C code for Calculation of the Loudness of Time-Varying Sounds

Brian R. Glasberg (bg12@cam.ac.uk), Brian C.J. Moore (bcjm@cam.ac.uk) and Josef Schlittenlacher (js2251@cam.ac.uk)

*Department of Experimental Psychology, University of Cambridge,
Downing Street, Cambridge CB2 3EB, England*

I. INTRODUCTION

This document describes use of the executable program “TVLBIN”, which calculates the loudness of monaurally and binaurally presented time-varying sounds (including sounds that differ at the two ears). The program is based on the model of Glasberg and Moore (2002) but is modified by incorporating the middle-ear transfer function specified by Glasberg and Moore (2006), the concept of “binaural inhibition” as described in Moore and Glasberg (2007), and Moore (2014), and the signal-processing flow described in Moore et al. (2016). The software implements the model described in Moore et al. (2016). The source code is copyrighted by Brian C. J. Moore, Brian R. Glasberg and Josef Schlittenlacher. The code is provided free for any research purposes. For commercial use please contact Brian C. J. Moore.

The program gives essentially the same results as the procedure described in ANSI S3.4-2007 (ANSI, 2007) and the procedure described in ISO 532-2 (2016) for steady sounds that are presented diotically (same sound at the two ears). However, the results differ from those given by the ANSI method for monaural presentation. TVLBIN allows calculation of loudness for sounds that differ at the two ears, in level, spectral content, or time pattern, for example sounds recorded using a dummy head.

If you received the program as part of a zip file, especially if sent via email, the executable program may have had its name changed from “TVLBIN.exe” to “TVLBIN.xxx”. Before you try to run the program you need to rename it as “TVLBIN.exe”.

The program should run on an IBM PC or compatible, using a “command prompt” or under “DOS”. The program uses a stereo “wav” file as its input. The two channels are referred to as “channel 1” and “channel 2”. The sample rate must be 32000 samples per second and the resolution must be 16 bit. For free-field frontal incidence or diffuse-field listening conditions, the signal may be picked up by a single microphone at the position corresponding to the centre of the listener’s head, giving a single-channel wav file. In this case a two-channel (stereo) wav file must be created, with the same signal in each channel. This can be done using software such as “Cooledit” or “Audacity”. These software packages can also be used to convert the sample rate to 32000 Hz if the original sample rate was different from that.

The first stage in the program is filtering to simulate the effect of the outer and middle ear. This can result in an increase in level, especially when the input signal has strong components close to 3 kHz. Hence, **it is recommended that the maximum level of the signal in the input file is 8 dB or more below the full-scale level**, in order to avoid “overload”.

II. RUNNING THE PROGRAM

To run the program, perform the following steps.

1) Open a “DOS” (command prompt) window.

2) Change the directory to the directory where the TVLBIN program is located. The input wav file should be in that same directory.

3) Type the name of the program (TVLBIN) followed by any command-line options (see below for details), followed by the name of the input file preceded by “–i”, followed by > outputfile.

For example, type

TVLBIN –i 1k100ms.wav –c (100,100) > 1k100ms.out

The program takes as input the file “1k100ms.wav” and saves the output of the loudness calculation in the file “1k100ms.out”. The command line option “–c (100,100)” indicates that the level of a full-scale sinusoidal signal is 100 dB for both channel 1 and channel 2. This allows calibration of the absolute level of the signal in the wav file. The level of a full-scale sinusoidal signal can be set to any desired value and can differ for channel 1 and channel 2.

By default, diffuse-field presentation is assumed. If it is desired to specify free-field presentation with frontal incidence, use the command-line option “–F (ff.32k, ff.32k); this indicates that presentation was free field frontal incidence for both channels. For example type:

TVLBIN –i 1k100ms.wav –F (ff.32k, ff.32k) –c (100,100) > 1k100ms.out

If the waveform is picked up via two microphones, one close to the right eardum and one close to the left, or via the two microphones in a dummy head, then it is appropriate to specify a “middle-ear only” transfer function: –F (midear.32k,midear.32k). For example type:

TVLBIN –i 1k100ms.wav –F (midear.32k, midear.32k) –c (100,100) > 1k100ms.out

Other command-line options are as follows:

–x This specifies a level, relative to full-scale level, below which spectral components are discarded before calculation of excitation patterns. This can speed up the computations. The default value is 60 dB. To change the default value to, for example, 80 dB below full scale, type “–x 80”.

