Interactive Classification Using Spectrograms and Audio Glyphs

Eren Cakmak Udo Schlegel Matthias Miller Juri Buchmüller Wolfgang Jentner Daniel A. Keim

Data Analysis and Visualization Group
University of Konstanz*

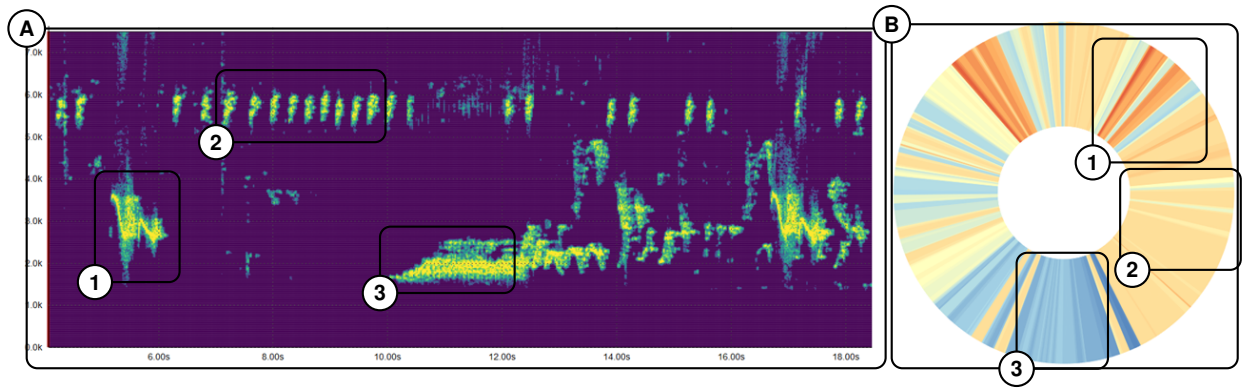

Figure 1: (A) Interactive spectrogram and (B) audio glyph – Both views (A) and (B) show the same preprocessed supposed audio sample of a pipit call. (A) shows an interactive spectrogram and (B) an audio glyph that maps the spectral roll-off of the audio sample to a global linear color scale. Three different bird calls (1-3) are visible in the spectrogram and the audio glyph.

Abstract

The VAST Challenge 2018 aims to clarify the situation of the *Rose-Crested Blue Pipit's* population in the *Boonsong Lekagul Wildlife Preserve*. We propose an interactive spectrogram and a representative novel audio glyph to support the classification of bird calls. The second visualization of our system helps to identify spatio-temporal patterns of bird species in the preserve. Using the system, we are able to solve the Mini Challenge 1 (MC1) tasks to classify claimed pipit calls and detect multiple spatio-temporal migratory patterns in the preserve. We find evidence that the pipit population is declining, but we can not identify any clear evidence that the accused company is responsible for the decrease of the pipit population.

Index Terms: Human-centered computing—Visualization—Visualization application domains—Visual Analytics;

1 Introduction

Mini Challenge 1 of the VAST Challenge 2018 succeeds the previous year tasks to unveil the impact of environmental pollution on the pipits in the Boosong Lekagul Wildlife Preserve. MC1 provides multiple data sources to clarify and reveal the current situation in the wildlife preserve. The tasks include the spatio-temporal analysis of the bird populations and their habitats in the preserve. Furthermore, the accused Kasios Furniture company provided pipit call recordings to prove their innocence.

We propose a Visual Analytics (VA) [1] system to solve the tasks of MC1. The tool provides two main components that combine machine learning techniques, visualization, and human interaction to allow the exploration of the provided datasets. The first component allows examining the temporal and spatial characteristics of the bird populations and habitats in a basic spatial scatter plot with filtering options for bird species and time to identify trends and anomalies.

*e-mail: firstname.lastname@uni-konstanz.de

The second component on audio analysis (Figure 1) helps investigating the Kasios recordings using interactive classification and a novel audio glyph. Using our tool, we are able to detect a decrease in the pipit population and multiple spatio-temporal migration patterns of different bird species. As well, we conclude that most of the bird calls provided by Kasios company are not of the pipit species.

