The Impact of Audio Effects Processing on the Perception of Brightness and Warmth

Gary Bromham*
Centre for Digital Music
Queen Mary University of London
g.bromham@qmul.ac.uk

David Moffat
Centre for Digital Music
Queen Mary University of London
d.j.moffat@qmul.ac.uk

Mathieu Barthet
Centre for Digital Music
Queen Mary University of London
m.barthet@qmul.ac.uk

Anne Danielsen
RITMO
University of Oslo
anne.danielsen@imv.uio.no

György Fazekas
Centre for Digital Music
Queen Mary University of London
g.fazekas@qmul.ac.uk

ABSTRACT

It is not uncommon to hear musicians and audio engineers speak of *warmth* and *brightness* when describing analog technologies such as vintage mixing consoles, multitrack tape machines, and valve compressors. What is perhaps less common, is hearing this term used in association with retro digital technology. A question exists as to how much the low bit rate and low-grade conversion quality contribute to the overall *brightness* or *warmth* of a sound when processed with audio effects simulating early sampling technologies. These two dimensions of timbre are notoriously difficult to define and more importantly, measure. We present a subjective user study of *brightness* and *warmth*, where a series of audio examples are processed with different audio effects. 26 participants rated the perceived level of *brightness* and *warmth* of various instrumental sequences for 5 different audio effects including bit depth reduction, compression and equalisation. Results show that 8 bit reduction tends to increase *brightness* and decrease *warmth* whereas 12 bit reduction tends to do the opposite, although this is very much dependent on the instrument. Interestingly, the most significant brightness changes, due to bit reduction, were obtained for bass sounds. For comparison purposes, instrument phrases were also processed with both an analogue compressor and an equalisation plugin to see if any subjective difference was noticed when simulating sonic characteristics that might be associated with warmth. Greater significance was observed when the sound excerpts were processed with the plugins being used to simulate the effects of bit depth reduction.

CCS CONCEPTS

• **Information systems** → **Speech / audio search**; • **Applied computing** → **Sound and music computing**; *Media arts*; • **Computing methodologies** → *Perception*; • **Software and its engineering** → *Semantics*.

KEYWORDS

Brightness, Warmth, Retro, Nostalgia, Music Perception, Timbre, Digital Audio Effects

1 INTRODUCTION

One area that dominates any discussion about retro digital technology is the lo-fi (low fidelity) aesthetics produced by early hardware samplers such as the Fairlight CMI, Akai Linn MPC 60 or Emu SP12. Our study investigates how much the low bit rate and low-grade conversion quality contributed to the sonic character of these iconic instruments. Two specific timbral dimensions that have become associated with this phenomenon are those of relative *brightness* and *warmth* perceived when processing audio with such processors. Since this has become common in colloquial language among audio professionals, it is valuable to measure whether there is agreement in the description of these effects in relation to different sounds and whether it can be quantified scientifically.

1.1 The Task

The task is to measure the impact of applying audio effects processing to different sound examples and to assess their influence on the perception of *brightness* and *warmth*. These two descriptors were chosen for several reasons. Besides their common use by the professional audio production community[33], these terms are also frequently used to characterise sounds in the academic literature as well as to tag sounds in sound libraries. They are also frequently used as search terms for sounds as demonstrated in [33], with brightness being the third and warmth being the 20th most commonly used descriptor in such contexts.

1.2 Study

The purpose of this study is to identify the perceived impact of a range of audio effect processing techniques on the semantic terms of 'brightness' and 'warmth'. We are specifically interested in how a reduction in bit depth alters the perception of these two dimensions, as they have become associated with the retro sound of early digital technologies. This was performed by manipulating seven different audio samples with four different effect processes. These include two forms of bit rate reduction, analog compression and equalization. A dry unprocessed example is included for reference purposes. A further task in our study is for participants to describe why they made specific choices. This may allow us to ascertain how listeners verbalise the most prominent perceptual effects of the transformations. A deeper analysis of the qualitative data is outside the scope of this paper, but it is discussed in the authors' related publication [4].

The rest of this paper is organised as follows. In Section 2, we discuss relevant works on the evaluation of brightness and warmth with a particular emphasis on the effects of bit rate reduction. In Section 3, we outline our research methodology, including the experiment design, audio material and participant data. In Section 4, we present our results while we discuss the main observations in Section 5. Section 6 presents our conclusions and finally we outline our future work in Section 7.

2 BACKGROUND

A wide range of published studies were considered in our research when designing the experiments discussed in Section 3. It has been shown that audio effect processing is an important field of study, and can greatly impact the perception of sounds, from music production [5] to sound synthesis [28]. Researchers have looked into the impact of audio effects, including dynamic range compression [5] and equalization [40]. An investigation into the use of a specific subjective descriptor, aggressive, attributed to the ubiquitous UREI 1176 compressor also proved useful in the paper by Moore and Wakefield where they adopt a similar approach to this study [32].

Low bit rate audio is discussed in the book 'Digital Audio Signal Processing' by Zölzer [46], where the effects of quantization noise (in Chapter 2) and AD/DA conversion (in Chapter 3) are introduced specifically. The website, DSP Illustrations, gives an excellent audio representation of the perceptual effects[1]. The effects of low bit depth are discussed in an article by Cascone when describing the deliberate use of the conversion process as an aesthetic statement [12].

