# ANALYZING THE LONG TERM AVERAGE SORTED SPECTRUM OF AUDIO UPLOADED TO YOUTUBE

by

ANDREW RYAN RABBIO

B.S., University of Colorado Denver, 2016

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Recording Arts Program

2019

This thesis for the Master of Science degree by

Andrew Ryan Rabbio

has been approved for the

Recording Arts Program

by

Catalin Grigoras, Chair

Jeff M. Smith

Cole Whitecotton

Date: May 18, 2019

Rabbio, Andrew Ryan (M.S., Recording Arts Program)

Analyzing the Long Term Average Sorted Spectrum of Audio Uploaded to YouTube

Thesis directed by Associate Professor Catalin Grigoras

## ABSTRACT

This thesis introduces and documents the collection of a general-purpose dataset of audio recordings from several recorders at all available settings. As a pilot study, this dataset was used in a study of YouTube effects on audio recompression. This is accomplished by analyzing the Long Term Average Sorted Spectrum (LTASS) of audio files before and after being uploaded to YouTube. Over 350 recordings are included in the dataset from a variety of recording devices and manufacturers to ensure a diverse dataset. All recordings were made under controlled conditions to ensure the results are easily comparable and reproducible. By analyzing the LTASS of audio uploaded to YouTube this paper will provide greater clarity to the effects YouTube has on audio compression across a variety of device settings.

The form and content of this abstract are approved. I recommend its publication.

Approved: Catalin Grigoras

**TABLE OF CONTENTS**

CHAPTER

## LIST OF TABLES

TABLE

## LIST OF FIGURES

FIGURE

## LIST OF ABBREVIATIONS

LTASS – Long Term Average Sorted Spectrum

SWGDE – Scientific Working Group on Digital Evidence

MP3 – MPEG-1 Audio Layer-3

WMA – Windows Media Audio

WAV – Waveform Audio File Format

ABR – Adaptive BitRate

DASH – Dynamic Adaptive Streaming over HTTP

HTTP – HyperText Transfer Protocol

MQD – Mean Quadratic Difference

CC – Correlation Coefficient

PCM– Pulse Code Modulation

LP – Long-Term Recording mode on Olympus WS-100

HQ – High Quality recording mode on Olympus WS-100

SP – Standard recording mode on Olympus WS-100

LCon – Low Cut Filter on

LCoff – Low Cut Filter off

MICSENS – Mic Sensitivity setting

VOR – Voice Operated Recording Setting on Sony PX-333

VSYNC – Voice Sync Setting on Olympus WS-853

VCVA – Variable Control Voice Actuator setting on Olympus WS-10

# CHAPTER I

# INTRODUCTION

YouTube has over 1.9 billion unique users that visit the site each month, watching over 1 billion hours of video content every day making it the most common video streaming service on the Internet [11]. With its widespread use, it is inevitable for videos depicting criminal activity, and confessions of crimes in some cases, to be present. There have been many instances where these recordings have been entered into courtrooms as evidence [12]. With this in mind, it is important to know what is happening to these videos when they are uploaded to YouTube and then downloaded using services such as Youtube-dl or other third party programs.

There have been previous publications analyzing the effects of YouTube on video compression but research is lacking for the analysis of the audio from these videos, which is where this paper looks to expand. With the growing popularity of handheld recorders such as small Olympus and Tascam devices being used by many governmental and private investigators as well as public consumers, these devices were used to make the exemplar recordings that comprise the dataset (Figure 1). Most of these devices do not record in a proprietary format, and instead record in standard formats such as MP3, WAV, and WMA, which poses problems for forensic investigators. These standard file formats are much easier to manipulate and re-encode into different formats making it harder to know if a recording is original.

This paper will provide information on the Long Term Average Sorted Spectrum of exemplar recordings taken from 4 different devices from popular manufacturers and compare those results to the LTASS of the recordings after they have been uploaded and downloaded from YouTube. By analyzing the differences between the original recordings and the uploaded/downloaded from YouTube versions, a clearer concept of how YouTube compresses audio from a variety of different file formats and device settings will be gained.


*Figure 1 Example of digital handheld recorder*

*(Olympus.co.uk)*

**Audio Compression**

While many in the public may not realize, audio compression is in use all around us every day from the music we listen to, what we hear and see on TV, to the sounds that are emitted from a child's favorite toy. The advent of audio compression began in the 1980's with the need to reduce the bit-rate requirement for Compact Disk (CD) without sacrificing noticeable decreases in audio quality. Since then a variety of audio compression algorithms, commonly known as codecs, have been introduced that exploit the way we hear to reduce the size of files. The main way that these codecs do this is called perceptual coding [7]. During this process the compression algorithm looks to exploit imperfections of human hearing with the goal of making the file size smaller while maintaining sound quality. Perceptual coding is a lossy compression technique that utilizes temporal and simultaneous masking to eliminate information that our ears are unable to hear. By eliminating this information and represent the digital audio signal with the least about of bits while maintaining transparent signal production.

## FFmpeg

FFMPEG is a powerful command line tool that allows for the conversion, creation, and streaming of digital video and audio files. It is offered for free from the developers website and allows the user to utilize its built in libraries. The FFmpeg framework is utilized in a variety of application for forensic and commercial purposes.

Audio by itself cannot be uploaded to YouTube and must be accompanied by some video component which is why for the experiments detailed in this paper FFmpeg was used to mux the exemplar recordings with a static video. Figure 2 below gives an example of the command used to combine the exemplar recordings with the test video, which is named "black.mp4". This command stores the original audio as an uncompressed 16bit pcm audio stream in the output

```
Andrews-MacBook-Pro-2:research andrewrabbio$ ffmpeg -i black.mp4 -i Olympus_WS10
0_HQ_MICSENS_high_2.WMA -c:v copy -c:a pcm_s16le -shortest Orig-Olympus_WS100_HQ
_MICSENS_high_2_Upload.avi
```

*Figure 2 FFmpeg command to combine exemplar recording with test video*

## YouTube-dl

Youtube-dl is a free cross-platform command line tool that allows the user to download video and audio from YouTube. The use of this tool is considered best practice compared to other third party programs that purport to download videos from YouTube as it allows for the download of all available file types and resolutions from YouTube's servers compared to other tools, which offer a limited download options. For the tests described in this paper this tool was used to download the audio from videos that were uploaded to YouTube. For the purposes of this test being more reproducible the account that was used to upload the videos was not used when the videos were downloaded using Yoube-dl.

## Previous Research

Currently because of the nature of YouTube being a video streaming service not dedicated to audio the bulk of research relating to this topic has emphasized the authentication of videos that have been re-encoded by YouTube, and source identification of high definition videos. In the Paper "YouTube Re-Compression Effects" Witecotton performs a variety of tests to figure out what information can be gleaned from files that are uploaded to YouTube using third-party options to analyze the files. In this paper he discusses the encoding algorithm that YouTube uses to re-encode videos. He arrives at the conclusion that regardless of the original container type the video stream data remains the same after download from YouTube's server using a variety of different third party applications such as youtube-dl, i.e. the algorithm is consistent across multiple containers [3].

