Progress in calculating tonality of technical sounds

Roland SOTTEK[1]

HEAD acoustics GmbH, Germany

ABSTRACT

Noises with tonal components, howling sounds, and modulated signals are often the cause of customer complaints when emitted from technical products. The perception and evaluation of sound events containing such components has become increasingly important, e.g., in the field of vehicle acoustics for the assessment of tonality due to alternative drives. Furthermore, Information Technology (IT) devices and products such as hard disk drives may emit tonal sounds. Despite their very low sound pressure levels, such noises are unwanted and should preferably be avoided or masked.

The psychoacoustic parameter tonality was introduced in order to quantify the perception of tonal content. However, existing methods for tonality calculation show problems when applied to technical sounds. Recently, a new approach to tonality calculation based on a hearing model was presented by Sottek, Kamp, and Fiebig. In accordance with recent research results, the calculation of tonality is therein performed upon the basis of the partial loudness of the tonal content.

This paper presents model validations exploiting the results of new listening tests using bandpass-filtered noise signals with varyingly steep filter slopes and model improvements, especially in order to adequately indicate the perceived tonality of technical sounds with low sound pressure levels.

Keywords: Tonality, Hearing model I-INCE Classification of Subjects Number(s): 63, 61

1. INTRODUCTION

Tone-to-Noise Ratio (TNR) and Prominence Ratio (PR) used as mandated by ECMA-74 (1) to quantify the tonality of identified discrete tones do not respond well or even at all to tonalities caused by narrow bands of noise or non-pure tones, and thus are particularly useless with many frequently-encountered tonalities. The very important topic of hard disk drive cover plate tonalities is an example. The latter involve combinations of elevated noise bands due to structural resonances and often also pure rotating-mechanism tones. Recently, a new approach to tonality calculation based on a hearing model of Sottek was presented (2). This approach has been proven and validated for many signals, including none-pure-tone tonalities like narrowband noise at medium levels. Now the model has been validated for bandpass-filtered noise signals with extremely steep filter slopes. Additionally, it has been improved for sounds with low sound pressure levels (near the threshold of hearing).

2. CONVENTIONALLY-USED TONALITY MEASUREMENTS

The widely-used procedures for evaluating tonalities of IT devices rely on tone-to-noise ratio and prominence ratio. Therefore, a short overview of the existing tonality measurement procedures and their limitations is given, followed by a robust and effective improvement considering the threshold of hearing.

2.1 Tone-to-Noise Ratio

In order to calculate the tone-to-noise ratio, first the tone candidates are extracted from the Discrete Fourier Transform (DFT) considering the following criteria as described in (3):

1. the level of the spectral line exceeds the corresponding lines of the smoothed spectrum (1/24 octave bands) by at least 6 dB,
2. the level of the spectral line is higher than the level of the two neighboring lines,
3. the level of the spectral line exceeds a threshold (such as the threshold of hearing).

[1] roland.sottek@head-acoustics.de

The next processing steps consist mainly of the calculation of relative tonal bandwidth (compared to critical bandwidth), the calculation of the tone level and a first estimation of the noise level based on the smoothed spectrum. A tone is considered to be audible if the level exceeds the noise level minus 4 dB, otherwise the tone candidate is discarded. Further processing is performed if there is more than one tone in one critical band. Depending on the frequency difference between the tones, their intensities are summed to one new tonal component. The noise intensity is calculated by subtracting the sum of the tone intensities within one critical band from the overall intensity (considering also some bandwidth corrections, for details see ECMA-74). According to ECMA-74, a tone is classified as prominent if the tone-to-noise ratio is higher than 8 dB for tone frequencies of 1 kHz and higher. For frequencies below 1 kHz, this threshold value is increased by 2.5 dB per octave.

2.2 (Specific) Prominence Ratio

When using the prominence ratio method according to ECMA-74, a tone is classified as prominent if the difference between the level of the critical band centered on the tone and the average level of the adjacent critical bands is equal to or greater than 9 dB for tone frequencies of 1 kHz and higher. For frequencies below 1 kHz, this threshold value is increased by 3 dB per octave. Only tones exceeding this prominence criterion are listed as single number values with their frequency and prominence. It is possible to calculate a specific prominence ratio (SPR, a spectrum) by performing the calculation at each frequency division of the Fourier transform which meets the criterion for formation of a flanking lower and upper critical band (for example the red curve in Figure 1). The SPR value at the frequency of a discrete tone agrees with the calculation convention described in ECMA-74, but it can be seen that the SPR may have values higher at nearby frequencies than at a discrete-tone frequency, which may have perceptual implications (Figure 1). The SPR also responds to bandwidth-related tonalities not caused by discrete tones.