Warmth and *Brightness* have been discussed in several publications. Robjohns [35] provides an overview on the subjective area of analogue warmth. The two dimensions of *brightness* and *warmth* are evaluated in some detail in works by Brookes and Williams [6], Barthet et al. [2] and Enderby and Stables [15] In their paper, they offer a non-linear method for manipulating *brightness* and *warmth*, through analysis of audio features [30], which can then be used to produce sound effect classification and hierarchical clustering [31]. The subject has also been formerly discussed in

the paper on additive synthesis techniques for modifications of *brightness* and *warmth* by Zacaharakis and Reiss [45].

Researchers have explored the use of semantic descriptors and their relationship to timbre perception [44]. Seminal works on the perception of musical timbre by Grey and McAdams can be found in [19, 27]. The perceptual realism of a range of synthesised sound effects was evaluated in [29]. The method for subjective evaluation was taken as inspiration for this research.

Processing audio examples by semantic descriptors and the inherent timbral transformations which take place are discussed in works by Caetano and Rodet in [8, 9]. A similar theme is explored by Kazazis et al. in their 2016 DAFx paper on sound morphing by audio descriptors in [26]. Timbral attributes for sound effects searching are also discussed by Pearce, Brookes and Mason [34]. Specific works on guitar and the timbral effect of adding distortion are evaluated by Tomczak and Stables [41] and Tsumoto et al. [42]. The background covers a diverse cross-section of existing knowledge in this subject area from different perspectives, including soundscapes, music and audio production, and automatic mixing.

A great deal of literary work exists on the subject of retro aesthetics. Elizabeth Guffey discusses the idea of retro being a connection with a recent past without nostalgic attachment in her book, 'Retro: The Culture of Revival' [20]. Brøvig-Hanssen and Danielsen in 'Digital Signatures' discuss the concept of early digital sonic limitations and how they shape our perspective of 'lofi' aesthetics [7]. Joseph Auner makes further reference to this in 'Making Old Machines Speak' [1].

Figure 1 shows a subjective scale proposed by Bob Katz [25] for describing semantic sound properties in terms of the attributes exhibited by the source at different frequencies. In the context of our research it is interesting that the sound properties are seen as opposites when adding or removing certain frequencies. Less bright can indeed mean more warmth in certain circumstances.

3 SUBJECTIVE EVALUATION EXPERIMENT

This section discusses the programme material used as basis for the experiment, the choice of effect processors used for producing the sound stimuli as well as the demographics of the participants, and outlines the procedure and methodology used in this study.

3.1 Stimuli

Seven different sounds samples, composed by the author, that were used for this test. They are Acoustic Bass; Acoustic Drums; Acoustic Guitar; Cinematic Strings; Electronic Drums; Synthesised Bass and Synthesizer keyboard pads. In each of these cases, participants were presented with four different processed samples along with a dry unprocessed sample which was included for reference purposes. The types of processing performed were Analogue Compression, Bit Crushing, Equalisation and 12 bit Retro Colour Plugin.

Four commercial audio processing plugins were selected due to their wide use in professional audio production and the samples were processed in a digital audio workstation. Due to the fact that the selected plugins are not available open source, it is challenging to disclose how the processing is performed precisely. However, this issue is partially addressed by analysing the performance of the plugins using standard test signals. Furthermore, the focus of this

---

[1]https://dspillustrations.com/pages/posts/misc/how-does-quantization-noise-sound.html
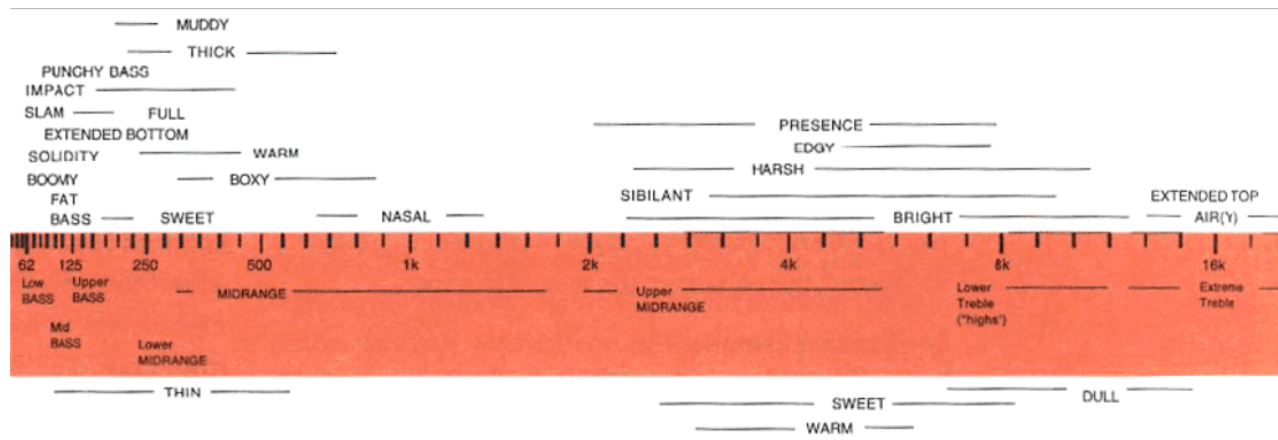
Figure 1: Subjective Terms we use to describe Excess or Deficiency of various frequencies. Taken from [25]

Figure 2: Decimort 2 Plugin

Figure 3: XLN Audio Retro Colour

Figure 4: Logic pro X EQ Plugin

Figure 5: Tubetech CL1B Analog Compressor

paper is to study the effect of typical audio processing tools that are commonly used by the professional audio community, which we hope to generalise from.

