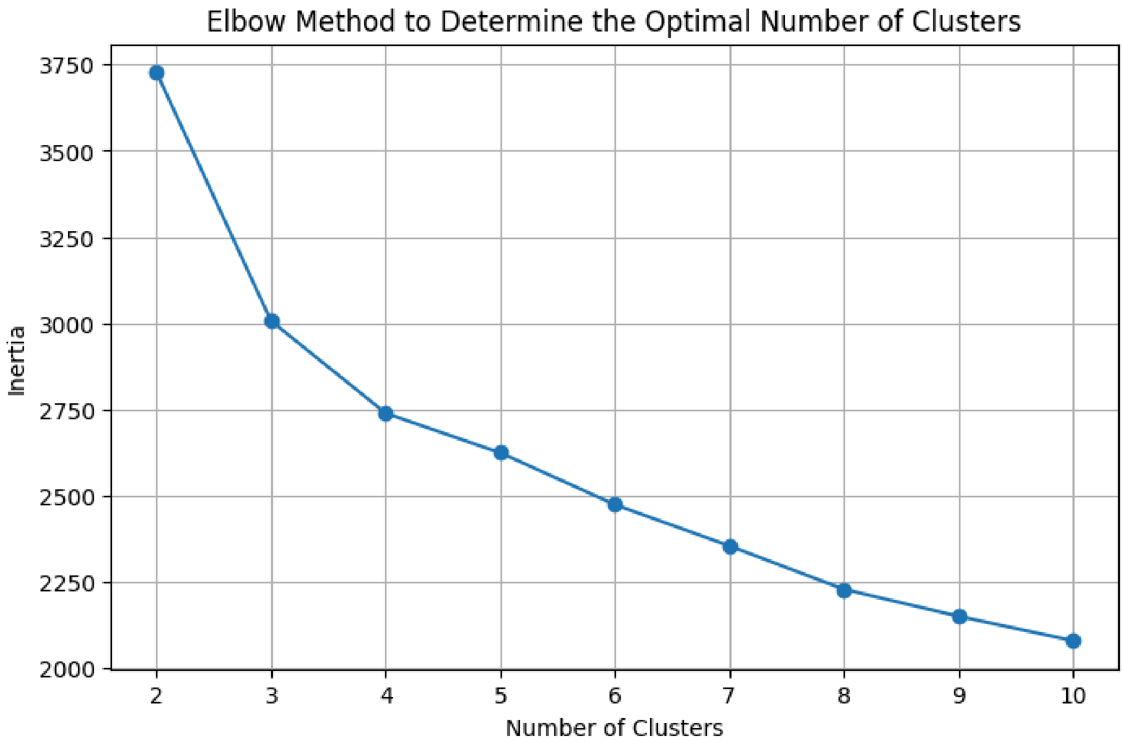
1. **Initialization**: Randomly assign k initial cluster centers (centroids).
2. **Assignment**: Assign each data point to the cluster whose centroid is nearest to it.
3. **Update**: Recalculate the centroids based on the mean of data points within each cluster.
4. **Repeat**: Continue the assignment and update steps until centroids stabilize or a stopping criterion is met.

Relevance of K-Means for Bird Calls

Uncover Patterns: By grouping bird calls with similar acoustic features (e.g., frequency, duration, pitch), K-Means can reveal trends related to bird behavior, habitat, or species-specific vocalizations.

Simplify Data: Clustering reduces the complexity of analyzing large datasets, making it easier to study overarching patterns and relationships in bird call data.

Aid Classification: K-Means provides unsupervised insights into the natural groupings within data, potentially helping to define classes for subsequent supervised learning models.



5.2 Results of Clustering Evaluation

The application of K-Means clustering to the bird call dataset resulted in the formation of four distinct clusters. A detailed evaluation of the clustering performance is presented below:

Cluster 1

- **Dominant Species**: *Magpie Goose*
- **Correct Assignments**: 22
- **Accuracy**: 62.86%
- **Interpretation**: The clustering algorithm demonstrated moderate success in grouping *Magpie Goose* calls. However, the relatively lower accuracy indicates overlapping features, possibly due to similarities in spectral characteristics such as bandwidth and energy with other species.

Cluster 2

- **Dominant Species**: *Penguin*
- **Correct Assignments**: 20

- **Accuracy**: 74.07%
- **Interpretation**: This cluster exhibited higher accuracy, suggesting that the algorithm effectively leveraged features such as Mel-Frequency Cepstral Coefficients (MFCCs) to capture the tonal attributes of *Penguin* calls.

Cluster 3

- **Dominant Species**: *Greater Prairie Chicken*
- **Correct Assignments**: 22
- **Accuracy**: 84.62%
- **Interpretation**: High clustering accuracy for *Greater Prairie Chicken* calls highlights the utility of spectral centroid and MFCCs in distinguishing their specific vocal features. The results demonstrate a clear separation from other species.

Cluster 4

- **Dominant Species**: *Red-tailed Hawk*
- **Correct Assignments**: 23
- **Accuracy**: 92.00%
- **Interpretation**: The clustering results for *Red-tailed Hawk* achieved the highest accuracy. This performance can be attributed to features like spectral contrast, which likely emphasized the unique frequency dynamics characteristic of this species.

