**DSSA PROJECT SUMMARY REPORT**

***Group 27***

**PROJECT:** Seperating Vocals from an Audio Track

**FINDINGS & DIFFICULTIES FACED:** First we tried to figure out how to treat the audio signal so as to remove the vocals. We came across many methods but one that we first chose was **REPET**(REpeating Pattern Extraction Technique).

**REPET:** REPET was based on the assumption that a song has two components. One was the Vocals and Second was the repeating Background music. Now the REPET method assumes that background music is a repeating sequence and uses cross correlation to find the period of the repeating sequence. It plots the cross correlation graph of the signal with many lagged versions of the signal to find the period. Once the period is obtained it takes all the sequences of that period from the song and recovers the common art which would be the background music using various operations on the signal like finding mean of the signals and other things.

**Why we discarded this method?**

The REPET method was based on the assumption that the background music is a repeating sequence which is not true for modern music which has several parts like the chorus the bridge and the build up in their songs. Also the results we obtained from the Repet method were pretty bad and had a lot of loss and noise in their output along with leftover vocals.

We researched for other methods but to no avail. But then we found how people use a music editor called “***Audacity”*** to remove vocals from songs. We decided to implement the same method for our project.

**Channel Seperation Method:** We learned from research that its a practice among modern music producers to centrally pan the vocals of the singer while the instruments are recorded differently on different channels. So we decided to use this method to remove the vocals from the song. In this method we take a stereo input and split it into two channels. Now considering that the vocals are common in both channels while the instrumentals are different we remove the common part which should be the vocals. So we should obtain the instrumental version of the input song.

**Problems Faced in this Method:** Since this method also assumes that the vocals are centrally panned(common in both channels) and the instrumentals aren’t, the method fails in cases when along with the vocals some other instruments are also panned in which case, the common instrument would also be removed.

**How we resolved those Problems:** On researching we found that along with the vocals the drums and the bass are also usually common on both channels. So to prevent the loss of certain instruments we used filters to filter out the instruments that are out of the human vocal frequency range to atleast preserve those instruments from getting removed. On trying out many songs we found the human vocal range to be approximately between 250Hz - 10kHz so any instrument out of this range is preserved.

**Unsolved Problems:** We were able to preserve the instruments that were outside the vocal range but as far as the instruments whose frequency range collides with the vocal range are concerned, we couldn’t derive a method to preserve them.