The four effects processors selected were:

- D16 Decimort 2 Plugin [Figure 2]
- XLN Audio Retro Colour [Figure 3]
- Logic Pro X Channel EQ [Figure 4]
- Tubetech CL1B Analog Opto Compressor [Figure 5]

Decimort 2, as seen in Figure 2, is a bit crusher plugin which was used to simulate the effect of 8 bit reduction. Retro Colour, as seen in Figure 3 is a lo-fi processing plugin used to simulate 12 bit processing. The two bit rate plugins were chosen because of their inherent ability to simulate the effects of bit depth reduction. Though neither plugins are open source they are very well respected in the audio industry. A similar approach in previous work by Bitzer et al. in 2006 when studying parameter estimations for dynamic range compressors can further justify this approach. The Logic Pro X Channel in Figure 4 EQ plugin is used to create a low mid lift around 350 Hz and high frequency roll off around 8 KHz. The intention was to approximate what could be perceived as an increase in the dimension of *warmth* as a reference. Finally, the Tubetech CL1B analogue compressor in Figure 5 was used with a 2:1 ratio, fast attack (0.5ms) and slow release (3 secs) and moderate gain reduction, also to see if smoothing transients might alter
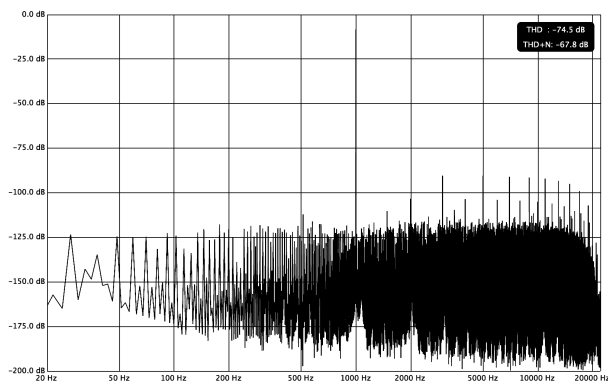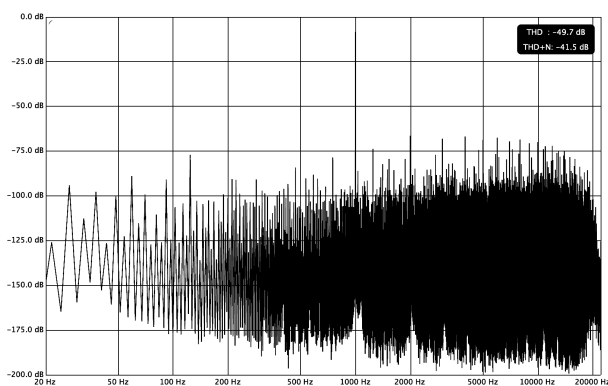
Figure 6: 12 bit audio burst processing

Figure 7: 8 bit audio burst processing

the perception of our two dimensions[5]. For additional reference purposes a Dry sound example was also included.

To further illustrate and justify the two plugins choices used specifically for bit rate reduction, and to quantify the influence of the instrument effect, Figure 6 and Figure 7 show the effect of passing a 1KHz sine wave through both the 8 bit setting in the Decimort 2 plugin and the 12 bit setting in the Retro Colour plugin. This test was conducted in the Plugin Doctor Utility made by DDMf Software[2]. The software acts as a VST host which facilitates the testing of audio plugins. The figures show a marked difference in total harmonic distortion and distortion components. The harmonic distortion displayed in the 8 bit example is almost 20 dB higher than that seen in the 12 bit example. This might offer at least some explanation as to the shift in spectral balance.

To prepare the stimuli for the listening experiemnt, all samples were loudness matched and normalized to -26 dB LUFS (Loudness Units Full Scale [21]) in accordance with [22]. Samples were recorded as WAV files, and were both processed and presented to the participants at a sample rate of 44.1kHz.

3.2 Participants

Twenty six participants completed the listening test (15 males, 11 females). These consisted predominantly of Music Production students with at least three years of recording and mix engineering experience. The responses of five subjects were discarded before the analysis on account of them not finishing the experiment. Among the test participants, there were two award-winning mix engineers who reported themselves as experts and more than half of the participants reported above average experience. The age range was between 20 and 51 years. Participants were told they were welcome to take regular breaks during the test. The average test duration was 18 minutes 12 seconds, and as such fatigue was not an issue [36]. The experiment was approved by the QMUL ethics committee.

3.3 Procedure and Methodology

The listening test was conducted using the Web Audio Evaluation Tool [23, 24].Participants were asked to rate each set of audio samples on their *warmth* and *brightness*, relative to all the other processed samples. Participants were provided with two continuous linear scales, one was labelled from *least warm* to *most warm*, and the other was *least bright* to *most bright*. In each case, all processed audio samples were compared to the dry samples as a hidden reference. All sounds were rated on two separate horizontal scales, to encourage inter-sample comparison. Participants did not have any other information regarding the samples. Both the ordering of sound samples and the ordering of effect processing of the samples within a category were randomised.

The experiment was conducted as an online listening test. Participants were asked to use a pair of high quality headphones such as the AKG K240 or Sennheiser HD 600 Open models to conduct the test. They were asked to adjust the volume to a comfortable listening level and refrain from adjusting the volume for the duration of the test. We acknowledge that choosing headphones to conduct the test is a potential limitation, however, it wasn't practical to conduct the test under lab conditions, due partly to easier access to students in a classroom setting and professionals in their usual work environment. By doing so, we were able to gather more participants, from a wider demographic of professional and semi-professional engineers and producers. The listening test was made available online[3].

