The influence of amplifier settings on the perception of 'heaviness' in guitar timbre

Dr. Dawn A. A. Black
Dr. Christopher Harte
School of Electronic Engineering and Computer Science
Queen Mary University of London
dawn.black@eecs.qmul.ac.uk
christopher.harte@eecs.qmul.ac.uk

Mr. Mincong Zhang
Beijing University of Posts and Telecommunications
jp092915@bupt.edu.cn

***Abstract*—In this paper we measure the relationship between the type of distortion applied to guitar sounds and the sense of heaviness imparted. How heavy a piece of music is considered to be is potentially genre defining and, although the product of a combination of factors, overall heaviness is strongly influenced by the level and type of guitar distortion. This paper seeks to formalise this relationship.**

A database of guitar riffs, each of over 10 seconds in length, was recorded and then distorted and equalised in a number of pre-determined ways to produce a listening set of distorted guitar riffs. Listeners were then asked to compare the same riff, with different distortions applied, and indicate which version of the riff was heavier.

Results show direct correlations between the perceived heaviness of a guitar and the distortion effects applied. This has interesting applications in both the fields of emotion or mood detection in music and genre detection.

I. Introduction

Previous research and our contribution.

The perceived heaviness of a guitar riff will be due to a number of factors, including melody, tempo, distortion due to amplifier settings, additional distortion applied using pedals etc. and playing style. In order to start to model heaviness as a multidimensional space, it is necessary to limit the parameters.

In this paper the aim is to establish the influence of distortion due to the amplifier settings on the perception of heaviness. We disregard all other forms of guitar signal distortion.

not finished ...

II. Amplifier settings

Guitar amplifiers have a number of controls that can be altered by the musician to achieve a desired sound. The format of these controls will vary between make and manufacturer but all will provide a way to set the gain applied to the guitar signal and the vast majority will also include some form of equalisation (EQ) to modify its frequency content.

Different amplifiers have different controls but all have three controls, known collectively as the EQ controls, that allow the musician to amplify or attenuate the level of bass, mid and treble frequencies. Amplifiers also have a gain or drive control. Although amplifiers will not have *both* a gain and drive control, as the two controls essentially serve the same purpose (to boost the signal), there exists a difference in the way the two terms are used. A signal is boosted by increasing the gain control, however, at some point the amplifier will start to saturate and the sound will distort. This is known as overdrive or drive. In this paper the term drive is used to indicate the level of signal amplification. Hereafter, when we refer to amplifier settings, we are referring to the bass, mid, treble and drive controls.

III. Establishing heaviness

In order to determine how the heaviness of the guitar tone is influenced by the amplifier settings, we conducted a listening test comprised of guitar riffs which were subjected to different combinations of amplifier distortion. The intention was to get participants to listen to two examples of the same riff distorted by different amplifier settings and to rate which one they thought was heavier. When designing such an experiment, one cannot ignore the influence of the underlying riff on perception or potential frequency masking phenomenon. We must also consider that the sample space for four variables ranging from 0 to 10 (hereafter normalised to 0 to 1) could quickly become unmanageably large. Each of these problems is addressed in the following paragraphs.

A. Guitar samples

When assessing heaviness in terms of amplifier distortion, the underlying guitar riff will have influence. Certainly, the listeners familiarity with the riff will introduce extra-musical associations that may influence their perception. We have therefore limited our test set to two, well known, heavy guitar riffs in an attempt to minimize this effect - 'Enter Sandman' by Metallica and 'Smoke on the Water' by Deep Purple.

These guitar riffs were recorded 'clean' with no distortion applied. The recording was performed using a DI box to capture the clean signal coming straight from the guitar. To make the guitarist more comfortable, and not influence playing style, the signal was also sent to a Fender Deluxe valve amplifier to get the right tone and feedback whilst he was playing. A single performer was used to record all guitar samples.

Distorted guitar riffs were generated for the listening tests using Reaper [1] and the Waves GTR3 VST plugin [2] capable of simulating a variety of well known amplifiers. Our experiments were conducted using the high gain 'monster' amplifier based on a Marshall 100W stack.

