Improvements in calculating the loudness of time varying sounds

Roland SOTTEK[1]

HEAD acoustics GmbH, Germany

ABSTRACT

Recently, a new ISO standard for loudness of arbitrary sounds ISO 532-1 (1) was proposed for the revision of ISO 532:1975 section 2 (method B) (2). It is based on DIN 45631/A1:2010 (3), which includes the widely used standard DIN 45631:1991 (4) for stationary sounds as a special case. ISO 532-1 eliminates uncertainties of existing standards by strictly defining the complete procedure of loudness calculation starting with the waveform of the time signal and ending with specific and total loudness vs. time functions.

The strict definition of the complete procedure is a step forward to comparability of calculated loudness results, and fully conforms to DIN 45631/A1:2010 for the sake of continuity.

However, although the results of this algorithm are in accordance with the results of many listening tests, there are still phenomena that are not covered by this method. For example, the calculated loudness of sweep signals shows fluctuations, whereas the perceived loudness does not. This is due to the implemented filter bank based on fixed, contiguous third-octave filters. As a possible solution a loudness calculation method is presented that is based on a hearing model (Sottek) using an aurally adequate filter bank of highly-overlapping asymmetric filters (5). In addition, the nonlinearity between specific loudness and sound pressure has been reconsidered in this model according to results of many listening tests (6).

Keywords: Time varying loudness, Hearing model I-INCE Classification of Subjects Number(s): 63.1, 81

1. INTRODUCTION

Many loudness calculation procedures can be described with a common scheme (Figure 1). The results vary because of different implementations of each processing block such as the filter bank (number of filters, bandwidth and shape of the filters), the frequency weighting and the nonlinearity between sound pressure and specific loudness.

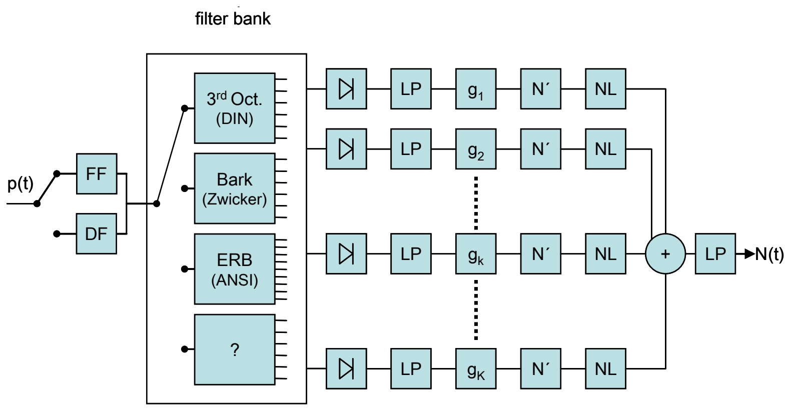

Figure 1 – General block diagram of loudness calculation procedures. Block NL (non-linearity) is used only for the loudness calculation of time varying sounds as in ISO 532-1 and DIN 45631/A1.

Loudness calculation begins with the microphone signal *p(t)*: a choice is made between free and diffuse sound fields (FF or DF). This is followed by filtering with a bank of bandpass filters, which in turn is followed by rectifying and lowpass filtering (envelope formation). The next two steps consist of a frequency-dependent weighting g_k and a nonlinear transform from sound pressure or intensity to

1 roland.sottek@head-acoustics.de

specific loudness (block N' in Figure 1). For time-varying sounds, the nonlinear decay of the human hearing system has to be modeled in detail. Furthermore, effects of the temporal summation and forward-masking must be taken into account (block NL in Figure 1). Finally, the total loudness *N(t)* is calculated by summing the specific loudness values. The lowpass filter at the end simulates that signals with a duration of 10 ms are perceived as approximately half as loud as signals with a duration of about 100 ms. Further details of each signal processing block can be found in (7).

