**SGN -14007**

**INTRODUCTION TO AUDIO PROCESSING**

**Separation of Drums from music signals**

**Group members**

**Vishal Gaur (281683)**Vishal.gaur@tuni.fi

**Ali Gohar (281668)**  
Ali.gohar@tuni.fi

Problem Definition

For detecting the drum signals, modifying the music components and multi pitch analysis, it is often required that the audio signal should be separated into various components like harmonic and percussive. Therefore, the focus of our project is to separate the music signal into percussive and harmonic components by designing an algorithm.

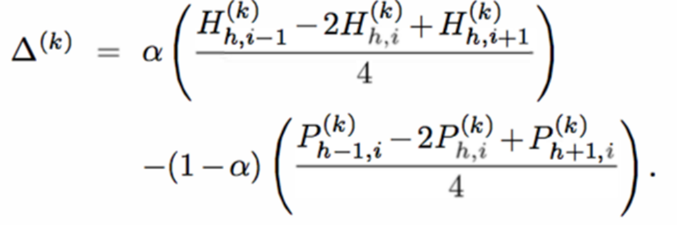
Implementation

To start with the problem, we follow the algorithm provided in the research paper -

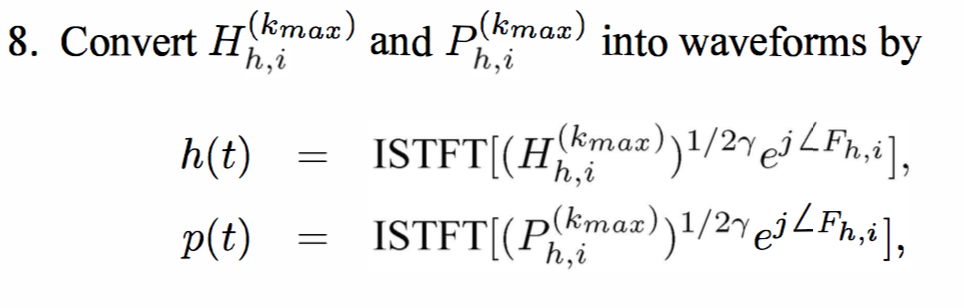
1. Calculate Fh,i ,the STFT of an input signal f (t).
2. Calculate a range-compressed version of the power spectrogram by-

Wh,i = |Fh,i|2γ (0 < γ ≤ 1).

1. Set the initial values to half of W for all h and i and set k=0.   
   i.e. H(0)h,I = P (0)h,I = ½ Wh,i
2. Calculate the update variables-



1. Update the Hh,i and Ph,i as-   
     
    H(k+1)h,i  = min(max(H(k)h,i + Δ(k),0),Wh,i),   
    P(k+1)h,i = Wh,i – H(k+1)w,i
2. Increment k. If k<Kmax -1 (Kmax: the max number of iterations), then, reiterate from step 4, else, go to step 7.
3. Binarize the separation result as-   
      
    (H(kmax)h,i ,P(kmax)h,i)  = {(0 , Wh,i) (H(kmax-1)h,i  < P(kmax-1)h,i)   
    = {(Wh,i , 0) (H(kmax-1)h,i >or =P(kmax-1)h,i)
4. Convert the H and P into waveforms by-



Evaluation

We evaluate the result by calculating the Signal-to-noise ratio (SNR) where,

ft = original signal, err = original minus separated

SNR =

In this case, SNR = 135.54 dB.

Workload distribution

First, we studied the paper individually and then arranged a meeting for discussing the main idea of this project and confusions we had during the reading session. Later, we tried to implement the algorithm together. For the final report, we arranged a single day to write all the necessary information and edit it together.

