Psychoacoustics in Imaging Localization of Sound Sources

Stefan Neugebauer, Benjamin Riebold, Dirk Döbler

Gesellschaft zur Förderung angewandter Informatik e. V., 12489 Berlin, Germany, email: neugebauer@gfai.de

Introduction

Delay-and-Sum-Beamforming came to be an important investigation technique for complex space and time variant sound sources delivering a sound pressure level based imaging evaluation. However often also psychoacoustic parameters like loudness, sharpness, tonality or roughness have to be taken into account to achieve a complete rating of sound quality. GFaI - the developer of the "Acoustic Camera" which provides acoustic mapping in 2D-virtual planes and 3D-models - implemented algorithms to calculate the most important psychoacoustic parameters which were integrated into the acoustic mapping. The values of the respective psychoacoustic parameters are represented by a color scale and superposed to the optical image of the investigated object.

The combination of both - imaging localization of sound sources and psychoacoustic evaluation - provides a powerful tool for enhancement of sound quality and efficient noise reduction.

Opportunities of application are demonstrated using typical examples.

Psychoacoustic Models

The aim of psychoacoustic models is to simulate those physical and neuronal processes that are important for sound perception and its psychological sensation. The models discussed here give a quantification of the sensation of:

- Loudness
- Roughness
- Sharpness

One common feature of all these models is processing of data in filter banks where the bandwidth of the filters is equidistant on the Bark Scale. Since the Bark Scale is a direct result of the anatomic characteristics of the cochlea of the human ear it provides the basis for many hearing models. Investigations using hearing tests showed that the human brain analyzes sound in relatively narrow frequency bands, called critical bands [1]. Frequencies of a critical band stimulate a specific segment of the cochlea so the frequency range from 0 Hz to 16 kHz is divided into 24 critical bands with constant width of 1 Bark corresponding to a non-constant width on the frequency scale.

This characteristic of the human ear is modeled in the following hearing models by a filter bank. The first filter-banks that were used and are still used for this purpose are third octave filter banks. Alternatively butterworth based filter-banks or a Gammtone filter bank as proposed by Patterson [3] can be used. For the Gammatone filter banks the transfer-function is given by Equation (1) for a centerfrequency $f_0$. $n$ is the order of the filter and $b$ is a constant that influences the width of the filter-band.

$$GT(f) \propto \left[1 + i(f - f_0)/b\right]^{-n}$$

After separating the signal into critical bands, every band can be analyzed using one of the following specific models.

Loudness

Loudness belongs to the category of intensity sensation. Its value says how much louder a sound is heard in relation to a reference sound [1]. This reference sound is a 40 dB 1 kHz tone. Its loudness is 1 sone. Detailed description of the model can be found in DIN 45631/A1 [6]. This model also simulates spectral and temporal masking.

Sharpness

Sharpness calculation is based on loudness calculation. But it uses a frequency dependent weighting function that is multiplied to the specific loudnesses of every critical band. So sharpness is dependent on the spectral envelope and the center frequency of the spectral bandwidth. It is relatively independent of level. 1.00 acum is the sharpness of a small band noise from 920 Hz to 1080 Hz at a level of 60 dB. The model we used in this case is according to DIN 45692 [7].

Roughness

Frequency dependent analysis of amplitude modulation is used to model the psychoacoustical sensation of roughness [4]. Roughness peaks at a modulation frequency of 70 Hz and is generated in a frequency range from 20 Hz to 300 Hz amplitude modulation. We implemented the hearing model by Sottek [2]. Unit for Roughness is asper. Roughness of 1 asper is produced by a 1 kHz sinusoidal at 60 dB with modulation frequency of 70 Hz.

Beamforming

Beamforming is a technique providing localization of sound sources in 2D-shapes and 3D-models by using a microphone array and an optical camera. A wide field of applications for sound pressure level analyses has been shown [5]. Furthermore this technology offers the possibility to reconstruct the time dependent function of the sound pressure for every coordinate in 2D or 3D. Once this time function of sound pressure has been reconstructed it is possible to calculate psychoacoustic parameters using the above models separately for every
source that contributes to the sound field. Results of this calculation will be shown in the following examples.

**Results**

The first example gives a comparison of a standard sound-pressure level based 2D mapping of sound sources (Figure 1) and a psychoacoustic image of the same scene using the roughness model in this case. As can be seen on the first picture the sewing machine is providing two clearly distinguishable sources: the motor (on the right) and the needle. The motor is the dominant source regarding its sound pressure whereas on the psychoacoustic image the motor is completely suppressed by the needle which is responsible for the roughness in the sum sound of this object.

![Figure 1: Comparison of sound pressure level in dBA (top) and calculated roughness in asper (bottom) for an engine with two separable sources.](image)

The second example (Figure 2) of a diesel engine shows two separable sources on the sound pressure map (top). The corresponding images of roughness and sharpness show in detail which sources are producing what kind of psychoacoustic perception. In case of the roughness calculation (middle) again the source is clearly dominant which has a lower sound pressure level. An especially interesting result can be seen on the sharpness mapping (bottom) where the two sources of sound pressure level are suppressed and sources at the cylinders which can’t be seen on the sound pressure map can be correlated to generation of sharp sounds.

![Figure 2: Sound pressure level in dBA (top), roughness in asper (middle) and sharpness in acum (bottom) mapping of a diesel engine.](image)

So the integration of psychoacoustic modeling into source mapping technique helps to optimize psychoacoustic sound quality by giving an exact split up of those sources leading to a psychoacoustic effect and those which do not.

**Conclusion**

We have shown that source mapping by sound pressure level often does not represent the psychoacoustic importance of sources correctly. We have also shown that in many practical cases in a complex set of sources only few individual sources correspond to a special kind of psychoacoustic sensation. On the other hand the calculation of psychoacoustic parameters of a signal which is a superposition of many sources doesn’t give a hint which sources should be manipulated to improve the overall sound. So the calculation of psychoacoustic maps as shown in the examples gives the opportunity to identify those sources that can be assigned to psychoacoustic parameters.

Further investigations will be needed to show applications where overall psychoacoustic parameters can be modified just by identifying and modifying single psychoacoustic relevant sources. This can be done by damping or gaining sources on the acoustic image by calculation as well as by direct modification of the measured object.

**Acknowledgement**

This research work has been funded by the German BMBF under project registration number 03WKP24A.
References


