Measuring Dynamics

Comparing and Contrasting Algorithms for the Computation of Dynamic Range

Jon Boley, LSB Audio LLC
Michael Lester, Shure Incorporated
Christopher Danner, University of Miami

Motivation

It has been argued that:

- An increased use of dynamic range compression is directly related to (or even a cause of) the "loudness war"
- Reduced dynamic range is related to poor audio quality

However, these statements are difficult to defend without a solid method for measuring dynamic range

Definition for this study

- Dynamic Range is
  - Not the range from lowest to highest energy *possible*…
  - But rather, the range that is actually *used* by the *audible* program material
- This definition should be relevant for any metric that claims to be related to some aspect of perceptual quality

Ways to measure dynamic range

- Loudness Range
  - EBU Technical Document 3342
- TT Dynamic Range Meter
  - Pleasurize Music Foundation
- Dynamic Spread
  - Vickers (AES 111th Convention)
- pfpf Software
  - Tollerton (http://audiamorous.blogspot.com)

Loudness Range (LRA)

EBU Technical Document 3342, published August 2010

- Based on ITU-R Bs.1770 loudness
- Integrated using a 3sec sliding window
- Values are gated at -20LU re the loudness level
  - (to focus on foreground sounds)
- A histogram is created
- LRA is the range from 10% to 95%
  - *"The lower percentile of 10%, can, for example, prevent the fade-out of a music track from dominating Loudness Range. "*
  - *"The upper percentile of 95%, ensures that a single unusually loud sound, such as a gunshot in a movie, cannot by itself be responsible for a large Loudness Range."*

Loudness Histogram

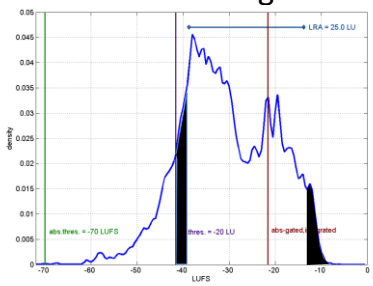


Figure 1. Loudness distribution, with gating thresholds and Loudness Range for the film The Matrix (DVD version). Adopted from Skovenborg & Lund (2009) "Loudness Descriptors to Characterize Wide Loudness-Range Material", 127th AES Conv.

TT Dynamic Range Meter

- For each 3sec window, calculate peak and RMS
- Sum the RMS over the top 20%
- DR = mean($Peak_2$ / ΣRMS)
- The authors argue that using just the top 20% :
  - allows them to compare a variety of program types (e.g., genres of music, speech, etc.), and…
  - is more likely to measure the effects of dynamic range compression since gain reduction is usually greatest for high-level signals.

Dynamic Spread

- Just use the P-norm of the signal

$$d = \left(\frac{1}{M} \sum_{i=0}^{M-1} \left| V_{dB}(i) - \overline{V} \right|^p \right)^{\frac{1}{p}}$$

- Vickers recommended p=1
  - Mean absolute deviation
- We also calculated p=2,3,4
  - Standard deviation, skewness, kurtosis

pfpf

- Based on BS.1770 loudness
- Three time scales
  - 10ms
  - 200ms
  - 3sec
- Histogram from 50% to 97.7%

Source Material

- 10 music clips from released albums
  - Chosen to represent a variety of dynamic ranges
  - Various genres (jazz, electronic, metal, pop, etc.)
- Each was 10-20sec long
- All normalized to the same BS.1770 loudness

Dynamic Range Metrics

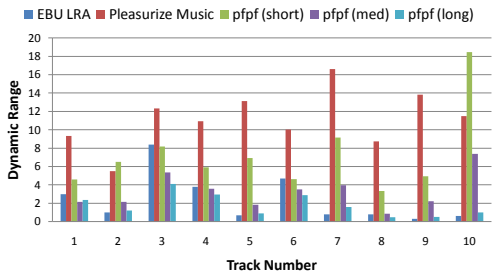
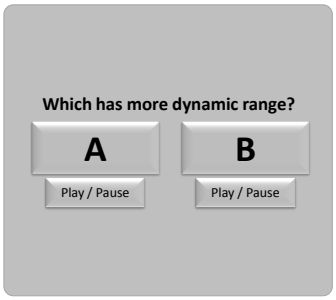


