# A REAL-TIME BINAURAL LOUDNESS METER

*Olli Tuomi and Nick Zacharov*

Nokia Research Center, Speech and Audio Systems Laboratory,
P.O. Box 100, FIN-33721 Tampere, FINLAND
e-mail:[olli.tuomi,nick.zacharov]@nokia.com

## ABSTRACT

The alignment of sound intensity is of great importance particularly in the field of psychometrics. When sounds have different spectral content, loudness alignment becomes a more suitable tool. However, traditional models assume the source either to reside in a free field at 0° azimuth or in a diffuse field. To assess or align sound source loudness at different locations or with different radiation characteristics is more problematic.

## 1. INTRODUCTION

When performing psychometric studies of sound, it is well known that normalisation of signals for equal loudness is of critical importance. This has often been illustrated as found by the work of Aarts and Bech [1-4]. Quite often intensity based alignment are suitable for simple signals, however, for signals with different spectral content, loudness alignment is know to be superior. Several models have been developed including the well known and standardised Zwicker loudness model [5-7]. In more recent times, Moore has developed [8] and refined loudness models [9]. Both are designed for use in either the diffuse field or free field with a free field microphone. In the latter case, it is assumed that the loudness spectrum measured represents a sound source situated in front of a human head at an azimuth and elevation of 0°. However, this is often not the case in practical situations particularly when researching spatial, multichannel or 3D reproduction sound systems. Furthermore, it is rare for a real room to exhibit either truly diffuse or free field characteristics.

In the situation where sources are in a free field at other angles than 0° azimuth and elevation, it is possible to apply a different free field to ear drum head related transfer function (HRTF) for that source angle for each ear. In this manner it becomes possible to simulate the free field directional loudness properties for a head. This has been performed for a one ear of a head and torso simulator (HATS) HRTF [10] and is illustrated in figure 1. It can be observed that there is a clear and significant directional effect, which supports the findings of Robinson & Whittle [11], and Sørensen et al [12].

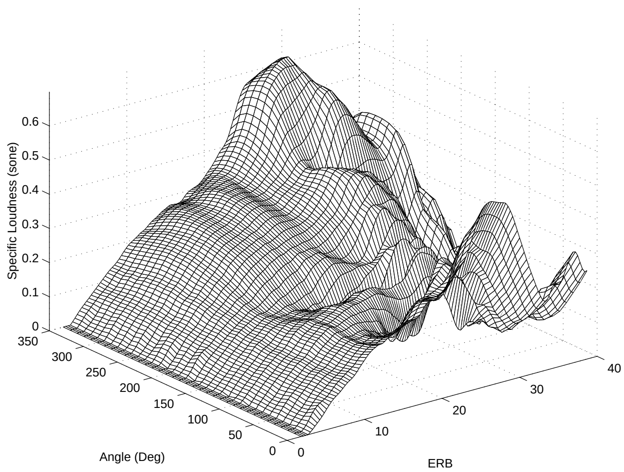


Figure 1 Simulated monaural specific loudness spectrum (left ear) of white noise as a function of azimuth for a head and torso simulator

According to Moore [9], it is possible to obtain an approximation of the binaural loudness response, simply by the summation of the monaural loudness spectra. This has been simulated and is illustrated in figure 2.

From these simulations, a clear need for a binaural loudness model is desired particularly for the loudness alignment of reproduction systems with multiple sound sources at varying locations.

Why the need for a real-time tool? Whilst alignment of sound pressure level can easily be performed with dB gain compensations, the mapping between the loudness and gain domain is non-linear and complex. By developing a real-time loudness meter it becomes possible to measure loudness in real-time and adjust gains for accurate and easy loudness alignment.

