Musical Transcription of Drum Patterns Using Main Audio Features in KNN

by

Charles Jayson L. Dadios

BS Computer Science

Introduction/Abstract

Drums is a prominent instrument in our musical culture. The most accurate way to learn how to play the drum parts of a song is by reading a drum musical chart. Learning and familiarizing to read drum notation can be an easy skill to pickup, but learning how to write charts takes more time and effort. This challenge encourages drummers to learn songs by ear, but people tend to forget the correct parts and play inconsistently.

Automatic Drum Transcription (ADT) can encourage drummers to read charts, and reading unlocks learning. An algorithm was developed to transcribe drum recordings into two formats: MIDI and PDF. KNN was used to classify what particular instrument of the drums sounded off. The classifier used the drum's main audio features as a 77-dimensional feature vector. The algorithm yielded promising results to offer solution to the ADT problem.

Introduction

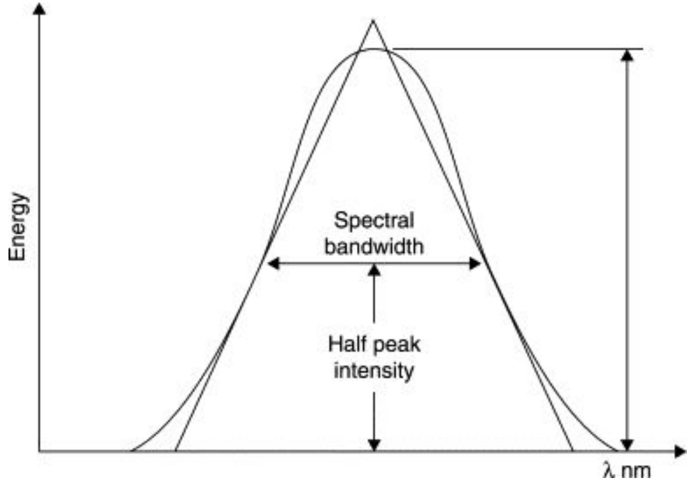
Automatic Drum Transcription (ADT) in simple terms is the process of converting a drum performance into a record usually as a drum notation printed as a music sheet.

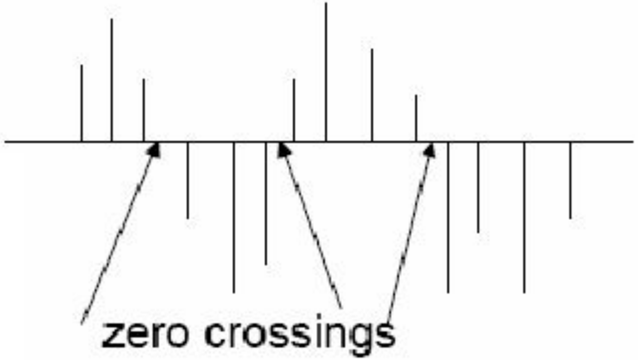
KNN is a classification algorithm which uses data points to find the k-number of closest classes as the basis of prediction.

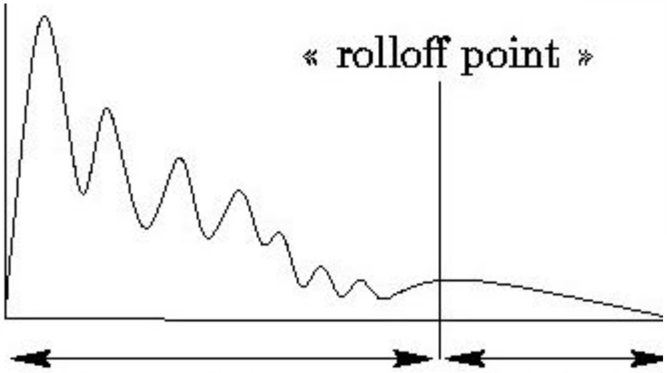
The feature vector used seven (7) main audio features

Audio Features

1. Spectral bandwidth -
2. Spectral contrast* - “ Spectral contrast is defined as the level difference between peaks and valleys in the spectrum”
3. Zero crossing rate -
4. MFCC* - “human perception graph”
5. Spectral centroid - spectral center of mass
6. Spectral flatness - tone-likeness
7. Spectral rolloff -