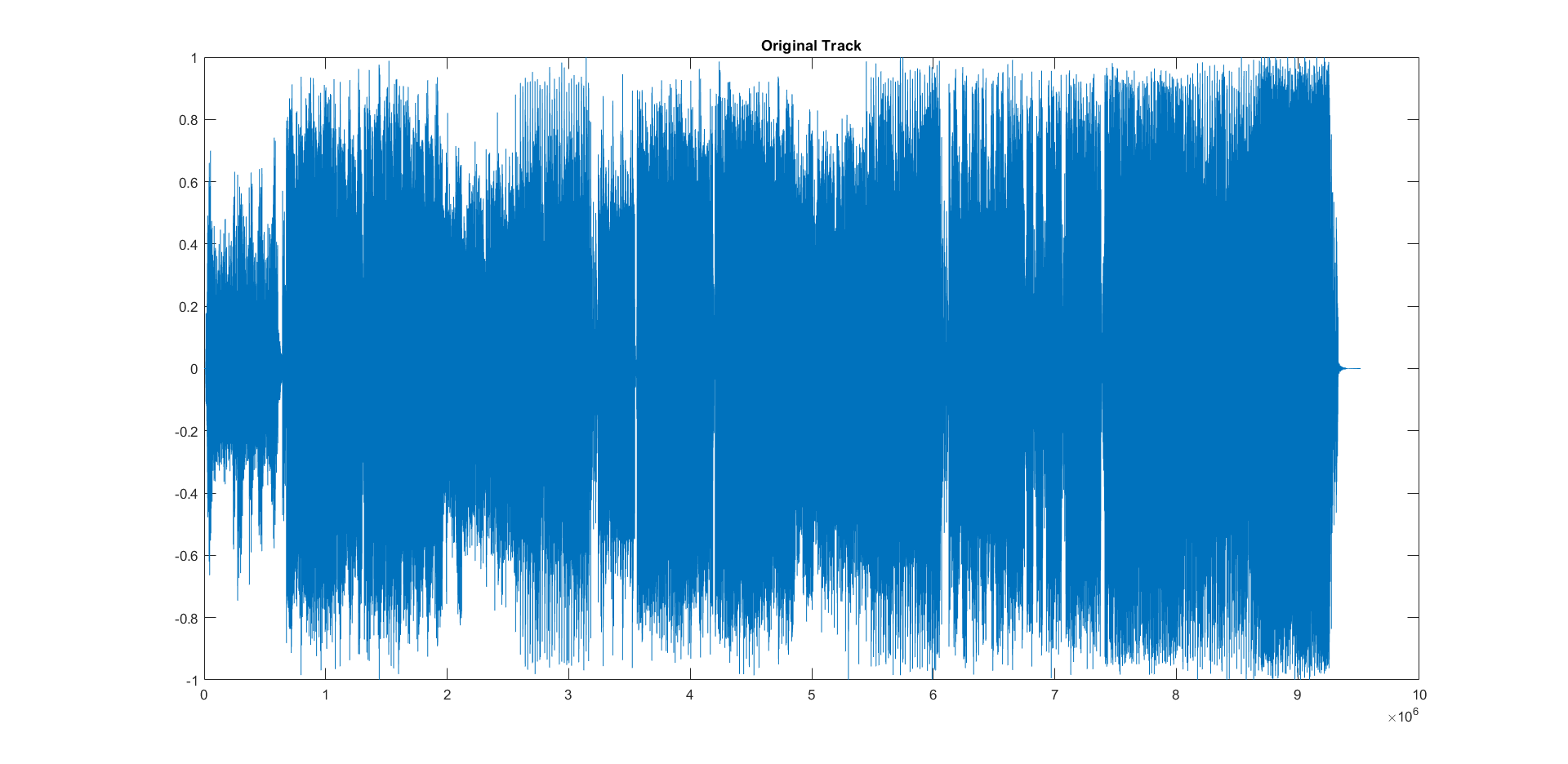
**ANALYSIS OF RESULT:**

To Analyse and Verify the output we used Cross Correlation

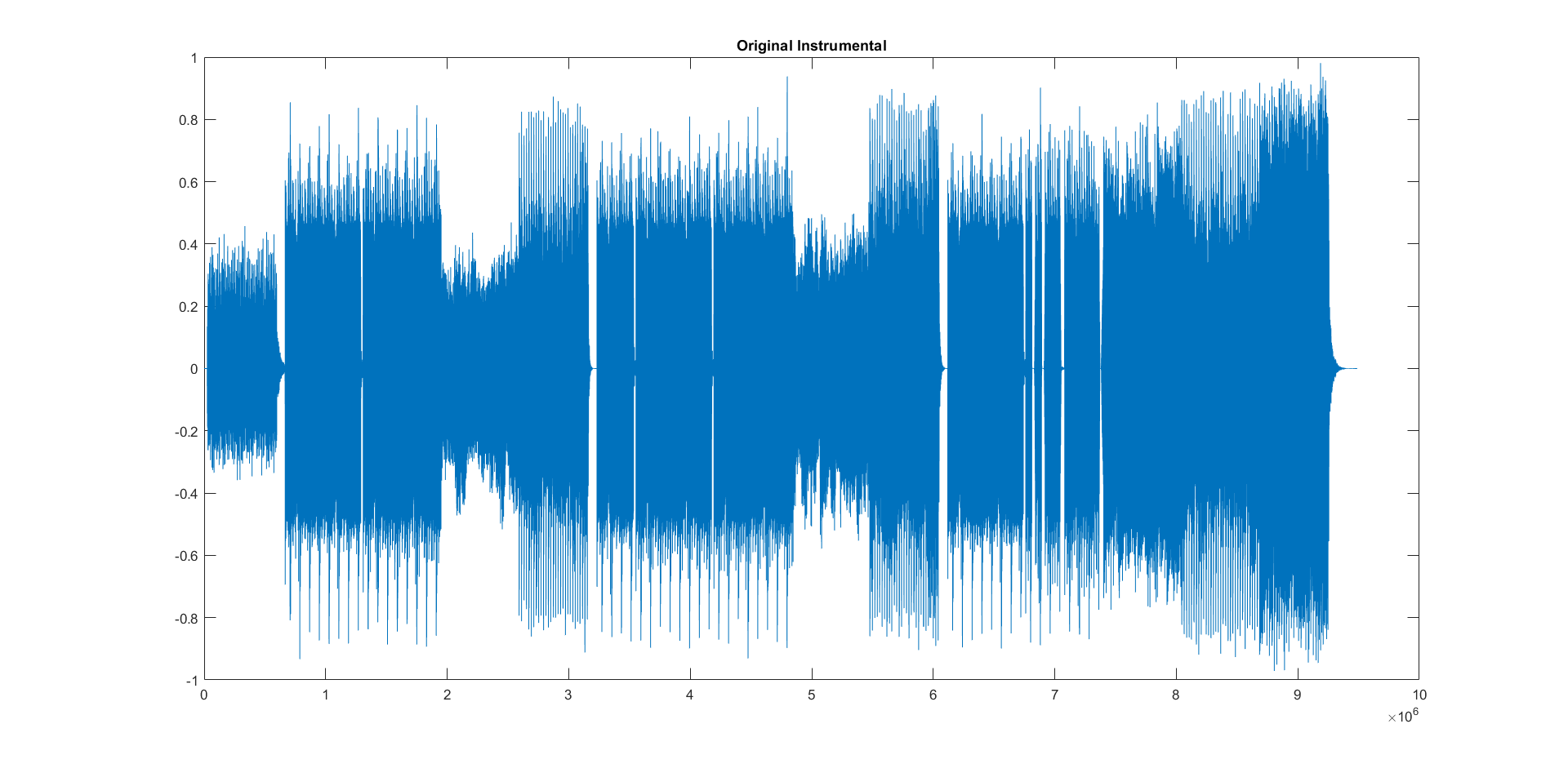
**Cross Correlation Graph:** The cross correlation graph is plotted between two signals. In cross correlation second signal traverses over the first signal and for each point a value is calculated which signifies the amount of overlap signal 2 had with signal 1. A graph of these value vs time(by which the second signal is shifted for the instant) is plotted. If the the graph peaks at the point ‘0’ then the two signals are considered to be similar. If it peaks at some other point the two signals are similar at that point.