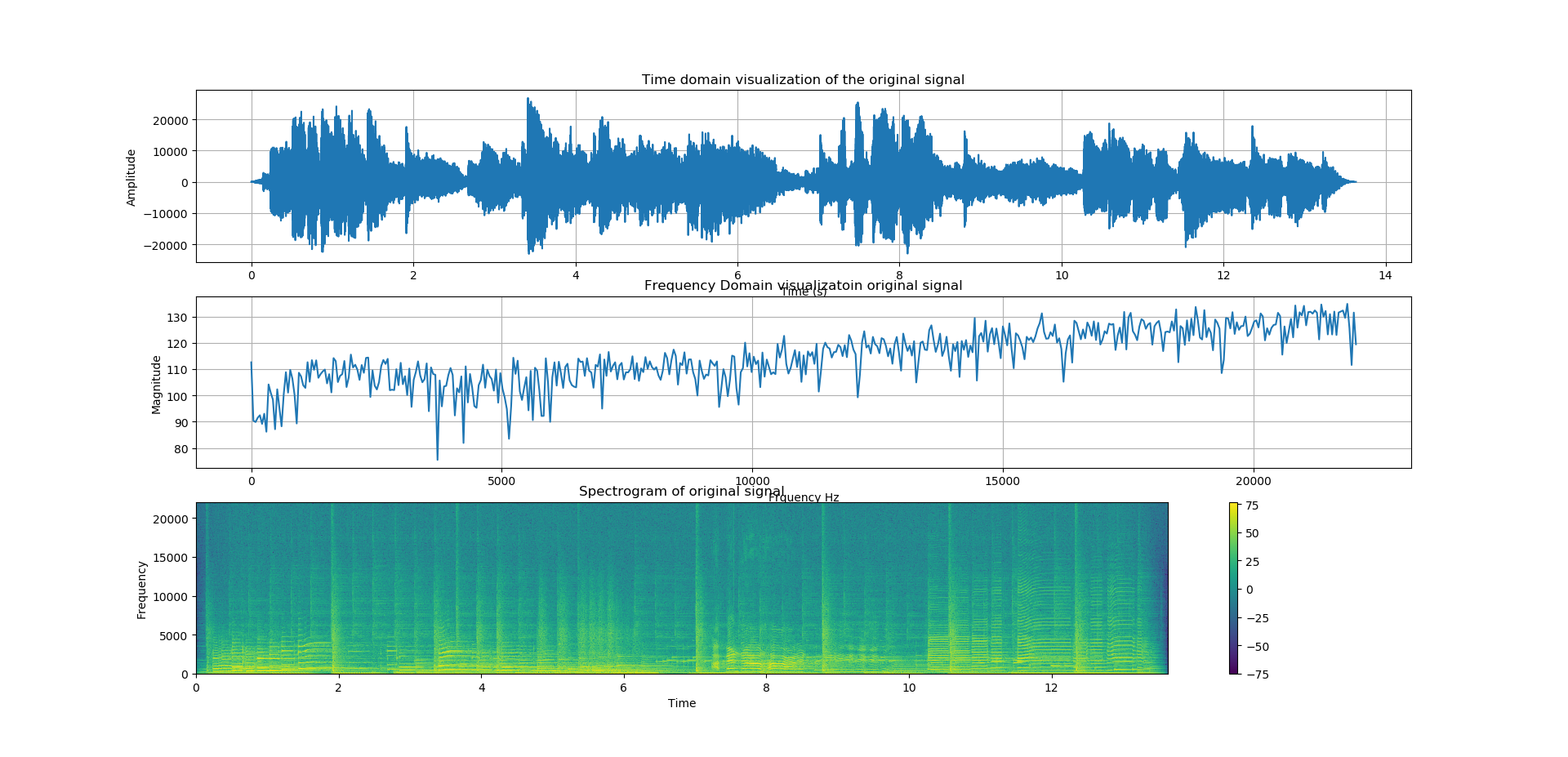
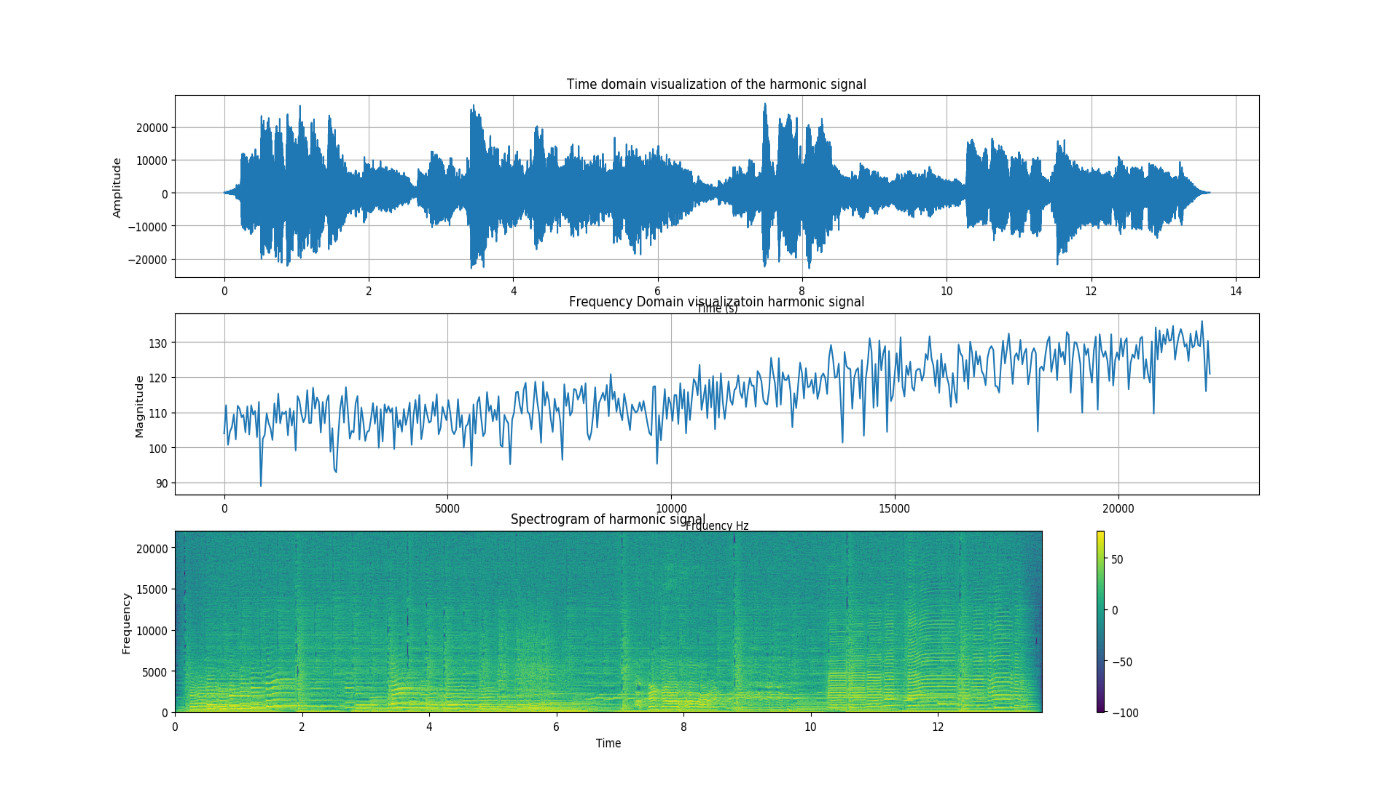
 Limitations