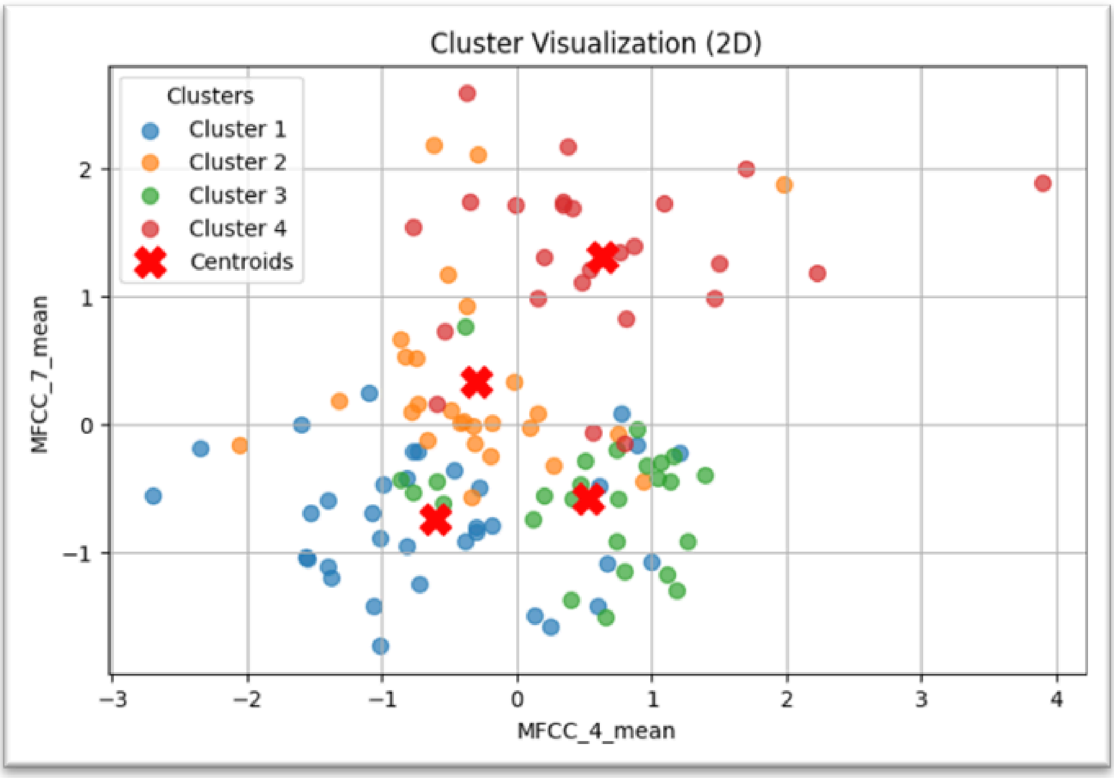
5.3 Evaluation of Feature Effectiveness

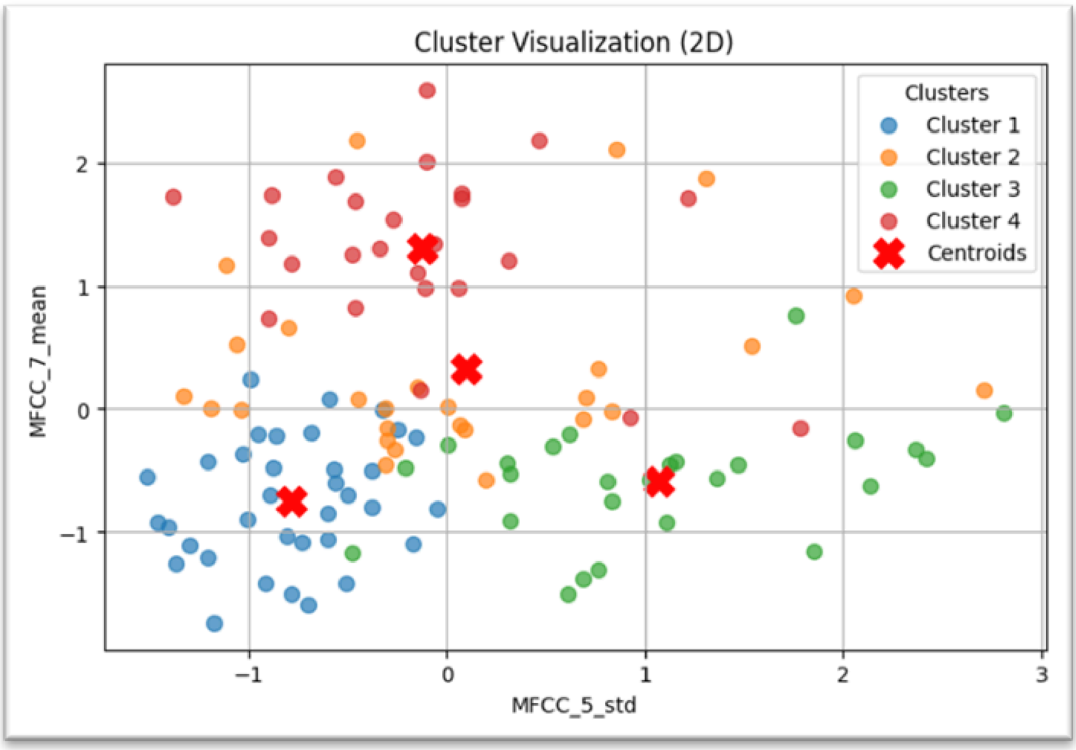
The selected features were pivotal in capturing the distinct characteristics of bird calls, contributing to the overall clustering performance. Key insights into feature effectiveness are detailed below:

