

# Kaleidoscope Pro 5

## Cluster Analysis de Base

Bienvenue sur Kaléidoscope Pro de Wildlife Acoustics.

Cette vidéo mettra l'accent sur la préparation à l'analyse de Clusters. L'analyse de cluster est un outil puissant pour trier et identifier les vocalisations basées sur des signatures acoustiques similaires.

L'analyse en clusters peut être utilisée pour les chauves-souris et d'autres animaux. Cette vidéo utilisera uniquement des oiseaux pour ses exemples.

Kaleidoscope Pro propose trois types d'analyses de clusters disponibles. Commençons par le Cluster d'analyse de base.

Sous l'onglet Lot, je vais parcourir pour définir un répertoire d'entrée. Pour cette démonstration, j'ai deux séries d'enregistrements qui ont été faits au même endroit à des moments différents. Je choisirai le dossier contenant le premier ensemble d'enregistrements à analyser. Ce dossier contient 36 fichiers audio et chaque enregistrement dure une minute ou moins. En une minute, il est possible d'enregistrer plusieurs oiseaux ou plusieurs chansons du même oiseau. Il est donc probable que ces enregistrements contiennent de nombreuses vocalisations d'oiseaux distinctes.

Je sélectionnerai un répertoire de sortie vide pour le premier ensemble de résultats d'analyse de cluster.

Kaleidoscope Pro n'a pas besoin de convertir ni de créer de fichiers audio pour la mise en cluster. Pour gagner du temps et de l'espace disque, je décocherai l'option permettant de créer ces fichiers de sortie.

Avant que l'analyse de cluster ne puisse être effectuée, Kaleidoscope Pro exécutera les fichiers d'entrée selon les conditions configurées sous l'onglet Paramètres de signalisation. Il existe différents paramètres de signal pour le mode analyse chauve-souris et le mode analyse non chauve-souris. Comme je travaille avec des enregistrements d'oiseaux, je m'assure que Kaléidoscope Pro est en mode d'analyse sans chauve-souris.

Les paramètres par défaut de l'onglet Signal Params sont généralement un bon point de départ. Ces paramètres décrivent le fait que Kaléidoscope Pro recherchera dans les fichiers des signaux dont la fréquence est comprise entre 250 Hz et 10 kHz. Le second paramètre recherche des signaux continus d'une durée comprise entre 0.1 et 7.5 secondes. Kaléidoscope Pro est actuellement configuré pour rechercher un écart de 0.35 seconde entre les syllabes. Si un signal

## Script Vidéo

---

remplit les deux premières conditions et est ensuite suivi de 0.35 seconde ou plus de silence, Kaléidoscope Pro considérera ce segment de l'enregistrement comme un signal détecté unique.

Je vais maintenant cliquer sur l'onglet Cluster Analysis. Dans le menu, je sélectionne les enregistrements d'analyse et de cluster pour créer cluster.kcs et cluster.csv. C'est la première étape d'une opération de clustering de base.

Pour le premier traitement par lots de l'analyse de cluster, je vais laisser la plupart des paramètres par défaut, mais je vais m'assurer que tous les signaux détectés seront inclus dans la sortie d'analyse de cluster. Je vais définir la distance maximale depuis le centre du cluster afin d'inclure les sorties de cluster.csv sur 2.

Maintenant, je vais appuyer sur le bouton Process Files et Kaleidoscope Pro se mettra au travail. Les fichiers d'entrée sont examinés à l'aide des paramètres de l'onglet Signal Params. L'analyse de cluster examine ensuite les vocalisations détectées et les trie en clusters.

Ok, la première étape d'analyse de cluster est maintenant terminée. La fenêtre Viewer and Results s'ouvre automatiquement. Je vais organiser ces fenêtres afin d'avoir une vue large du spectrogramme à gauche et une liste des vocalisations détectées à droite. Ce que nous voyons sont des signaux individuels ou des vocalisations détectées qui ont été dérivés des fichiers d'entrée. Vous rappelez-vous que nous avons commencé avec 36 fichiers d'entrée? Si je fais défiler l'écran vers le bas, je constate que Kaléidoscope Pro a détecté près de 1500 vocalisations distinctes au sein de ces 36 fichiers.

Je constate également que 30 clusters au total ont été formés, numérotés de zéro à vingt-neuf. Sur les 36 fichiers d'origine, 1 500 vocalisations ont été trouvées et ces 1 500 vocalisations ont été regroupées en un total de 30 clusters distincts.

Les clusters ont été nommés avec des nombres. Le nombre le plus bas représente le cluster ayant les vocalisations les plus similaires.

Les vocalisations au sein de chaque cluster ont été triées de sorte que la voix la plus proche du centre du cluster se trouve en haut de la liste. La vocalisation en haut de la liste est le meilleur exemple du cluster le plus courant. Je peux utiliser la visionneuse pour examiner les vocalisations. Si je passe à la prochaine vocalisation de la liste, vous allez constater qu'elle est très similaire à la précédente.

Je peux continuer à avancer à travers les vocalisations dans ce cluster et vous pouvez voir qu'elles sont toutes très similaires.

Je vais appuyer sur le bouton Play pour écouter rapidement le premier signal détecté. Je pense que c'est une mésange. Je peux vérifier l'horodatage et je vois que cette vocalisation a été enregistrée le 30 avril 2017 à 13h50. Je vérifierai une vocalisation à partir d'un fichier d'entrée différent, mais à l'intérieur du même cluster. Cette vocalisation a été enregistrée le 30 avril 2017 à 14h. Donc, cela pourrait être la même mésange dix minutes plus tard. Voici une autre

## Script Vidéo

---

vocalisation du même cluster, mais un fichier d'entrée différent. Je peux voir que cet enregistrement a été réalisé le 14 avril 2017 à 6h30. Ce pourrait être la même mésange ou ce pourrait être une mésange différente. Mais maintenant, je commence à comprendre qu'un oiseau commun dans ces enregistrements est la mésange, au moins avec cette vocalisation spécifique et sur la base des informations contenues dans les fichiers d'entrée originaux et les clusters trouvés.

J'ai un ensemble de boutons séparé qui permet de passer au premier fichier du cluster suivant.

Si je continue d'avancer dans les clusters, je constate que chaque cluster représente un oiseau différent ou des variations du même oiseau. Si une seule espèce possède une variété de chansons, celles-ci peuvent être classées en différents clusters.

C'est le premier niveau d'analyse de cluster. Kaleidoscope Pro a détecté des vocalisations dans des fichiers audio et les a triées par similarité. La fenêtre Résultats représente un fichier cluster.csv sous-jacent. Le fichier cluster.csv peut être ouvert dans un tableur pour une analyse ultérieure. Nous n'avons pas encore identifié d'espèce particulière, à l'exception peut-être de la mésange, mais nous avons regroupé les vocalisations similaires et, selon ce que je cherche à obtenir pour les résultats finaux, c'est peut-être tout ce dont j'ai besoin pour obtenir un aperçu de la diversité des vocalisations détectées sur ce site.

Merci d'avoir regardé la première partie du Quickstart pour la mise en cluster de vidéos. La deuxième partie explique comment utiliser un classificateur simple afin de pouvoir classer et étiqueter les vocalisations de nouveaux fichiers d'entrée.