–b The beginning time in the file in ms from which the loudness is to be calculated. For example, type “–b 20” to start calculating loudness at 20 ms into the file. Parts of the waveform from 0 to 20 ms are effectively ignored.

–e The time in the file at which to stop the loudness calculation. For example, type “–e 200” to tell the program to ignore the waveform in the file after 200 ms from the start of the file.

–v This turns on verbose mode, which leads to more information in the output file (and a larger file size).

–I Gives information about how to use the program, including command-line options. To use this, type “TVLBIN –I”.

III. THE OUTPUT FILE

The description given below applies when the verbose mode is not activated.

The first part of the output file gives some general information, including the input file name, the initial filter that was used for each channel (diffuse field, free-field or middle-ear only, to allow for different listening conditions), statistical information about the signal in each channel, and the calibration levels (levels of full-scale signals) for each channel. The “P56LevelMeasure” is based on recommendation ITU-T P.56 (2011) of the International Telecommunications Union (ITU): “Objective measurement of active speech level” and it may be useful when the signal is speech. Here is a quote from the ITU recommendation: “Active speech level is measured by integrating a quantity proportional to instantaneous power over the aggregate of time during which the speech in question is present (called the active time), and then expressing the quotient, proportional to total energy divided by active time, in decibels relative to the appropriate reference.”

Following that, the results of the loudness computations are presented, updated every 1 ms.

The first column shows the time in ms.
The second column shows the instantaneous loudness in sones.
The third column shows the instantaneous loudness level in phons.
The fourth column shows the short-term loudness (average 1) in sones.
The fifth column shows the short-term loudness level (average 1) in phons.
The sixth column shows the long-term loudness (average 2) in sones.
The seventh column shows the long-term loudness level (average 2) in phons.

The instantaneous loudness is assumed to be an intervening variable, not available for conscious perception.
The short-term loudness corresponds to the loudness of a brief segment of a sound, e.g., a specific syllable in speech or note in music.
The long-term loudness corresponds to the overall loudness impression of a relatively long piece of sound, e.g., a sentence or a musical phrase.

At the end of the file the maximum values of the short-term loudness (average 1) and long-term loudness (average 2) are given both as loudness in sones and as loudness level in phons.

IV. ADDITIONAL OUTPUT

In addition to generating a plain text file with the results of the loudness computations, the program generates some wav files. The program starts by dividing the two-channel (stereo) input wav file into two mono files, which are saved as “mono_left.wave” and “mono_right.wav”. The program then filters each mono wav file with the filter selected to account for transmission through the outer and middle ear (df.32k, ff.32k, or midear.32k). The wav files resulting from this filtering are stored as “mono_left_fir.wav” and “mono_right_fir.wav”. As noted earlier, the level may be higher after filtering than before, which is why it is recommended to have 8 dB of “headroom” in the original input file. If overload (clipping) occurs, i.e. the full-scale level is exceeded after filtering, the file containing the results of the loudness computations will give a message about “overflows” and “underflows” and will indicate how often they occurred. The files “mon_left_fir.wav” and “mono_right_fir.wav” can be inspected using software such as “Cooledit” or “Audacity” to check where overload (clipping) occurred.

Note that the wav files generated by the program are over-written each time the program is re-run, so if it is desired to retain any of the files they should be re-named.

REFERENCES

ANSI (**2007**). *ANSI S3.4-2007. Procedure for the computation of loudness of steady sounds* (American National Standards Institute, New York).

Glasberg, B. R., and Moore, B. C. J. (**2002**). "A model of loudness applicable to time-varying sounds," J. Audio Eng. Soc. **50**, 331-342.

Glasberg, B. R., and Moore, B. C. J. (**2006**). "Prediction of absolute thresholds and equal-loudness contours using a modified loudness model," J. Acoust. Soc. Am. **120**, 585-588.

ISO 532-2 (**2016**). *Acoustics - Methods for calculating loudness - Part 2: Moore-Glasberg method* (International Organization for Standardization, Geneva).

Moore, B. C. J. (**2014**). "Development and current status of the "Cambridge" loudness models," Trends Hear. **18**, 1-29.

Moore, B. C. J., and Glasberg, B. R. (**2007**). "Modeling binaural loudness," J. Acoust. Soc. Am. **121**, 1604-1612.

Moore, B. C. J., Glasberg, B. R., Varathanathan, A., and Schlittenlacher, J. (**2016**). "A loudness model for time-varying sounds incorporating binaural inhibition," Trends Hear. (submitted).