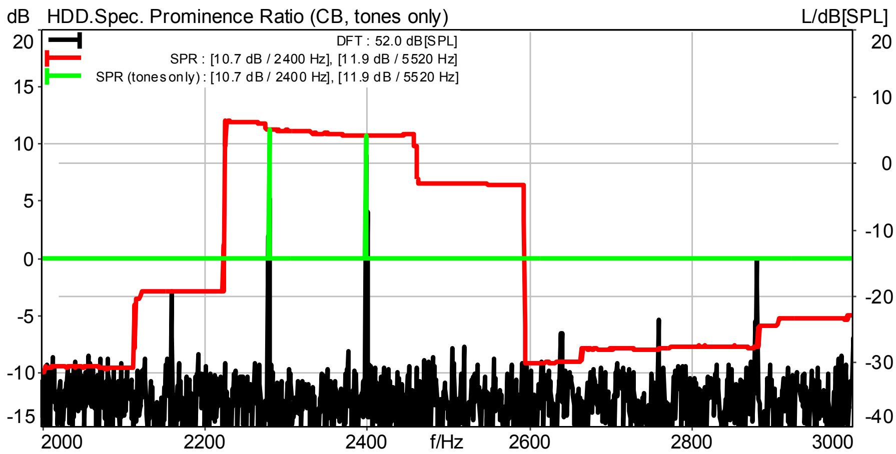

Figure 1–Specific prominence ratio (SPR) analyses (left axis) of a hard disk drive (HDD) noise signal: The red curve shows the SPR for each frequency of the DFT (black curve, right axis, DFT resolution: 65536, sampling rate: 48 kHz).The green curve displays the results only at frequencies where possible prominent tones have been detected based on the DFT according to the criteria 1-3 in section 2.1. The detected prominent tone within one critical band corresponds to the candidate with the highest level (at 2.4 kHz).

2.3 New method for determining tonality, including a hearing threshold compensation

Threshold of hearing data based on the audibility of pure tones are used to calculate an additional internal masking noise level. According to ECMA-74 the 1-percentile distribution $P_1(f)$ (essentially, the "lower limit" of the hearing threshold LTH) is more suitable than the 50-percentile distribution. It is assumed that the masking noise level exceeds the tone level by 4 dB at the threshold. Thus L_{int} = LTH+4 dB is calculated at the frequency of interest. By default L_{int} is added to the level of the masking noise. In the case of the PR method it is added to the average level of the adjacent critical bands and in the case of the TNR method to the predicted noise level.

This method is robust and easier to implement than adding a special noise to the time signal prior to any analysis in order to decrease PR or TNR values (4). Figures 2 and 3 show the influence of compensating the effect of the hearing threshold on tonality for PR and TNR, respectively, using the example of a hard disk drive (HDD). The noise of the HDD has two more or less tonal components around 2.4 kHz and 5.5 kHz with very low energy.

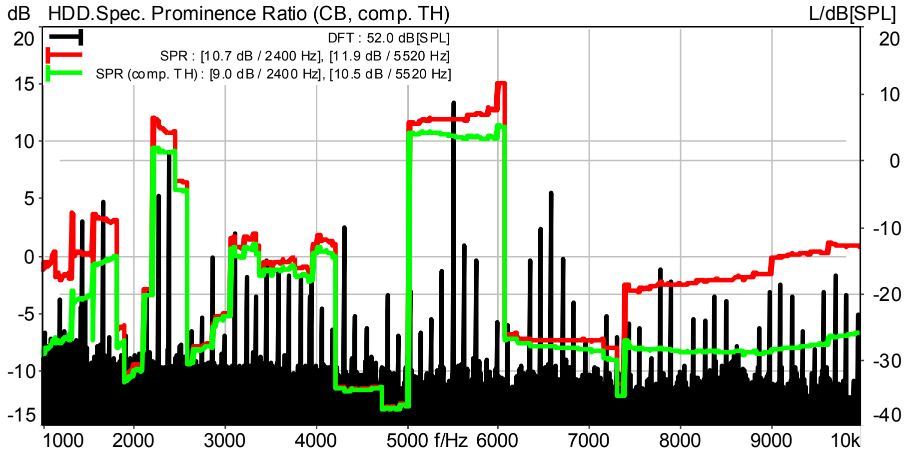

Figure 2– Specific prominence ratio (SPR) analyses (left axis) of a hard disk drive (HDD) noise signal: The red curve shows the SPR for each frequency of the DFT (black curve, right axis, DFT resolution: 65536, sampling rate: 48 kHz). The green curve displays the results considering the threshold of hearing.