In the paper "Source Identification of High Deffinition Videos: A Forensic Analysis of Downloaders and YouTube Video Compression Using a Group of Action Cameras" Giammarrusco gives a thorough explanation of YouTube's framework and how it re encodes videos which was integral to the research done in this paper. He looks at a series of third party downloader tools, which at the time of publication were most popular when it came to downloading videos from YouTube, analyzing the resulting structure and metadata after each tool was used to download a video from YouTube.  This analysis lead to the understanding that these tools and the re encoding that YouTube automatically does manipulated much of the metadata information in the header of the file, which has strong implications for forensic analysis and authentication purposes [5].

## CHAPTER II

## YOUTUBE FRAMEWORK

A firm understanding of YouTube's framework is essential to understand what is happening when videos are uploaded and downloaded, either from the site or using third party applications. When a video is uploaded to YouTube, regardless of the original format, size, and dimensions, it is automatically re encoded which may have certain ramifications for authentication and analysis purposes. With this is mind it is important to note that the file that is uploaded will be used to create a variety of different video streams at different resolutions. As Gimmarrusco puts it, the file you upload is like a master file that is used to create different video streams, "Simply stated, the better the quality of file that is uploaded, to YouTube, the better quality that will be received upon download" [5]

These different video streams are created using adaptive dynamic streaming over HTTP (DASH) during playback, which is an Adaptive BitRate (ABR) video streaming technology. ABR allows YouTube to switch the video and audio quality based on the users available connection. Añorga et al. explains that "The main outcome of this feature is that if on *YouTube's* player quality parameter is set on '*auto*', *YouTube* can adapt the bitrate of the video based on the client's available download bandwidth, so the video streaming is adapted to a dynamic environment." [9].

### Available Formats

In order to see the different audio qualities that YouTube offers during playback based on the users available bandwidth youtube-dl can be used to query for all available audio and video formats. Figure 3 below shows an example of this query of available formats of a 192kbps mono Mp3 file.

```
[info] Available formats for gQQ6pqX6zXc:
format code  extension  resolution note
139          m4a        audio only DASH audio   49k , m4a_dash container, mp4a.40.5@ 48k (22050Hz)
140          m4a        audio only DASH audio  130k , m4a_dash container, mp4a.40.2@128k (44100Hz)
134          mp4        640x360    DASH video   10k , mp4_dash container, avc1.4d4016, 6fps, video only
136          mp4        1280x720   DASH video   19k , mp4_dash container, avc1.4d401f, 6fps, video only
278          webm       256x144    DASH video   95k , webm_dash container, vp9, 6fps, video only
160          mp4        256x144    DASH video  108k , mp4_dash container, avc1.4d400b, 6fps, video only
242          webm       426x240    DASH video  220k , webm_dash container, vp9, 6fps, video only
133          mp4        426x240    DASH video  242k , mp4_dash container, avc1.4d400c, 6fps, video only
243          webm       640x360    DASH video  405k , webm_dash container, vp9, 6fps, video only
244          webm       854x480    DASH video  752k , webm_dash container, vp9, 6fps, video only
135          mp4        854x480    DASH video 1155k , mp4_dash container, avc1.4d4014, 6fps, video only
247          webm       1280x720   DASH video 1505k , webm_dash container, vp9, 6fps, video only
43           webm       640x360    medium , vp8.0, vorbis@128k, 284.40KiB
18           mp4        640x360    medium  107k , avc1.42001E, mp4a.40.2@ 96k (44100Hz), 348.43KiB
22           mp4        1280x720   hd720  148k , avc1.64001F, mp4a.40.2@192k (44100Hz) (best)
```

*Figure 3 Example of Youtube-dl query for available formats of a 192kbps mono Mp3 file.*

# CHAPTER III

## TEST RECORDING FRAMEWORK

Developing and testing forensic methods for digital audio requires a dataset that is acquired under certain basic criteria that allows it to be diverse and reproducible by others. For the tests detailed in this paper four different devices were used to make the test recordings that were ultimately uploaded and then downloaded to YouTube to analyze the resulting LTASS. These devices include the Tascam-DR07, Olympus WS-100, Olympus WS-853, and the Sony ICD PX-333. From these devices a list of all the recording settings were made and recordings were made using every possible combination of all the settings on each device. Some of the settings used include VOR, VSYNC, Low Cut Filter, mic sensitivity level, and a variety of preprogrammed recording settings such as lecture, interview, meeting, and voice notes settings (Table 1). To control variability each combination of settings was recorded a total of three times. In total the dataset for this test consists of over 350 recordings. In each recording the make and model of each device, serial number, encoding algorithm, bitrate, mono/stereo setting, and all other applicable recording settings were listed.

*Table 1 List of Device Settings*

| Recorder Name | Device Setting |
|---|---|
| Socy ICD PX333 | Low Cut Filter<br>Mic Sensitivity (Low. Medium, High)<br>VOR<br>Lecture<br>Meeting<br>Voice Notes<br>Interview<br>8kbps Mono Mp3<br>48kbps Mono Mp3<br>128kbps Mono Mp3<br>192kbps Mono Mp3 |
| Olympus WS 853 | Low Cut Filter<br>VCVA<br>Vsync (1,2,3, and 5 seconds)<br>8kbps Mono Mp3<br>64kbps Mono Mp3<br>128kbpsStereo Mp3 |
| Olympus WS100 | HQ, LP, SP<br>Mic Sensitivity (Low, Medium, High)<br>VCVA |

*Table 1 cont'd*

| | |
|---|---|
| Tascam DR07 | 16 bit wav 48KHz |
| | 16 bit wav 44.1KHz |
| | 24 bit wav 48KHz |
| | 24 bit wav 44.1KHz |
| | 32kbps MP3 48KHz |
| | 32kbps MP3 44.1KHz |
| | 64 kbps MP3 48KHz |
| | 64kbps MP3 44.1KHz |
| | 96kbps MP3 48KHz |
| | 96kbps MP3 44.1KHz |
| | 128kbps MP3 48KHz |
| | 128kbps MP3 44.1KHz |
| | 192 kbps MP3 48KHz |
| | 192 kbps MP3 44.1KHz |
| | 256kbps MP3 48KHz |
| | 256kbps MP3 44.1KHz |
| | 320kbps MP3 48KHz |
| | 320kbps MP3 44.1KHz |

## Uploading and Downloading From YouTube

Since YouTube does not allow for the upload of just audio with no video content all test recordings were combined with the same static video using FFmpeg as discussed earlier in Chapter 1 (Figure 2). The resulting files were then manually uploaded to YouTube using the host platform. After all the videos had been uploaded Youtube-dl was used to query each file for the best quality DASH audio available (Figure 3). Youtbe-dl was then use to download the best quality audio available from each video using the command shown below.

*% youtube-dl -f 140 –ict https://youtu.be/vid-ID-string*

In this example 140 would be replaced by the audio format code that represents the highest quality available from the query list. Before the test recordings were downloaded the YouTube account that was used to upload the videos was signed out to ensure the versions that were downloaded weren't the original files, which are only accessible to the owner of the account used to upload them. Each of the downloaded files was then relabeled to identify them as the downloaded version.

**MATLAB Analysis**

In order to analyze the LTASS MATLAB was used to perform the analysis. MATLAB is multi-paradigm computing environment that is a commonly used tool for many forensic analyses. In this test the similarities between original exemplar recordings and their YouTube downloaded versions were compared to one another using Correlation Coefficient (CC) and Mean Quadratic Difference (MQD) of the LTASS and Power Spectral Density (PSD). The information gained from these plots show the level of compression that YouTube imparted on each of the test recordings after they were re encoding and downloaded using Youtube-dl.