IV. Listening test to determine Just Noticeable Difference

In order to constrain our sample space we need to determine the step-size for each variable. For example, in order to accurately judge the influence of the treble on the perception of heaviness, we need to determine a suitable step-size for the treble control. The optimum step-size will allow us to keep the number of necessary samples to a minimum whilst still accurately representing the full range of the treble setting. This step size needs to be established for each amplifier setting separately and will be equivalent to the Just Noticeable Difference (JND) for each amplifier setting.

The JND for an amplifier setting is the smallest difference in that setting that can still be perceived by a listener. To determine the JND for the amplifier settings a listening test was conducted. Listeners were required to alter a single amplifier setting so that, in their opinion, two samples of the same guitar riff sounded the same.

Participants were played a reference sample which consisted of a guitar riff that had been distorted using amplifier settings that were unknown to them. They were then played an adjustable sample which consisted of the same riff but the distortion on one amplifier setting - the target amplifier setting - was different. The participant could adjust an unmarked slider to change the target setting until the adjustable sample sounded equivalent to the reference sample. All samples could be replayed. Samples were shuffled so listeners heard a random selection of amplifier settings.

How many participants? What were the EQ settings? List the absolute specifics.

A. JND results

A total of 20 listeners took the test. Each listener was played twenty pairs of samples. In each reference and adjustable sample pair, three of the EQ settings were set to values which were randomly selected according to a Gaussian distribution. The fourth 'target' setting in the reference sample was given a value of x where $x = 0.1 + (0.2n)$ and $0 \leq n \leq 4$, $0 \leq x \leq 1$. Results are presented in figure 1 and the RMS error and error standard deviation are presented in table I.

The JND of any one amplifier setting can be considered to be the average amount of error listeners made. Figure 1 shows a box and whisker plot for the four amplifier settings. The mid and treble results appear linear and so will be discussed first.

Examining figure 1 the median lines on the mid boxplot are all within 0.05 of the target value, showing that listeners were able to estimate the amount of mid present in the signal. This observation is supported by table I which shows the rms error for the mid for all target values to be around 0.1 with small

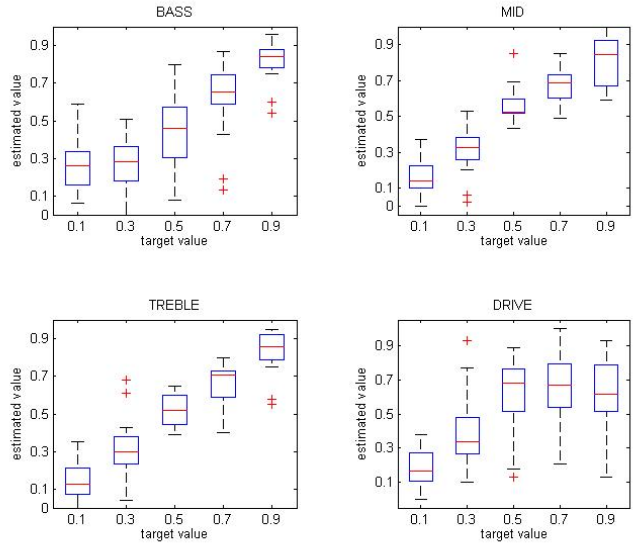

Fig. 1. JND listening test results. The central mark on each box is the median, the box edges the 25^{th} and 75^{th} percentiles, the whiskers extend to the most extreme datapoints not considered to be outliers. Outliers are marked as '+'.

EQ setting	error measure	target value 0.1	0.3	0.5	0.7	0.9
Bass	rms	0.2063	0.1389	0.1967	0.2041	0.1283
	std	0.1304	0.1408	0.1876	0.1884	0.1041
Mid	rms	0.1177	0.1227	0.1078	0.0940	0.1624
	std	0.1036	0.1251	0.0905	0.0910	0.1423
Treble	rms	0.1144	0.1525	0.0842	0.1102	0.1260
	std	0.1056	0.1556	0.0829	0.1066	0.1098
Drive	rms	0.1228	0.2245	0.2372	0.1963	0.3539
	std	0.1002	0.2124	0.2153	0.1965	0.2166

TABLE I

JND listening test error (rms = root mean square error, std = standard deviation of the error).

standard deviation. We can therefore conclude that a workable JND for the mid is 0.1.