Figure 8 shows an example of one of the test pages [24] from the experiment. The Web Audio Evaluation tool is an evolution of the Audio Perception Evaluation toolbox (APE) developed in Matlab by De Man and Reiss [13].

4 RESULTS

The data was analysed in both Matlab and Python. Matlab was used to generate plots whilst statistical significance of pairwise comparisons was completed in Python. The distribution of the data was not normal, therefore the non-parametric Friedman test was chosen for analysis. Bonferroni correction was also applied.

The Friedman test between the perceived *brightness* of samples showed there was a significant difference between samples, presenting a test statistic of $\chi^2 = 85.14$ ($p = 1.418e - 17$), and similarly for *warmth*, a significant difference is also reported, presenting a test
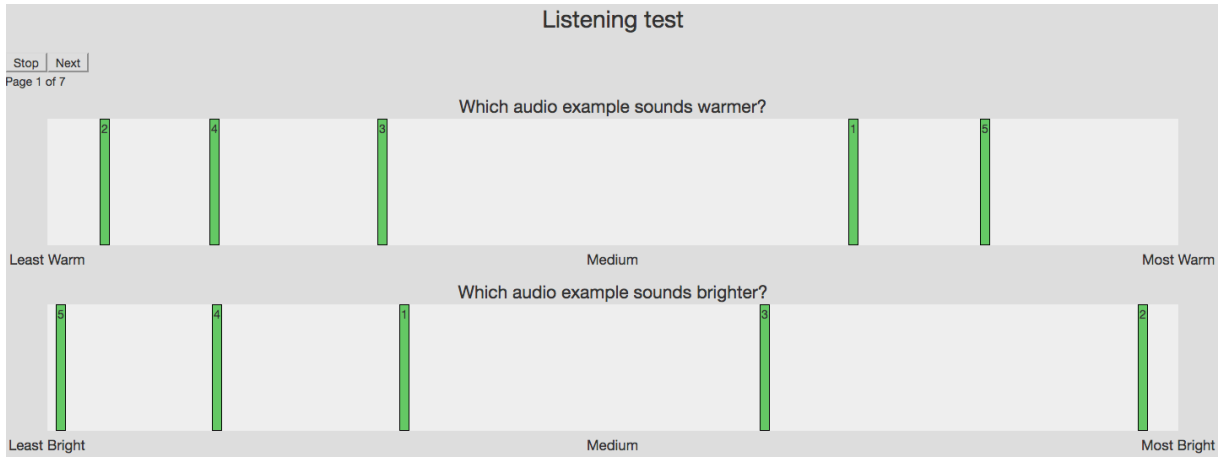

Figure 8: Screenshot of the listening test interface, constructed with the Web Audio Evaluation Tool [24]

	Compressor		Bit Crusher		Equaliser		Retro		Dry	
	Brightness	Warmth	Brightness	Warmth	Brightness	Warmth	Brightness	Warmth	Brightness	Warmth
Compressor	-	-	o	***	***	***	***	**	***	*
Bit Crusher	o	***	-	-	***	***	***	***	o	***
Equaliser	***	***	***	***	-	-	o	o	***	*
Retro	***	**	***	***	o	o	-	-	***	*
Dry	***	*	o	***	***	*	***	*	-	-

Table 1: Pairwise Comparison of Significance, with post-hoc Wilcoxon signed-rank test, with bonferoni correction, for each Processor Type. *o = p>0.05, * = p<0.05, ** = p<0.01, *** = p<0.001*
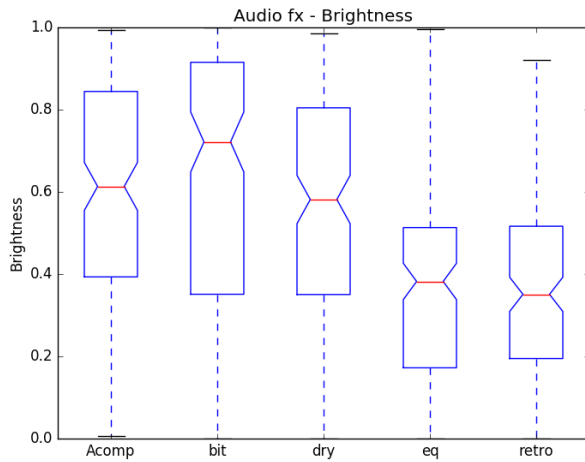

Figure 9: Boxplot of the brightness rating of every processed sample
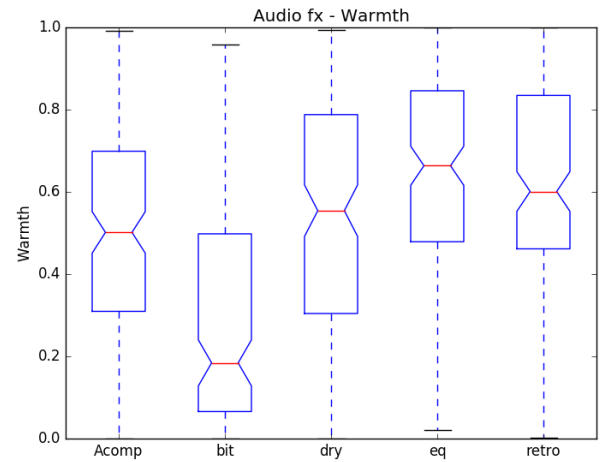

Figure 10: Boxplot of the warmth rating of every processed sample

statistic of $\chi^2 = 108.6$ ($p = 1.427e - 22$). As these results are significant, multiple comparison post-hoc analyses using the Wilcoxon signed-rank test with Bonferoni correction were performed on data for both the *brightness* and *warmth* ratings. Significant results are presented in Table 1.