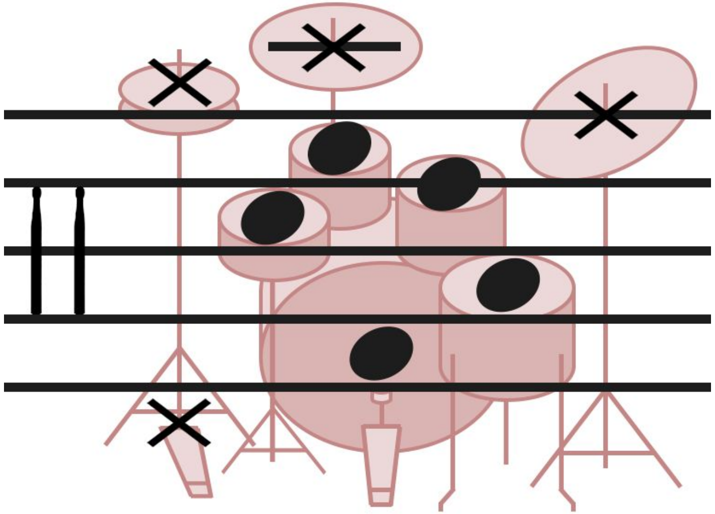
Objectives

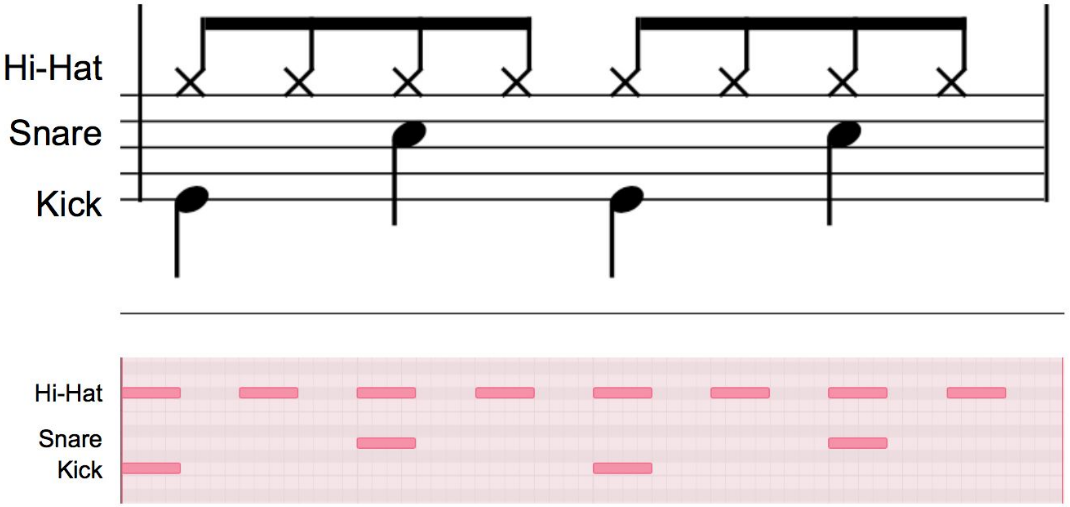
The general objective of this proposal is to develop an application which can produce a musical transcription of drum beat patterns from digital drum audio samples.

Specific objectives:

1. To produce physiologically realistic drum sound combinations using downloaded drum samples
2. To extract main audio features from the data set
3. To classify the sounding instruments at given time intervals
4. To evaluate the accuracy of the produced transcription

Billy Jean





1 & 2 & 3 & 4 &

Methodology

Create samples for each class

Extract audio features

Create a classifier

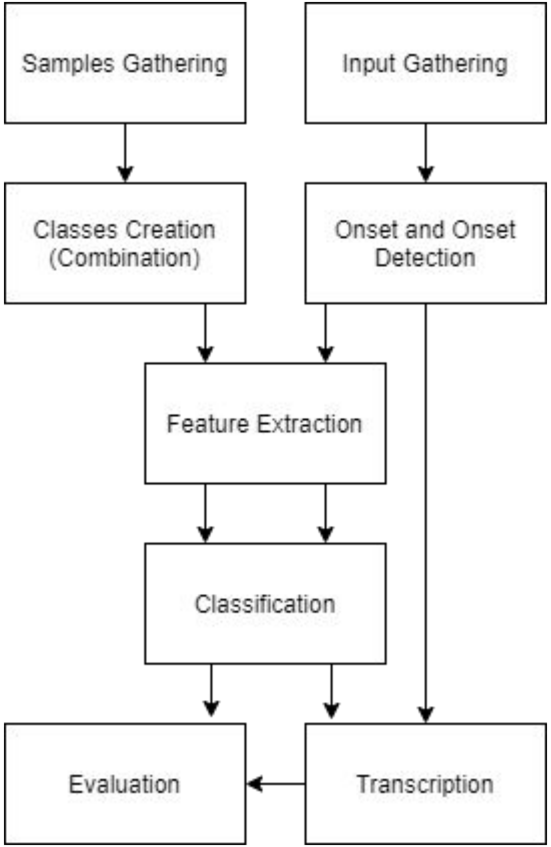
Evaluate classifier

Capture onsets and identify the timing of the input

Predict onset classes

Produce transcription

Evaluate transcription



Samples Gathering

Input Gathering

Classes Creation (Combination)