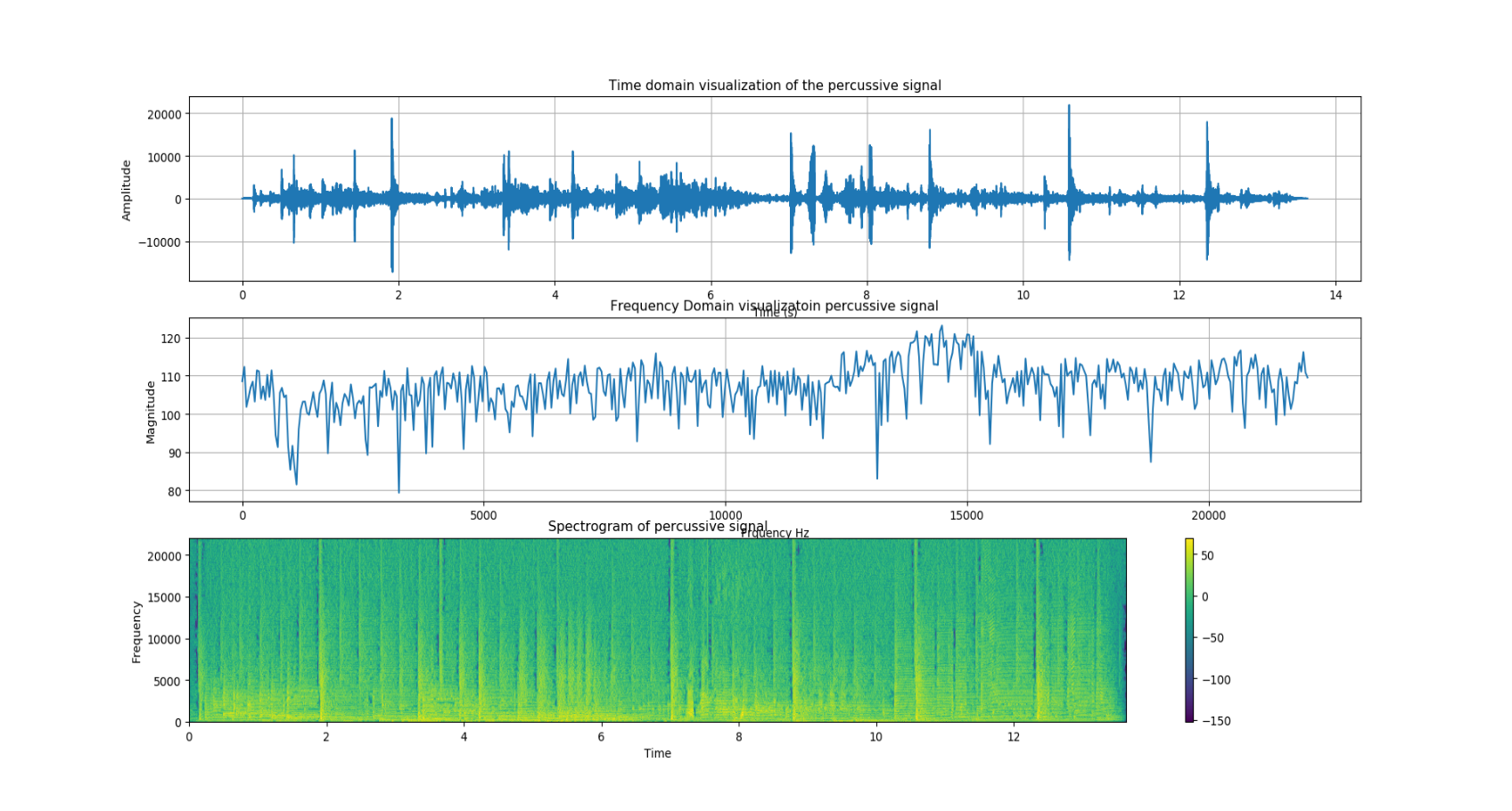
The limitation of the algorithm is that it works only for mono signals. As provided in the research paper, they worked only on mono signals because stereo signals used multiple channels. Various research institutes around the world have worked on both mono and stereo signals. Specifically, in the University of York, the researchers have been concentrating on mono signals and separate it into different tracks. Now they are focusing on stereo signals in similar way [2]. It is also shown in the research how signals from a single channel are separated into multiple channels. The problem in stereo signals is that there must be another complex algorithm to synchronize the various channels after processing of the original signal.

