

Draft Report for MIR Project

Date: 17 March 2019

Title: Tabla bol recognition

Authors: Blazej Kotowski and Ninad Puranik

Dataset used : <https://compmusic.upf.edu/tabla-solo-dataset>

Abstract:

The tabla is a pair of drums known as ‘tabla’ and ‘bayan’ commonly used in accompaniment and solo performances in Indian Music. A variety of sounds can be produced using different kinds of strokes on the tabla-bayan. A mnemonic syllable or bol is associated with each of these strokes. We address the problem of recognizing the bol performed in an audio recording of a tabla solo performance with a harmonium accompaniment. We use a supervised learning approach to classify the ‘bols’ using the timbral features of audio and the manual annotations from the compmusic tabla-solo dataset.

Previous work:

S. Gupta et.al. [1] have developed the dataset comprising audio recordings, scores and time aligned syllabic transcriptions for 38 tabla solo compositions of different forms in tīntāl (a metrical cycle of 16 time units). The compositions are from the instructional video DVD [Shades Of Tabla](#) by [Pandit Arvind Mulgaonkar](#), who is among the most renowned contemporary tabla maestros. They have approached the problem using a standard speech recognition model with HMM based pattern recognition. Gillet and Chordia have worked with a dataset with pure tabla sounds without harmonium accompaniment. Gillet has used power spectra of sounds while Chordia has used 31 manually selected timbral features (which include MFCCs)

We approach the problem differently by extracting large number of features and then training different classifiers to get the best model for bol transcription. The feature extraction is done using essentia. In addition to the standard mfccs, we also extract the first and second derivatives of the mfccs to help us get features which captures the time varying spectral changes.

Critical view of previous work and dataset :

Dataset: Dataset is well organized and documented. Onsets of tabla bols seem to be accurately marked and annotated. Mapping of similar sounding bols with different names is accurately provided. The dataset reflects a close to real life scenario due to the harmonium accompaniment.

The audio files are all 44100 fps 16-bit PCM Mono. This makes it difficult to do work with the separate tabla or tabla mixed with harmonium. Distribution of bols across categories is not uniform. Some bols are not sufficiently represented

Code: The code developed by S. Gupta et. al is accessible at <https://github.com/swapnilgt/percPatternDiscovery>. Code could not be installed and executed as per the instructions provided in the github. The code may be complete and

reproducible but was not easy for the authors to fully understand in the available time-frame for this work. This was one of the reasons to adopt a different methodology rather than extend the work of by S. Gupta et. al.

Methodology :

- 1) Collect 100 ms slices of audio starting from each annotated onset representing a bol.
- 2) For each audio slice extract variety of features. Use essentia extractor for this purpose. Extract HTK model based MFCC features using own extractor
- 3) Create a new dataset with the ground-truth 'bol' annotation from the tabla-solo-dataset and the extracted features.
- 4) Train a classifier for this dataset for bol-recognition task.
- 5) Identify best features using feature selection models and obtain best possible classification, as measured by standard classification metrics like 'correctly classified instances' and 'precision, recall and F-measure'

The code and results obtained using this approach can be accessed at

https://github.com/PNinad/tabla_bol_recognition_MIR_Project/

Summary of results:

Features, classifier and parameters	Percent accuracy*
KNN	60
SVM	72
Decision Trees	55

Conclusion: The best classification was obtained using SVM classifier .

Future Work:

Use onset detection coupled with this trained model to do an end-to end tabla transcription.

References :

- S. Gupta, A. Srinivasamurthy, M. Kumar, H. A. Murthy, X. Serra. Discovery of Syllabic Percussion Patterns in Tabla Solo Recordings. In Proc. of the 16th International Society for Music Information Retrieval Conference (ISMIR), 2015.
<http://compmusic.upf.edu/ismir-2015-tabla>
- P. Chordia. Segmentation and recognition of tabla strokes. In Proc. of the 6th International Conference on Music Information Retrieval (ISMIR), pages 107–114, London, UK, September 2005. <http://ismir2005.ismir.net/proceedings/1137.pdf>

*Draft report, table to be updated with exact parameters