Research Paper Advanced Academic Skills PRA2015

Academic year 2017-2018

Sophie Dobbie i6105435, Kai Hendriks i6106037

Word count 1148

**Fourier Analysis as a Reliable Approach to Automated Bird Identification**

**Abstract**

Avian identification in ornithology is subjected to human error and decreased efficiency (Kogan & Margoliash, 1998). Methods of identification automation include two sections: pitch examination and pattern recognition. Previously attempted methods emphasise pattern recognition and carry explicit limitations concerning noise, management of large datasets, and reliance upon manual input (Stowell & Plumbley, 2014). As such, there is not yet an established industry-wide method of automated identification (Stowell & Plumbley, 2014). This may be helped by the implementation of a generalised pitch examination program. This research sought to determine the reliability of Fourier analysis as a method of generalised pitch examination within automated bird identification. The Fast Fourier Transform (FFT) was used to provide a database containing the average frequencies of a species, to which the average frequency of an unknown birdsong was compared. This comparison was an assessment of compatibility between the frequency of an unknown birdsong and frequencies of species within the database. By extension, this results in a possible identification. The reliability of the program was assessed by testing the compatibility of known species' calls with their true species within the database. The program correctly matched 75% of the test vocalisations. Of this fraction, 77% had multiple compatibilities. This indicates Fourier analysis to be a sufficient platform from which pattern recognition methods may be developed. The limitations concern low sample-size, noisy audio files, and the contamination of the focus bird's frequency by other species. Possible improvements include the increase in sample size and the insertion of an additional filter.

**Keywords:** Bird Identification, Fourier Analysis, Pitch Examination, Automation

**1. Introduction**

As avian identification in ornithology is assessed through human effort, it is subjected to human error and excessive time consumption (Kogan & Margoliash, 1998). These issues can be rectified through the automation of bird identification; one approach being audio recognition. Recognition of this manner has been researched since the 1990’s, where analysis is reliant upon pattern recognition (Stowell & Plumbley, 2014). Previously attempted methods carry limitations concerning noise, management of large datasets, and reliance upon manual input (Stowell & Plumbley, 2014). At this time, there is an absence of accurate methodology of automated bird vocalisation recognition. This absence may be helped by an emphasis on pitch examination.

Currently, pitch examination is only used to determine the note segmentation within vocalisations (Stowell & Plumbley, 2011). However, only pitch examination could differentiate between a singular, flat note vocalised by two species. Additionally, pitch selection would decrease the number of species need referring to in the pattern recognition section, thereby lowering required processing power. As such, this research seeks to establish a generalised pitch examination program to provide a foundation to automated classification. Fourier analysis of bird vocalisations will be used to compare average frequencies between known bird vocalisations to a database of species. Ideally, the signals would be consistently compatible with their known database matches. As this program only tests for the average frequency, similar frequency birds are expected to be compatible with each other, resulting in multiple matches for a single input. If successful, Fourier analysis will provide the foundation upon which pattern recognition methods can be developed.

**2. Methods**

*2.1 Fourier-Based Bird Vocalisation Identification*

Any signal can be linearly approximated as a Fourier series composed of linear combinations of weighted sine and cosine functions (equation 1) (Lay, Lay, & Mcdonald, 2015). The Fast Fourier Transform (FFT) allows for the examination of the time and frequency domains of this series (ex. figures 1 and 2 in the appendix) (Brigham, 1988). This is possible as the FFT decomposes the series into sinusoids of different frequencies (equation 2). The result provides insight into the frequency components of the original signal. Therefore, performing FFT on an avian vocalisation will provide its constituent frequencies. This will be used as a means of comparison of the averages of constituent frequencies between audio files.

*Equation 1. Fourier series.*

*Equation 2. Fourier Transformation. S(f) is the Fourier transformation of the signal s(t).*

*2.2 Data Acquisition*

The signals were acquired from the crowdsourced bird vocalisation sharing website xeno-canto (2005). Only audio files between 10 and 90 seconds long featuring the bird of choice were used. The selected audio files were converted from stereo MP3 to mono WAVE files to allow for MATLAB processing.

*2.3 Main Approach*

The pitch examination program was sorted into two functions with two separate purposes.

1. Addtodatabase Function

The Addtodatabase Function’s goal is to compute a summary database composed of the average frequencies and associated standard deviations of various avian species’ vocalisations. This is constructed first by the design of a preliminary collective database. This collective database contains all recordings inputted into the addtodatabase function. Each column is representative of a single species, where each row contains the average frequency of a single recording. These average frequencies are obtained through FFT performed upon an input WAVE file. The function returns the average frequency of the loudest source within the file (i.e. the bird of choice). This frequency is subsequently added to an empty row within the chosen bird’s column. The mean frequency and associated standard deviation per species is saved into the summary database for comparison within the IDbird function.

1. IDbird Function

The IDbird function calculates the average frequency of the primary source within an audio file and compares this value to those within the summary database. It performs an FFT upon an input vocalisation to determine the constituent frequencies. It then filters frequencies lower than 50% of the maximum detected volume (power), ideally leaving only the primary caller’s frequency. Once the average frequency of the loudest audio-source is calculated, it is compared to each column within the summary database. Upon fulfilling equation 3, the function returns the compatible species' name. As the unknown vocalisation is compared to all species within the database, multiple compatibilities are possible.