# CHAPTER IV

# RESULTS

After analyzing the results from the LTASS plots and comparing their corresponding Correlation Coefficient (CC) values a few observations were made. When looking at the LTASS plots the most common loss in quality after recompression by YouTube was at and above the 15KHz range. This was not the case for all the plots but a majority of them showed these affects (see tables 4,6,9, and 10). In some cases the results between two recordings with different bit rate settings but otherwise identical device settings were more pronounced such as in Figure 4 and Figure 5 below which represent the plots of an 8kbps mono MP3 file and a 64kbps MP3 file from the same recorder (highlighted in Table 1). Similar results are also shown in Figure 6 and Figure 7 (highlighted in Table 1).

Another observation that was made is the relationship between the bit rate of the original recordings and their correlation to the CC values. The general trend across most of the recorders was the higher the bit rate of the original recording, lower CC values were found. This trend is clearly illustrated in Figure 15, which compares the bit rate and CC values of the Tascam DR-07 recorder and in Figure 11, which shows the bit rate and CC comparison for all recordings in the dataset. It was also observed that certain recorders showed more variability in their CC results and exhibited a lower average CC value, such as the Olympus WS853 as seen in Figure 16 which compares 8kbps recordings across all devices that record at that bit rate. This is also exhibited in Figure 12, which shows the Bit Rate vs. CC for the Olympus WS100.