2. LOUDNESS PERCEPTION OF SYNTHETIC SOUNDS

Loudness calculation procedures provide single number values (loudness and specific loudness vs. time functions, representative values like percentiles) that can be used in many scientific and technical applications to estimate the perceived loudness and loudness level of a sound without conducting separate listening tests. Nevertheless, listening tests are required to validate the procedures.

2.1 Loudness adjustment of pink noise using 1 kHz pure tones

Schlittenlacher et al. (8) performed an experiment where the loudness of pink noise was adjusted to that of a 1 kHz pure tone at 9 loudness levels from 43 phon to 83 phon at a spacing of 5 phon. The pure tones were presented in quasi-random order (Figure 2, blue markers), whereby the tone with 73 phon always was the first and the last presented sound (red marker). After that, pink noise at levels of 43 dB and 63 dB was used as the reference stimulus and the 1 kHz pure tone should be adjusted (black markers).

The results achieved by ISO 532-1 (method for stationary sounds) do not meet the target values exactly, but its calculations are always within the interquartile range. The result of the test with the pink noise at 63 dB as the reference stimulus (upper right black marker) is lower than expected. According to (8) a psychological reason for this could be the fact that the participants concentrate particularly on the adjustable stimulus, which in this case means that the participants focus on frequencies around 1 kHz.

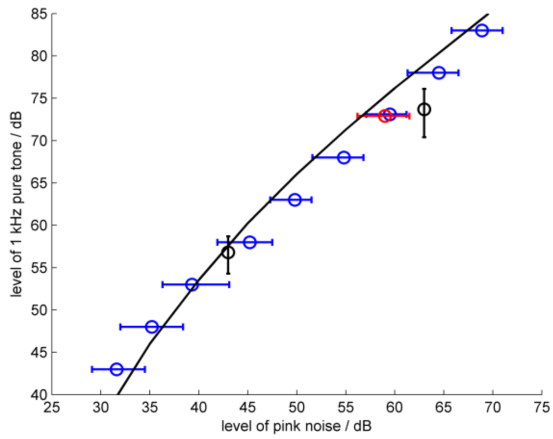

Figure 2 – Level of a 1 kHz pure tone that is perceived as loud as pink noise (20 Hz – 20 kHz). Circles indicate medians, whiskers the interquartile range (8). The solid line shows the results achieved by ISO 532-1 (method for stationary sounds).

2.2 Loudness of logarithmic sine sweeps in pink noise

For a sine sweep without superimposed sound (curve 5 in Figure 3) one would expect, in terms of hearing sensation, a "smooth" course of the loudness pattern. However, the actual course obtained by the calculation using ISO 532-1 shows a certain amount of variations. These are caused by the way the signals are processed in the third-octave filter bank. Due to the fact that the edge steepness of the third-octave band filters is finite, there will be overlapping ranges when the sine signal passes from one third-octave band to the next. Depending on the characteristic of the implemented filters, deviations from the "ideal value" differing in size and magnitude can occur in the transition frequency ranges when adding up the partial loudnesses.

In principle, it is possible to implement special filters avoiding this effect (Figure 8, results of the loudness calculation based on the hearing model of Sottek). However, the requirements for the third-octave filter bank on the input side are predetermined in ISO 532-1 for the sake of continuity.

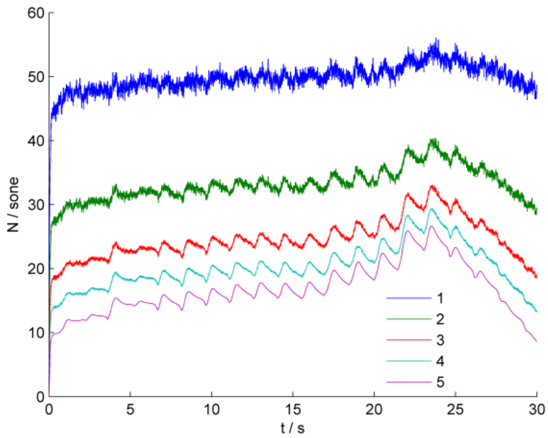

Figure 3 –Calculated loudness of logarithmic sine sweeps (80 dB, 100 Hz-10 kHz) in pink noise (20 Hz-20 kHz) with different levels (curve 1: 80 dB, curve 2: 70 dB, curve 3: 60 dB, curve 4: 50 dB, curve 5: without pink noise), results achieved by ISO 532-1 (method for time varying sounds).

