SoundExpert
www.soundexpert.info

Infinite Grade Impairment Scale

Serge Smirnoff
ssmirnoff@soundexpert.info

During the last decade almost all audio turned into digital audio. While creation, processing and distribution of audio content have changed dramatically, providing of appropriate audio quality level throughout the whole chain is a real problem still. The reasons are well known:

- Any objective parameters like THD, S/N, IMD … or even a set of them no longer correlate reliably to the perceived audio quality. To be true, they did not even before digital revolution, but now this problem becomes evident.

- Subjective quality assessments, being today the only reliable method of judging audio quality, are expensive. Their results loose reliability when sound artifacts are near or beyond the threshold of human perception.

Extensive search for new methods of judging audio quality has led to appearance of so called "objective perceptual measurements". Basic idea of those methods is to measure some significant parameters of a signal under test (output of a device under test) and to analyze them according to some model of human perception. Choosing appropriate parameters and defining the model more precisely one could expect to get reliable audio quality estimates. PEAQ method is the most developed of its class. This measurement scheme, being actually a kind of an expert system with its own "pros" and "cons", seems to be a promising one.

Infinite grade measurement scheme (IGMS) described below differs essentially from the methods above as it combines subjective and objective measurements into a single procedure, taking the best from both methods. It is simple and reliable like an objective parameter and it takes into account human perception by means of specially designed listening tests.

BASIC CONCEPT

In order to judge audio quality IGMS uses input and output signals of a device under test. If sound artifacts in output signal are clearly audible a classical listening test is held and final subjective score is located on standard 5-grade impairment scale. In case of the artifacts near or beyond threshold of human audibility the output signal is substituted by series of test stimuli (usually 3-5) with artifacts amplified (unmasked) to different levels. Then artificial test stimuli are graded as usually according to 5-grade impairment scale. Final subjective score is calculated analytically on the basis of these "servo-results" and located above the 5th grade on the "virtual" part of the scale.

DETAILS[†]

Let's have a closer look at some details are essential for the method.

Sound artifacts amplification (SARTAMP), used to create test stimuli, could be divided into three operations:

- Time alignment of input and output signals;

- Subtraction of some portion of input signal from output one;

- Post filtering of resulting signal in order to reject parasitic spectral components absent in output signal; normalization.

Difference level. The level of artifacts in test stimuli is controlled by objective parameter – Difference level (Df). It is calculated for two signals – X and Y as:

$$Df(X,Y)=\sqrt{1-|\rho(X,Y)|}\ ,$$

where ρ is correlation coefficient.

The parameter shows how mach the shape of test signal (stimulus) differs from the shape of input one:

- Df ≡ *0* ≡ *0 [%]* ≡ *-∞ [dB],* if sound signals have exactly the same shape;

- Df ≡ *1* ≡ *100 [%]* ≡ *0 [dB],* if signal shapes are completely different.

Other characteristics of the parameter:

- Difference level does not depend on amplitudes and DC offsets of both signals.

- Physical essence of Difference level is normalized level of difference signal:

$$Df(X,Y)=\frac{P^{RMS}_{(Y-X)}}{P^{RMS}_{X}\sqrt{2}}$$

  Being a dimensionless value Df is the ratio of RMS levels of two signals – difference *(Y-X)* and input *(X)*.

- In case of sinusoidal input signals Df is almost linearly dependent on THD. For small Df and THD (Df ≤ *7%,* THD *≤10%)*:

$$Df=\frac{THD}{\sqrt{2}}$$

- Difference level can be computed for signals both in time and frequency domains.

- Difference level is seriously affected by time alignment errors. They are caused either by constant time or phase shift or by time stretch/shrink. In order to fix the problem precision controlled resampling has to be applied to signals beforehand.

Difference level can be used as independent instrumental audio parameter for quantitative estimation of signal degradation during signal processing or transmission. In this case it could be considered as an extension of THD for signals of any nature. With this parameter, for example, one could analyze a device under test in details simply modifying and combining input signals. Measurement procedure always stays the same and its results can be compared with each other. White noise has to be mentioned specially. Its spectral complexity makes it to be the most severe test for any audio equipment. Probably it can be assumed that a tested device under white noise introduces all possible types of distortion. Therefore parameter Df_{WN} could be helpful for end users of audio equipment as it indicates the equipment potential under worst case conditions while THD shows it under most favorable ones and useful mostly for developers.

Listening tests. As amplified artifacts in test stimuli are clearly audible listening tests could be of simple design and the results less depend on participants and their listening environments.

Psychometric curves. Extensive listening tests with various test stimuli created by means of SARTAMP technology reveals that the relationship between Difference level and subjective scores is a monotonic function well consistent with 2nd order curve. Two examples of such psychometric functions are shown on **Fig.1**.

