

*If one does not begin with a right attitude,
there is little hope for a right ending.
— Kung Fu Meditation*

From an audio consumer perspective, the most valid test vector for measuring audio quality is a real-life music signal - the signal that audio equipment is designed for. A real-life music signal makes it possible to compare objective measurements to subjective evaluations. Degradation of this signal is the cause of all perceived differences in sound. Measurement and research of a signal degradation in relation to corresponding sound degradation is the basis of the proposed audio metric - the df-metric.

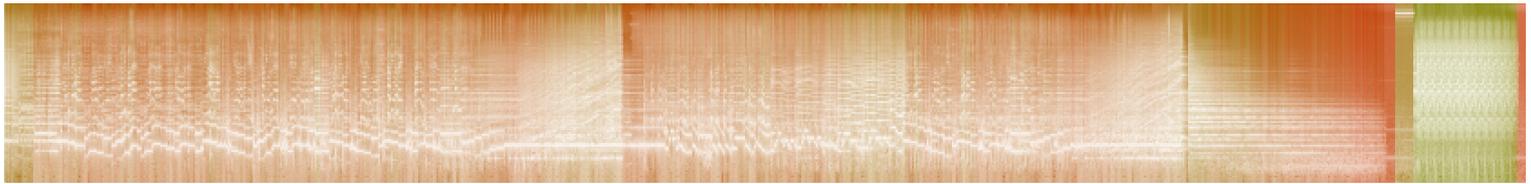
Single audio parameter. The metric has only one measurable parameter, the Difference Level (Df, dB), which measures differences between waveforms of two signals, the input and the output of some device under test (DUT) [1]. A specially-designed time-warping algorithm removes linear pitch and phase shifts from the output signal with any predefined accuracy. The algorithm finds linear pitch and phase transformation in the output signal, such that the latter has maximum correlation with the input signal. This maximum correlation R_{max} determines the value of Df :

$$Df [dB] = 10\lg(1-|R_{max}|)$$

The time-warping operation is deterministic and produces repetitive results for any signals. Hence, Df values are objective measurements (for time-accurate DUTs, time warping is not required). Df levels can be computed with arbitrary time windows: 50ms, 400ms, 3000ms, etc. For visualization purposes, the scale of the Df values is associated with a color mapping:



A sequence of Df values, computed for some DUT and test signal, shows degradation of the signal as it varies with time. A visual representation of such sequences is called a diffragram [3]. For convenience, the diffragram is combined with a spectrogram, where intensity is coded with luminosity of color:



Diffragram (400ms) showing degradation of the test signal (A Day in the Life - The Beatles) in the DUT (HTC Desire C). Different parts of the signal are distorted to different levels.

Multiple test signals. The Df-metric can be established with an unlimited number of test signals, including real-life music signals. Levels of degradation can then be compared with one another. An audio researcher can design any waveform and measure the level of its degradation in some DUT. The performance of that DUT can be examined in detail using arbitrarily elaborate sets of test signals. This new dimension of audio research and experimentation is unattainable with standard audio metrics, which are mostly sinusoid-based. It should be noted that the df-metric has the exact analog of the THD+N parameter - Df with a pure sine signal. With respect to perceived audio quality, THD+N/Df(sine) can be a useless audio measurement parameter.

Artifact signature. A sequence of Df values, computed with various music material, shows in detail how the test signal was distorted in a DUT, such a sequence can be considered as a signal artifact signature of the DUT. The Df-metric distinguishes two types of signature artifacts:

- sound artifact signature or just *sound signature*, which is revealed by listening to a device
- signal artifact signature or just *artifact signature*, which is revealed by measurements

Devices that have similar artifact signatures distort the musical signal in a similar way. This results in similar listening experience for these devices. They sound similar or ... they have similar sound signatures. So, these two types of signatures are interrelated. In addition, artifact-signature similarities, in most cases, indicate similarity of software/hardware audio solutions used.

Artifact signatures can be compared with one another by means of cluster analysis [6]. Early research (still in development [4]) shows that for DUTs with close artifact signatures their df-measurements correlate well to their perceived quality estimations. This relation holds for any type of audio - analog, digital, psychoacoustic processing... Thanks to this relation, some audio research "magic" is possible:

- reusing of results of already finished listening tests for grading new similar DUTs
- amplification of artifacts (SARTAMP [5]) for DUTs with small impairments; in case of such amplification, listening tests of less-strict design can be performed.