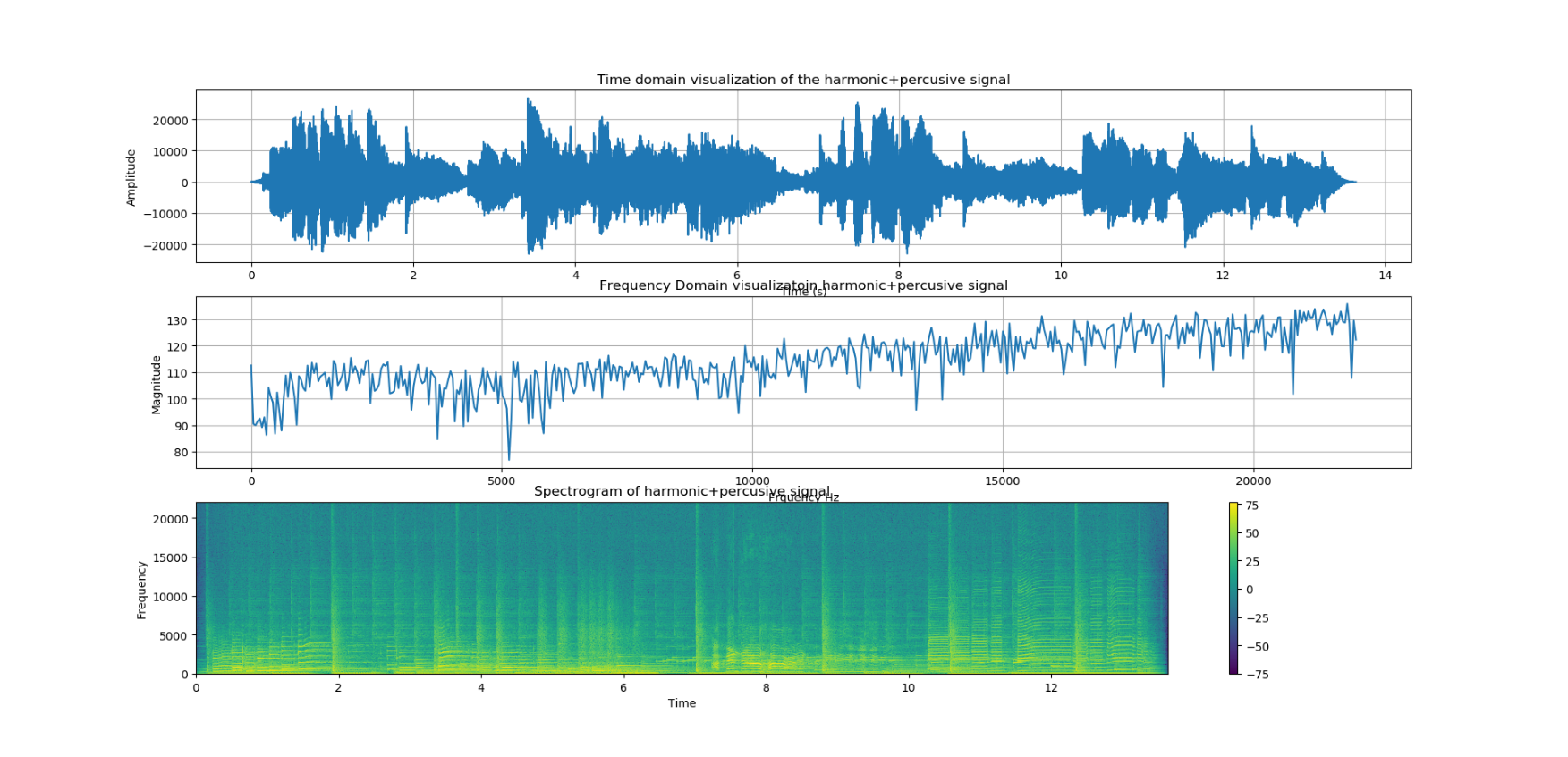
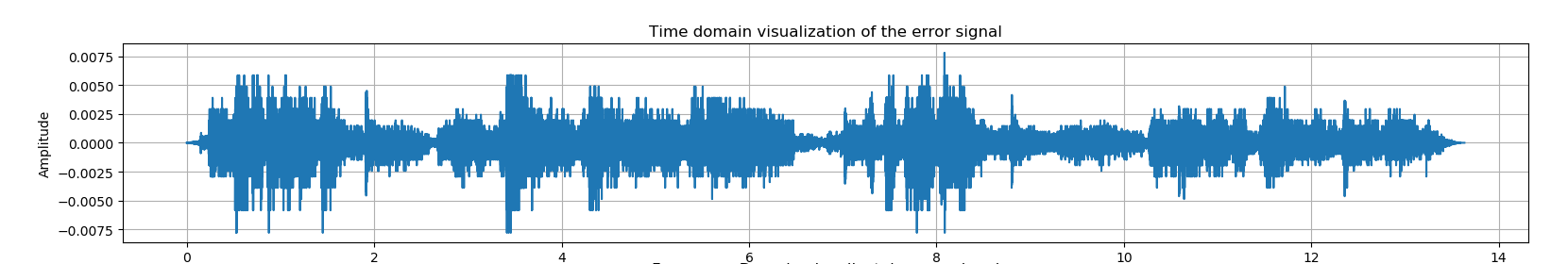
Separation quality measurement and assessment