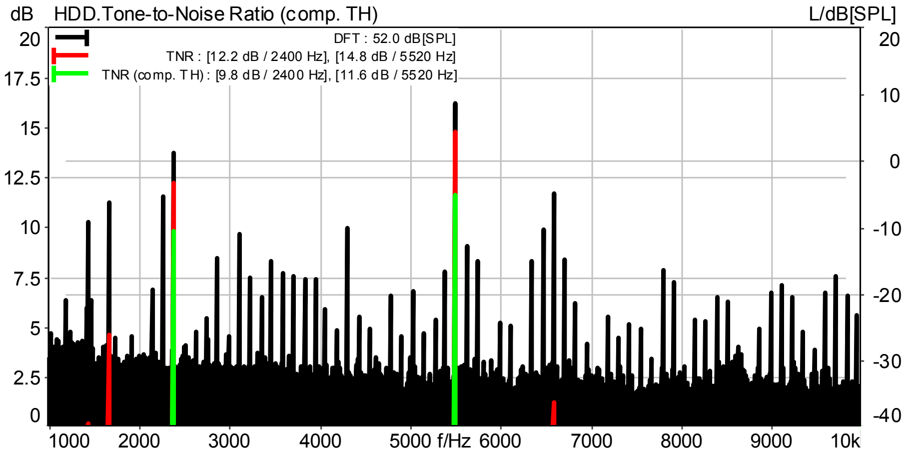

Figure 3–Tone-to-Noise Ratio (TNR) analyses (left axis) of a hard disk drive (HDD) noise signal: The red curve shows the TNR for the detected tones based on the DFT (black curve, right axis, DFT resolution: 65536, sampling rate: 48 kHz). The green curve displays the results considering the threshold of hearing.

TNR and PR methods as conventionally used do not respond to narrowband or non-pure-tone tonalities and to tonality phenomena with no distinct spectral maxima at all, e.g., to the tonality of filtered noise with very steep filter slopes (see section 4). In the following section, a new approach to tonality calculation based on a hearing model is presented that overcomes the drawbacks of the simplified approaches.

3. A HEARING MODEL APPROACH TO TONALITY

Recent research results show a strong correlation between tonality perception and the partial loudness of tonal sound components (5-7). Therefore a new hearing model approach to tonality on the basis of the perceived loudness of tonal content has been developed (2). Detailed listening tests on synthetically designed sounds have been carried out in order to evaluate this approach and provide reference data for model optimization. The applicability of the model was investigated for technical sounds and compared to established methods of tonality calculation (2).

With the intention to cover the most important aspects of tonality perception (as given in Zwicker and Fastl (8) and Hansen et al. (5)), pure sinusoids of different level, frequency and signal-to-noise ratio between a tone and broadband pink background noise have been generated. Furthermore, the superposition of two sinusoids with various frequency differences, multi-tone complexes with a different number of harmonics and narrow-band noise signals of different bandwidth have been investigated (2).

In the listening tests, 27 participants with unimpaired, normal listening abilities rated the tonality of the sounds mentioned above in random order on a 13-point category scale labeled with consecutive numbers. The verbal identifier (label) “no tonality” was assigned to the zero-point of the scale. The upper scale border was specified with the verbal identifier “extreme tonality”. The remaining verbal category identifiers were labeled “very low tonality”, “low tonality”, “medium tonality”, “high tonality” and “very high tonality”, each providing a subcategory between consecutive categories.

The results of the listening tests show a high correlation between tonality and the loudness of tonal components. Most cases show only insignificantly small discrepancies between the two parameters. One exception was found for the case of pure tones that are embedded in pink background noise with low noise level (2).