2 Data and Processing

MC1 provides four datasets. The ground truth dataset contains more than 2000 verified bird calls which were collected during the last 35 years. Metadata about the bird calls, e.g., date and time, species, vocalization type (call or song), quality, and geo-spatial position complements the audio data. A bitmap of the preserve for the geographic locations provides the spatial references. The most important task was to verify the 15 claimed pipits call recordings from the Kasios company. The analysis of these data sources is challenging since the recordings vary in quality (e.g., purity, background noise), and in frequency fluctuations in each specific bird species. Further, it is difficult to interpret the spatial patterns due to the variations in the sampling strategy and vocalization types over the years.

To increase the quality of the bird calls, we preprocess the files by removing noise and silent moments from the recordings. Also, we normalize the bird calls using amplification to prepare the bird calls for the classification.

3 Analysis Approach

Our analysis approach includes the extraction of more than 280 million fingerprints from the cleaned calls of the bird library. In general, many audio querying tools use fingerprints to classify audio files, which correspond to a given query snippet, such as Shazam [2].

We incorporate the user in an interactive classification process to categorize the claimed pipit calls of the Kasios company. First, the user selects a claimed pipit call. Consequently, the system displays the selected recording as an interactive spectrogram (see Fig. 1 (A)) that can be used to replay the bird recording. The user can then explore the spectrogram of the claimed pipit call using zooming and panning. The spectrogram enables the user to classify

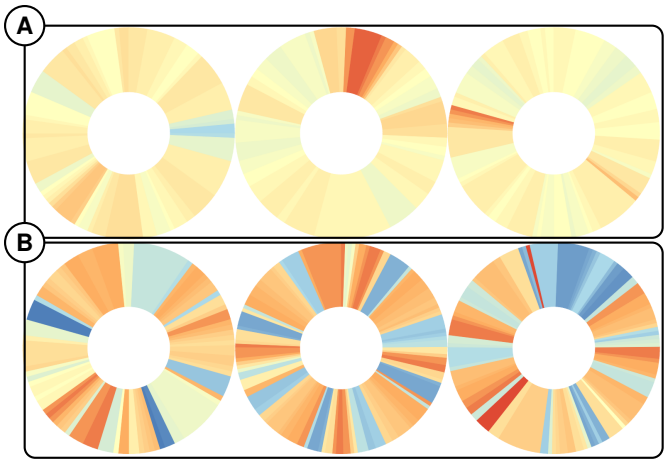

Figure 2: Audio glyphs for two bird species visualizing the spectral roll-off extracted from the cleaned bird calls. (A) shows three pipit audio glyphs and (B) three audio glyphs for the bird species *Qax's*. The audio glyph enables to distinguish bird species from each other.

segments of the audio file using the generated fingerprints of the ground truth. Afterwards, we visualize the classification result in a second interactive spectrogram. The user is then able to examine the classification result replaying both bird calls and comparing the spectrograms to increase the confidence of the classification result. If the result does not match the claimed pipit call, the user can dismiss this classification and try the same or other segments again blacklisting the classification result in the ground truth. For instance, the approach enables us to classify the 8. Kasios file as the *Lesser Birchbeere* bird species.

To further support of the classification and exploration extraction of the bird calls, we extract multiple audio features. We use the spectral centroid (average frequency weighted by the frequencies energy), the spectral bandwidth (variance in the frequency), and the spectral roll-off (the .85 quantile of the frequency sorted by energy) to obtain meta-features of the bird calls. We display the extracted audio features in an audio glyph that encodes the audio features in a radial view. Figure 1 (B) shows the spectral roll-off, but we can use any other meta feature as well. We use the spectral roll-off and a global color scale from to color the segments of the audio glyph, mapping the minimum frequency to blue and the maximum to red. Figure 1 (B) illustrates how different birds (1), (2) and (3) are visible in the spectrogram and the audio glyph. The advantage of the audio glyph is that it aggregates and encodes meta information about the bird calls using less space than the spectrogram while facilitating the detection of commonalities and differences between bird species. For instance, Figure 2 visualizes three audio glyphs (recordings have the same length) for two different species. The differences between these two types of bird species are visible. The pipits (A) are primarily yellow and only slightly red, while Qax (B) has a higher variance in its spectral roll-off. To characterize movement patterns and habitats of the bird species in the preserve over time, we propose a map view using the preserve's bitmap as the background and the location of the bird calls (Figure 3). The view visualizes the spatial evolution of bird species for the identification of patterns and characteristics of each species to detect clusters anomalies. The habitats of the bird populations are all over the map. Some species such as *Vermillion Trillian*, *Queenscoat*, and *Orange Pine Plover* have two non-overlapping settling areas. Other species live in a single area such as *Carries Champagne Pipit*, *Ordinary Shape*, or *Canadian Cootamum*. The majority of the species are residing in the west of the wildlife preserve with overlapping habitats.