Figure 9 shows overall *brightness* across all audio examples whilst Figure 10 shows *warmth*. The boxplots generated in both examples show that there is high agreement across all participants about the effects of the 8-bit processing (shown as 'bit' in the plot) being *most bright* and *least warm*. The converse is true in the case of the 12-bit
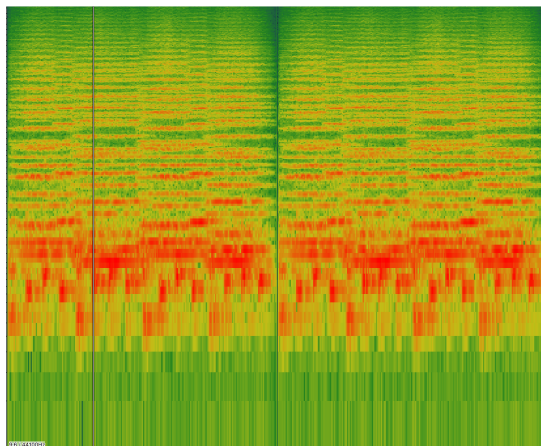
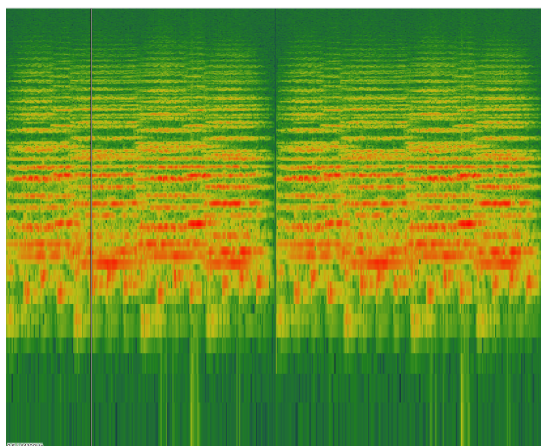Figure 11: Spectrogram of unprocessed dry Cinematic Strings

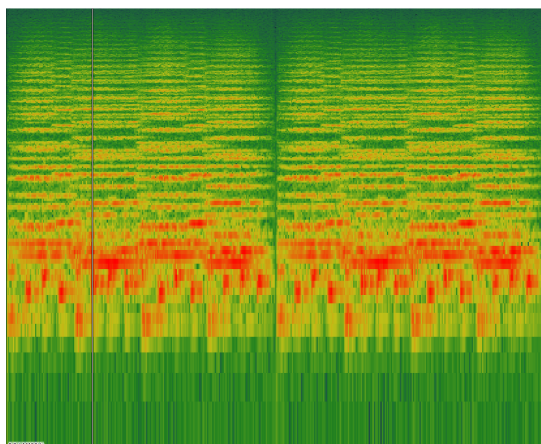Figure 12: Spectrogram of processed 8 bit Cinematic Strings

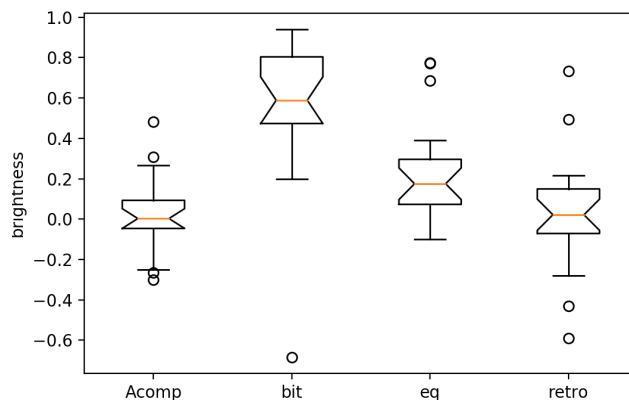Figure 13: Spectrogram of processed 12 bit Cinematic Strings

Figure 14: Boxplot of the change in brightness rating for synthesised bass when comparing the dry unprocessed with the various processed examples
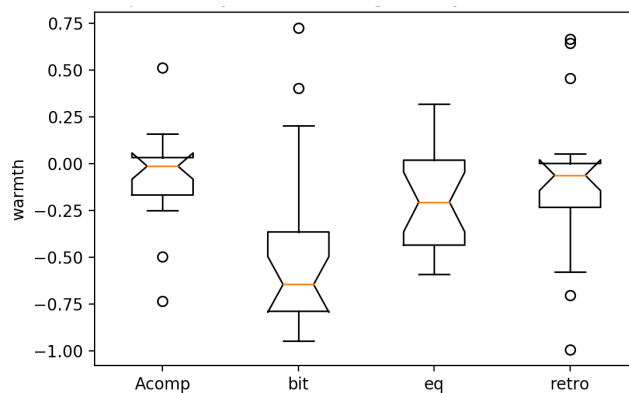

Figure 15: Boxplot of the change in warmth rating for synthesised bass when comparing the dry unprocessed with the various processed examples

processing (shown as 'retro' in the plot). Interestingly, the effects of equalising the sound to approximate 'warmth' as shown in Fig 4 creates a similar effect to retro.

