

Kaleidoscope Pro 5

Simple Classifiers

Welcome to Kaleidoscope Pro from Wildlife Acoustics.

This video is the second in a three-part series that describes Cluster Analysis.

There are three types of Cluster Analysis available in Kaleidoscope Pro. The first video demonstrated basic cluster analysis. Basic cluster analysis is used to identify and sort similar vocalizations detected from recordings made by Wildlife Acoustics Song Meters or other recorders.

This second video in the series describes how to create and customize a Simple Classifier for sorting and labelling of new input files. A simple classifier is easy to make and useful for general surveying of present species. The third video in this series will explain how to make an Advanced Classifier. An Advanced Classifier is a little more work to create, but the results are more accurate identification of species with less false positive IDs.

We'll start off by looking at the results of the initial cluster analysis batch process. 36 files were analyzed and the result was 1500 detected vocalizations, which were sorted into 30 separate clusters.

Based on auditioning the vocalizations from the initial cluster analysis scan, I'm going to create customized names for the clusters. I'll look through the first few vocalizations of each cluster to get a sense of what has been recorded. Ok, this vocalization looks like it could be a robin. I'll play the vocalization to confirm. Initially, this cluster is named cluster number 2. I'll click on the Results window, go to the File menu and choose Edit Cluster Name. I'll type in Robin and click ok. Now all the vocalizations in this cluster have been relabeled as robin.

I'll keep going through the clusters. This cluster looks like it could be a cardinal. I'll name that cluster. Then I've got what sounds like some sort of jay. The jay is very distinctive because of the extra harmonics it generates. I'll name the jay cluster.

And here is another cluster that also looks like a jay. Kaleidoscope Pro has found enough differences in this jay compared to the previous vocalization that a new cluster has been created for this variation of the jay. I know it's a jay so I'll name it as such.

If I keep on going through the clusters I see different types of birds and then I get to a cluster that sounds like a cardinal but with a different song than we heard in the earlier example. I'll name that cluster.

It's actually a good idea to look at the first few detections in a cluster, which are closest to the cluster center, to get an idea about the general pattern. Yes, this cluster appears to be all about jays.

Video Script



Here's another version of the Jay. I'll name that cluster also.

And this last cluster looks a little unusual. If I audition I hear the sound of a human voice. I'll label that cluster accordingly.

When I'm finished naming the clusters I care about, I'll go to the Results window File menu and choose Save. The Results window represents an underlying cluster.csv file that was created in the initial batch process. A .csv file is a standard spreadsheet file, containing metadata from the original files and from the cluster analysis scan. The .csv file can also be updated with manually added IDs.

A cluster.kcs file was also created in the initial scan and now that's also been updated with the cluster names.

I'll go back to the batch tab and select my second group of input recordings. I'll select a new output directory. Under the Cluster Analysis tab, I'll select Use existing .kcs file to sort new recordings and create new cluster.csv file. I can press the browse button and direct Kaleidoscope Pro to the cluster.kcs file that was created and then edited from the first scan. The purpose of the cluster.kcs file is that it can be used as a classifier when analyzing new recordings.

Let's pause and review what we've done so far. An initial batch of files was analysed for vocalizations. Vocalizations from the files were sorted into clusters, based on similarity. This was the basic cluster analysis.

Next, we examined the clusters created from the first batch process. We found different species of birds and we renamed the clusters to represent those species. The newly edited cluster names were written to a .csv spreadsheet file, and also to a cluster.kcs file. The cluster.kcs file is used as a simple classifier to sort new input files.

What I have now done is selected the new batch of input files and I've set the Cluster Analysis mode to use the edited .kcs file as a classifier. My goal is to sort the vocalizations in the new batch of input files using the simple classifier I created based on the first batch of files.

I'll press the Process Files button and Kaleidoscope Pro analyzes the second set of files using the cluster.kcs simple classifier. Once the process is complete the Viewer and Result window open. These listed vocalizations have been created from the second set of files. They've been grouped into clusters based on the clusters created from the first set of input files. The new clusters also have the names that we assigned.

Here is a cluster named cardinal. If I look through the cardinal vocalizations I see Kaleidoscope Pro did a pretty good job at finding all the cardinals in this new batch of recordings. Here's the cluster I named as Human. And if I keep scrolling through the vocalizations I eventually find vocalizations labelled as robins. You can also see the clusters I didn't bother to name from the first batch. The clusters are listed alphanumerically. I can now compare for example, how many robins were found in the second batch of recordings as compared to the first batch. My simple classifier has already saved me a great deal of time sorting the vocalizations from the second batch of files and any other files that may come after

Video Script



that.

Thank you for watching part two of the Quickstart for Clustering videos. Part three will address how to create and tune an Advanced Classifier to discriminate and find specific vocalizations from new input files.