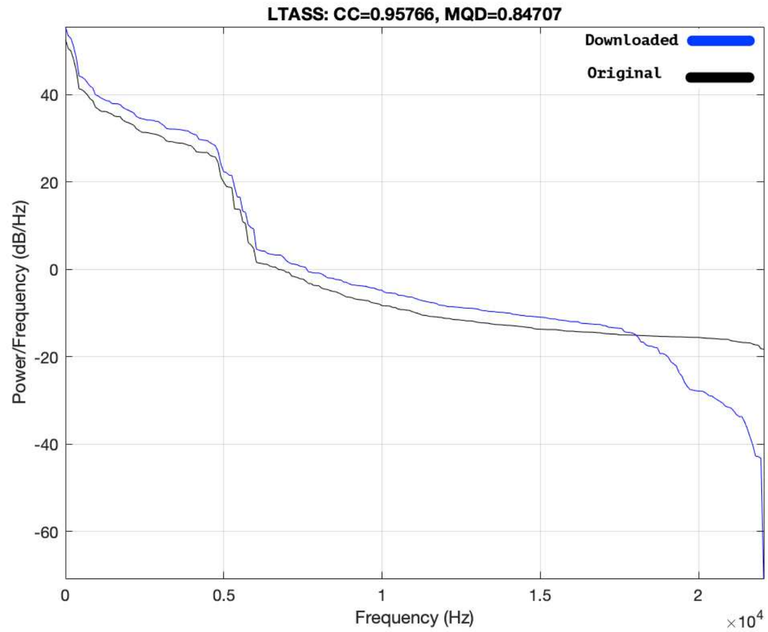


*Figure 4 LTASS Results for "Olympus_WS853_8kbps_MonoMP3_LCoff"*

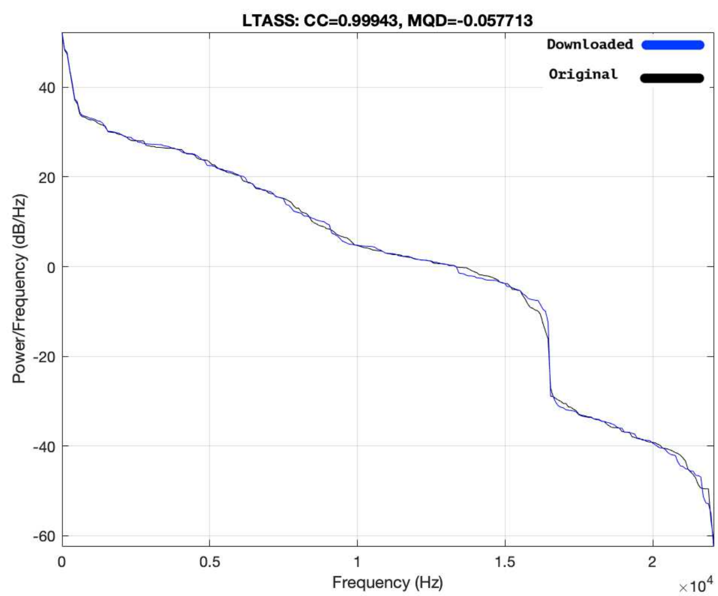


*Figure 5 LTASS Results for "Olympus_WS853_64kbps_MonoMP3_LCoff"*

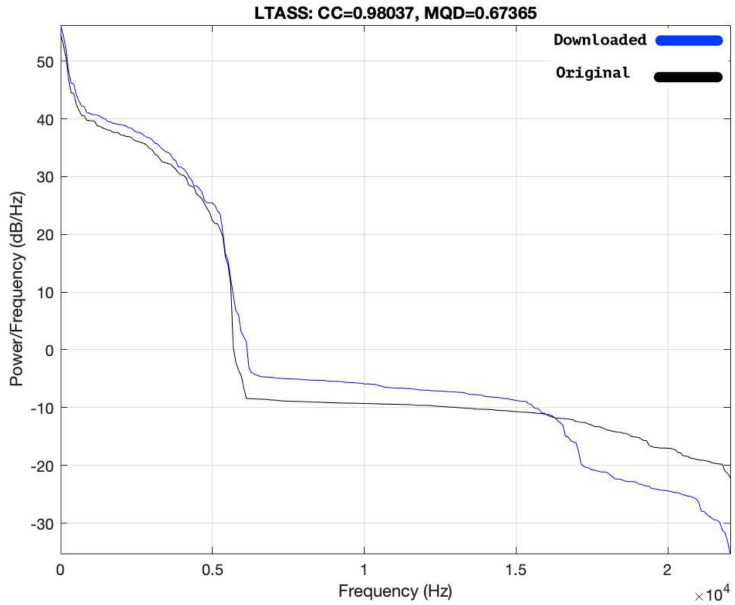


*Figure 6 LTASS Results for "Olympus_WS853_8kbps_MonoMP3_VSYNC_5sec_03"*

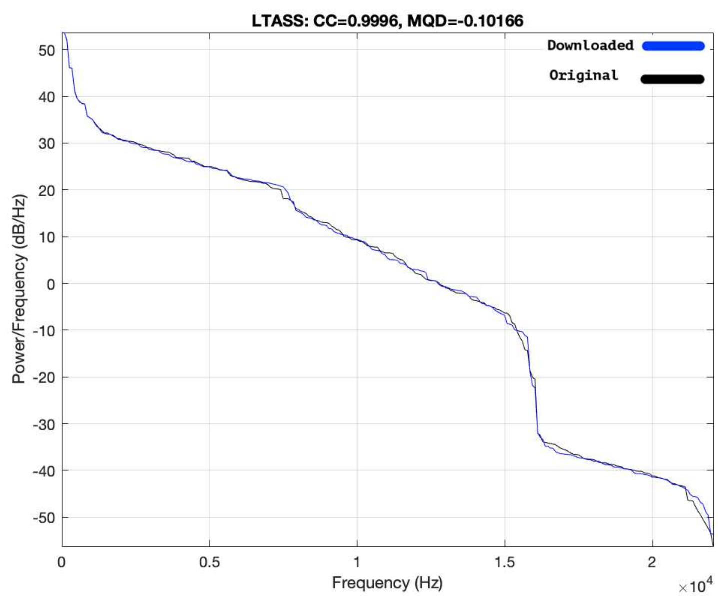


*Figure 7 LTASS Results for "Olympus_WS853_64kbps_MonoMP3_VSYNC_5sec"*

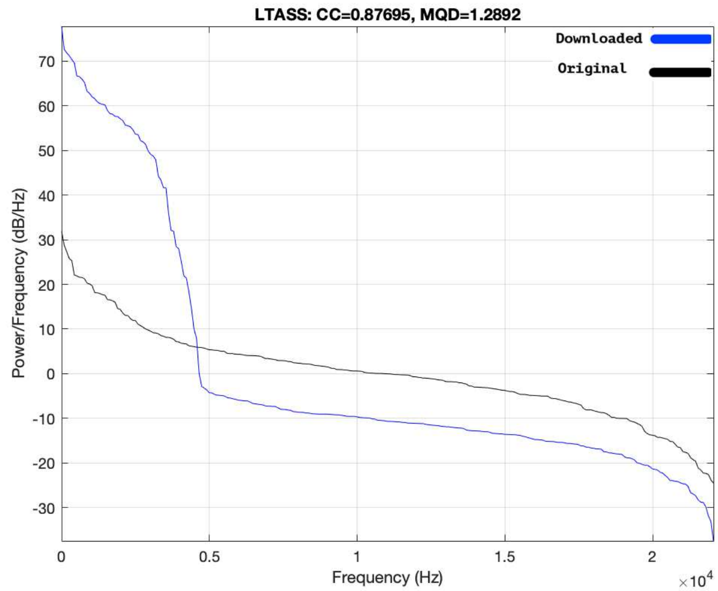


*Figure 8 LTASS Results for "Olympus_WS100_LP_MICSENS_high"*

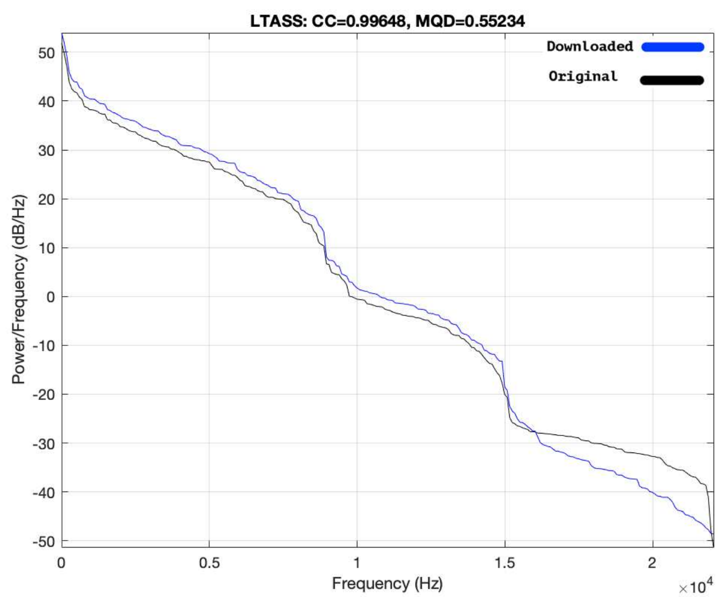


*Figure 9 LTASS Results for "Sony ICD PX333 48kbps Mono MP3 LCon MICSENS med"*

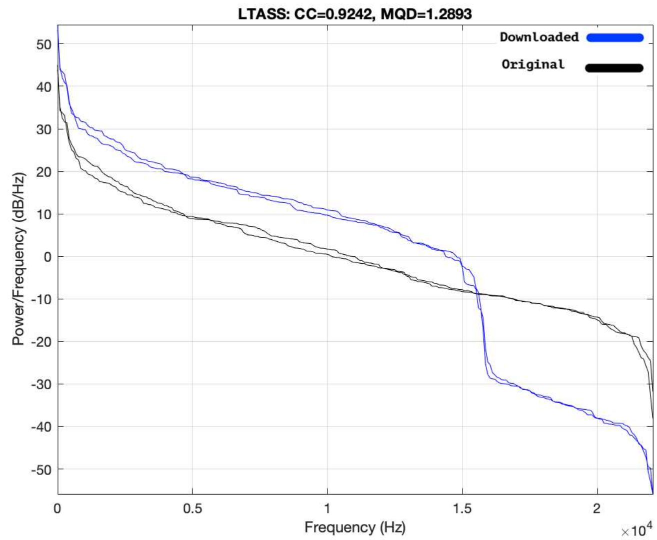


*Figure 10 LTASS Results For "Tascam DR07 16bit Stereo Wav 44.1kHz"*

*Table 2 Outline of every file in the dataset along with the CC, MQD, sample rate before/after download, and the bitrate before/after download.*

| Test Recording File Name | Correlation Coefficient | Mean Quadratic Difference | Original Bitrate | Bitrate After Download | Original Sample Rate | Sample Rate After Download |
|---|---|---|---|---|---|---|
| Sony ICD PX333 8kbps MONO MP3 LCon MICSENS high | 0.99864 | 0.17642 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 LCon MICSENS low | 0.99937 | 0.072384 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 LCon MICSENS med | 0.99932 | 0.1571 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 LCon VOR MICSENS high | 0.99936 | 0.05124 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 LCon VOR MICSENS low 02 | 0.99922 | 0.11038 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 LCon VOR MICSENS med | 0.99904 | 0.15531 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |

*Table 2 cont'd*

| | | | | | | |
|---|---|---|---|---|---|---|
| Sony ICD PX333 8kbps MONO MP3 MICSENS high | 0.99905 | 0.10662 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 MICSENS low | 0.9995 | 0.1592 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 MICSENS med | 0.99948 | 0.1995 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 VOR MICSENS high | 0.99868 | 0.1865 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 VOR MICSENS low | 0.99858 | 0.28357 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 8kbps MONO MP3 VOR MICSENS med | 0.99553 | 0.3844 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Sony ICD PX333 48kbps Mono MP3 LCon MICSENS high | 0.99095 | 0.62022 | 8kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps Mono MP3 LCon MICSENS low | 0.99944 | -0.0052299 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps Mono MP3 LCon MICSENS med | 0.99648 | 0.55234 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps MONO MP3 LCon VOR MICSENS high | 0.99437 | 0.55642 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps MONO MP3 LCon VOR MICSENS low | 0.99969 | -0.094572 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps MONO MP3 LCon VOR MICSENS med | 0.99938 | 0.064902 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps MONO MP3 MICSENS high | 0.99955 | 0.0049564 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps MONO MP3 MICSENS low | 0.99908 | 0.015947 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps MONO MP3 MICSENS med | 0.99019 | 0.67009 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |

*Table 2 cont'd*

| | | | | | | |
|---|---|---|---|---|---|---|
| Sony ICD PX333 48kbps MONO MP3 VOR MICSENS high | 0.99943 | 0.59887 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps Mono MP3 VOR MICSENS low | 0.9992 | 0.077577 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 48kbps MONO MP3 VOR MICSENS med | 0.98432 | 0.78206 | 48kbps | 96kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 Lcon MICSENS high | 0.99417 | 0.67404 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 LCon MICSENS low | 0.98685 | 0.70185 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 LCon MICSENS med | 0.99665 | 0.44112 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 LCon VOR MICSENS high | 0.99854 | 0.43344 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 LCon VOR MICSENS low | 0.98677 | 0.74819 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps Mono MP3 LCon VOR MICSENS med | 0.9875 | 0.67995 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 MICSENS high | 0.99444 | 0.53507 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 MICSENS low | 0.98641 | 0.72433 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 MICSENS med 02 | 0.99147 | 0.67318 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 VOR MICSENS high | 0.99536 | 0.54886 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 VOR MICSENS low | 0.98707 | 0.75738 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 128kbps MONO MP3 VOR | 0.99269 | 0.72581 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |

*Table 2 cont'd*

| | | | | | | |
|---|---|---|---|---|---|---|
| MICSENS med | | | | | | |
| Sony ICD PX333 192kbps MONO MP3 LCon MICSENS high | 0.96018 | 0.90092 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 LCon MICSENS low | 0.92841 | 1.1361 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 LCon MICSENS med | 0.95081 | 0.95192 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 LCon VOR MICSENS high | 0.99037 | 0.82 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 LCon VOR MICSENS low | 0.94887 | 1.1081 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 LCon VOR MICSENS med | 0.96923 | 0.80867 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 MICSENS high | 0.98345 | 0.69052 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 MICSENS low | 0.94022 | 1.1037 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 MICSENS med | 0.95941 | 0.90362 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 VOR MICSENS high | 0.9887 | 0.81066 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 VOR MICSENS low | 0.92653 | 1.1491 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 192kbps MONO MP3 VOR MICSENS med | 0.94719 | 0.96923 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 Lecture | 0.92848 | 1.0274 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 Meeting | 0.95538 | 0.90807 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony ICD PX333 VoiceNotes | 0.98523 | 0.76471 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Sony_ICD_PX333_Interview | 0.97418 | 0.73401 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 16bit Wav | 0.9242 | 1.2893 | 1536kbps | 128kbps | 44.1KHz | 44.1KHz |

*Table 2 cont'd*

| | | | | | | |
|---|---|---|---|---|---|---|
| 44.1kHz | | | | | | |
| Tascam DR07 16bit Wav 48kHz | 0.92193 | 1.2913 | 1536kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 24bit Wav 44.1kHz | 0.92523 | 1.2673 | 2304kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 24bit Wav 48kHz | 0.91475 | 1.2681 | 2304kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 32kbps MP3 44.1kHz | 0.99877 | 0.21071 | 32kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 32kbps MP3 48kHz | 0.99279 | 0.60881 | 32kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 64kbps MP3 44.1kHz | 0.99696 | 0.4163 | 64kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 64kbps MP3 48kHz | 0.97568 | 0.94223 | 64kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 96kbps MP3 44.1kHz | 0.99499 | 0.52883 | 96kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 96kbps MP3 48kHz | 0.97591 | 0.88385 | 96kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 128kbps MP3 44.1kHz | 0.99677 | 0.41362 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 128kbps MP3 48kHz | 0.9825 | 0.80648 | 128kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 192kbps MP3 44.1kHz | 0.99535 | 0.50164 | 192kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 192kbps MP3 48kHz | 0.98009 | 0.83141 | 192kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 256kbps MP3 44.1kHz | 0.99814 | 0.31085 | 256kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 256kbps MP3 48kHz | 0.97793 | 0.84058 | 256kbps | 128kbps | 48KHz | 44.1KHz |
| Tascam DR07 320kbps MP3 44.1kHz | 0.99662 | 0.4307 | 320kbps | 128kbps | 44.1KHz | 44.1KHz |
| Tascam DR07 320kbps MP3 48kHz | 0.97987 | 0.92321 | 320kbps | 128kbps | 48KHz | 44.1KHz |
| Olympus WS100 HQ MICSENS high | 0.99252 | 0.60786 | 32kbps | 705.6kbps | 44.1KHz | 44.1KHz |
| Olympus WS100 HQ VCVA MICSENS Low | 0.99908 | 0.17131 | 32kbps | 705.6kbps | 44.1KHz | 44.1KHz |
| Olympus WS100 LP VCVA MICSENS high | 0.90977 | 1.2368 | 5kbps | 705.6kbps | 8KHz | 44.1KHz |
| Olympus WS100 LP VCVA MICSENS Low | 0.87338 | 1.289 | 5kbps | 705.6kbps | 8KHz | 44.1KHz |

*Table 2 cont'd*

| | | | | | | |
|---|---|---|---|---|---|---|
| Olympus WS100 SP MICSENS high | 0.84984 | 1.4207 | 16kbps | 705.6kbps | 22.05KHz | 44.1KHz |
| Olympus WS100 SP MICSENS Low | 0.86082 | 1.3502 | 16kbps | 705.6kbps | 22.05KHz | 44.1KHz |
| Olympus WS100 SP VCVA MICSENS high | 0.86514 | 1.4098 | 16kbps | 705.6kbps | 22.05KHz | 44.1KHz |
| Olympus WS100 HQ VCVA MICSENS high | 0.99916 | 0.4349 | 32kbps | 705.6kbps | 44.1KHz | 44.1KHz |
| Olympus WS100 LP MICSENS high | 0.87695 | 1.2892 | 32kbps | 705.6kbps | 8KHz | 44.1KHz |
| Olympus WS100 LP MICSENS Low | 0.89148 | 1.179 | 5kbps | 705.6kbps | 8KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 Lcoff | 0.95766 | 0.84707 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 Lcon | 0.99957 | -0.11693 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 LCon VCVA | 0.9801 | 0.73817 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 LCon VSYNC 1sec | 0.99839 | 0.27923 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 LCon VSYNC 2sec | 0.99889 | 0.25384 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 LCon VSYNC 3sec | 0.99828 | 0.33971 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 LCon VSYNC 5sec | 0.99832 | 0.2308 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 VCVA | 0.99923 | 0.27807 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 VSYNC 1sec | 0.99915 | 0.28521 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 VSYNC 2sec | 0.99814 | 0.17792 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 VSYNC 3sec | 0.99669 | 0.47638 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 8kbps MonoMP3 VSYNC 5sec | 0.98037 | 0.67365 | 8kbps | 96kbps | 11.025KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 LCoff | 0.99914 | 0.028172 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps | 0.9993 | -0.012722 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |

*Table 2 cont'd*

| | | | | | | |
|---|---|---|---|---|---|---|
| MonoMP3 LCon | | | | | | |
| Olympus WS853 64kbps MonoMP3 LCon VCVA | 0.99772 | 0.28113 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 LCon VSYNC 1sec | 0.99959 | -0.086612 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 LCon VSYNC 2sec | 0.99946 | -0.030395 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 LCon VSYNC 3sec | 0.99896 | 0.11884 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 LCon VSYNC 5sec | 0.99957 | -0.091456 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 VCVA | 0.99573 | 0.34471 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 VSYNC 1sec | 0.99888 | 0.13636 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 VSYNC 2sec | 0.99907 | 0.090133 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 VSYNC 3sec | 0.99951 | -0.03995 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 64kbps MonoMP3 VSYNC 5sec | 0.9996 | -0.10166 | 64kbps | 96kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 128kbps StereoMP3 LCoff | 0.98064 | 0.8661 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 128kbps StereoMP3 LCon | 0.9819 | 0.84225 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 128kbps StereoMP3 LCon VCVA | 0.98842 | 0.77031 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 128kbps StereoMP3 LCon VSYNC 1sec | 0.99152 | 0.74156 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 128kbps StereoMP3 LCon VSYNC 2sec | 0.99344 | 0.67421 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |
| Olympus WS853 128kbps StereoMP3 LCon VSYNC 3sec | 0.99256 | 0.67793 | 128kbps | 128kbps | 44.1KHz | 44.1KHz |