The three spectrograms in Figure 11, 13 and 12 were created in Sonic Visualiser [10] using the Cinematic Strings sample. We clearly see a degradation in the spectral balance of the sounds when either 12 and 8 bit processing are applied. There is a direct correlation between increased *brightness* and reduced *warmth* perceived by the participants as the sound deteriorates.

A further important observation can be made when comparing the processed sounds with the dry, unprocessed, example. This makes itself particularly noticeable in the case of the synthesised bass sound where there is a significant increase in the perception of *brightness* for the 8 bit sound but a corresponding decrease for the dimension of *warmth*. As can be seen in figures 14 and 15 the difference from the dry sound is far less prevalent for all other audio effect processing. Values greater than zero on the 'Y' axis in figure 14 show an increase in brightness whist values less than zero

show a decrease. Similarly, figure 15 shows an increase in warmth for values above zero and a decrease for values below.

Table 1 shows direct pairwise comparisons between each processor type. Most interestingly, the largest significant differences are seen when comparing the 8 bit processed sound with the dry example. This is less significant for the 12 bit retro sound. This gives some indication of the participants ability to differentiate between the two types of bit reduction.

5 DISCUSSION

The results show that there is an interaction effect on both *brightness* and *warmth* between audio effect processing and instrument selection. Interestingly, 8 bit reduction tends to increase *brightness* but the converse is true for 12 bit processing. This could be explained by the fact that the most significant effect is seen in the case of bass sounds. Results also show that there is a significant effect of changing the audio effect processing on the dimension of *brightness* and *warmth*. Whilst 12 bit sounds are shown to be diminished in *brightness*, 8 bit ones tend to show the opposite effect. This may be explained by the distortion type effect, as a result of the 8 bit, bit reduction processing and the effects of band limiting on the low frequency content. This may also be a result of quantisation noise becoming more prevalent and more audible. The distortion style effect of the 12 bit processor is less pronounced but the effects of band limiting are noticeable enough to be perceived as warmer. Less *brightness* can often be perceived as increased *warmth* as is seen from our data. This is commonly known as a 'tilt' effect in equalisation where making a sound brighter is sometimes achieved by removing bass or lower mid frequencies. Figures 6,7,11-13 clearly show a change in the spectral balance.

Better understanding of the perceptual effects of audio processing tools may have several beneficial applications. Firstly, the use of semantic descriptors in music production tools is becoming increasingly common and widespread. Plugins by Waves, Izotope and XLN Audio are actively embracing this method for manipulating sound. Identifying audio effect processes which exemplify and most easily demonstrate the use of descriptors such as *warmth* and *brightness* is important for this area of research. In intelligent music production research, audio effects processors responding to the descriptive language of a user have been discussed in some detail by Cartwright and Pardo in their paper on Social-EQ [11]. Similarly, the subject of adaptive audio effects, where configuration is determined by high-level semantic metadata, has been discussed by Wilmering et al. [43]. An understanding of perceived effects of audio transformation may also inform the design of production related ontologies [16], while such data can be encoded semantically to aid production tools [17, 39] and help, for instance, the design or assessment of intelligent tools for automatic control of audio effects[38].

Our paper builds on existing knowledge and offers further evidence that effects which limit and change bandwidth will often accentuate changes in dimensional perceptual attributes. In our case, bit reduction effects could be considered commensurate with such processes as equalisation and dynamic range compression.

There were several limitations of this experiment which could be addressed in the future. The fact that the bit-reduction plugins were

not open source was acknowledged as a potential problem, but the fact that the two plugins selected for the experiment are widely used in studio production should provide some degree of justification for their use. Additionally, the fact that the experiment was conducted using headphones and not under lab conditions introduced a further variable. The motivation behind this was purely to accumulate more professional and expert participants for whom it would not have been possible to extricate from their busy work schedules.

6 CONCLUSION

The aim of this paper was to measure perceived differences in *brightness* and *warmth* for changes in audio effects. As a sound is transformed through bit depth reduction, the inherent timbral qualities will change. There will be a resultant change in high and/or low frequency content which can be perceived as more or less *warm* or *bright*. Depending on the sound example used for the listening test, the subject will experience varying degrees of this phenomenon. The results from our experiment clearly show a significant change between 8 and 12 bit processing, particularly in the case of bass sounds or materials with greater low frequency content. Some of this may be explained by the fact that more aliasing is present, which is itself, introduced as the bit depth is progressively reduced. The incremental change in aliasing may be perceived as an increase in distortion to the listener, this may in turn be perceived as increased *brightness* and/or reduced *warmth*.

We may also conclude from our experiment that there is still something of a disconnect between tools used by music professionals and those used by researchers which could be construed as problematic for future research. In our specific case, the two audio effects used to simulate bit reduction, Decimort 2 and Retro Colour, though acknowledged to be very effective by audio professionals, were not recognised academically, due to their lack of open source access. Though understandable, this precludes tools from being used which might otherwise contribute to existing knowledge.

Retro digital sampling technology was not viewed as having sonic limitations when it was introduced. Lo-fi aesthetics produced by early digital samplers are now revered in a nostalgic way, often referred to as warm, despite once being criticised as cold and sterile. Audio effects that attempt to replicate this condition are often measured in terms of their ability to 'transport the listener back in time'. Relative, and comparative, *warmth* and *brightness* are effective scales of evaluating such phenomena.