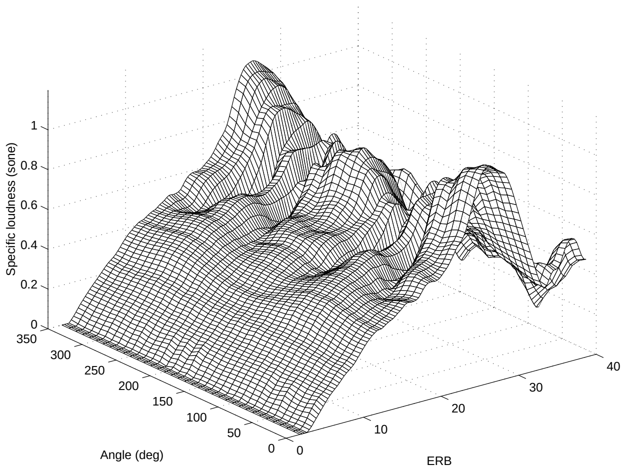


Figure 2 Simulated binaural specific loudness spectrum of white noise as a function of azimuth for a head and torso simulator

## 2. TOOL DESIGN

### 2.1. STARTING POINT

The starting point for the tool development was the Fortran version of Moore's model, which had been translated into a version running in the Matlab environment. This version was the basis for the development of the C-language version of the algorithm. C was selected as the implementation language due to the limited speed of the Matlab in-built functions and structures. Matlab has a means of making the use of C-language routines within native Matlab code possible, called MEX routines.

### 2.2. BASIC TOOL STRUCTURE

The C-code part of the model was divided into two parts in an attempt to reach a more efficient structure for the code. In simple implementations of the model, a great portion of time is spent in creating the large data structures required by the model. This has been avoided by taking out the data structure creation from the computational core and placing it in a separate module.

This approach brings in additional problems with memory management. As the data structures are created separately from the actual computation, in a different process, a way to share these structures had to be devised. Different processes generally would need operating system calls to be able to share data structures. The solution selected was to use C-code to create the data structures in the Matlab workspace. This allowed easy access of the structures and verification of the data from within Matlab, and greater computational efficiency.

The basic structure of the data structures had to be altered, because all the data had to be pre-calculated. This results in a data structure of greater than 10 megabytes, when the original model only needed a fraction of this. The memory requirements would have been even greater, had special techniques not been applied to compress the data.

As a result of the above, the speed of the model would be increased considerably. In comparison to a version of the model implemented in pure Matlab code, the divided C-approach would run approximately 100 times faster. The model could be used to calculate the specific loudness of 8192 sample window in 20 milliseconds. This was fast enough to build a real-time application with the model.

### 2.3. PRACTICAL ISSUES

In the full version of the model, the biggest latencies are no longer due to the computation, but are a result of the plotting processes. Because of this, an option to turn off the plotting was implemented. This removes relevant information from the user interface, but increases the efficiency of the model tenfold. When the UI is stopped, this information is updated in the UI. This makes it possible to run the model efficiently even on a slow pentium–class machine. An alternative approach to this would be to implement the plotting routines in a more efficient manner using some external graphics library, but this was not done, because the required processing power was available.

### 2.4. THE USER INTERFACE

The structure of the tool is based on the action driven UI model supported by the Matlab graphical user interface generator. In the development of the user interface special care was taken to ensure easy and reliable operation. All relevant controls are displayed in the main screen, and the user has access to all data as the measurement is performed.

The tool offers OS independent operation with sample rates up to 48kHz, 16bit samples, mono and stereo operation. The use of any kind of signal source is supported, the model supports diffuse field, free field and a special Brüel & Kjær (B&K) HATS measurements mode. The tool provides a constant display of overall loudness in sones, phons and sound pressure level in dB with plotting of specific loudness or frequency domain graphical data. Results can be saved and several takes can be averaged with the associated averaging tool.

The user interface of the tool is depicted in figure 3. The main window shows the specific loudness of the signal shown in the lower part of the interface.

## 2.5. SIGNAL FLOW

The signal routing through the system is presented in figure 4. The signal is first routed through the sound hardware and operating system routines, which forward it to the data acquisition engine. The model reads data using the DAQ toolbox functions frame by frame, first applying a gain block as a way of aligning the input data to a known scale. Calibration is achieved using an external calibration signal and consequently adjusting the gain of the calibration block.