1. **Spectral Centroid and Bandwidth**: These features proved essential for differentiating high-pitched, tonal calls from broader, noisy ones, thereby facilitating species-specific clustering.
2. **MFCCs**: MFCCs were instrumental in capturing subtle nuances in timbre and harmonic structure. Their effectiveness was particularly evident in the clusters with high accuracy, such as *Greater Prairie Chicken* and *Red-tailed Hawk*.
3. **Spectral Rolloff**: This feature was useful in segregating high-frequency calls, such as those of *Penguin*, from low-frequency calls, including those of *Red-tailed Hawk*.
4. **RMS Energy**: RMS energy effectively distinguished louder calls, characteristic of *Magpie Goose*, from softer calls of other species.
5. **Spectral Contrast**: By emphasizing the dynamic and harmonic content, spectral contrast played a critical role in identifying species-specific vocal patterns, particularly in clusters with higher accuracy.

The results of the K-Means clustering analysis validate the effectiveness of the chosen features in representing the acoustic complexity of bird calls. Clusters with high accuracy, such as those dominated by *Greater Prairie Chicken* and *Red-tailed Hawk*, underscore the importance of MFCCs, spectral centroid, and spectral contrast in achieving precise separation. The overlap observed in clusters with lower accuracy, such as *Magpie Goose*, indicates potential areas for improvement, either through refinement of the feature set or enhancement of preprocessing techniques.

Overall, the feature set provided a robust foundation for clustering bird calls, demonstrating its potential application in bioacoustics and automated species monitoring systems.





6. Conclusion

This project successfully demonstrates the application of advanced signal processing techniques for analyzing bird calls, achieving the outlined objectives. By adhering to a systematic methodology that combines data acquisition, preprocessing, feature extraction, and clustering, the study has provided a robust framework for bird call analysis.

Key Achievements:

1. **Signal Importation and Processing**:
    - Leveraged the Xeno-canto API to access a diverse collection of bird call recordings.
    - Focused on a varied dataset including species such as Penguin, Red-tailed Hawk, Greater Prairie Chicken, and Magpie Goose.
    - Employed effective preprocessing methods, including noise reduction, temporal centroid extraction, and bandpass filtering, to ensure high-quality input for further analysis.
    - These steps ensured the audio signals were clean and adequately prepared for feature extraction and subsequent analytical tasks.
2. **Feature Extraction**:
    - Feature extraction emerged as the cornerstone of this study, with a deliberate focus on identifying and isolating key acoustic properties of bird calls.
    - Extracted a comprehensive set of temporal features, such as RMS Energy and Zero Crossing Rate, which capture dynamic aspects of the calls.
    - Spectral features, including Spectral Centroid, Bandwidth, Rolloff, and MFCCs, provided detailed insights into frequency domain characteristics essential for distinguishing bird species.
    - This multi-faceted approach to feature extraction ensured a holistic representation of the audio signals, combining both time and frequency domain information.
    - Incorporated manual feature validation to confirm the relevance and precision of the extracted attributes, further strengthening the reliability of the analysis.
3. **Unsupervised Learning**:
    - Applied K-Means Clustering to classify bird calls based on extracted features.
    - Utilized the Elbow Method and Silhouette Score to determine and validate the optimal number of clusters, ensuring meaningful acoustic categorization.
    - The results of the K-Means clustering analysis validate the effectiveness of the chosen features in representing the acoustic complexity of bird calls. Clusters with high accuracy, such as those dominated by Greater Prairie Chicken and Red-tailed Hawk, underscore the importance of MFCCs, spectral centroid, and spectral contrast in achieving precise separation.

6.1 Future Perspectives

1. Integration of Supervised Learning

- Extend the project by training classification models (e.g., Random Forest, Neural Networks) using labeled data.
- Enable automatic bird species identification from new recordings.

2. Dynamic Feature Expansion

- Experiment with time-frequency representations like spectrograms or wavelet transforms for deeper insights.

3. Dataset Augmentation

- Expand the dataset to include more species and environmental conditions for greater generalizability.
- Explore real-time audio recording and processing for field applications.

4. Advanced Clustering Methods

- Investigate clustering algorithms like DBSCAN or Gaussian Mixture Models (GMMs) for more nuanced groupings.
- Utilize deep learning approaches (e.g., autoencoders) for feature reduction and unsupervised learning.

5. Ecological Insights

- Correlate acoustic features with ecological data (e.g., habitat type, time of day) to explore environmental and behavioral patterns.