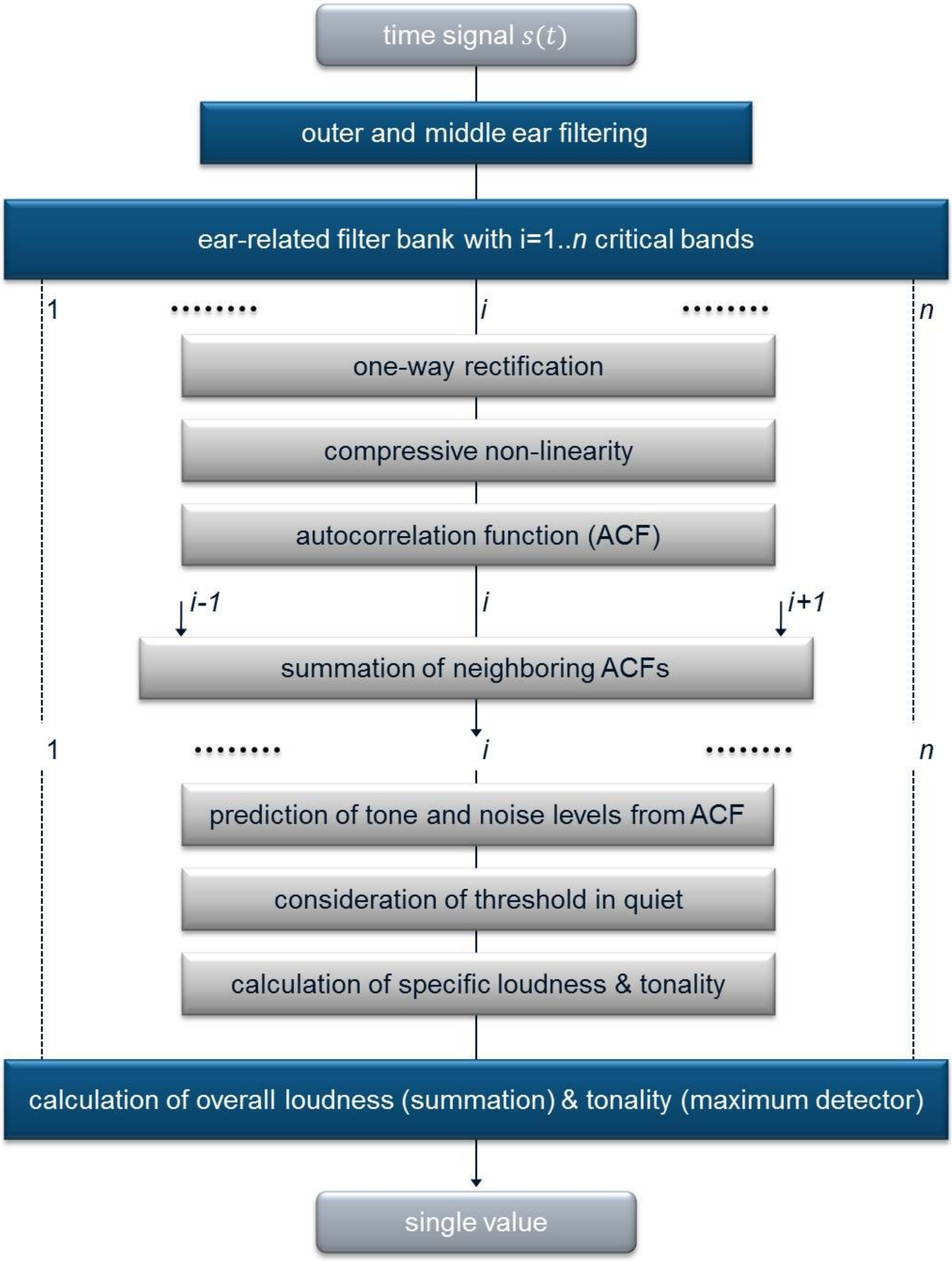

Figure 4–Model structure for the determination of loudness and tonality.

In early publications, Licklider assumed that human pitch perception is based on both spectral and temporal cues (9). According to (9), the neuronal processing in human hearing applies a running autocorrelation analysis of the critical band signals. Under this assumption, psychoacoustic phenomena like difference tone perception or the missing fundamental phenomenon can be explained.

This work has inspired the idea to use the sliding autocorrelation function as a processing block in the hearing model for the calculation of roughness and fluctuation strength (10) and later for other psychoacoustic quantities like tonality (2). Figure 4 displays the model structure. The many existing models differ mainly in three points:

1. the frequency weighting, which is the main cause for differences in modeling equal loudness contours (especially at low frequencies: modeling the outer and middle ear transfer function, the input signal *s(t)* is filtered by a filter representing the equal loudness contour at 100 phon),
2. the frequency scale (Bark or ERB) meaning the frequency-dependent bandwidth of the implemented m-channel filter bank (in this model to decompose the input signal into *n* critical bands, the envelope of each sub-band signal is calculated by one-way rectification),
3. the nonlinear relation between sound pressure and specific loudness (a strongly compressive non-linear function in combination with the calculation of the autocorrelation function in each sub-band).

The nonlinearity of this hearing model uses power functions with different exponents for different level ranges (10, 11). Such a nonlinearity function has proven applicable to predict many phenomena like ratio loudness, just-noticeable amplitude differences and modulation thresholds as well as the level dependence of roughness.

Calculating the autocorrelation function (ACF) of the bandpass signals provides a possibility to separate tonal content from noise. The autocorrelation function of white Gaussian noise is characterized by a Dirac impulse. Any broadband noise signal has at least a non-periodic autocorrelation function with high values at low delay times, whereas the autocorrelation function of periodic signals shows also a periodic structure. Thus, the loudness of the tonal component can be estimated by analyzing the ACF at a certain range with respect to the delay time, and also the loudness of the remaining (noisy) part (2). As a consequence, the presented tonality model may serve also for an improved loudness prediction with a special weighting of the loudness of tonal and non-tonal components (12).

(Specific) tonality depends on a kind of 'tone-to-noise ratio' as described in (2), which is not calculated as a ratio of intensities but as a ratio of nonlinearly transformed quantities ('specific loudness values'). In order to consider the threshold of hearing, a frequency dependent 'specific loudness threshold' is predicted based on the level of an 'internal noise' L_{int} = LTH (P_1) +4 dB (see section 2.3) (by applying the compressive nonlinearity to the corresponding sound pressure of L_{int}).

