AUTOMATIC MUSIC TRANSCRIPTION 2016

Nora Huang
Department of Computer Science
University of Victoria
Victoria, B.C, Canada
norah@uvic.ca

Aazim Lakhani
Department of Computer Science
University of Victoria
Victoria, B.C, Canada
aazimlakhani@uvic.ca

Parul Smaddar
Department of Computer Science
University of Victoria
Victoria, B.C, Canada
author3@ismir.edu

ABSTRACT

The objective of this project is to evaluate the accuracy of various models to perform automatic music transcription by extracting features from the audio music and applying different classifiers to get the note. Through this initiative we plan to focus at transcription on certain music instruments. Furthermore, only the main instrument of the polyphonic music will be transcribed.

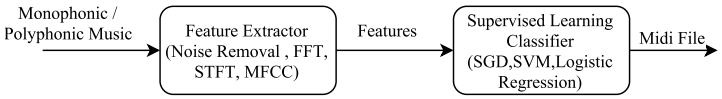
1. INTRODUCTION

The dataset for this project will collected through MIDI Aligned Piano Sounds (MAPS) , Isophonics Annotations , KSN Database to name a few. The sources claim the data is used for transcriptions of both monophonic & polyphonic instruments. The datasets are annotated & we would evaluate the data to use music relevant for this project.

Then acoustics features such as Zero Crossing Rate, Autocorrelation, YIN metrics for time domain & FFT , STFT , Spectrogram , cestrum metrics, spectral flux for the frequency domain would be used to extract notes from audio music.

The extracted features will be used as input to the classifiers for note transcription. Different classifiers, such as K-NN, Logistic Regression, Stochastic Gradient Descent, Support Vector Machines (SVM) would be evaluated for accuracy. Features would be combined together to give better accuracy on real world data.

Marsyas would be used as the framework for acoustic features extraction. Python & Scikit-learn would be used to prepare the learning model, which would eventually, transcribe.



System Flow Diagram

1.1 Timeline

Dataset collection by
Features and Classifier selection based on papers by
Features extraction implementation by
Classifier implementation by
System integration for training by Run the test and get result by
Report should be update after every previous steps

1.2 Role of team member

Nora: Architecture design, team management, coding
Aazim: Training & testing with different combinations of classifiers & features
Parul:

2. DATASET

We obtained data for both monophonic and polyphonic sounds through MAPS. We're also in the process of having data from Isophonics Annotations & KSN Database.

3. ACOUSTIC FEATURES FOR TRAINING

Time Domain : Zero Crossing Rate, Autocorrelation, YIN metrics Frequency Domain : FFT , STFT , Spectrogram , cestrum metrics, spectral flux.

4. CLASSIFIERS

Stochastic Gradient Descent, Logistic Regression, Support Vector Machines

5. RESULT

This would comprise of the collection of features, classifiers used to give the highest level accuracy. Various combinations would be tried, but only the one's that generalize well, would be mentioned in the result.

6. CONCLUSION

Basic on the result we should be able to figure out which combination is best for automatic music transcription.

7. REFERENCES