7 FURTHER WORK

There are several areas where the research could be developed. The dimension of warmth suggests many things in terms of nostalgia and retro that have been investigated as an aesthetic expression or cultural phenomenon, but not necessarily from a technical perspective. The shaping and enveloping of sound also produces perceptual attributes which may be connected with warmth and brightness, these may in turn provide clues which deserve a more scientific approach to evaluation. This field has already been explored in previous work by the author, but which also require further investigation [5]. When softening transients or slewing time constants this can sometimes be perceived as sounding more retro or warm.

We also acknowledge that the auditory perceptual dimensions such as *warmth* and *brightness* may also be skewed by a bias towards skeuomorphism. Though this is beyond the scope of this study, it would be valuable to investigate the impact of introducing a visual element. The effects and influence of a graphical user interfaces need further investigation. A plugin may rely on metaphors or traditional paradigms which suggest reference points in the past, colours, for example, such as blue or orange may suggest attributes associated with *brightness* or *warmth*.

Finally, we can emphasise that systematic assessment of the perceived effects of different types of audio transformations and their rating along commonly used perceptual dimensions could be a valuable contribution to several fields, including intelligent music production and automatic mixing [14], sound synthesis [37], and a broad range of semantic audio applications [18].

Acknowledgements

REFERENCES

[1] Joseph Auner. 2000. Making old machines speak: images of technology in recent music. *ECHO: a music-centered journal* 2, 2 (2000), 2.
[2] Mathieu Barthet, Philippe Guillemain, Richard Kronland-Martinet, and Sølvi Ystad. 2010. From clarinet control to timbre perception. *Acta Acustica united with Acustica* 96, 4 (2010), 678–689.
[3] Joerg Bitzer, Denny Schmidt, and Uwe Simmer. 2006. Parameter estimation of dynamic range compressors: models, procedures and test signals. In *Audio Engineering Society Convention 120.* Audio Engineering Society.
[4] Gary Bromham, Anne Danielsen, and Gyorgy Fazekas. 2020. 'Digital Warmth': The Retro in Digital. *Journal of the Art of Record Production (In Press)* (2020).
[5] Gary Bromham, Dave Moffat, Mathieu Barthet, and György Fazekas. 2018. The impact of compressor ballistics on the perceived style of music. In *Audio Engineering Society Convention 145.* Audio Engineering Society.
[6] Tim Brookes and Duncan Williams. 2007. Perceptually-motivated audio morphing: Brightness. In *Audio Engineering Society Convention 122.* Audio Engineering Society.
[7] Ragnhild Brøvig-Hanssen and Anne Danielsen. 2016. *Digital signatures: The impact of digitization on popular music sound.* MIT Press.
[8] Marcelo Caetano and Xavier Rodet. 2013. Musical instrument sound morphing guided by perceptually motivated features. *IEEE Transactions on Audio, Speech, and Language Processing* 21, 8 (2013), 1666–1675.
[9] Marcelo Freitas Caetano and Xavier Rodet. 2010. Automatic timbral morphing of musical instrument sounds by high-level descriptors. In *International Computer Music Conference.* 11–21.
[10] Chris Cannam, Christian Landone, and Mark Sandler. 2010. Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the 18th ACM international conference on Multimedia.* ACM, 1467–1468.
[11] Mark Brozier Cartwright and Bryan Pardo. 2013. Social-EQ: Crowdsourcing an Equalization Descriptor Map.. In *ISMIR.* 395–400.
[12] Kim Cascone. 2000. The aesthetics of failure:"Post-digital" tendencies in contemporary computer music. *Computer Music Journal* 24, 4 (2000), 12–18.
[13] Brecht De Man and Joshua D Reiss. 2014. APE: Audio perceptual evaluation toolbox for MATLAB. In *Audio Engineering Society Convention 136.* Audio Engineering Society.
[14] Brecht De Man, Joshua D Reiss, and Ryan Stables. 2017. Ten years of automatic mixing. In *Proceedings of the 3rd Workshop on Intelligent Music Production.*
[15] Sean Enderby and Ryan Stables. 2017. A nonlinear method for manipulating warmth and brightness. In *Proceedings of the 20th International Conference on Digital Audio Effects, Edinburgh, UK.*
[16] G. Fazekas and M. Sandler. 2011. The Studio Ontology Framework. In *Proc. of the 12th International Society for Music Information Retrieval (ISMIR'11) conference, 24-28 Oct., Miami, Florida, USA.* 24–28. http://ismir2011.ismir.net/papers/PS3-20.pdf
[17] G. Fazekas and M. Sandler. 2012. Knowledge Representation Issues in Audio-Related Metadata Model Design. In *Proc. of the 133rd Convention of the Audio Engineering Society, San Francisco, CA, USA.* https://secure.aes.org/forum/pubs/conventions/?elib=16507
[18] G. Fazekas and T. Wilmering. 2012. Semantic Web and Semantic Audio Technologies. Tutorial presented at the 132nd Convention of the Audio Engineering