*Equation 3. Compatibility Range formula. avgF is the average frequency, std.dev is the standard deviation, DB is database, and unknown refers to the unknown bird call.*

*2.4 Analyses Performed*

The reliability of the program was determined by a test of known compatible and incompatible vocalisations. Five audio files from four species each were added to the database. Each selected species’ vocalisations covered different ranges of pitch. In order of increasing pitch, the chosen birds were: *Columba palumbus* (common wood pigeon), *Cyanistes caeruleus* (Eurasian blue tit), *Passer domesticus* (house sparrow), and *Alcedo atthis* (common kingfisher). Three audio files from each species were input into the IDbird function, acting as the unknown audio files. The accuracy of the program depends upon the success in identifying an input signal as compatible with its known species. Precision of the program is determined by the instances of multiple compatibilities.

**3. Results & Discussion**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Common Kingfisher | Wood Pigeon | House Sparrow | Eurasian Blue Tit |
| Average frequency (Hz) | 5368 | 462 | 4032 | 5495 |
| Standard deviation (Hz) | 992 | 41 | 377 | 1303 |
| Compatibility Span (Hz) | 4376 – 6360 | 421 – 503 | 3655 – 4409 | 4192 – 6798 |
| Compatibility Range | 1984 | 82 | 754 | 2606 |

*Table 1: Calculated column average frequency and standard deviation of the summary database. Compatibility Span is calculated using equation 3. Compatibility Range describes how large the span is.*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Kingfisher 1 | Kingfisher 2 | Kingfisher 3 |
| Compatible with | Kingfisher, blue tit | Kingfisher, blue tit | Kingfisher, blue tit |
| Average frequency (Hz) | 5978 | 5425 | 6026 |
|  | **Wood pigeon 1** | **Wood pigeon 2** | **Wood pigeon 3** |
| Compatible with | None | Wood pigeon | None |
| Average frequency (Hz) | 574 | 480 | 402 |
|  | **House sparrow 1** | **House sparrow 2** | **House sparrow 3** |
| Compatible with | House sparrow | None | House sparrow, blue tit |
| Average frequency (Hz) | 3912 | 3251 | 4310 |
|  | **Blue tit 1** | **Blue tit 2** | **Blue tit 3** |
| Compatible with | Kingfisher, blue tit | Kingfisher, blue tit | Kingfisher, blue tit |
| Average frequency (Hz) | 5253 | 5612 | 5703 |

*Table 2: Input audio files with their average frequencies and corresponding matches.*

|  |  |
| --- | --- |
| Percentage of correct matches (Accuracy) | 75% |
| Percentage of calls with multiple compatibilities, within correctly matched calls category (Precision) | 77% |

*Table 3: Lists the accuracy and the precision of the program, where the program correctly matched species 75% of the time. Of that 75%, 77% of them had multiple compatibilities.*

The results outline a definite direct proportionality between the compatibility range and the correct identification of a species. With a compatibility range of 82 possible frequencies, the wood pigeon test had only one correct match. There is an evident inconsistency of ~400Hz between the average frequency of House Sparrow 2 (table 2) and the compatibility span minimum (table 1). This is despite the house sparrow’s compatibility range covering 754 possible frequencies. Still, House Sparrow 1 and 3 both identified correct matches. Conversely, the kingfisher's and blue tit’s ~2000Hz database compatibility range resulted in correct matches amongst each of their three test signals. At the same time, these large compatibility ranges engendered the multiple compatibilities seen in 77% of the 75% correctly matched calls (table 3).