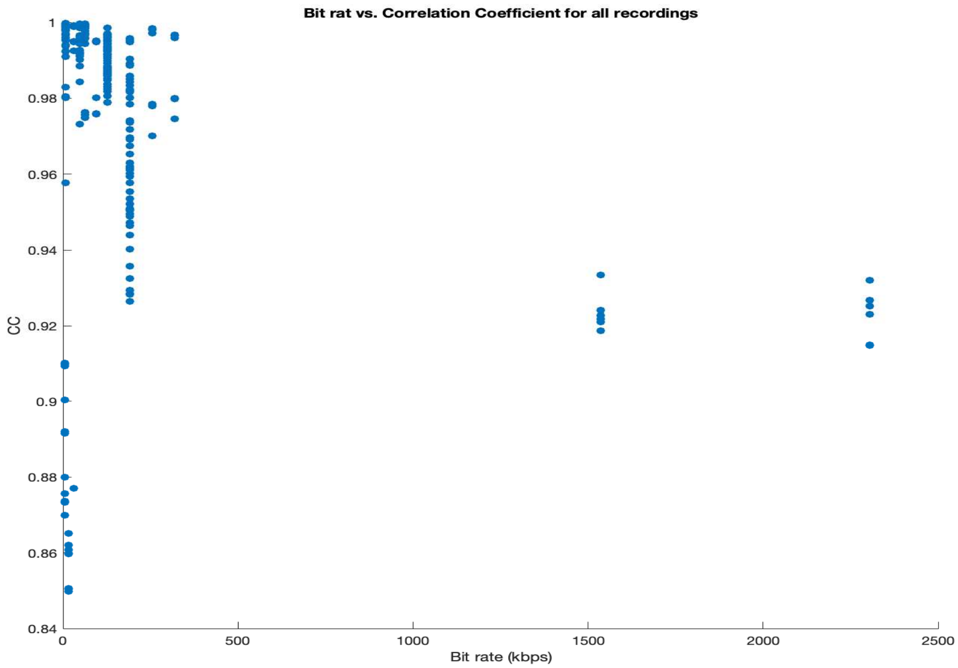


*Figure 11 Plot of Bit Rate vs. Correlation Coefficient For All Recordings*

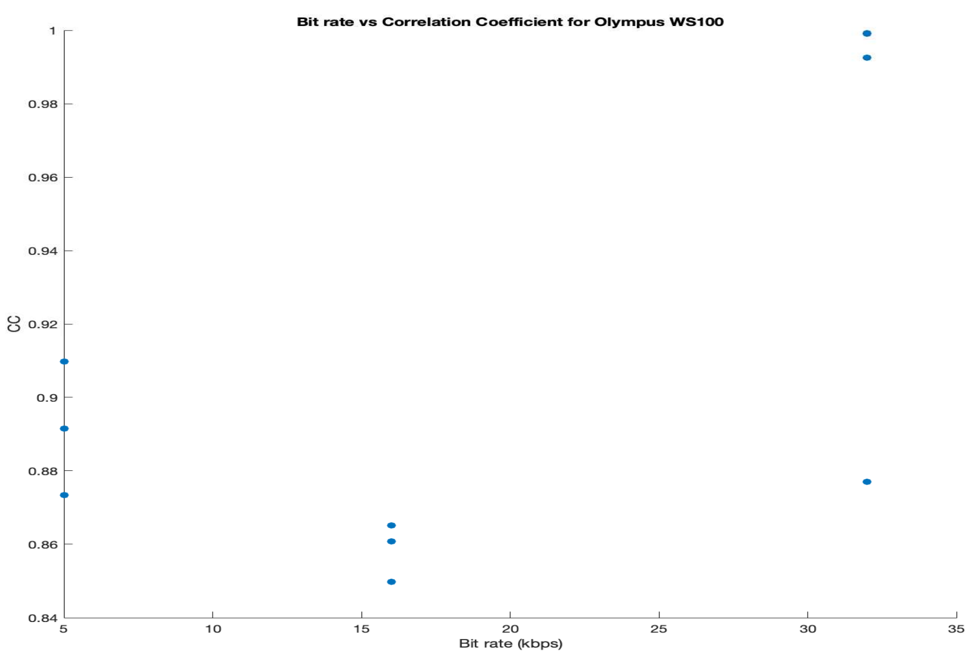


*Figure 12 Plot of Bit Rate vs. Correlation Coefficient For Olympus WS100*

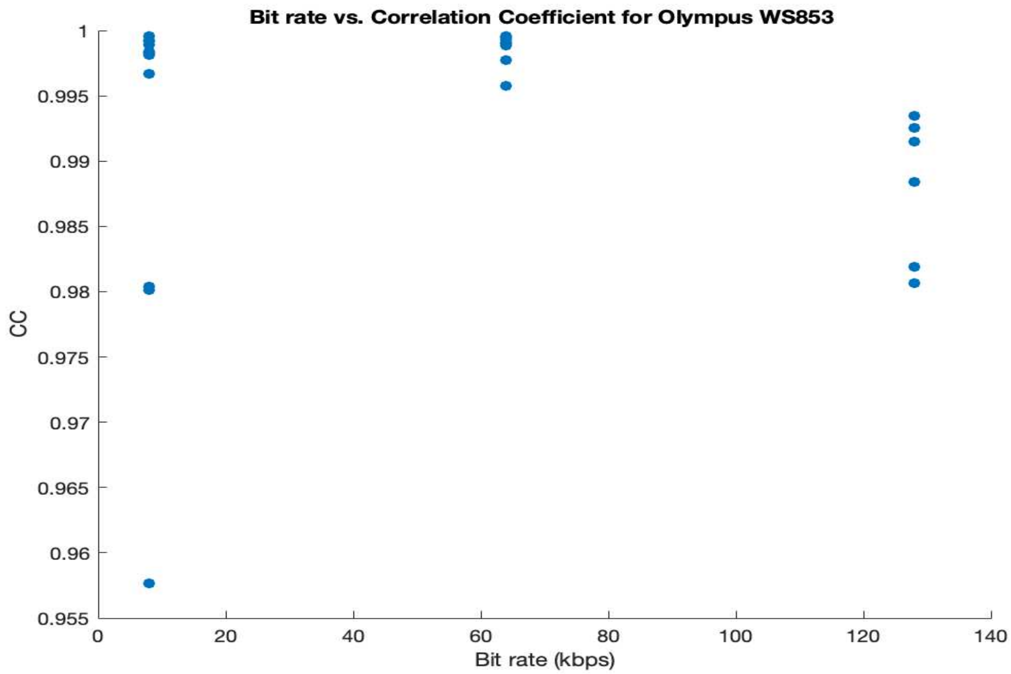


*Figure 13 Plot of Bit Rate vs. Correlation Coefficient For Olympus WS853*

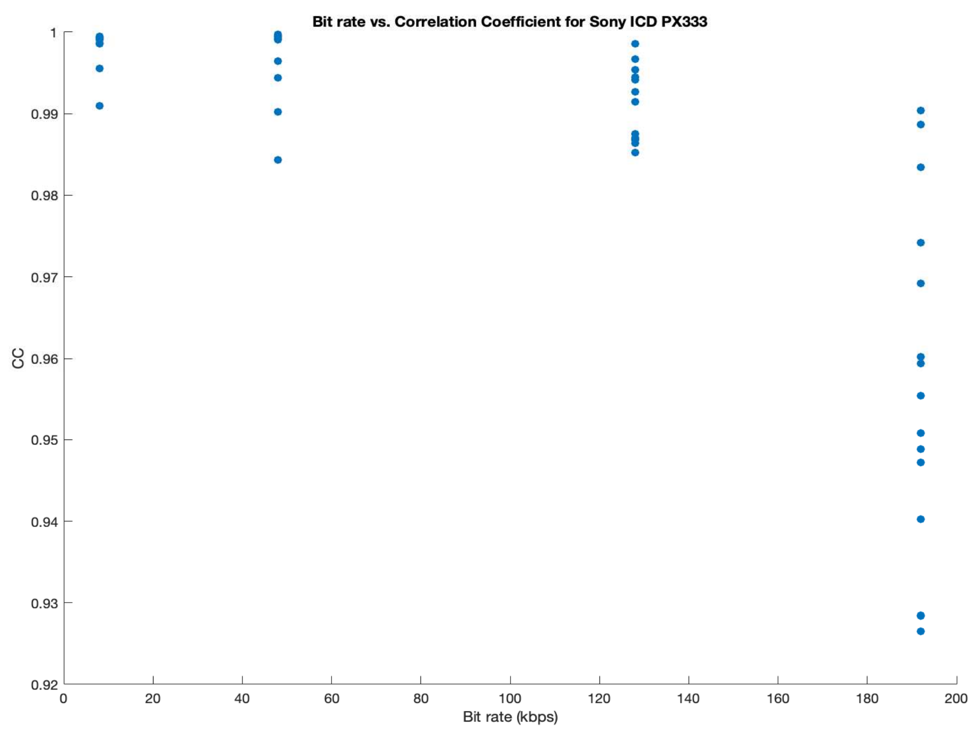


*Figure 14 Plot of Bit Rate vs. Correlation Coefficient For Sony ICD PX333*

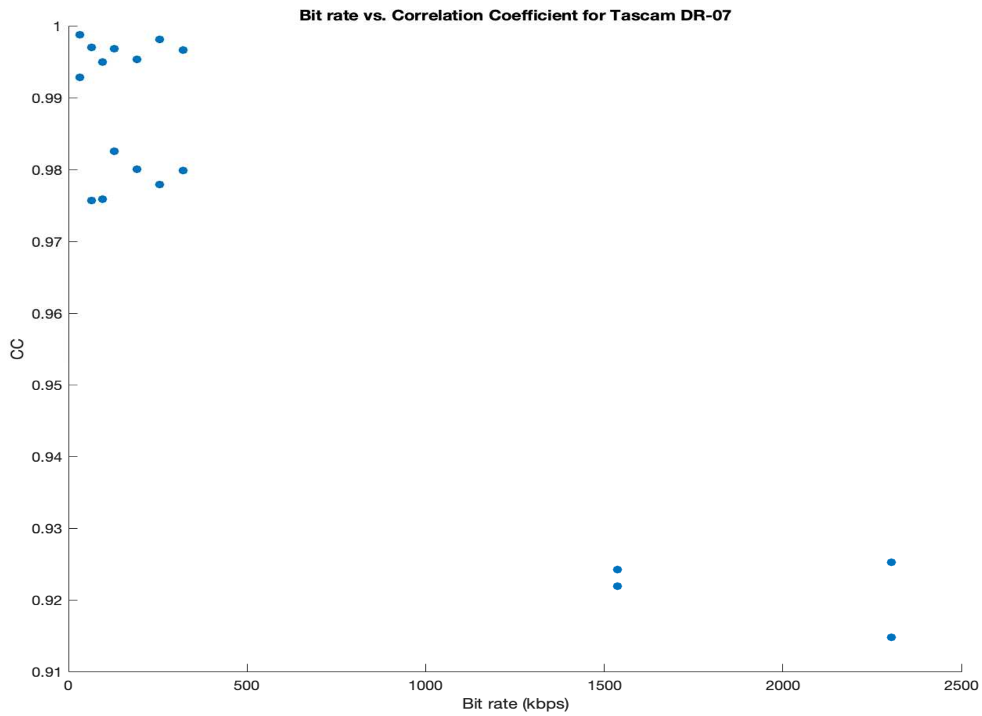


*Figure 15 Plot of Bit Rate vs. Correlation Coefficient For Tascam DR-07*

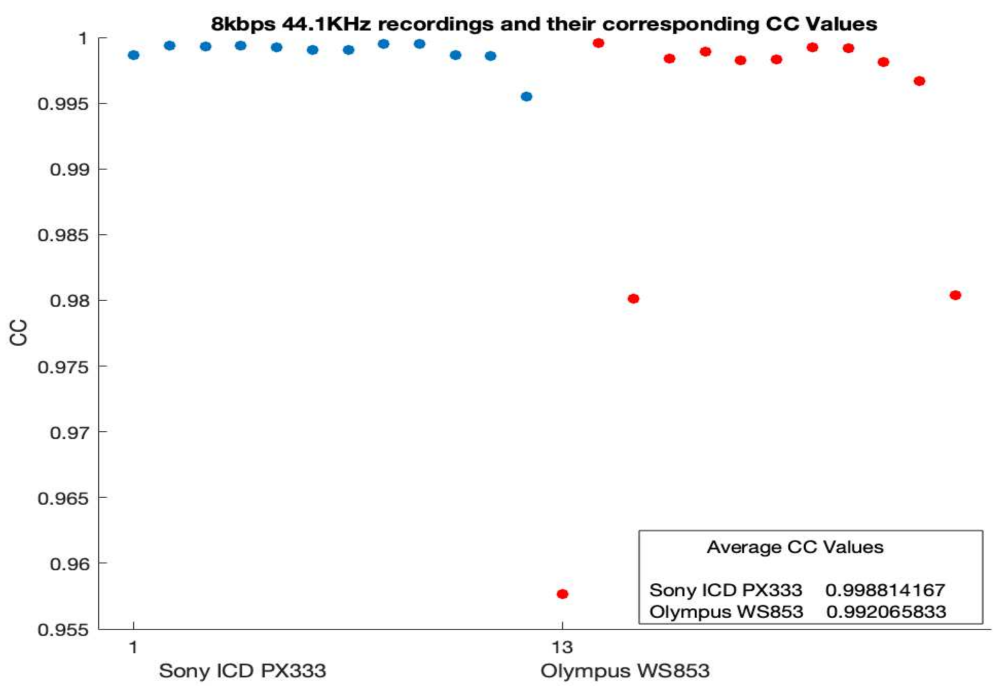


*Figure 16 Plot of 8kbps 44.1KHz Recordings And Their Corresponding CC Values*

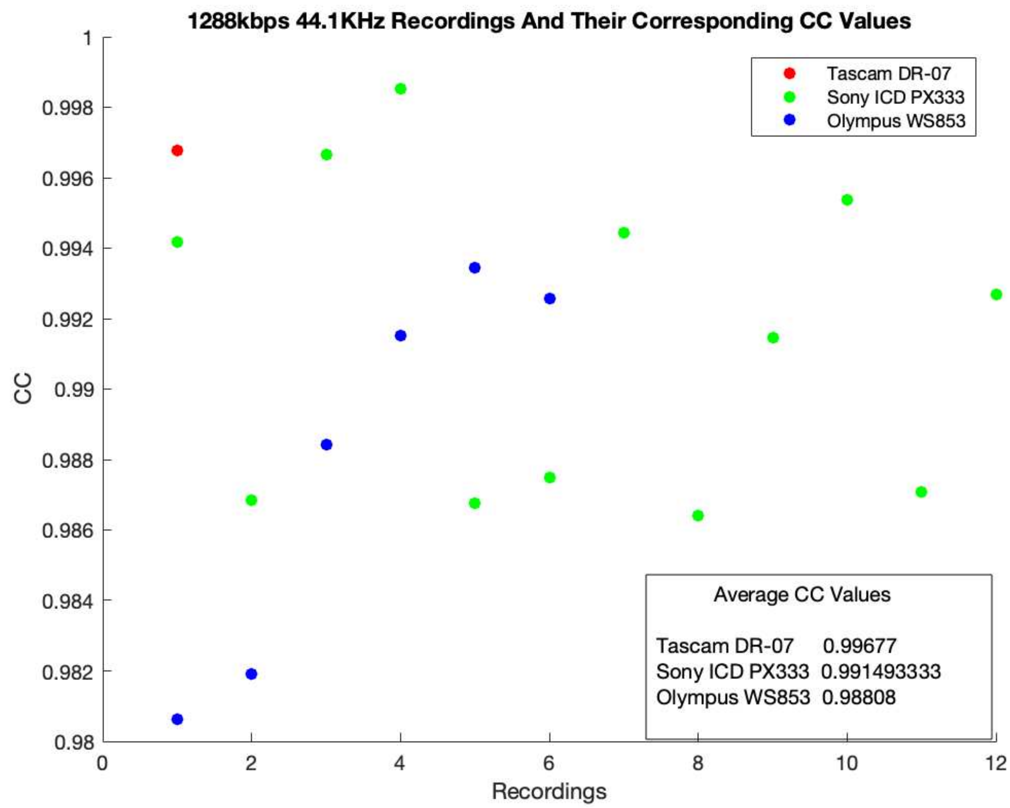


*Figure 17 Plot of 128kbps 44.1KHz Recordings And Their Corresponding CC Values*

# CHAPTER V

# CONCLUSION

This thesis introduced and documented the collection of a general-purpose dataset of audio recordings from several recorders at all available settings. As a pilot study, this dataset was used in a study of YouTube effects on audio recompression. This was accomplished by analyzing the Long Term Average Sorted Spectrum (LTASS) of audio files before and after being uploaded to YouTube. Over 350 recordings were included in the dataset from a variety of recording devices and manufacturers to ensure a diverse dataset. All recordings were made under controlled conditions to ensure the results are easily comparable and reproducible.

In this study the most noteworthy finding was the relationship between the bit rate of the original recordings and their correlation to the CC values. The general trend across most of the recorders was the higher the bit rate of the original recording, lower CC values were found. This means that when higher bit rate recordings are uploaded and downloaded from YouTube using a tool such as Youtube-dl, more compression is detected compared to lower bit rate recordings. This study also revealed that the most common loss in quality after recompression by YouTube was at and above the 15KHz range. This was not the case for all the recordings but a majority of them showed these results.

Some possible areas of interest for future research could be: a deeper look into different audio compression analysis methods for audio that has been uploaded and then downloaded from YouTube to see if different results of what the re encoding process that YouTube automatically performs upon upload does to the audio files; a look into different source devices that record audio such as digital video cameras that record audio and videos from cell phones to see how these files compare to those of handheld recorders after being re encoded by YouTube.