[†] More details are in the paper: Smirnoff S, Difference level: an objective audio parameter. *118th AES Convention*, Barcelona, Spain, 2005 May, preprint #6493

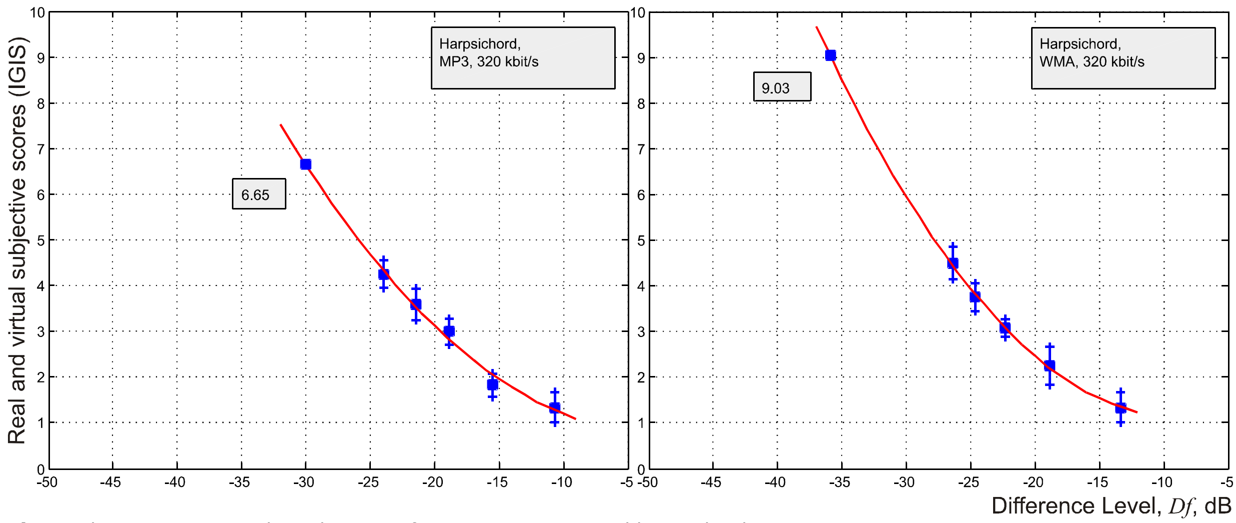

Fig.1 Subjective scores and psychometric functions approximated by 2-nd order curves.

The consistency with 2nd order curve is not ideal even in case of perfectly designed listening tests. Non-linear nature of human hearing and masking thresholds makes perception of gradually unmasked artifacts uneven. Level of that inconsistency depends on type of input/difference signals.

Infinite grade impairment scale (IGIS). After results of servo-tests were approximated by 2nd order polynomial, it is not hard to find unknown "subjective" score of real output signal - the polynomial has to be solved for the difference level of that output signal. In order to locate the found score, standard 5-grade impairment scale was continued to infinity and thus transformed into infinite grade impairment scale. Its "virtual" part joins the real one seamlessly.

Infinite grade measurement scheme (IGMS) is shown schematically on **Fig.2**.

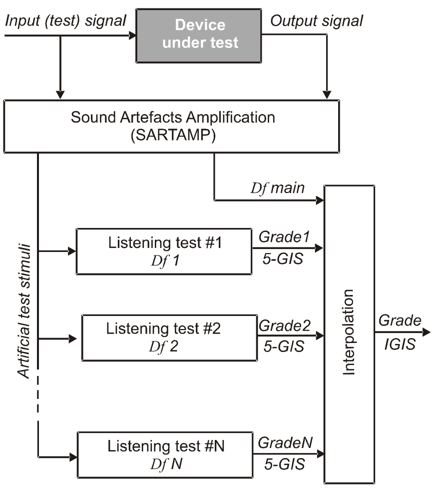

Fig.2 Infinite Grade Measurement Scheme

Device under test could be of any kind – software or hardware, digital or analog. The latter needs high precision DA/AD converters and effective time alignment procedure. SARTAMP performs time alignment of input and output signals, their subtraction in various proportions and filtering of resulting test signals. Listening tests could be of simple design with participation of ordinary listeners. Interpolation block calculates unknown virtual subjective scores with confidence intervals.

This infinite grade measurement scheme has something in common with already mentioned PEAQ method. The main difference is subjective part, which is formalized and incorporated inside PEAQ and, on the contrary, exists in the form of specially designed listening tests in IGMS. Thus virtual subjective score of IGMS is combination of objective measurement of signal change by Difference level and subjective estimation of that change by means of psychometric function in listening tests.

APPLICATION

Low sensitivity of IGMS to listening tests design makes it attractive for carrying out listening tests through the internet with broad participation of ordinary listeners. Such experimental project **SoundExpert** has been already launched at www.soundexpert.info. Its testing engine is entirely based on IGMS. Listening tests are designed in accordance with double stimulus / hidden reference / double blind methodology. In order to take part in testing a visitor has to:

- download an audio file with two samples of the same sound excerpt;
- grade basic audio quality of a sample which has degraded audio quality;
- feedback his judgment.

Participants have no information either about order of samples or about device tested. SoundExpert testing engine automatically "unblinds" received data, computes subjective scores and presents them live in the form of ratings on infinite grade impairment scale. Additionally raw data access is reserved in the system for independent analyses of collected grades by anyone who cares. Thus active participation of internet community in these listening tests will help both to find out subjective ratings of audio equipment and to collect valuable data for further research of human hearing.