4. TONALITY AND PITCH OF BANDPASS-FILTERED NOISE SIGNALS

The loudness of differently bandpass-filtered uniform exciting noise signals was studied as a function of bandwidth and level (12). During these experiments it could be confirmed that filtering with 'infinitely' steep spectral slopes (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) evokes a strong tonal character, for almost any bandwidth (8, chapter 5.5). Narrow filters elicit a stronger tonality, also recognizable in the case of 4^{th}- or 6^{th}-order filters. To study this phenomenon in more detail, the loudness experiments were extended by evaluating the tonality of these sounds using a categorical scaling method.

In the listening tests, 14 participants with unimpaired, normal listening abilities rated the tonality of the uniform exciting noise signals centered around 1.5 kHz as a function of bandwidth using three different filter types (4^{th}-order, 6^{th}-order and 'infinitely' steep) in random order on a 7-point category scale. For the experiment the verbal category identifiers "no tonality" "very low tonality", "low tonality", "medium tonality", "high tonality" "very high tonality" and "extreme tonality" were used. The calculated loudnesses according to ISO 532-1 (13) were 14.9 sone, 12.3 sone, and 11.0 sone, respectively. The signals with a bandwidth of 3 Bark were generated with a level of 70 dB.

Figure 5 shows the mean of the ratings and the confidence intervals (95%). The noise signals with 'infinitely' steep spectral slopes all show a very high tonality. The 6^{th}-order filtered signals show high or medium tonality for a bandwidth of 1 Bark and 2 Bark, respectively. The 4^{th}-order filtered signals show medium tonality only for a bandwidth of 1 Bark. All other signals are judged to have no or a very low tonality. The 'inf'-order filtered signals do not change their sound character very much with increasing bandwidth: they remain tonal sounds, but with decreasing pitch as shown in the following.

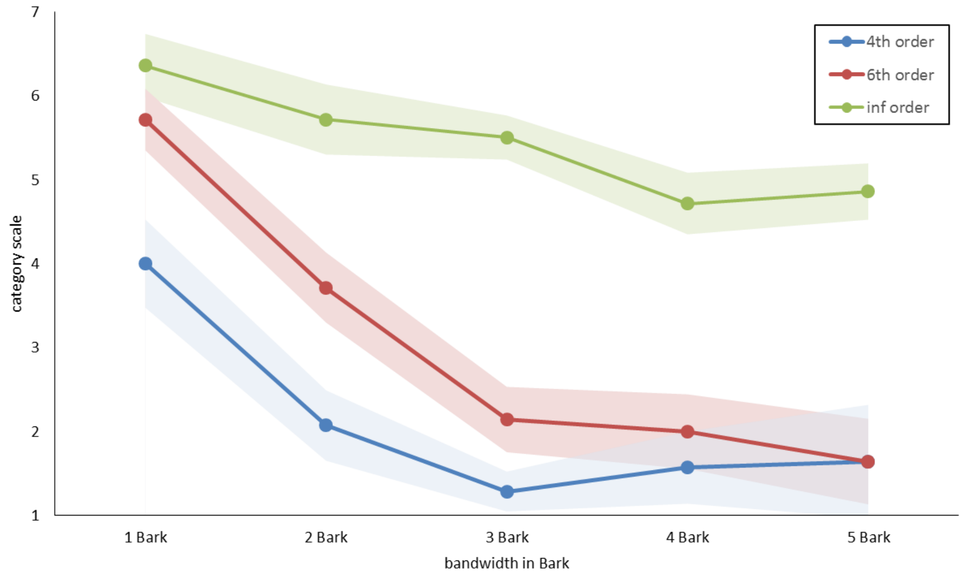

Figure 5–Categorical scaling of tonality on a 7-point category scale for uniform exciting noise signals centered around 1.5 kHz as a function of bandwidth using three different filter types (as indicated). The arithmetic mean of the ratings of 14 participants and the confidence intervals (95%) are shown.

Table 1 displays the dominance matrix (relative preference in % of col. vs row with respect to pitch) for a full paired comparison test using the sounds with 'infinitely' steep spectral slopes as a function of bandwidth. For example, in 89.29% of the comparisons, the pitch of the signal with a bandwidth of 1 Bark was rated higher than the pitch of the signal with a bandwidth of 2 Bark. Obviously, the pitch is decreasing with increasing bandwidth.