The quality of signal can be calculated by calculating the error signal between the original signal and the signal after addition of harmonic and percussive signals. The error is calculated in such a way that the signal x(t) contains the monophonic mixture of instruments and y(t) is the mixture of separated harmonic components. We calculated the error using the formula:   
 error = original signal in time domain - percussive components – harmonic components   
 err = ft – p [: length(ft)] – h [: length(ft)]   
The error came out to be -1.52 \* 10-6 which is almost zero.   
Ideally, the error should be close to zero and the sound to noise ratio should be high.   
  
While playing the original audio along side the harmonic+percussive audio, we could spot a slight noise variation, but most of the sound signal remained unaltered.   
Whereas while playing both the harmonic and percussive signals separately we could imitate the sound from the examples provided.

Result visualization

Time domain waveforms and spectrogram from the original and results are shown on the Figure 1 to Figure 4 below -  Figure 1- Visualization of the original signal   
  
 Figure 2- Visualization of the Harmonic part

 Figure 3- Visualization of the percussive part

 Figure 4- Visualization of Harmonic + Percussive signal   
  
 Figure 5 – Visualization of Error

Conclusion  
  
From the results obtained, it is seen that the harmonic and percussive part, when played separately, are clearly perceived by human ear. Once they are combined, there is a slight variation in original and the processed audio signal. The spectrograms also show the horizontal lines in the harmonic spectrogram and vertical lines in Percussive spectrogram. The error came out to be almost zero and the SNR ratio is comparatively high.

REFERENCES

[1] K. M. J. L. R. H. K. a. S. S. Nobutaka Dno, «Separation of a monaural audio signal into

harmonic/percussive components by complementary diffusion on spectrogram,» в 16th

European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, 2008.   
[2] <http://www-users.york.ac.uk/~jes1/Audio_Signal_Processing_Research.html>