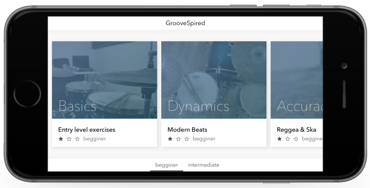
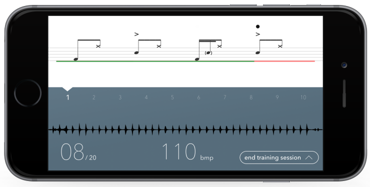
BRNO UNIVERSITY OF TECHNOLOGY FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2017
http://excel.fit.vutbr.cz

GrooveSpired - Application for Drums Training with Beat Recognition

Bc. Tomáš Štrba





Abstract
The main goal is implementation of mobile application for drums training. The application must be capable to analyze and evaluate drumming skills of a drummer using drum beats recognition of recorded audio. Other features of the application are displaying drum notation of different grooves from various music styles and playing audio examples of those grooves. For time-frequency analysis discrete wavelet transform is used in this paper. Results show a true positive rate above 96% with reasonable tolerance in the onset detection.

Keywords: mobile application, iOS, drum beats recognition, discrete wavelet transform, audio analysis, drum training, GrooveSpired

Supplementary Material: App Store Link

xstrba01@stud.fit.vutbr.cz, *Faculty of Information Technology, Brno University of Technology*

Every person has experienced how hard is it to learn new ability at least once in a life. We go through the same feelings when we learn to play musical instruments. There is a lot of frustration and nervousness. In case of other people's attention, for example the one of a music teacher, feelings are even stronger.

The main idea of the project is to design and implement a mobile application, which is capable of replacing drum teacher at all. The app should provide enough materials for exercises, analyze drum play in a real time and give a feedback to user.

From the conducted study, it was determined that there is no similar learning mobile application with drum beat recognition in the world. Therefore, there is a big opportunity to leave fingerprints in this area of business.

Important thing within the human learning process is feedback. If GrooveSpired wants to give a feedback to a drummer, it needs to record and analyze his play on the drum set. The second important task is a presentation of this analysis. Great user experience was number one priority in designing and prototyping of the application.

1. Beat Detection using Discrete Wavelet Transform

This section explains drum beats detection on basic parts of the drum set: bass drum, snare drum and hi-hat.

Bass drum and snare drum are membranophones constructed from corpus and drum heads. Heads are connected to the top and the bottom of the drum. There are plenty of different corpus sizes that lead to different drum pitch. Pitch is influenced also by tension of the drum head. An important fact for our drum record analysis is that membranophones have most of their spectral energy contained in regions of the frequency spectrum under 1000 Hz. Bass drum has the lowest

Figure 1. Bass drum is the biggest drum in front of the set. Hidden behind him is snare drum and hi-hat cymbal is on the right side of the drum set.

spectrum[1]. On the other hand, cymbals produce the highest frequency signal from all musical instruments.

1.1 Discrete Wavelet Transform

Discrete wavelet transform (DTW) is signal transformation used for time-frequency analysis. It is a good alternative to Short Time Fourier Transform (STFT). The key difference between these two transformations is an information about the time. STFT provides uniform time resolution for all frequencies. On the other hand, DWT provides high time resolution and low frequency resolution for high frequencies and high frequency resolution and low time resolution for low frequencies[2].

This approach is more effective and used in statistical pattern recognition quite often where feature vectors are computed from DWT coefficients. The other usage of DWT is presented in this work.

Discrete wavelet transform is defined [2]:

$$W(j,k) = \sum_j \sum_k x(k) 2^{-j/2} \Psi(2^{-j}n - k) \tag{1}$$

where $\Psi(t)$ is function called ”morlet wavelet”.

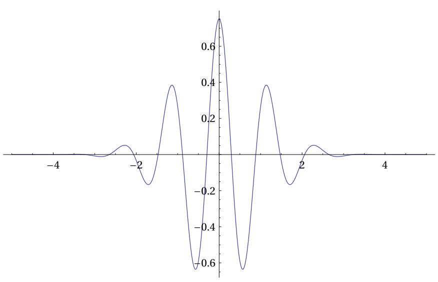


Figure 2. Morlet wavelet function

The result of DWT is a decomposition of the input signal into detailed and approximation coefficients. This is achieved by highpass and lowpass filtering of the signal time domain described in following equations:

$$y_{high}[k] = \sum_n x[n] g[2k - n] \tag{2}$$

$$y_{low}[k] = \sum_n x[n] h[2k - n] \tag{3}$$

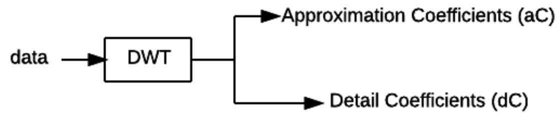


Figure 3. Illustration show how DWT transforms data into detail and approximation coefficients

Coefficients from discrete wavelet transform are a compact representation of the input signal in time and frequency.