Table 1 – Dominance matrix for a full paired comparison test with respect to the pitch of uniform exciting noise signals ('infinitely' steep spectral slopes centered around 1.5 kHz) as a function of bandwidth

bandwidth	1 Bark	2 Bark	3 Bark	4 Bark	5 Bark
1 Bark		10.71%	7.14%	3.57%	0.00%
2 Bark	89.29%		14.29%	7.14%	3.57%
3 Bark	92.86%	85.71%		14.29%	21.43%
4 Bark	96.43%	92.86%	85.71%		32.14%
5 Bark	100.00%	96.43%	78.57%	67.86%	
mean	96.64%	71.43%	46.43%	23.21%	14.29%

The noise signals have been processed using the hearing model approach for tonality calculation shown in Figure 4.

Figure 6 shows as an example the specific tonality vs. time distributions of the signals with the steepest spectral slopes for the lowest and highest bandwidth. The location of the maximal specific tonality is shifted from 13 Bark to 8.5 Bark with increasing bandwidth. The maxima of the two distributions are 0.91 and 0.64, respectively, relative to the tonality of the reference sound (1 kHz-tone with a level of 60 dB plus a pink noise with a level of 60 dB).

Figure 7 shows the calculated specific tonality vs. time distributions for the 6^{th}- and 4^{th}-order filtered noise signals with the smallest bandwidth. The maxima of the two distributions are 0.4 and 0.005, respectively.

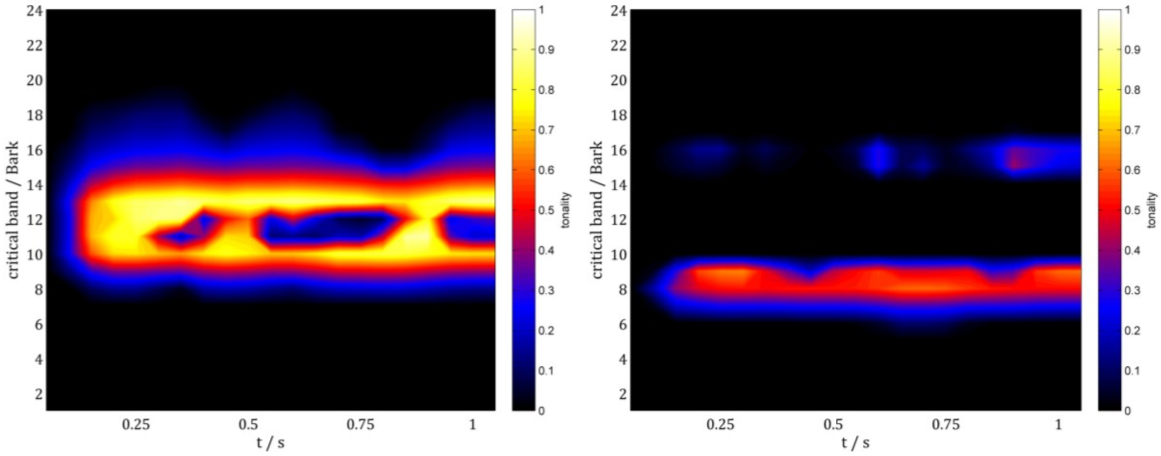

Figure 6– Specific tonality vs. time distributions for bandpass-filtered uniform exciting noise signals ('infinitely' steep spectral slopes, left: bandwidth=1 Bark, right: bandwidth=5 Bark).

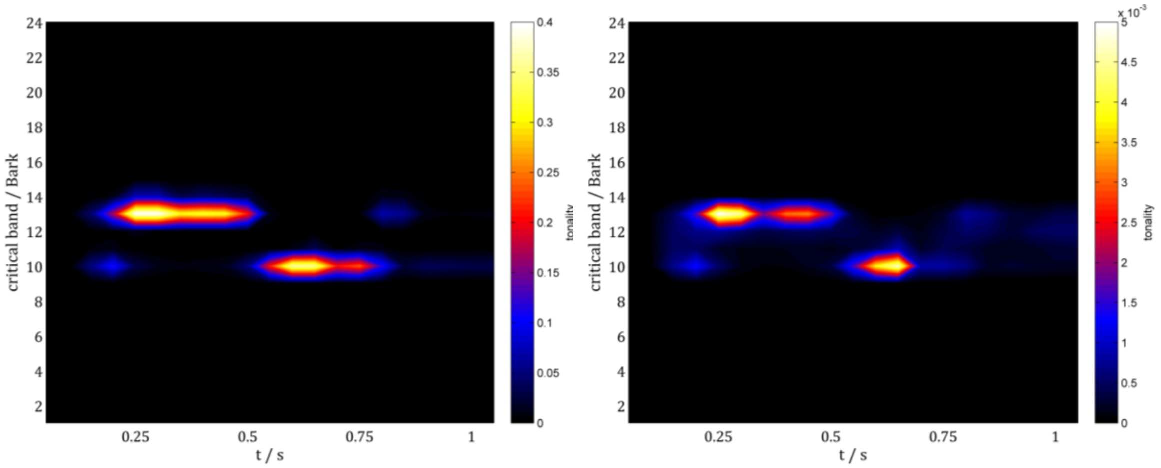

Figure 7– Specific tonality vs. time distributions for bandpass-filtered uniform exciting noise signals (bandwidth = 1 Bark, left: 6^{th}-order filtered, right: 4^{th}-order filtered). Note the different scaling factors.