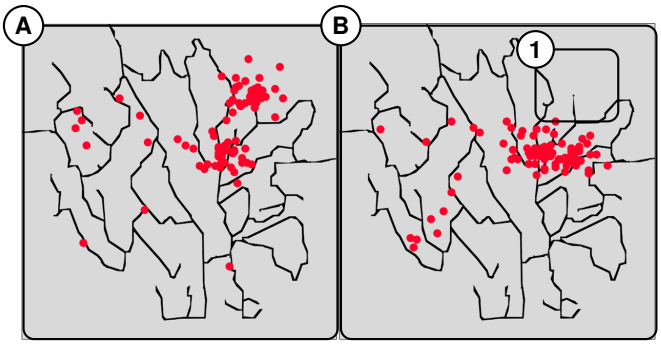

Figure 3: The Rose-Crested Blue Pipits migrated southwestwards in the Wildlife Preserve away from the contaminated area (1). (A) The habitat of the pipits until 2014. (B) The habitat of the pipits after 2014.

Multiple interactions such as highlighting and zooming allow the exploration of the bird data in its spatio-temporal context. The application enables the user to focus on selected bird species by time intervals of interest to visualize the vocalization type, the convex hull for each species, and a heat map. We propose visual mappings that support the analysis of the evolution of changes in bird species characteristics over time in the preserve. For instance, Figure 3 shows that the primary habitat of the pipits was in the north-east until 2014 after which it moved southwestwards. We assume that the pipits are migrating due to the contamination in this area. Figure 3 also illustrates that some pipit recordings are in the other areas in the preserve.

4 Conclusion

The proposed system supports the exploration of all available data sources of the Mini Challenge 1 in an interactive manner. After exploring and analyzing all datasets, we are not able to identify a clear cause for the decline of the pipits. Yet, we are able to identify multiple patterns of the bird populations and their habitats. For instance, the decrease of the pipits and migration away from the contaminated area in the preserve. Overall, the audio recordings from Kasios do not back up their claim that the pipits are thriving and are not affected by any environmental pollution, since most of the claimed pipit call recordings belong to different bird species or provide a recording location outside the known pipit habitats. The fact that most of the alleged pipit recordings belong to other species does not support the claim that Kasios is responsible for the decrease of the pipits. Nevertheless, we suspect that Kasios may have provided misleading information about the thriving of the pipits across the wildlife preserve to distract from other potential negative impacts of their environmental pollution. Using our system we are able to solve the tasks of the challenge. The main contribution of our proposed tool is the interactive classification of the claimed pipit calls using the interactive spectrogram and the novel audio glyph. The audio glyph facilitates the identification of differences and commonalities between bird species. Gathering and investigating further and more exact data is required to obtain deeper insight into the current situation of the bird populations in the preserve and to strengthen the hypothesis that Kasios is responsible for the decline of the *Rose-Crested Blue Pipit*.

References

[1] D. A. Keim, F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler. *Visual Analytics: Scope and Challenges*, pp. 76–90. Springer Berlin Heidelberg, 2008. doi: 10.1007/978-3-540-71080-6_6

[2] A. L. Wang. An Industrial-Strength Audio Search Algorithm. In *Proceedings of the 4th International Conference on Music Information Retrieval*, 2003.