2.3 Loudness of filtered uniform exciting noise

Recently, Hots et al. published two papers about loudness of subcriticial sounds as a function of bandwidth, center frequency, and level (9, 10): "the level at equal loudness was close to 0 dB for the narrowest bandwidth, increased with bandwidth for bandwidths smaller than the critical bandwidth, and decreased for bandwidths larger than the critical bandwidth. For bandwidths considerably larger than the critical bandwidth, the level difference was negative." (10)

Some years ago, an own experiment on loudness comparison between a tone and bandpass-filtered pink noises showed similar results with respect to the loudness perception of the tone and the narrowband noises, but stronger effects of the loudness summation of bandpass noises with larger bandwidths (11). For the generation of the sound samples, 4th-order bandpass filters were used, whereras in (9, 10) the spectra of the sounds were set to zero outside the frequency band of interest and time signals were generated by means of inverse Fourier transform (12), producing, in effect, infinite side-slopes.

These findings motivated new experiments on loudness perception of differently bandpass-filtered uniform exciting noise. Filtering was performed using filters with infinite steep slopes (named as 'inf order', the technique used by Hots et al.) as well as 6th- and 4th-order filters as a function of bandwidth at two levels (50 dB and 70 dB) and only at the center frequency of 1500 Hz. For the level of 50 dB the bandwidth was varied from 0 to 5 Bark, for the level of 70 dB, the bandwidth was between 2 and 6 Bark. An adaptive loudness matching procedure (2I-2AFC) was performed by 10 normal-hearing listeners with a 3 Bark-wide reference sound at a level of 50 dB, and a 4 Bark-wide reference sound at a level of 70 dB. One half of the subjects were students with experience in listening tests and the other half were acoustic experts.

Figures 4 and 5 show box plots with medians (red lines), upper and lower quartiles as well as some outliers (+) from the results of the loudness matching procedure together with predictions based on ISO 532-1 and the algorithm based on the hearing model of Sottek (5, 6).

In general, the model predictions correspond well with the results of the matching procedure, especially for bandwiths larger than one critical bandwidth. The effect of loudness summation is much stronger when using less steep filters. The results from the experiments using 6th-order filters (not shown) confirms this finding. The difference between the highest and lowest level at equal loudness as a function of bandwidth is the smallest in the case of 'inf'-order filtering: for the level of 70 dB the differences between the levels for 2 and 6 Bark are 3.5 dB ('inf' order), 7.0 dB (6th order), and 8.2 dB (4th order). The results of the matching experiments show smaller interquartile ranges when using less steep filters.

The model predictions show larger deviations for tones and to a smaller extent for the noise with a bandwidth of 1 Bark (Figure 4). To improve the model accuracy, current work is focussing on a weighted summation of the loudness of tonal and non-tonal components for an overall loudness. The hearing model of Sottek allows for a separation of these two components based on an evaluation of correlation functions in the different frequency bands (13).

Note that the 'inf'-order filtered signals have a special sound character: they evoke a very high tonality perception. This will be discussed in an accompanying paper (14).