Audio singularity point. There is a special point on the Df scale, which corresponds to some Df level with a real music signal. Achieving this low level of difference effectively causes an audio device to disappear from the audio path and its presence cannot be discovered by any listening tests. At this *s-level*, the amount of all possible distortions is so small that discovering/researching them makes no sense; psychoacoustics does not work below this point. Warmth, harshness, openness, brightness, scene depth ... are then the author's instruments. Playback devices are just a communication channel aimed to deliver all those characteristics of sound untouched. Thus the problem of delivering high quality audio can be turned into a pure engineering task - to provide the required level of accuracy for real music signals that can be effectively controlled by the single parameter - the Df level. Thanks to the audio singularity, the most accurate DUTs in the audio similarity space form a dense (cold) core while other less accurate DUTs (warmer) are gathered around this core.

Having a transparent audio path, the hard work of both artists/musicians and audio engineers/producers becomes clearly audible. Numerous audio effects for creative listening (tube/vinyl distortions, etc.) can be additionally inserted into that path. The current level of technology in the audio industry is sufficient for producing cheap consumer audio devices operating below the audio singularity level. The market of audio playback devices can be commoditized right now.

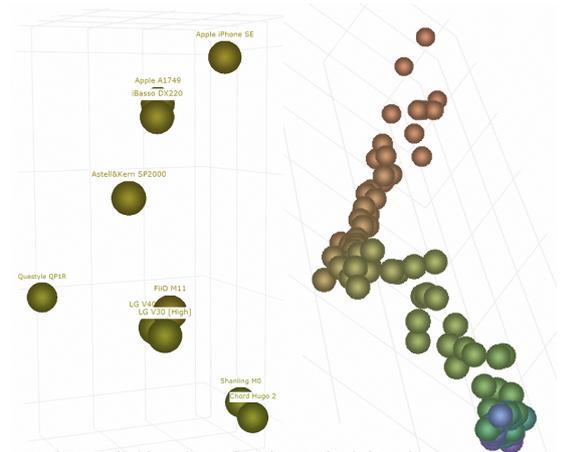
Use cases

- objective audio measurements of portable players - <http://soundexpert.org/portable-players>
- objective audio measurements of Bluetooth audio codecs - <http://soundexpert.org/articles/-/blogs/audio-quality-of-sbc-xq-bluetooth-audio-codec>
- SE listening tests of high bitrate codecs (128+ kbit/s) use SARTAMP technology since 2005 - <http://soundexpert.org/encoders>

How to measure

- in most cases it suffices to have a good quality digital audio recorder - <http://soundexpert.org/portable-players#how2measure>
- all Df values are computed from such recordings with the following Matlab code - <http://soundexpert.org/articles/-/blogs/visualization-of-distortion#part3>

Relation to traditional audio metric. The Df-metric is based on a generalized method of distortion measurement having THD+N as a special case when the test signal is a pure sinusoid. The ability to measure distortions of various test signals helped to discover the main reason why the classic set of audio parameters correlates so badly to perceived quality estimations. It is often assumed that if some DUT has less distortion with some technical signal, then all other signals will be less distorted as well by this DUT. Df measurements of real-world DUTs clearly show that this is not true - in the general case, levels of distortion of various signals may be independent. Hence, an assessment of sound quality by means of any objective measurements using purely technical signals has limited validity and should be phased out. The only relevant signal for such measurements is real music.



Real DUTs (left) and simulated ones (right) in the **audio similarity space**; distances between DUTs are computed based on similarity of their artifact signatures.

Serge Smirnov
Research Engineer
ssmirnov@soundexpert.org

References

- [1] Smirnov S, "Difference Level: An Objective Audio Parameter", in Proc. AES 118th Conv., 2005, http://soundexpert.org/documents/10179/11017/DiffLevel_AES118.pdf
- [2] Kindlmann G, Reinhard E, Creem S, "Face-based Luminance Matching for Perceptual Colormap Generation", <http://www.cs.utah.edu/~gk/papers/vis02/FaceLumin.pdf>
- [3] Smirnov S, "Diffragram: visualization of signal differences in audio research", <http://soundexpert.org/articles/-/blogs/visualization-of-distortion>
- [4] Smirnov S, "Objective difference measurements to predict listening test results?", <http://soundexpert.org/articles/-/blogs/objective-difference-measurements-to-predict-listening-test-results->
- [5] Smirnov S, "Infinite Grade Impairment Scale", http://soundexpert.org/documents/10179/11017/se_igis.pdf
- [6] Smirnov S, "The concept of artifact signatures", in discussion at ASR forum, <https://www.audiosciencereview.com/forum/index.php?threads/alternative-method-for-measuring-distortion.10282/page-3#post-284457>