

70Hz. Historically roughness has been used in the calculations of annoyance metrics and also to quantify sound quality in a large number of noise evaluation applications.

$$R = cal \cdot \int_0^{24bark} f_{mod} \cdot \Delta L \cdot dz \quad (2)$$

Where *cal* is a calibration factor, f_{mod} is the frequency of modulation and ΔL is the perceived masking depth. The main difficulty in roughness calculation is to obtain an accurate quantification of ΔL . Because of this there are several proposed methods of calculation.

2.2. Perceptual models

As mentioned above, these methods are based on the simulation of psychoacoustic effects responsible of sound quality perception. Perceptual methods were developed to assist assessment of sound quality in speech and music compression systems since the classical subjective test are usually more expensive and time-consuming. These methods have been developed in two ways: explicit simulation of masking processes and simulation of physiological and psychoacoustical effects in hearing system.

2.2.1. Rnonlin

The Rnonlin metric was developed as an extension of the DS metric developed by Moore et. al. in [5]. The Rnonlin metric analyzes the difference between the input test signal and its distorted output. A coherence analysis is performed by taking the cross-correlation between the input and distorted output waveforms. Moreover, the metric algorithm uses a model of the frequency analysis performed in the peripheral auditory system including the filtering produced by the outer and middle ear. Figure 1 shows a block diagram of the steps involved in the human auditory periphery model applied in order to obtain the velocity signals for each frequency band.

Figure 2 shows the entire block diagram that summarizes the calculation process of the Rnonlin value. Both, the input (reference) and the distorted waveforms are time aligned to remove delays caused by the nonlinear system. These waveforms are filtered to simulate the response of the outer and middle ear using 4097 FIR filters, as described by Glasberg and Moore [6].

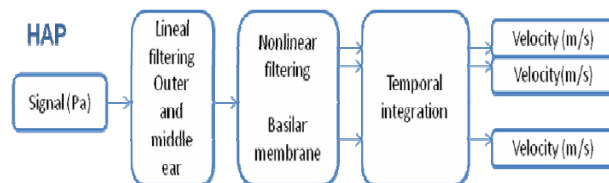


Figure 1 Block diagram of the human auditory periphery model.

Both waveforms are then filtered by an array of 40 gammatone filters with a bandwidth of 1-ERBN. This filtering provides a modeling of the auditory filtering mechanism [7].

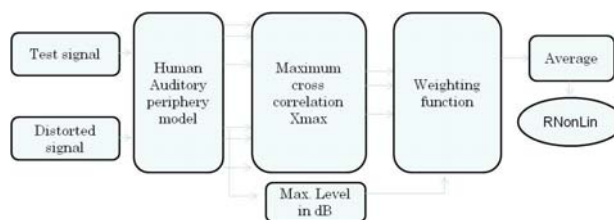


Figure 2 Block diagram for the Rnonlin calculation process.

After the human auditory periphery model, the input and output signals are split into 30ms. The next step consists in calculate the maximum value of the normalized cross-correlation between the input and output signals, X_{max} . For each 30 ms frame, the X_{max} values are summed across all filters. Finally, the X_{max} values are averaged over all the frames resulting in the single valued, Rnonlin .

3. PROCEDURE

3.1. Test signals

The following five test signals have been used in order to make the different approaches: one single 1 kHz tone used to obtain the total harmonic distortion of each system; a 1-4 kHz multitone signal to measure intermodulation; one 1kHz signal AM modulated with with $f_m=70$ Hz and a modulation depth of 100%. This signal is used for roughness, tonality and sharpness calculations. Finally, two music passages have also been employed for Rnonlin calculations. Both music passages were presented to the listening test subjects.