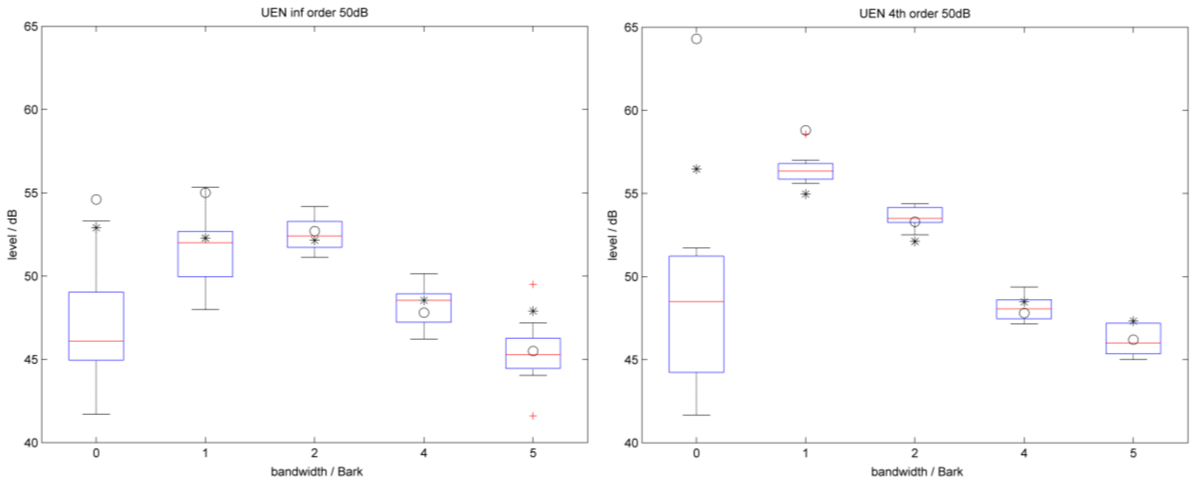

Figure 4 –Boxplots of loudness matching experiments using uniform exciting noise (UEN) as a function of the bandwidth in Bark. Reference signal: UEN with a bandwidth of 3 Bark and a level of 50 dB. Additionally, results of model predictions are shown: based on ISO 532-1 (‘star’) and the hearing model of Sottek (‘circle’).

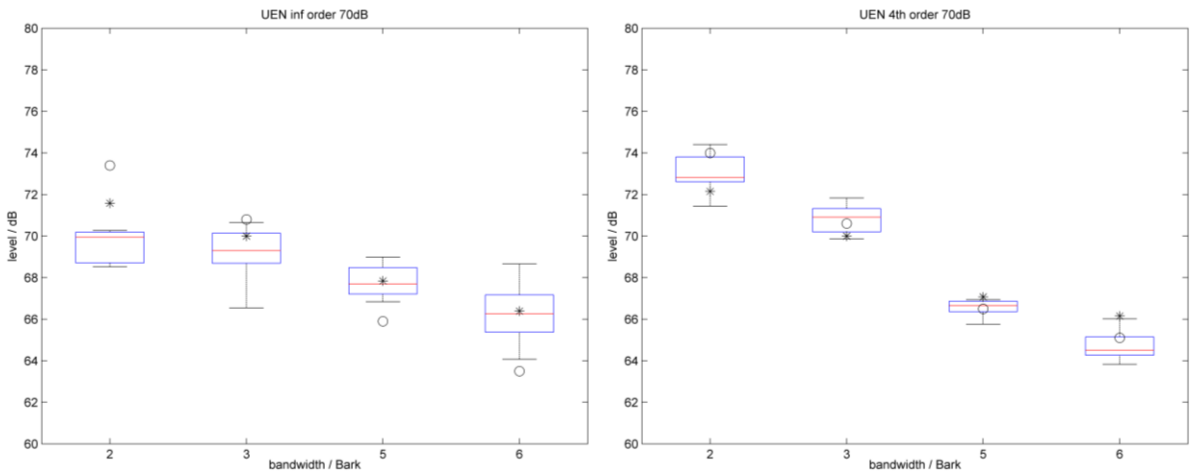

Figure 5–Boxplots of loudness matching experiments using uniform exciting noise (UEN) as a function of the bandwidth in Bark. Reference signal: UEN with a bandwidth of 4 Bark and a level of 70 dB. Additionally, results of model predictions are shown: based on ISO 532-1 (‘star’) and the hearing model of Sottek (‘circle’).