Hamming windowing is applied to the data, in preparation for the FFT algorithm. The output from the FFT is excessive for input to the model, so that a rescaling of the data was required prior to the model. The algorithm applied to create the nonlinear scaling was specifically designed in C for speed, large windows and different frequency grids, such as the ERB and Bark scales.

The signal is then forwarded to the computational routines of the loudness model described by Moore [9], but our specification required that an additional correction mode be built into the model. This mode would bypass any outer-ear correction, but still use the middle-ear correction block, and would be used in conjunction with the B&K HATS.

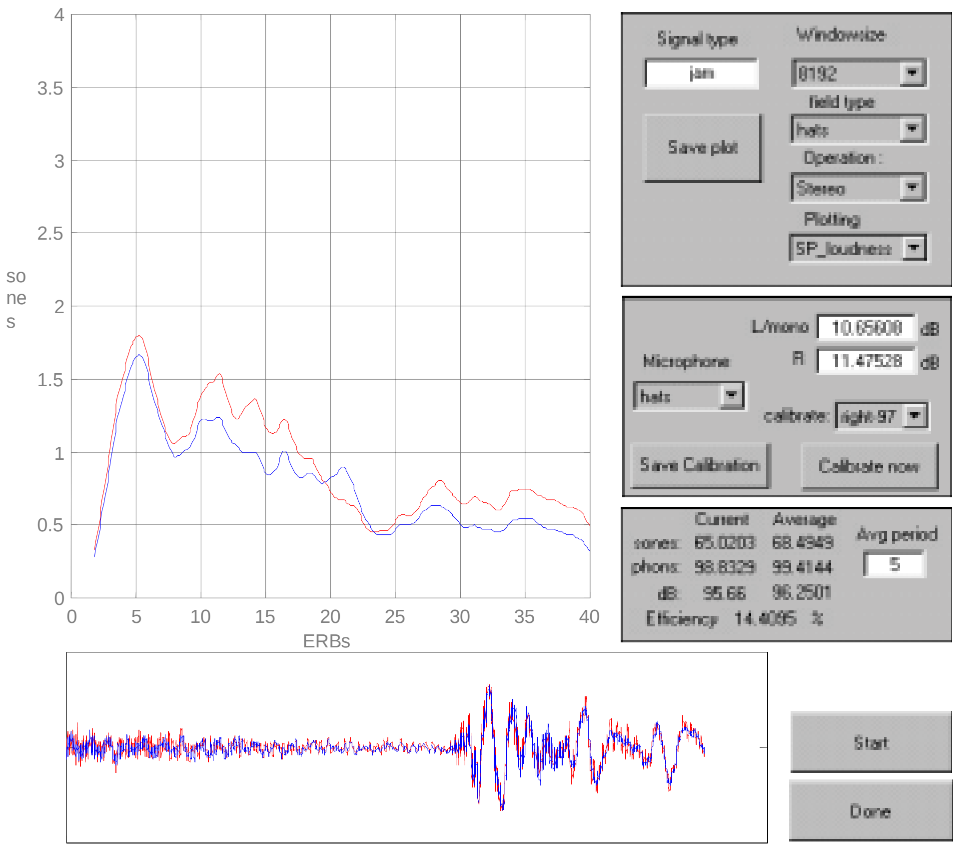


Figure 3 Example user interface

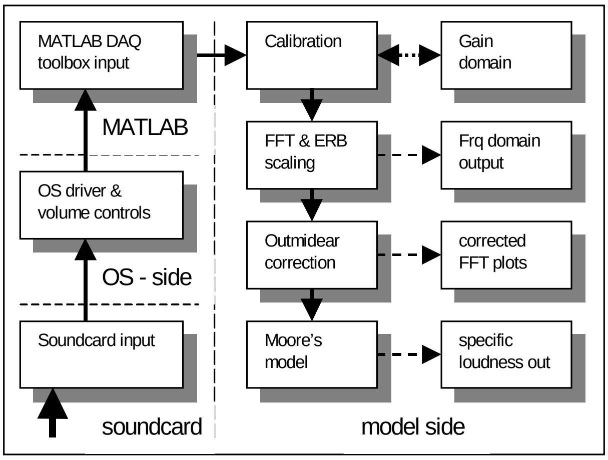


Figure 4 Model structure

## 3. VERIFICATION TESTS

The HATS binaural model was to be verified by comparison of performance with respect to the free field microphone model under ideal conditions in the original model. The range of conditions considered included:

- Free field condition, 0° azimuth source
- Diffuse field conditions
- A range of test signals (See table 1.)