The treble boxplot in figure 1 shows that listeners were also able to estimate the amount of treble present in the signal. The median of the estimated values are all within 0.04 of the target value. Table I shows that the rms error for all target treble values is in the region of 0.1 with a standard deviation of 0.1. We can therefore conclude that a workable JND for the treble is 0.1.

The bass results do not appear to be linear. The results suggest that listeners were less able to estimate the bass levels than the mid or treble, which would be consistent with known psychoacoustic properties of the human ear. The rms error for the bass is in the region of 0.2 with a standard deviation of 0.2 giving us a working JND of 0.2

The drive setting behaves differently to the other three settings. Figure 1 shows that the estimation accuracy decreases significantly as the drive increases and table I shows the rms error increasing from 0.1 to 0.4 with a standard deviation of around 0.2. It is clear that further research is necessary to fully gage the human ear's response to the the drive setting, however, for this paper a working drive JND of 0.2 is used for drive values below 0.5, and a JND of 0.3 for values ≥ 0.5 is used.

Having determined working values for the JND of our amplifier settings, we can now select suitable step-sizes for our main listening test. The step-size should not be equal to the JND, otherwise listeners might be able to only just notice the difference between two samples if they were one step apart. The step-size needs to be greater than the JND, so that sample differences are perceptible, but not so great that there are insufficient data points in the sample space to be meaningful. In consideration of these factors, the following step-sizes were used: bass = 0.2, mid = 0.2, treble = 0.2 and drive = 0.2 when the drive value < 0.5 and drive = 0.3 when the drive value ≥ 0.5.

V. Listening Test to Determine Heaviness

Need to specify here how many listeners participated, how many versions of the same riff they were asked to rate, and what the EQ settings were.

In order to discover the influence of distortion on the perception of heaviness, a further listening test was performed. Having established the JND for all four amplifier settings, a test set of distorted guitar riffs was generated using the following values for amplifier settings: bass, mid, treble = [0.1, 0.3, 0.5, 0.7, 0.9] drive = [0.1, 0.3, 0.5, 0.8]. Participants in the listening test were presented with two versions of the same riff that differed by one amplifier setting and asked to indicate which version they thought was heavier, or whether they perceived both to be equally heavy. This simple test allowed us to determine which amplifier setting has the greatest influence on the perception of heaviness.

A. Guitar sample sorting algorithm

The step sizes determined by the JND test give us a possible 500 samples to compare using a possible ($500^2/2 - 500 = 124500$) comparisons. If we constrain our listening test to ten minutes, giving each listener the time to make approximately ten comparisons, we would require 50 listeners to take the test even if every sample was to be listened to only once. In order to reduce the number of samples that need to be listened to, we can consider that there will be certain amplifier settings that would produce a sound that is not heavy at all. Whilst it would be good to eliminate such samples before the test starts, we need to avoid making any assumptions about the amplifier settings and the perception of heaviness. We have therefore introduced a sorting algorithm to the listening test. In essence we can infer that, if most listeners perceive sample A to be heavier than sample B, and B to be heavier than C, there is little point in comparing A and C. This is explained as pseudo-code in the code box below. Audio samples are assigned a weight and these are used to rank the audio in terms of perceived heaviness. Listeners are asked to compare two audio samples that have the same weight. If no two samples with the same weight are found, then two samples with different weights are randomly selected.

```
weightA = 0; weightB = 0;
    if A is heavier
        if weightA ≡ weightB
            increment weightA
        if weightA > weightB
            no change
        if weightB > weightA
            weightA = weight B + 1
         }
    }
```

Pseudo-code to illustrate the sorting algorithm

B. Listening test results

References

[1] Cockos Incorporated. Reaper: digital audio workstation. http://www.reaper.fm.

[2] Waves Audio Ltd. Waves: developer of signal processing solutions. http://www.wavesgtr.com.