University of Huddersfield Repository

Fenton, Steven and Lee, Hyunkook

Towards a Perceptual Model of 'Punch' in Musical Signals

Original Citation


This version is available at http://eprints.hud.ac.uk/26288/

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

http://eprints.hud.ac.uk/
Research Background

- Part of a wider study on finding new metrics to quantify elements of a complex mix.
- Stress ‘Towards’.
- Application in MIR, auto-mixing and auto-mastering tools
- Tools that are meaningful to both the engineer and musician alike.
Punch - A ‘formal-ish’ definition

- Semantics (from Ancient Greek: σημαντικός sēmantikós, "significant") is the study of meaning. It focuses on the relation between signifiers, like words, phrases, signs, and symbols, and what they stand for; their denotation.

- A term often thrown around but opinions differ on its meaning in a musical sense.

- ‘Thump’ is another one, the clue is in the name

- “Specifies whether the strokes on drums and bass are reproduced with clout, almost as if you can feel the blow.” [1]

Blown-Away Guy - Maxell
Punch - A more formal definition?

- A characteristic related to dynamics is known as ‘punch’. Punch can be described as a short period of significant change in power in a piece of music or performance. In essence, productions that do not possess any transient information cannot possess punch. Thus, punch is both related to transient change and the energy density at a particular moment in time and duration [2]. Furthermore, dynamic change in particular frequency bands contribute to the overall perception of punch perceived by the listener and this is inherently affected by the overall average loudness level at that time [3].

Evaluation / Goals

- Evaluate listeners perception of punch.
- Objective measure.
- Intelligent Processing rather than brute force.
- Another ‘Search’ parameter that can be utilised in music selection.
- Produce a real-time meter - More production tools.

Audio Example 1

Mr Brightside – The Killers.
Mr Brightside – The Killers.

Short term loudness is -10 LU for both excerpts.
Some Measures to start with..

- Crest Factor - Peak To Average
- PLR - Peak to Loudness
  - Can be used to indicate some microdynamic content, eg spikes.
- Dynamic Spread - A statistical approach to quantify the general spread of a time varying loudness signal. [6]. Calculation of the variation about the mean.
- All of these measures are great but have limitations with respect to application to punch. They’re not weighted, they integrate, they rely on peak levels..

Better Measures

- Octave Range CF (TT Dynamics Meter).
- IBR (Inter Band Dynamics). [5]
- Loudness range - Quantifies the variation in a time-varying loudness measurement. Based on analysis of the ‘Short Term’ loudness data. (Formerly ‘consistency’). Statistical approach eliminating top 5% and bottom 10%.
- Microdynamic -LDR - Developed to quantify the microdynamic loudness variations of an audio signal [6].
- Percentile of difference values between slow and fast loudness level.
- Testing relied on ‘inherent’ notion of microdynamics by the listeners. Still based on ‘loudness’ model albeit with smaller window sizes.

- What is punch anyway?

Elicitation of punch.

- Key frequencies of interest
- Nature of the material - Instrumentation
- Temporal / Frequency interaction
- Extraction of key modeling parameters
- Identify any other existing parameters that may also correlate
Dynamics = Punch = Transients

• “Thus, punch is both related to transient change and the energy density at a particular moment in time and duration”
Evaluate listeners perception of punch.

- Controlled experiment using 45 pink noise bursts
- The samples were arranged over 9 octave bands, with temporal attack times of 0, 5, 10, 20, 60ms. A fixed offset time of 40ms was used for all samples.
- The samples were presented in 16bit, mono WAV format.
- Loudness normalisation on the test samples was applied.
Loudness Normalisation

• Samples processed with temporal and frequency weighting. The frequency weighting curve was an inverse modified ‘K’ Weighting filter (BS.1770-3)

• Modifications were made to shelving filter $G = 10\text{dB}$ rather than $4\text{dB}$ and $F_c = 1\text{kHz}$ instead of $1.6\text{kHz}$.

• These modifications were based on recommendations made by Pestana et al. [7] and through testing.

Loudness Normalisation

- In addition - temporal compensation was applied.

- Based on the centre frequency of each octave band and the pulse duration.
An informal listening test took place prior to the main test to briefly evaluate the effectiveness of the loudness normalization algorithm.

Of the 11 expert listeners that took part, 7 agreed on equal loudness.

Differences they were hearing were primarily as a result of timbral differences rather than loudness.

A formal listening test is planned to evaluate the loudness normalisation.

In comparison, an ITU-R 468-4 filter model was also tested however it was found that the 2kHz-8kHz octave bands were perceptually significantly louder on playback than the lower bands.
Test Interface & Results

Mean Punch Scores per octave band re. 1Khz band 0mS onset

Octave Bands:
- Band 1 (1KHz)
- Band 2 (1.25KHz)
- Band 3 (1.6KHz)
- Band 4 (2KHz)
- Band 5 (2.5KHz)
- Band 6 (3KHz)
- Band 7 (4KHz)
- Band 8 (6.3KHz)

Punch Scores:
- Punch Score 0mS Onset
- Punch Score 5mS Onset
- Punch Score 10mS Onset
- Punch Score 20mS Onset
- Punch Score 60mS Onset
Results

- Upon interviewing the participants after the experiment— the upper bands had been scored higher as a result of their timbral weight or presence. This being particularly relevant to the 4kHz band.

- Through linear regression an Estimated ‘Punch’ score algorithm was derived.
Objective Model
A multistage approach was adopted based on the ‘punch’ definition outlined.

The first stage separates the component parts of the signal.

The approach adopted is based on median filtering [8].

Ignores the steady state portion of the signal, unlike a standard integration based meter.

Temporal Data Extraction

- MIR tool box was utilised to extract temporal data relating to attack times and note onsets. [9]

- Utilised within the weighting stage to produce the summed ‘Punch’ score.

Weighting Stage

- The 0dB output represents full scale.
- Level output is similar to that of the standard loudness model such that if the input stimulus is a full scale digital broad band pink noise burst, the output of the model would be -3dB.
Punch Measure
Punch Measure
Punch Profiling

- The previous output measure could be used in conjunction with a loudness meter to aid in mixing.

- For track comparison, a statistical approach can be adopted, based on a frame by frame analysis of the punch coefficient.
60% of the frames analyzed exhibit moments of punch $\leq -30$ dB.

20% of frames @ $\geq -26$ dB.
Punch Profile Examples

Relatively low level of punch detected 100% @≤-35dB. Consistency throughout, but no knees and wide dynamic. 50% of signal punch score is <45dB, clustering around -40dB above this point.
25% of frames are >= -20dB. Knee around 60% suggests dynamics present due to punch fluctuation. More so than the last example due to greater range.
Punch Profile Examples

10% of frames are >= -20dB
Plot suggests a consistency in punch and wide range.
Punch Profile Examples

@3% of frame moments of punch $\geq -20$dB, wide dynamics with some consistency.
Further Work

- Bring the residual/harmonic into the equation.
- Smooth the detection envelope / ignore spurious spikes of little interest.
- Test with various median filter window sizes.
- Compare to LDR and other promising models.
Thankyou