The first results look very promising although the last calculated values seem to be a bit small (Figure 7). The main model parameters have not been changed compared to (2), except the consideration of the threshold of hearing. A possible optimization will be proven carefully in the near future by, among others, using more test cases.

5. PERCEPTION OF TONALITY OF TECHNICAL SOUNDS

In this section, two practical examples for tonality evaluation and prediction of technical sounds are given: howling vehicle sounds and hard disk drive noise.

5.1 Tonality of howling vehicle sounds

The tonality of howling sounds has 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 tonality of the reference sound was defined to be 100 and the subjects were asked to give a corresponding number for the tonality of the second sound.

Figure 8 shows the geometric mean of the ratings and the confidence interval (95%) together with calculated ratios based on the algorithms of Sottek et al. (2) and Terhardt et al. (12). Both methods correspond well to the perceptual data; the algorithm of Sottek et al. lies even within the confidence interval.

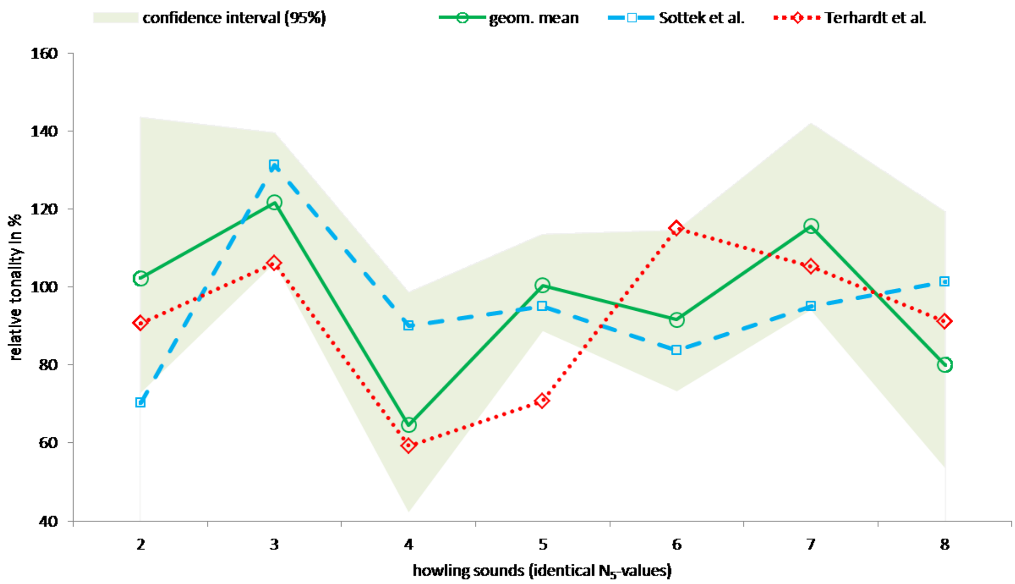

Figure 8–Magnitude estimation of the tonality of howling sounds. The geometric mean of the ratings of 18 students (age between 23 and 29 years) and the confidence interval (95%) are shown and compared to the results based on the tonality calculation according to Sottek et al. (2) and Terhardt et al. (14). All signals have been adjusted to the same loudness (N_5-value) according to ISO 532-1 (12).

5.2 Tonality of hard disk drive noise

IT devices and products like hard disk drives may emit tonal sounds. Despite their very low sound pressure levels, such noises are unwanted and should preferably be avoided or masked if they are perceived as prominent. Figures 1 to 3 show SPR and TNR analyses of a hard disk drive noise signal. Two candidates of tonal components were detected around 2.4 kHz and 5.5 kHz (about 15 Bark and 20 Bark, respectively). The component at 5.5 kHz was evaluated as the more prominent of the two; the other one was almost at the threshold. Figure 9 shows the tonality results achieved by the application of the hearing model without compensating for the threshold of hearing (2) and with compensation.