### 3.1 TEST SETUP

In order to verify the correct operation of the tool, and that it produced correct results, a series of verification tests was planned. These tests would include tests involving all the functional parts of the model in different environments.

The test were conducted on a computer equipped with a Pentium III processor of 450MHz, 128Mbytes of memory, operating system Windows NT 4 service pack 5, Matlab version R11 and using a Crystal integrated soundcard.

Microphones: B&K 4133 and B&K 4128 HATS.

Sound source: Genelec 1030A

Artificial test signals were created with the Sonic Foundry Sound Forge32, which were then transferred to an audio CD to be played with a CD player. The test signals used are shown in table 1.

| SIGNAL TYPE | FREQUENCIES |
|---|---|
| pure sine waves | 250Hz, 500Hz, 1kHz, 2kHz, 5kHz |
| 2 sines, same ERB | 190Hz + 200Hz |
| 2 sines, different ERBs | 190Hz + 450Hz |
| noise | pink noise |
| noise | white noise |

Table 1 Primary test signals employed for verification

In all tests, the signal was amplified using the B&K NEXUS conditioning amplifier set to 3.16 V/Pa. The Windows mixing panel was used to adjust the signal levels to maximize the dynamic range of the soundcard, prior to calibration.

All signals were generated with a DENON professional CD player type DN-8680, and the signal was routed through a YAMAHA digital mixing console type 01V.

Prior to the verification tests, the system gains and microphones are calibrated via an in-built calibration routine. The free-field 4133 microphone was calibrated to 94dB with the calibrator B&K type 4231 calibrator. The 4128 HATS is calibrated to 97.1dB employing the UA-1546 adapter.

Pure sine waves were selected to verify the basic operation of the tool under simple signals. Signal intensities and frequencies are easy to calculate using simple signals. White noise was employed due to its flat spectral characteristics whilst pink noise is more psychoacoustically motivated and broadly excites the audible bandwidth. The combination of two sine waves in the same or different critical bands was selected to verify that the basic masking properties of the signals were visible from the output of the tool.

In the test, the inputs of the 4133 and HATS are compared. In all resulting figures, the overall loudness is calculated for a binaural signal, but the individual plots are the monaural responses. The binaural loudness in sones for monaural input as in the case of 4133 is calculated as 2 * sones, for ease of comparison.

### 3.2 FREE FIELD

The source was set up in the anechoic chamber at 2m from the microphones. The responses were first measured with the calibrated B&K 4133 microphone at the effective centre of the non-present HATS. The procedure was repeated for the HATS. The results of these measurements are illustrated in figures 5-8.

Overall, it can be stated that there is a good correspondence between the microphone and HATS results, both in terms of the loudness spectrum and the resulting overall loudness values. However, it can be noted, particularly at high frequencies, that there are some differences between the results. This may be cause by three factors. Firstly, the sound source is not an ideal point source and thus suffers from diffraction, and directivity associated with a two-way speaker design. Additionally, some differences may exist between the free field to ear drum transfer function employed in the original microphone and that of the HATS. Thirdly, there is a slight horizontal offset between the microphone position and the entrance of the HATS ear canal (~9cm).

### 3.3 DIFFUSE FIELD

Six point sources were set up in the standard reverberation chamber to create a diffuse field into the chamber. Measurements were made in a similar fashion to the free field case. However, to improve the diffuse field approximation, measurements were made in eight different positions and then results were averaged across these positions. The results are illustrated in figures 9-12.

Once again good agreement can be found between the two sets of responses. However, the variations between the two responses are greater than in the free test, particularly at high frequencies. In additions to the error sources discussed earlier, which also apply here, the diffusion of the chamber may provide some additional error. Firstly, such standard chamber are intended to be diffuse only in a 1/3rd octave sense and with an upper cutoff of ~10kHz. Secondly, microphone positional variations may also lead to some error. The directivity of the HATS vs. the microphone will also contribute to differences. This is considered to be the source of variation in figure 9, which is real.