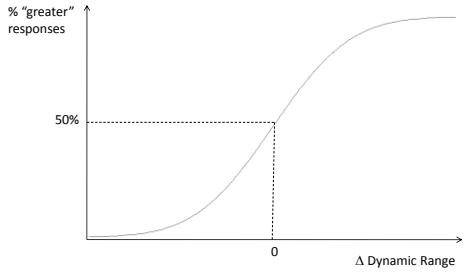
Table of Paired Comparisons

	1	2	3	4	5	6	7	8	9	10
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										

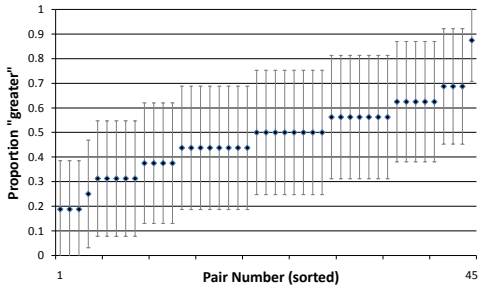
Psychoacoustics GUI



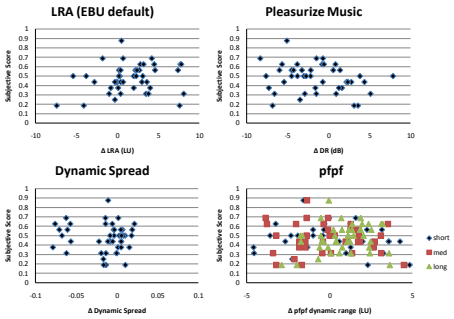
Psychometric function



Perceptual Results (N = 16 listeners)



Perception vs. Metrics



LRA permutations

- Window size
  - 100ms
  - 200ms
  - 400ms
  - 800ms
  - 1500ms
  - 3000ms
- Percentile range
  - Every combination, in steps of 5%
  - (e.g., 0-5%, 0-10%, …, 55-60%, …, 5-100%, 0-100%)

Perception vs. LRA

Top Ten Correlations:

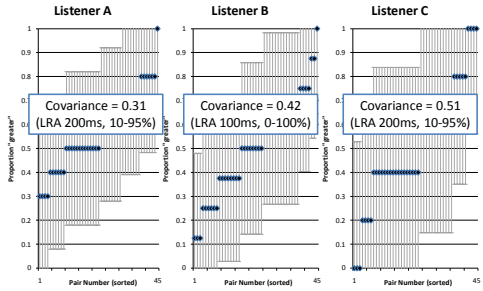
1	0.409604	3000ms	30-35%
2	0.399333	1500ms	35-40%
3	0.341461	1500ms	35-45%
4	0.307979	1500ms	35-50%
5	0.305513	3000ms	30-40%
6	0.297876	1500ms	35-55%
7	0.294927	3000ms	30-45%
8	0.290928	1500ms	30-40%
9	0.272833	1500ms	35-60%
10	0.272404	1500ms	30-50%

Perception vs. LRA

Top Ten Covariances:

1	0.16575	400ms	0-100%
2	0.155111	1500ms	0-100%
3	0.154583	1500ms	0-95%
4	0.154028	400ms	0-95%
5	0.15275	1500ms	0-90%
6	0.148806	400ms	0-90%
7	0.148083	1500ms	0-85%
8	0.146167	1500ms	0-80%
9	0.145556	800ms	0-100%
10	0.145444	800ms	0-95%

Individual Listeners



Concerns to be addressed

- Listening tests have too much variability
  - More extensive listening tests needed
- Musical clips may have been too short (10-20s) to get a meaningful metric from some of these algorithms
  - However, there certainly was a meaningful dynamic range for our listeners
- Use material with more DR