3. LOUDNESS PERCEPTION OF TECHNICAL SOUNDS

3.1 Loudness of howling sounds

In daily life, technical sounds must be evaluated, e.g., vehicle sounds in the automotive industry. Considering the fact that most of these sounds are time varying, a model of time varying loudness is preferable. At the moment, DIN 45631/A1 is the sole standardized loudness method for time varying sounds, applicable also for stationary sounds. The proposed standard ISO 532-1 eliminates uncertainties of DIN 45631/A1. The following example describes how different howling sounds have been judged by 18 students (age between 23 and 29 years) using a magnitude estimation technique. The subjects first listened to a reference howling sound and then to another howling sound. The loudness of the reference sound was defined to be 100 and the subjects were asked to give a corresponding number for the loudness of the second sound. Figure 4 shows the geometric mean of the ratings with the confidence interval (95%) in comparison to the results based on the loudness calculation according to ISO 532-1 and to the equivalent continuous A-weighted sound pressure level. The correlation of the N_5-ratios calculated with ISO 532-1 for time varying sounds with the geometric mean is 0.98, whereas the correlation with the ratio of the transformed A-weighted sound pressure levels (using the formula for the conversion from loudness level to loudness: $N = 2^{0.1\cdot(L(A)/\mathrm{dB}-40)}$, see ISO 532-1) is only 0.17 and also lies outside the confidence interval.

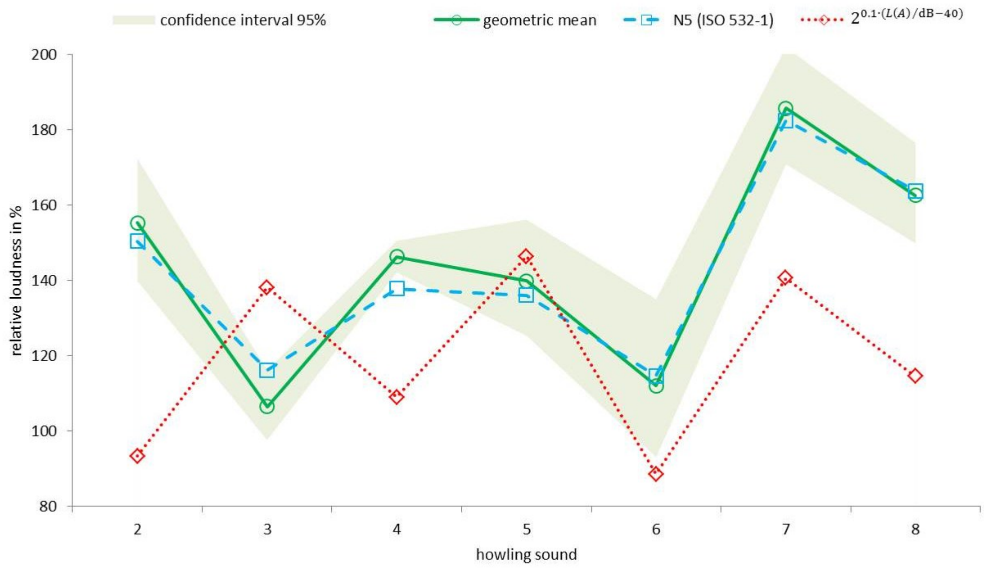

Figure 6 –Magnitude estimation of the loudness of howling sounds. The geometric mean of the ratings of 18 subjects and the confidence interval (95%) are shown and compared to the results based on the loudness calculation according to ISO 532-1 and to the transformed equivalent continuous A-weighted sound pressure level (using the formula for the conversion from loudness level to loudness, see ISO 532-1).