**REFERENCES**

[1] B. E. Koenig; D. S. Lacey, "Forensic Authentication of Digital Audio Recordings," Journal of the Audio Engineering Society, 09/2009, Volume 57, Issue 9

[2] N.C. Donnangelo, W.S. Kuklinski, R. Szabo, R.A. Coury, G.R. Hamshar, "Identification and Exploitation of Inadvertent Spectral Artifacts in Digital Audio," *Journal of Digital Forensics, Security & Law*. 2015; 10(3): 35-57.

[3] Whitecotton, C. M. (2017). "*YouTube: Recompression effects," M.S., University of Colorado Denver. United States – Colorado, 2017*

[4] "SWGDE." [Online]. Available: https://www.swgde.org/documents/Current%20Documents/SWGDE%20Technical%20Notes%20on%20FFmpeg. [Accessed: 10-January-2019].

[5] Z. P. Giammarrusco, "Source identification of high definition videos: A forensic analysis of downloaders and YouTube video compression using a group of action cameras," M.S., University of Colorado at Denver, United States -- Colorado, 2014.

[6] "ISO/IEC 23009-1:2014(en), Information technology— Dynamic adaptive streaming over HTTP (DASH)— Part 1: Media presentation description and segment formats." [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso-iec:23009:-1:ed-2:v1:en. [Accessed: 21-February-2019].

[7] T. Painter, A. Spanias, "Perceptual Coding of Digital Audio," *Department of Electrical Engineering Telecommunications Research Center Arizona State University*. Proceedings of the IEEE ( Volume: 88 , Issue: 4 , April 2000 )

[8] Sinha, Deepen. "Audio Compression." *AccessScience*, McGraw-Hill Education, 2009.

[9] Añorga, J., Arrizabalaga, S., Sedano, B. et al. "Analysis of YouTube's traffic adaptation to dynamic environments" Multimed Tools Appl (2018) 77: 7977.

[10] Koenig BE, Lacey DS, Reimond CE. Selected Characteristics of MP3 Files Re-encoded With Audio Editing Software. Journal of Forensic Identification. 2014;64(3):304-321.

[11] "YouTube For Press," Google Developers. [Online]. Available: https://www.youtube.com/yt/about/press/. [Accessed: 12-Jan-2019].

[12] Video: Suspect Confesses Crime on YouTube. (2012, December 6). *Local Broadcast Video Content.* Available: http://link.galegroup.com.aurarialibrary.idm.oclc.org/apps/doc/A311113718/AONE?u=auraria_main&sid=AONE&xid=eb0f79b0. [Accessed: 1-Feb-2019].