1.2 Audio Analysis of Drum Beats with Discrete Wavelet Transform

This subsection describes extraction of the input signal for different parts of drum set. To decompose the signal into different frequency bands, we use a cascade of DWTs [3]. Cascade is shown in picture 4.

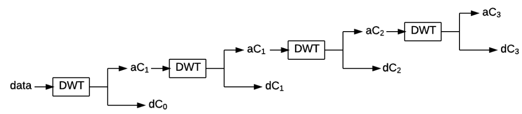


Figure 4. DWT filters cascade

Picture 5 shows audio diagram and music notation of basic drum groove.

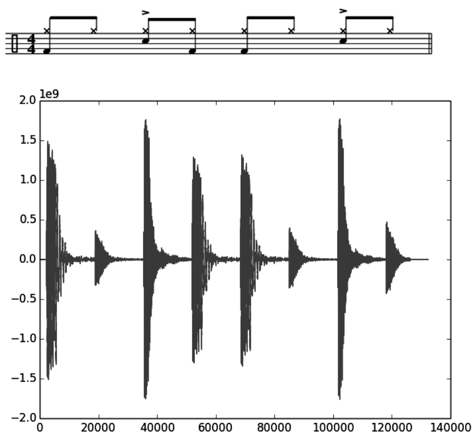


Figure 5. Audio record with its music notation. Hi-hat strokes are marked with an “X”. The bass drum is represented in the last row and snare drum is between them.

After application of DWT cascade we get detailed and approximation coefficients for each frequency band. Some frequency bands may contain a signal of the drum set instrument. Relation between band and instrument is defined by experiments. For hi-hat cymbal it is a frequency band with interval *11025 Hz - 22050 Hz*, for snare drum *689 Hz - 1378 Hz* and bass

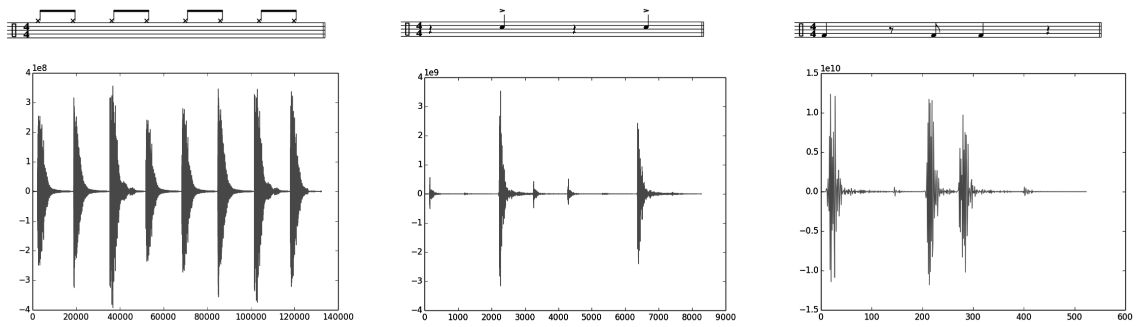


Figure 6. The first picture shows audio signal produced by hi-hat cymbal, in the middle we can see two snare drum strokes separated and last picture contains bass drum strokes.

drum *43 Hz - 86 Hz*. To convert the signals back to the time domain, coefficients for all the bands except the chosen one are set to zero and then we apply the Inverse Wavelet Transform. Picture 6 shows a separation of the basic drum groove presented above.

1.3 Onset detection

Exact times of drummer hits are computed with an onset detection. For every signal we must consider acoustic properties of instruments, environment, various decays and also signal noise.

Steps of onset detection:

1. Signal rectification

$$y[n] = abs(x[n]) \tag{4}$$

2. Sub-sampling

$$y[n] = x[kn] \tag{5}$$

3. Sliding window with moving average

$$y[n] = 1/k \sum_{k} x[n+k] \tag{6}$$

4. Thresholding t

$$t = \max[x] * 0.7 \tag{7}$$

5. If these conditions are true:

$$x[n] < t \wedge x[n+1] > t \tag{8}$$

   there is onset on position n

1.4 Experiments

At first, method presented in this paper was implemented in Python scripting language. Several MIDI groove files of different music styles (funk, rock and metal) were created for testing. The first module of this testing script synthesized all MIDI grooves into the audio files. It used real drum sounds and added a hall effect to make records as real as possible. Next module of the testing script analyzed record files and the last module compared hit times of MIDI files and detected hit times with variable time tolerance.

Following plots display accuracy and true positive rate for funk, rock and metal grooves. The x axis represents tolerance of onset detection in hundredths of a second. Variables are computed from true positives, false positives and false negative values. True negative values are not considered in this work, because silence is not important for this application.