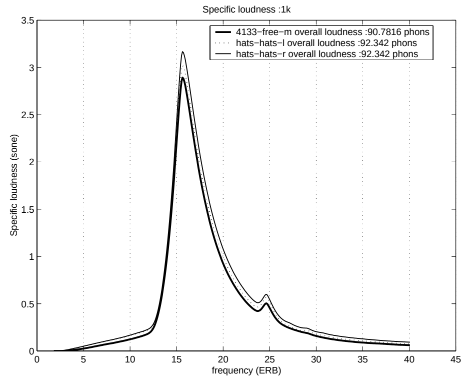


Figure 5 Free field specific loudness spectrum for a 1kHz sine wave (0° azimuth)

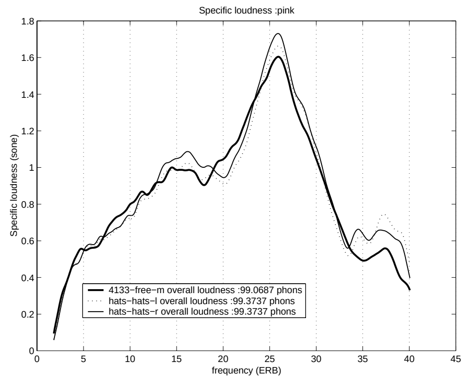


Figure 6 Free field specific loudness spectrum for pink noise (0° azimuth)

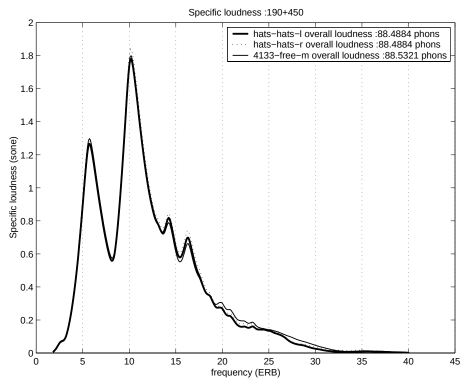


Figure 7 Free field specific loudness spectrum for a 190 and 450 Hz sine wave signal (0° azimuth)

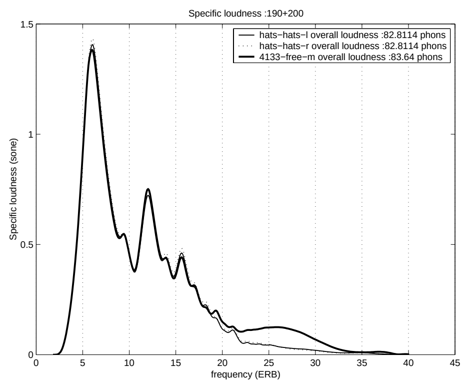


Figure 8 Free field specific loudness spectrum for a 190 and 200 Hz sine wave signal (0° azimuth)

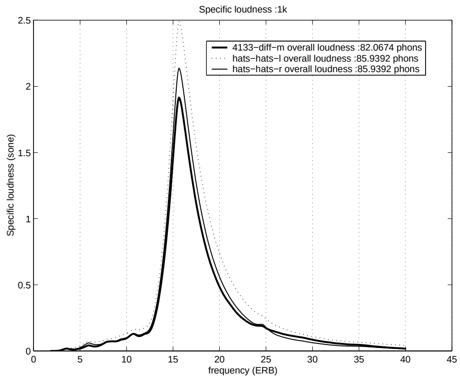


Figure 9 Diffuse field specific loudness spectrum for a 1kHz sine wave (diffuse field)