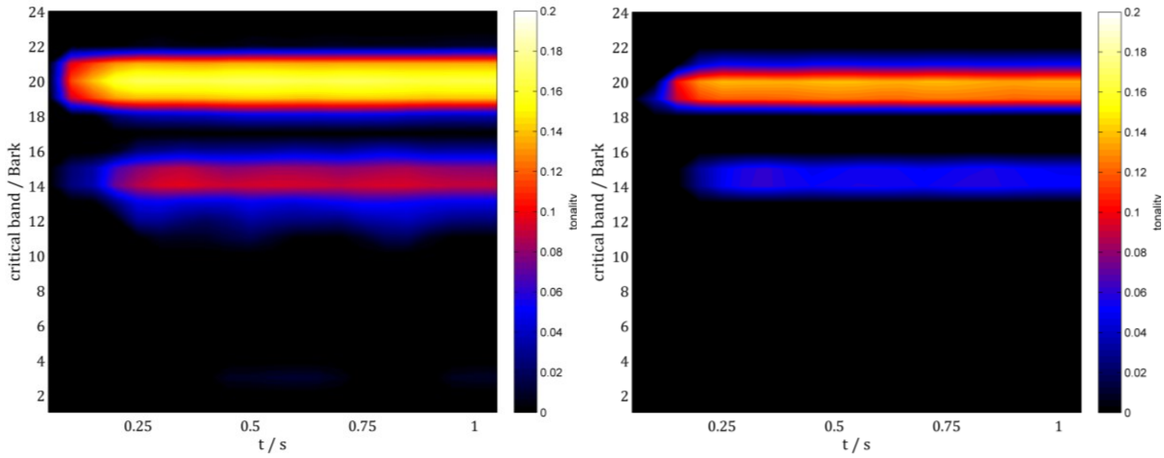

Figure 9– Specific tonality vs. time distributions for a hard disk drive noise signal; results are achieved by applying the hearing model without (left) and with (right) compensation for the threshold of hearing.

6. CONCLUSIONS

This paper began with an overview of existing *tonality measures* (TNR and PR) and has introduced an effective method to compensate for the threshold of hearing. Then an approach has been presented to describe the perception of tonality as a *psychoacoustic sensation* upon the basis of a hearing model (2). Tonality calculation is therein performed by estimating the loudness of tonal components. The separation of tonal content and noisy background is provided by an analysis of the autocorrelation function in the different frequency bands. The parameters for the tonality model have been used according to (2), where they have been derived as a result of an optimization process using extensive data from listening tests. New listening tests showed that this model is also applicable for predicting the tonality of bandpass-filtered noise signals with varyingly steep spectral slopes. The model has been extended to consider the threshold of hearing by applying a specific loudness threshold based on the LTH. This is a promising approach to predict the prominence of tonalities near the threshold of hearing. Further experiments will be performed to validate the proposed method.

ACKNOWLEDGEMENTS

The author wishes to thank his colleagues, Wade R. Bray and Georg Caspary, for many fruitful discussions.

REFERENCES

1. ECMA-74 12th Edition/December 2012, D.7. Tone-to-Noise Ratio method; ECMA International, Rue du Rhône 114, CH-1204 Geneva, Switzerland, 2012.
2. Sottek R, Kamp F, Fiebig A. A new hearing model approach to tonality, Proc. Internoise, Innsbruck, 2013.
3. Bray W, Caspary G. Automating prominent tone evaluations and accounting for time-varying conditions, Sound Quality Symposium, SQS 2008, Detroit, 2008.
4. Rafaelof M. Accounting for the Absolute Threshold of Hearing when Analyzing Low Level Sound. Noise-Con 2013, Denver, 2013.
5. Hansen H, Verhey JL, Weber R. The Magnitude of Tonal Content. A Review, Acta Acustica united with Acustica, Vol. 97, pp. 355-363, 2011.
6. Hansen H, Weber R. Zum Verhältnis von Tonhaltigkeit und der partiellen Lautheit der tonalen Komponenten in Rauschen, Deutsche Jahrestagung für Akustik, DAGA 2010, Berlin 2010.
7. Verhey JL, Stefanowicz S. Binaurale Tonhaltigkeit, Deutsche Jahrestagung für Akustik, DAGA 2011, Düsseldorf, 2011.
8. Fastl H, Zwicker E. Psychoacoustics – Facts and Models, Springer Verlag Berlin, 2007.
9. Licklider JCR. A Duplex Theory of Pitch Perception, Cellular and Molecular Life Sciences, Vol. 7 (4), pp. 128-134, 1951.
10. Sottek R. Modelle zur Signalverarbeitung im menschlichen Gehör. Doctoral Thesis, RWTH Aachen University, 1993.
11. Bierbaums T, Sottek R. Modellierung der zeitvarianten Lautheit mit einem Gehörmodel. Deutsche Jahrestagung für Akustik, DAGA 2012, Darmstadt, 2012.
12. Sottek R. Improvements in calculating the loudness of time varying sounds, Proc. Internoise, Melbourne, 2014.
13. ISO 532-1. Methods for calculating loudness, Part 1: Zwicker method, in preparation.
14. Terhardt E, Stoll G, Seewann M. Algorithm for extraction of pitch salience from complex tonal signals, J. Acoust. Soc. Am. 71, no. 3, pp. 679-688, 1982.