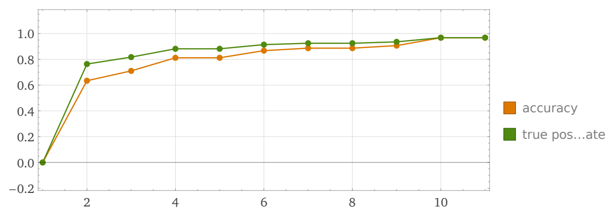


Figure 7. Funk grooves results

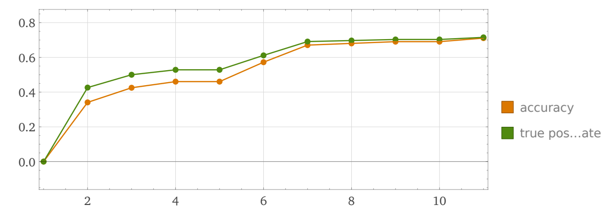


Figure 8. Metal grooves results

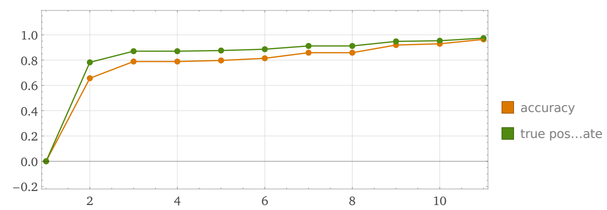


Figure 9. Rock grooves results

These results shows that method is suitable for beat recognition of grooves containing basic parts of the drum set. Experiments also suggested reasonable tolerance - 0.1 of a second - which will be used in iOS mobile application. Metal grooves shown the worst results, because of fast bass drum strokes. The best result is an accuracy value 0.97% for funk music style with tolerance 0.11 of a second.

2. User Interface of the GrooveSpired

This section describes user interface of iOS mobile application. Due to the content size limit no implementation details will be specified.

Basic view of the whole app is menu which is shown in the picture 10. It contains groove sets specialized in different music styles or drumming techniques. Every groove set has its own difficulty.

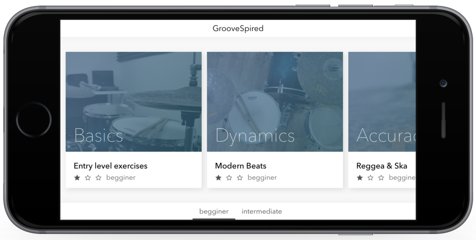


Figure 10. Application menu

After choosing the groove set from the menu user can see and hear containing grooves. There is also an option of changing the tempo. Groove notation is located in the upper area of the exercise view. It is rendered from an internal representation of grooves and animates according the play.

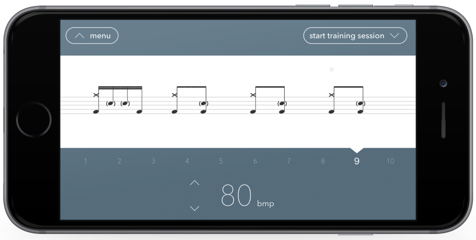


Figure 11. Groove set preview

"Start training session" button activates a training session. It starts with counting of 4 beats so the drummer has appropriate time for preparation. While the application is in the training session mode, it records an audio from the microphone, analyzes it in background threads and calculates good and bad parts of the groove. These areas are displayed in form of background color under the drum notation as feedback.

Figure 12. Training session mode

In the end of training session results view is presented to the user. It contains information about his score in accuracy, speed, difficulty and stamina.

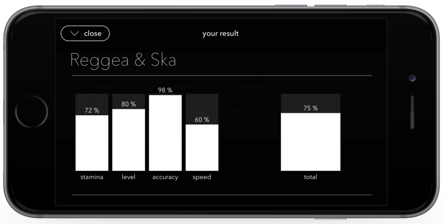


Figure 13. Results view

3. Conclusion

This work describes iOS mobile application for drums training with beat recognition. It involves a method used for analysis of recorded audio too.

Beat recognition algorithm was tested using MIDI synthesis with best result of accuracy above 96% in funk music style with onset detection tolerance 0.1 of a second.

After these tests whole algorithm was implemented in iOS environment (Objective-C and Swift) and packed to the simple application. In this application a drummer can choose between various exercise sets, look at the notations, listen exercises and most important practice these exercises with application real time feedback.

Future steps of development process will be focused on the drummer's personality. The GrooveSpired will contain a profile section with history of exercises. Important step will be also marketing and expansion of the application to wide audience of drummers. As application name suggests, we need to inspire by the groove.

Acknowledgement

The GrooveSpired is also my master thesis and I would like to thank my thesis advisor Ing. Igor Szőke, Ph.D. for his willingness, time, help, patience and steering me in the right direction.

References

[1] N.H. Fletcher and T. Rossing. *The Physics of Musical Instruments*. Springer New York, 2008.

[2] George Tzanetakis, Georg Essl, and Perry Cook. Audio analysis using the discrete wavelet transform. In *in Proc. Conf. in Acoustics and Music Theory Applications. WSES*, 2001.

[3] George Tzanetakis, Ajay Kapur, and Richard I. Mcwalter. Subbandbased drum transcription for audio signals. In *in IEEE International Workshop on Multimedia Signal Processing*, 2005.