4. LOUDNESS MODELING USING THE HEARING MODEL OF SOTTEK

The loudness modeling using the hearing model of Sottek is described in detail in (5). In addition, the nonlinearity between specific loudness and sound pressure has been reconsidered in this model according to results of many listening tests (6). Further improvements have been achieved by introducing a nonlinearity function according to formula (1).

$$y = x^{\nu_0} \cdot \prod_{i=1}^{N} \left(1 + \left(\frac{x}{x_{s_i}} \right)^{\alpha_i} \right)^{\frac{\nu_i - \nu_{i-1}}{\alpha_i}} \tag{1}$$

with $\nu_0 = 1$ and N further exponents ν_i. For $N = 8$ the exponents have been achieved by applying a nonlinear optimization procedure using the hearing model to predict loudness matching data of uniform exciting noise as a function of bandwidth and level. The thresholds x_{s_i} were set to 15 dB, 25 dB, 35 dB, 45 dB, 55 dB, 65 dB, 75 dB, and 85 dB; α_i was set to 1.5.

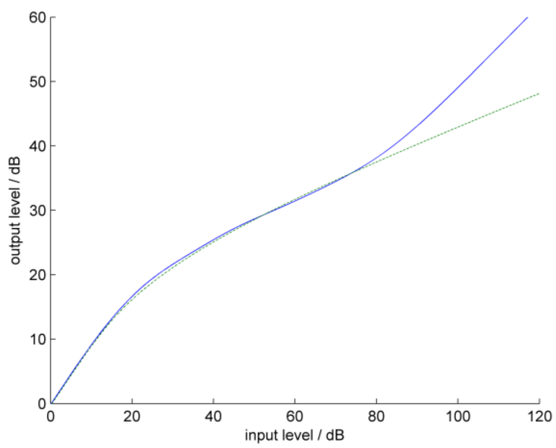

Figure 7 – Nonlinearity between specific loudness and sound pressure according to formula (1) (solid line) with $\nu_0 = 1$ and 8 exponents ν_i as a result of an optimization procedure. The dashed line shows the nonlinearity described in (5) using only 2 exponents.

The data for the optimization algorithm were results of an adaptive loudness matching procedure using 4^{th}-order bandpass-filtered uniform exciting noise centered around 1 kHz. Two noise signals with a bandwidth of 3 Bark and 5 Bark, respectively, were matched in loudness as described in (6) by 24 normal-hearing listeners (8 females/16 males, age between 24 and 50 years). Other tests with smaller bandwidths were also performed, but were not used for the determination of the nonlinearity because of the tonality of the narrowband signals (leading to a larger variance in the perceptual data). In a second experiment with 35 normal-hearing listeners (13 females/22 males), the results of the first experiment were examined: A paired comparison test was performed for each level set, using always the reference sound, the compared sound at the matching level (median value), and at slightly changed levels (adjusted in steps of 1 dB). The thresholds found in the first experiment could be confirmed and slightly corrected (6). In addition to these data, the results of the new listening test described in section 2.3 using signals with a bandwith wider than 1 Bark have been used to derive the nonlinearity shown in Figure 7.

As a consequence of using another nonlinear relation between specific loudness and sound pressure than in current standards, the output values differ of course, e.g., from the results of ISO 532-1. A new nonlinearity means a new loudness scaling. But, there are many reasons for the need to update the model for the nonlinearity of human hearing (5, 15).

Figure 8 shows as an example the calculated loudness of sweep signals with additive pink noise using the hearing model of Sottek. The unit $sone_{HM}$ indicates the use of the hearing model for loudness calculation. Obviously, the certain amount of 'unwanted' loudness variations (shown by the ISO 532-1 approach in Figure 3) is not present due the application of an aurally adequate filter bank of highly-overlapping asymmetric filters.