[19] Society, Budapest, Hungary. https://doi.org/10.5281/zenodo.1324499
[19] John M Grey. 1975. An exploration of musical timbre. *Ph. D dissertation Stanford University* (1975).
[20] Elizabeth E Guffey. 2006. *Retro: The culture of revival.* Reaktion Books.
[21] Rec ITU-R. 2011. ITU-R BS. 1770-2, Algorithms to measure audio programme loudness and true-peak audio level. *International Telecommunications Union, Genewa* (2011).
[22] ITU-R BS.1387-1. 1998. *BS. 1387, Method for objective measurements of perceived audio quality.* Technical Report. ITU-R.
[23] Nicholas Jillings, Brecht De Man, David Moffat, Joshua D. Reiss, and Ryan Stables. 2015. Web Audio Evaluation Tool: A Browser-Based Listening Test Environment. In *Proceedings of the International Sound and Music Computing Conference.* Maynooth, Ireland.
[24] Nicholas Jillings, David Moffat, Brecht De Man, and Joshua D. Reiss. 2016. Web Audio Evaluation Tool: A framework for subjective assessment of audio. In *Proc. 2nd Web Audio Conference.* Atlanta, Georgia, USA.
[25] Bob Katz and Robert A Katz. 2003. *Mastering audio: the art and the science.* Butterworth-Heinemann.
[26] Savvas Kazazis, Philippe Depalle, and Stephen McAdams. 2016. Sound Morphing by Audio Descriptors and Parameter Interpolation. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16).* Brno, Czech Republic.
[27] Stephen McAdams and Bruno L Giordano. 2009. The perception of musical timbre. *The Oxford handbook of music psychology* (2009), 72–80.
[28] Lucas Mengual, David Moffat, and Joshua D. Reiss. 2016. Modal Synthesis of Weapon Sounds. In *Proc. Audio Engineering Society Conference: 61st International Conference: Audio for Games.* Audio Engineering Society, London.
[29] David Moffat and Joshua D. Reiss. 2018. Perceptual Evaluation of Synthesized Sound Effects. *ACM Transactions on Applied Perception (TAP)* 15, 2 (March 2018).
[30] David Moffat, David Ronan, and Joshusa D. Reiss. 2015. An Evaluation of Audio Feature Extraction Toolboxes. In *Proc. 18th International Conference on Digital Audio Effects (DAFx-15).*
[31] David Moffat, David Ronan, and Joshusa D. Reiss. 2017. Unsupervised Taxonomy of Sound Effects. In *Proc. 20th International Conference on Digital Audio Effects (DAFx-17).* Edinburgh, UK.
[32] Austin Moore and Jonathan Wakefield. 2017. An Investigation into the Relationship Between the Subjective Descriptor Aggressive and the Universal Audio 1176 FET Compressor. In *Audio Engineering Society Convention 142.*
[33] Andy Pearce, Tim Brookes, and Russell Mason. 2016. Hierarchical ontology of timbral semantic descriptors. *Audio Commons Deliverable* (August 2016). https://www.audiocommons.org/materials/
[34] Andy Pearce, Tim Brookes, and Russell Mason. 2017. Timbral attributes for sound effect library searching. In *Audio Engineering Society Conference: 2017 AES International Conference on Semantic Audio.*
[35] Hugh Robjohns. 2010. Analogue warmth: the sound of tubes, tape and transformers. *Sound On Sound* (2010).
[36] Raimund Schatz, Sebastian Egger, and Kathrin Masuch. 2012. The impact of test duration on user fatigue and reliability of subjective quality ratings. *Journal of the Audio Engineering Society* 60, 1/2 (2012), 63–73.
[37] Rod Selfridge, David Moffat, Eldad J. Avital, and Joshua D. Reiss. 2018. Creating Real-Time Aeroacoustic Sound Effects Using Physically Informed Models. *Journal of the Audio Engineering Society* (2018).
[38] D. Sheng and G. Fazekas. 2019. A Feature Learning Siamese Model for Intelligent Control of the Dynamic Range Compressor. In *Proc. of the International Joint Conf. on Neural Networks (IJCNN), July 14-19, Budapest, Hungary.* https://arxiv.org/pdf/1905.01022.pdf
[39] R Stables, B De Man, S Enderby, JD Reiss, G Fazekas, and T. Wilmering. 2016. Semantic description of timbral transformations in music production. In *Proc. ACM Multimedia, Oct. 15-19, Amsterdam, Netherlands.* 337–341. https://doi.org/10.1145/2964284.2967238
[40] Spyridon Stasis, Ryan Stables, and Jason Hockman. 2016. Semantically controlled adaptive equalisation in reduced dimensionality parameter space. *Applied Sciences* 6, 4 (2016), 116.
[41] Maciek Tomczak and Ryan Stables. 2015. *The salience of MFCC semantic classification on electric guitar recordings.* Master's thesis. Birmingham City University.
[42] Koji Tsumoto, Atsushi Marui, and Toru Kamekawa. 2016. The effect of harmonic overtones in relation to "sharpness" for perception of brightness of distorted guitar timbre. In *Proceedings of Meetings on Acoustics 172 ASA*, Vol. 29. Acoustic Society of America.
[43] Thomas Wilmering, György Fazekas, and Mark B Sandler. 2012. High-level semantic metadata for the control of multitrack adaptive digital audio effects. In *Audio Engineering Society Convention 133.* Audio Engineering Society.
[44] Asterios Zacharakis. 2013. *Musical timbre: bridging perception with semantics.* Ph.D. Dissertation. Queen Mary University of London.
[45] Asteris Zacharakis and Joshua Reiss. 2011. An additive synthesis technique for independent modification of the auditory perceptions of brightness and warmth. In *Audio Engineering Society Convention 130.* Audio Engineering Society.
[46] Udo Zölzer. 2008. *Digital audio signal processing.* John Wiley & Sons.