**SONG USED : Dark Horse - Katy Perry**

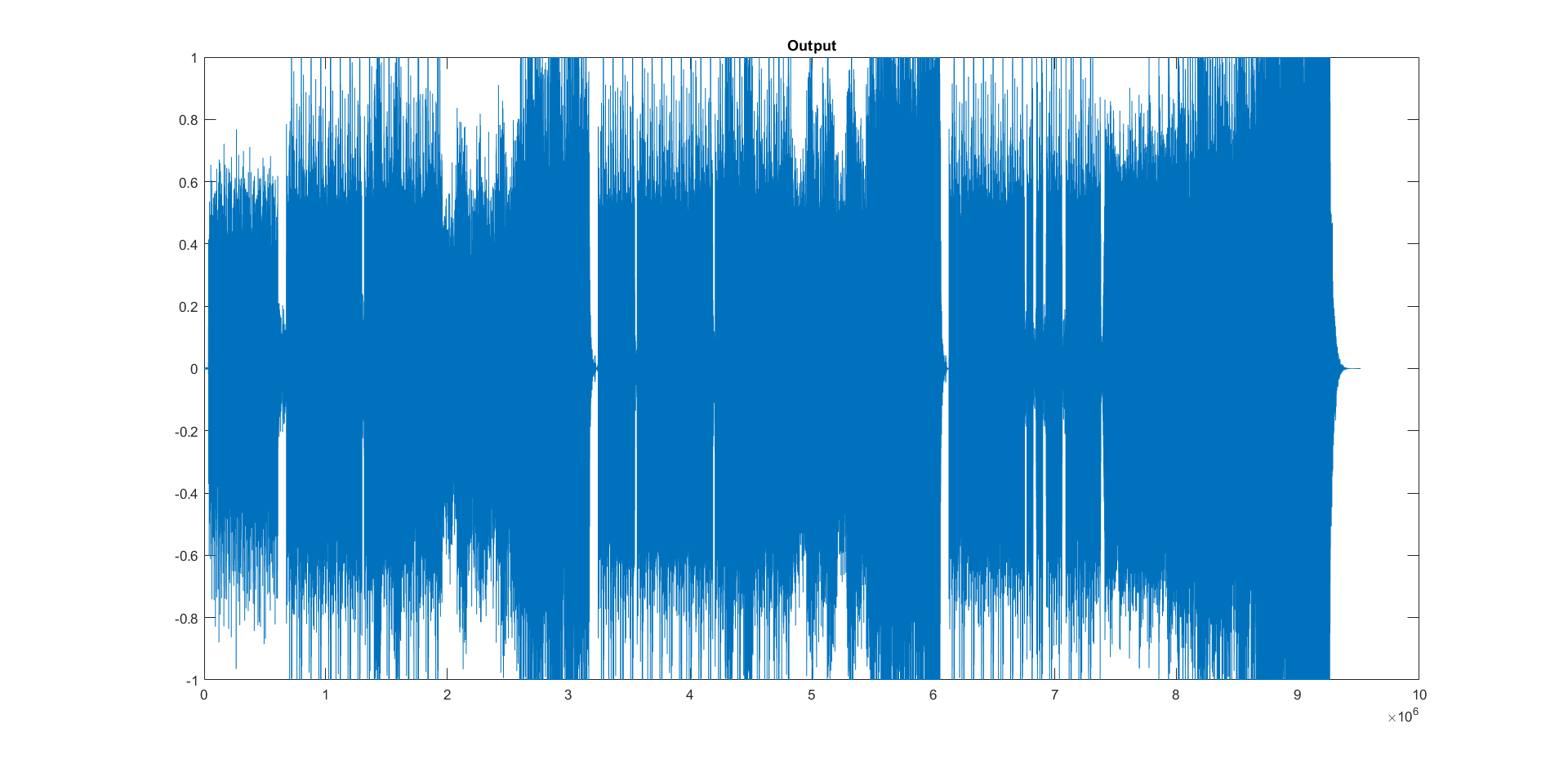
**Original Track:**

****

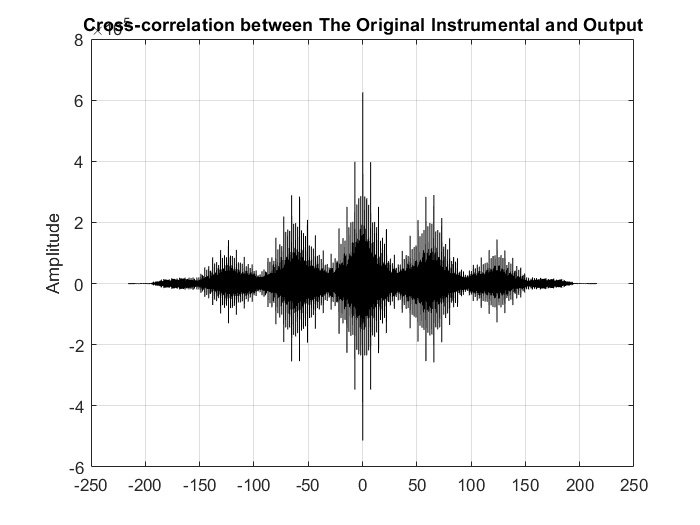
**Original Instrumentals:**

****

**Our Output:**

****

**Verifications:**

****