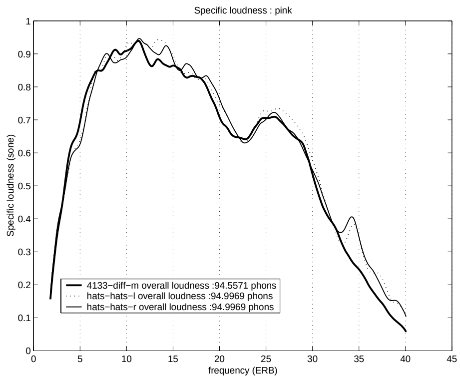


Figure 10 Diffuse field specific loudness spectrum for pink noise (measured in a diffuse field)

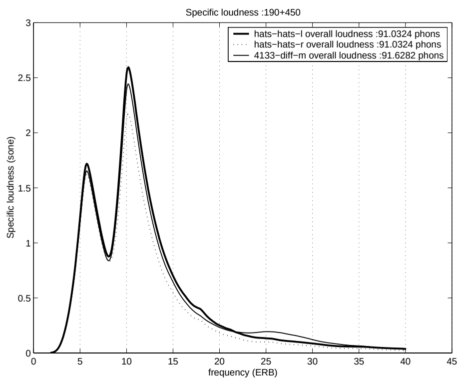


Figure 11 Diffuse field specific loudness spectrum for a 190 and 450 Hz sine wave signal

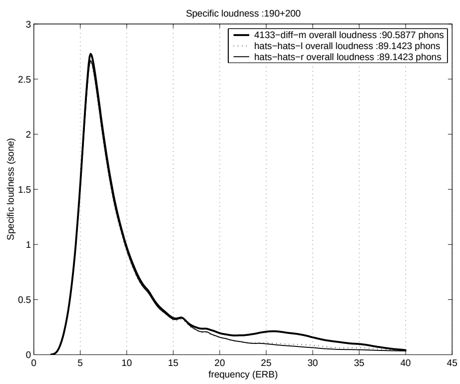


Figure 12 Diffuse field specific loudness spectrum for a 190 and 200 Hz sine wave signal

## 4. DIRECTIONAL LOUDNESS PERFORMANCE

Having established that the performance of the binaural implementation is in line with the common and standard models, it is now of interest to verify the directional loudness performance compared to simulations presented earlier.

To achieve this the HATS was mounted on an automated turntable. White noise was reproduced by a source at 2m distance in an anechoic chamber and the loudness spectrum data was collected in the horizontal plane in 10° increments. The monaural results are illustrated in figure 13 below and can be directly compared to the simulated version found in figure 1. The binaural measurements are shown in figure 14 and compare favorably to simulations shown in figure 2. A clear similarity can be found in both cases.

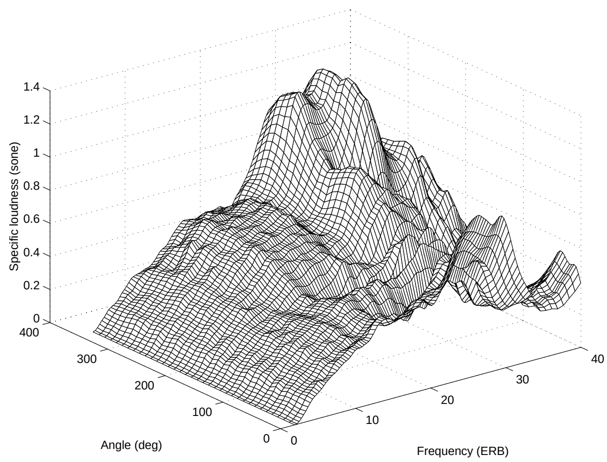


Figure 13 Measured monaural specific loudness spectrum (left ear) of white noise as a function of azimuth for the B&K 4128 HATS

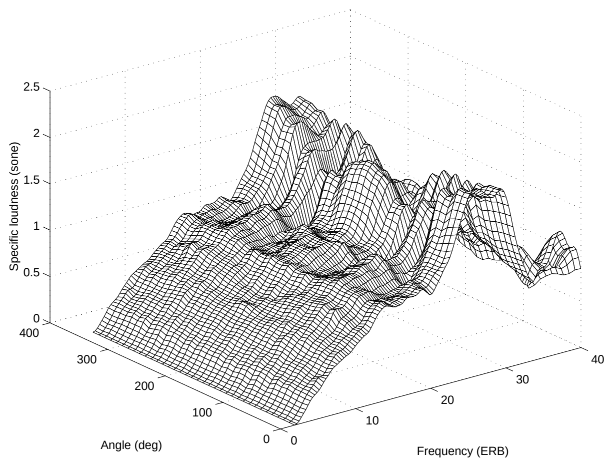


Figure 14 Measured binaural specific loudness spectrum (left ear) of white noise as a function of azimuth for the B&K 4128 HATS

