Instrumenten-Klassifizierer

# Idee / Aktueller Stand

Ein Audiostream (Mikrofon-Input im Live-Modus oder ein .WAV-File) wird in einem 1s-Raster klassifiziert. Jedem Zeitintervall (Window) wird dann genau ein vermeintlich erkanntes Instrument zugeteilt, sobald das Signal eine bestimmte Lautstärke-Schwelle (Threshhold) übersteigt.

# Umsetzung

Für das erste Projekt wählten wir Gitarre und Schlagzeug, eine Kombination, von der wir uns erhofften eindeutige Unterschiede in den Signalen erkennen zu können.

## Daten

Als Trainingsdaten wurden für jedes Instrument jeweils 100 Audiofiles in Studioqualität aus eigenen Projekten und Sample-Libaries verwendet. Somit sind bei der relativ kleinen Menge an Daten schon unterschiedliche Umgebungen vertreten. Es ist jeweils nur ein Instrument zu hören, nie beide. In einem zugehörigen CSV-File sind die Labels (in diesem Fall „Drums“ und „Guitar“) für alle 200 Dateien gespeichert.

## Feature extraction

Für jedes 1s-Fenster wurden 27 charakteristische Features des Signals extrahiert, die für die anschließende Klassifizierung dienten. Neben üblichen Signaleigenschaften wurden speziell auch Mel Frequency Cepstral Coefficients (MFCC) aus der Spracherkennung verwendet. (TODO: genauere beschreibung von MFCC oder egal?)

Es wurden stets alle 27 Features verwendet. Dabei ist anzunehmen, dass nicht alle eine für die Unterteilung sinnvolle Information enthalten. Eine Reduktion auf wenigere aussagekräftige Merkmale sollte noch bessere Ergebnisse erzielen.

## Klassifizieren

Mit der *Classification Learner-App* aus der *Statistics and Machine Learning Toolbox* konnten verschiedene Algorithmen verglichen und anschließend die Modelle exportiert werden. Die besten Ergebnisse lieferten Decision Trees und Support Vector Machines. Für die Klassifizierung neuer Daten werden wieder die gleichen Features exportiert und der predict-Function des Models übergeben.

Ein möglicher nächster Schritt wäre hier ein Vergleich mit Neural Networks für die Mathworks mit der *Deep Learning Toolbox* Implementierungen bietet.