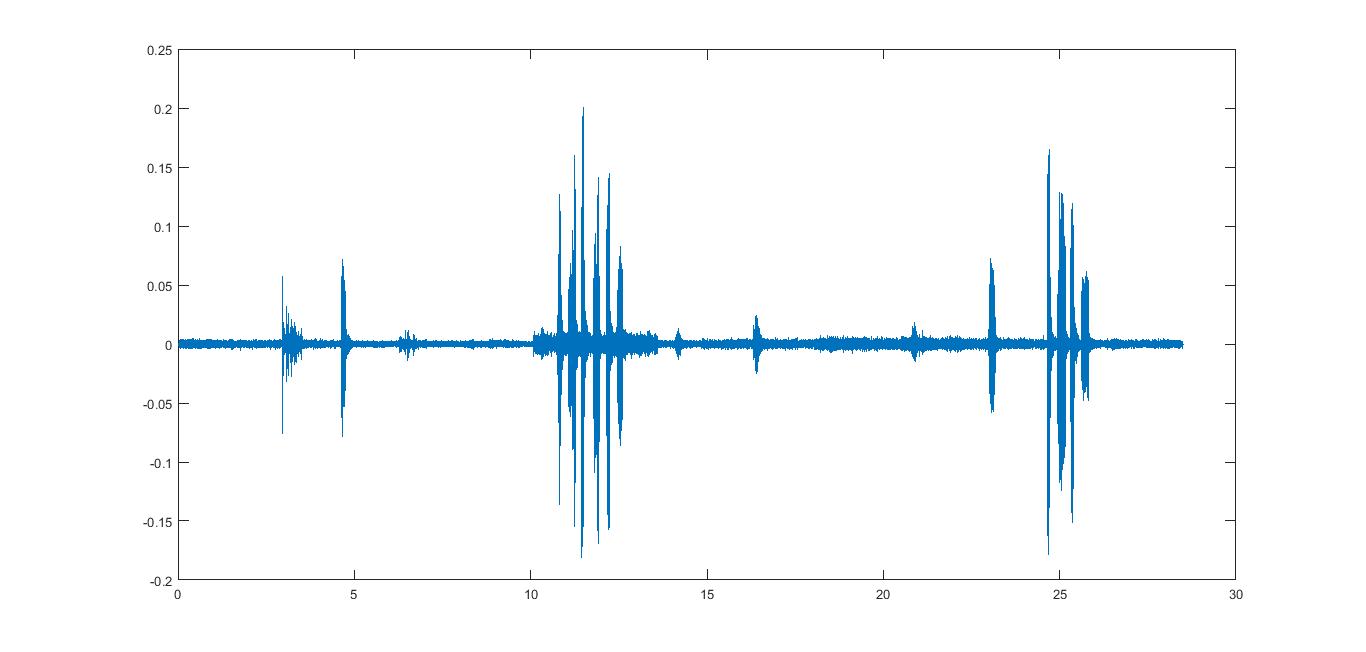
The comparison of the average frequencies was expected to yield multiple compatibilities, as pattern recognition would be the only solution to differentiating between species of similar frequencies. As for the 25% accuracy discrepancy, the false identifications are suspected to stem from an inaccurate depiction of mean and standard deviation values, particularly in the house sparrow tests. These inaccuracies may be a result of noise or contamination by loud undesired vocalisations within an audio file. Additionally, the low sample-size of both database signals and input signals may not be representative of all possible vocalisation pitches of a species, resulting in in the misidentifications in both the house sparrow and wood pigeon tests. The implementation of additional noise filters specified to each frequency range as well as an increased sample-size would increase the resultant accuracy. As such, Fourier analysis paired with these improvements would be a reliable method of generalised pitch examination.

**4. Conclusion**

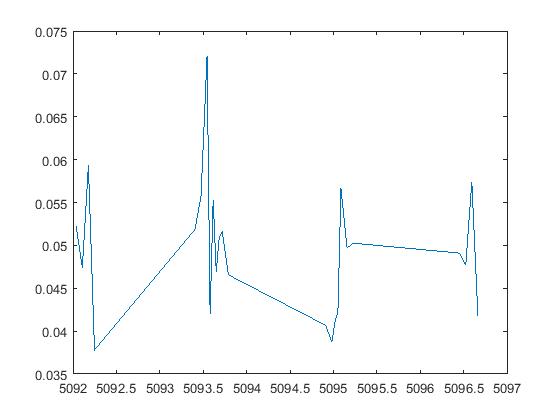
This research sought to establish the reliability of Fourier analysis within automated bird classification. The program correctly identified 75% of the input vocalisations, of which and 77% had multiple compatibilities. The low precision was expected to occur between similar pitched species. While 75% accuracy is a moderately successful outcome, it may be improved by the implementation of increased sample size and additional noise filters so as to precisely record the desired frequency. Nonetheless, Fourier analysis is indicated to be a reliable method from which automated bird identification programs can be developed.

**References**

* + Brigham, E. (1988). *The fast Fourier transform and its applications* (1st ed.). Englewood Cliffs, N.J.: Prentice Hall.
  + Kogan, J., & Margoliash, D. (1998). Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study. *The Journal Of The Acoustical Society Of America*, *103*(4), 2185-2196. http://dx.doi.org/10.1121/1.421364
  + Lay, D., Lay, S., & McDonald, J. (2015). *Linear algebra and its applications* (5th ed.). Boston: Pearson.
  + Stowell, D., & Plumbley, M. (2011). *Birdsong and C4DM: A survey of UK birdsong and machine recognition for music researchers*. London: Centre for Digital Music, Queen Mary, University of London. Retrieved from <http://c4dm.eecs.qmul.ac.uk/papers/2010/Stowell2010-C4DM-TR-09-12-birdsong.pdf>
  + Stowell, D., & Plumbley, M. (2014). Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *Peerj*, *2*, e488. http://dx.doi.org/10.7717/peerj.488
  + *xeno-canto:: Sharing bird sounds from around the world*. (2005). *Xeno-canto*. Retrieved 15 October 2017, from http://www.xeno-canto.org

**Appendix**

*Figure 1: Original signal of the song of a common kingfisher. The x-axis indicates time, the y-axis indicates intensity of the signal.*



*Figure 2: Power-frequency graph of a sound file recording the song of a common kingfisher (calculated average frequency 5094,3 Hz). Post-FFT and post removal of frequencies lower than 50% of the maximum power.*