These tools have been employed for the loudness alignment of multiple sound reproduction systems with different number of sources in a standard listening room. Mono, stereo and multichannel systems were loudness aligned with this tool and the informal subjective alignment was found to be very satisfactory.

## 5. SUMMARY AND CONCLUSIONS

A real-time binaural loudness meter has been efficiently implemented for use in the Windows PC environment. The system employs the systems native soundcard for use with measurement microphones and allows for calibration of all gains within the input chain. The performance has been shown to be in line with the original loudness model, though with real-time performance. Directional loudness properties have also been shown to compare favorably with the simulated data.

## 6. FUTURE WORK

Future work with the model includes creating a standalone version of the tool, that could be used without the Matlab environment, greatly lowering the price of the tool. As the Zwicker model is still in widespread use in the audio engineering community, it will be implemented into the model as an alternative.

If the tool is to be used with transient signals, triggering will have to be implemented. Furthermore, a better way of approximating temporal loudness integration is required, as proposed in [13].

## 7. ACKNOWLEDGEMENTS

Prof. Brian Moore is thanked for his comments and assistance in the development of this tool. Pekka Suokuisma and Miikka Vilermo are thanked for their assistance with implementations. Th second author expresses thanks to the Nokia Foundation for funding support during finalising this study.

## 8. REFERENCES

[1] Aarts R. M., Calculation of the Loudness of Loudspeakers during Listening Tests, Journal of the Audio Engineering Society, Vol. 39, pp.27-38, January/February 1991.

[2] Aarts R. M., Comparison of Some Loudness Measures for Loudspeaker Listening Tests, Journal of the Audio Engineering Society, Vol. 40, pp.142-146, March 1992.

[3] Aarts R. M., On the design and psychophysical assessment of loudspeaker system, Proefschrift Technische Universiteit Delft, 1995.

[4] Bech, S., Calibration of relative level differences of a domestic multichannel sound reproduction system, J. Audio Eng. Soc., vol. 46, pp 304 – 313, April 1998.

[5] Paulus E., Zwicker E., Programme zur automatischen Bestimmung der Lautheit aus Terzpegeln oder Frequenzgruppenpegeln, Acustica, Vol. 27. pp. 253-266, 1972.

[6] Zwicker E., Fastl H., Psychoacoustics, Facts and Models, Springer-Verlag, 1990.

[7] ISO Rec. R. 532, Method for calculating Loudness Level, Method B, International Organization for Standardization, Geneva, Switzerland, 1966.

[8] Moore B. C. J., Glasberg B. R., A revision of Zwicker’s Loudness Model, Acustica, Vol.82, pp. 335-345, 1996.

[9] Moore B. C. J., Glasberg B. R., and Baer T., A model for the prediction of thresholds, loudness, and partial loudness, J. Audio Eng. Soc., vol. 45, pp 224 – 239, April 1997.

[10] N. Zacharov, An overview of multichannel level alignment, in Proc. of the AES 15th Int. Conf., pp. 174–186, Audio Eng. Soc., 1998.

[11] Robinson, D. W., Whittle, L. S., The loudness of directional sound fields, Acustica, vol 10, (1960), pp 74 – 80

[12] Sørensen M. F., Lydolf M., Frandsen P. C., Møller H., Directional dependence of loudness cues and binaural summation, proceedings of the 15th international congress on acoustics, Trondheim, Norway, pp. 293-296, June 1995.

[13] Stone, M. A., Moore, B. C. J., and Glasberg, B. R., A real-time DSP based loudness meter, Seventh Oldenburg symposium on psychological acoustics, Germany, pp. 587-601, 1997.