Onset and Onset Detection

Feature Extraction

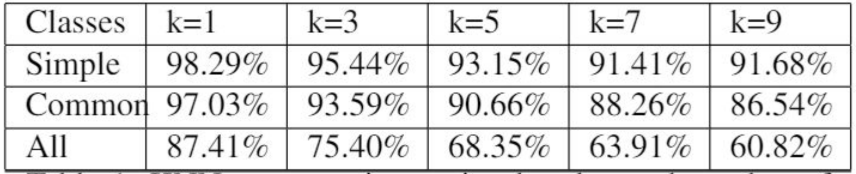
Classification

Evaluation

Transcription

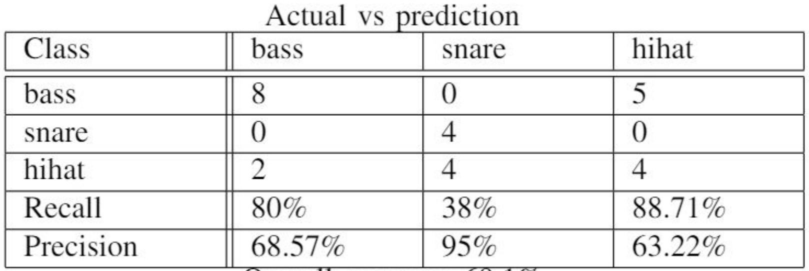
Results

Classifier (1) and Transcription (2, 3)



Classes	k=1	k=3	k=5	k=7	k=9
Simple	98.29%	95.44%	93.15%	91.41%	91.68%
Common	97.03%	93.59%	90.66%	88.26%	86.54%
All	87.41%	75.40%	68.35%	63.91%	60.82%

Table 1: KNN accuracy in varying k value and number of classes

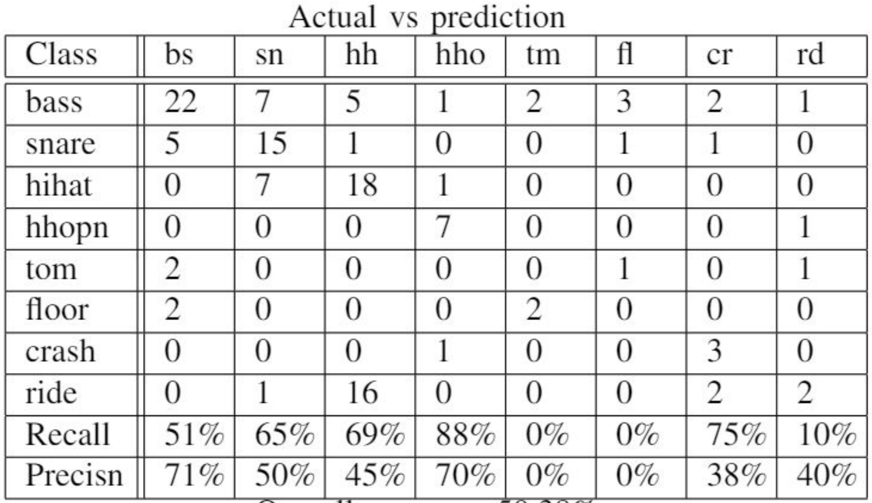


Actual vs prediction

Class	bass	snare	hihat
bass	8	0	5
snare	0	4	0
hihat	2	4	4
Recall	80%	38%	88.71%
Precision	68.57%	95%	63.22%

Ovarall accuracy 69.1%

Table 2: Confusion matrix of bass, snare and hihat transcriptions



Actual vs prediction

Class	bs	sn	hh	hho	tm	fl	cr	rd
bass	22	7	5	1	2	3	2	1
snare	5	15	1	0	0	1	1	0
hihat	0	7	18	1	0	0	0	0
hhopn	0	0	0	7	0	0	0	1
tom	2	0	0	0	0	1	0	1
floor	2	0	0	0	2	0	0	0
crash	0	0	0	1	0	0	3	0
ride	0	1	16	0	0	0	2	2
Recall	51%	65%	69%	88%	0%	0%	75%	10%
Precisn	71%	50%	45%	70%	0%	0%	38%	40%

Ovarall accuracy 50.38%

Table 3: Confusion matrix for commonly used classes

Conclusion

Using main audio features in KNN is effective in Automatic Drum Transcription (ADT), with the highest overall accuracy achieved when k=1, even with the feature vector dimension of 77, in this case.

Recommendation: Just like speech recognition, ADT may improve with the help of other machine learning techniques like Hidden Markov Model, or Neural Networks.

References

Spectral bandwidth image - https://ars.els-cdn.com/content/image/3-s2.0-B9780857092298500061-f06-15-9780857092298.gif

Zero crossing image - https://www.researchgate.net/profile/Buket_Barkana/publication/259828967/figure/fig1/AS:299377940811777@1448388671414/Definition-of-zero-crossings-rate.png

Spectral rolloff image - https://www.researchgate.net/profile/Julien_Pinquier/publication/278631244/figure/fig5/AS:669382405545993@1536604610274/Definition-du-Spectral-Rolloff-Point.ppm