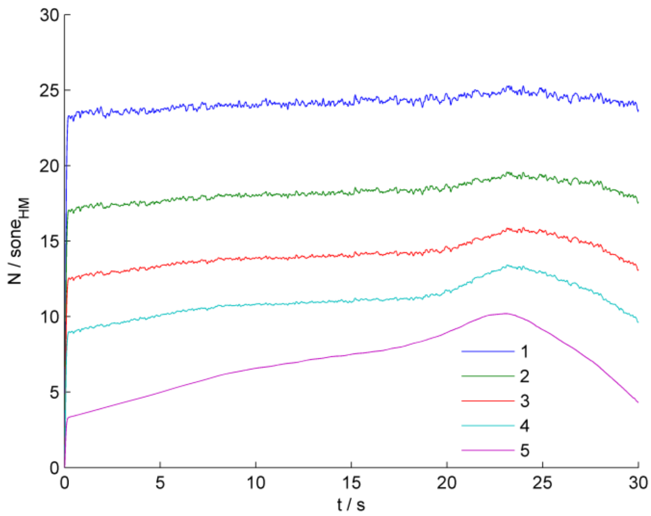

Figure 8 – Calculated loudness of logarithmic sine sweeps (80 dB, 100 Hz-10 kHz) in pink noise (20 Hz-20 kHz) with different levels (curve 1: 80 dB, curve 2: 70 dB, curve 3: 60 dB, curve 4: 50 dB, curve 5: without pink noise), based on the presented hearing model of Sottek.

5. CONCLUSIONS

The paper informs about the update of the loudness standard ISO 532 and shows the applicability and validity of the method for stationary and time varying synthetic and technical sounds. It points also to limitations of the standard which can be overcome by the application of the hearing model of Sottek. The introduction of an improved nonlinearity between specific loudness and sound pressure allows for predicting the results of loudness matching experiments better. An aurally adequate auditory filter bank of highly-overlapping asymmetric filters serves, e.g., for calculating a "smooth" course of the loudness pattern of sweep signals. This hearing model is a holistic model to analyze and model many psychoacoustic sensations, not only loudness, but also tonality, roughness, fluctuation strength, and impulsiveness.

ACKNOWLEDGEMENTS

The author wishes to thank his colleague, Wade R. Bray, for many fruitful discussions.

REFERENCES

1. ISO 532-1. Methods for calculating loudness, Part 1: Zwicker method, in preparation.
2. ISO 532: 1975. Acoustics – Methods for calculating loudness level, 1975.
3. DIN 45631/A1:2010. Calculation of loudness level and loudness from the sound spectrum - Zwicker method - Amendment 1: Calculation of the loudness of time-variant sound, Beuth Verlag, 2010.
4. DIN 45631:1991. Procedure for calculating loudness level and loudness, Beuth Verlag, 1991.
5. Sottek R. Modelle zur Signalverarbeitung im menschlichen Gehör. Doctoral Thesis, RWTH Aachen University, 1993.
6. Bierbaums T, Sottek R. Modellierung der zeitvarianten Lautheit mit einem Gehörmodel. Deutsche Jahrestagung für Akustik, DAGA 2012, Darmstadt, 2012.
7. Sottek R, Genuit K. “Progresses in standardizing loudness of time-variant sounds”, Deutsche Jahrestagung für Akustik, DAGA 2013, Meran, 2013.
8. Schlittenlacher J, Hashimoto T, Fastl H, Namba S, Kuwano S, Hatano S. Loudness of pink noise and stationary technical sounds. Proc. Internoise, Osaka, pp. 2314–2318, 2011.
9. Hots J, Rennies J, Verhey JL. Loudness of sounds with a subcritical bandwidth: A challenge to current loudness models? J. Acoust. Soc. Am. 134, EL334–EL339, 2013.
10. Hots J, Rennies J, Verhey JL. Loudness of subcritical sounds as a function of bandwidth, center frequency, and level, J. Acoust. Soc. Am. 135 (3), pp. 1313-1320, 2014.
11. Sottek R. Loudness models applied to technical sounds, Noise-Con 2010, Baltimore, 2010.
12. Hots J. Private communication (04.08.2014).
13. Sottek R, Kamp F, Fiebig A. A new hearing model approach to tonality, Proc. Internoise, Innsbruck, 2013.
14. Sottek R. Progress in calculating tonality of technical sounds, Proc. Internoise, Melbourne, 2014.
15. Sottek R, Genuit K. Models of signal processing in human hearing, Int. J. Electron. Commun. (AEÜ) 59, pp. 157-165, 2005.