Acoustical Impulse Response Measurement with ALIKI

Fons ADRIAENSEN
fons.adriaensen@skynet.be

Abstract

The Impulse Response of an acoustical space can be used for emulation of that space using a convolution reverb, for room correction, or to obtain a number of measures representative of the room's acoustical qualities. Provided the user has access to the required transducers, an IR measurement can be performed using a standard PC equipped with a good quality audio interface. This paper introduces a Linux application designed for this task. The theoretical background of the method used is discussed, along with a short introduction to the estimated measures. A short presentation of the program's features is also included.

Keywords

Acoustics, impulse response, convolution, reverb.

1 Introduction

Equipment for measuring the acoustical parameters of an environment has traditionally been the realm of a small group of highly specialised electronics manufacturers. During the last decade, the ready availability of mobile computers and of high quality portable audio interfaces has resulted in a move towards mainly software based solutions.

Early programs just emulated the standardised (hardware based) procedures to measure e.g. reverb time. The computing power being available today enables the use of other methods, such as a direct measurement of a room's impulse response, from which all interesting information can be derived. Several methods to capture impulse responses have been developed, and these will be discussed below.

It seems that very little free and Linux-based software is available for this task. The Digital Room Correction package from Denis Sbragion [1] includes some scripts to perform an impulse response measurement. It is possible to obtain very good results with these scripts, but they are not easy to use.

In the Windows based world, a number of solutions have been available for some time. Most of these are based on the use of pseudo-random sequences. The MLSSA system from DRA Laboratories [2] (requiring special hardware) was one of the first using this method, and is well known. As an example of a package using more advanced methods, the Aurora Plugins [3] from Angelo Farina should be mentioned.

This paper introduces a new Linux based integrated system [4] developed by the author, and available under the terms of the GPL. ALIKI will capture impulse responses in up to eight channels simultaneously. The recorded data can be used directly for e.g. digital room correction, edited and prepared for use in a convolution based reverb, or used to compute acoustical parameters such as reverb time and various energy ratios.

The following sections will describe the measurement and analysis methods used in this software.

2 IR measurement methods

The impulse response (IR) of a system is the output signal it produces for an input consisting of a single Dirac pulse. The mathematical definition of a Dirac pulse requires zero width and unit energy, which is not possible in the real world, so in practice finite-width impulses compatible with the required bandwidth are used. In a sampled system in particular, the Dirac impulse is a signal consisting of one sample of unit amplitude followed by all zeros. It contains all frequencies from zero to the Nyquist limit with equal energy and a known phase.

Provided the system is linear and time-invariant, the IR contains all information there is about its behaviour, and permits the calculation of the system's response to any input signal.

[1] http://drc-fir.sourceforge.net

[2] http://www.mlssa.com

[3] http://farina.eng.unipr.it/aurora/home.htm

[4] http://users.skynet.be/solaris/linuxaudio/aliki.html

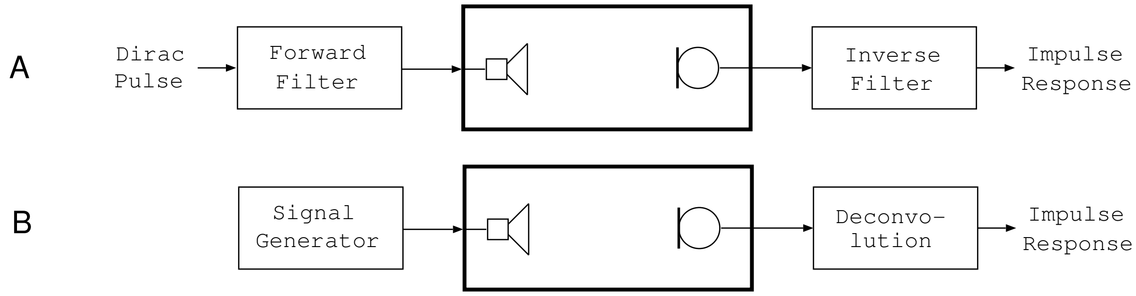

Figure 1: IR measurement using filtered Dirac pulse. A: theoretical model, B: practical realization.

The main problem with using Dirac pulses in an acoustical measurement is that as a result of their very short duration and finite amplitude, they contain very little energy, and measurement accuracy will be limited by the signal to noise ratio of the equipment used and of the system itself. While it is possible to use Dirac pulses reproduced by a loudspeaker in the controlled environment of an acoustics laboratory, this is all but infeasible in most real life situations, e.g. for measuring a room or concert hall, where there will always be background noises of some sort.

There are basically two ways to overcome this difficulty: either generate a high energy impulse directly as a sound, or find some method to spread the test signal over a longer time and to undo this operation after the measurement.

For the first approach, various methods have been used by acoustics engineers, ranging from exploding balloons and starter's pistols to very expensive special equipment to generate short high amplitude sound pulses. While such methods can be used e.g. to measure the reverb time of a concert hall, they still require a very large dynamic range in the measurement system, and they are not accurate and repeatable enough to obtain an IR to be used for room correction or for a convolution reverb.

The second solution is based on the following idea. Suppose we have a filter $\mathbf{H}$ with complex frequency response $H(\omega)$. If the filter has a non-zero gain at all frequencies, we can find an inverse filter $\mathbf{R}$ with frequency response $R(\omega) = z^{-n}/H(\omega)$. The z^{-n} is a pure delay required to make such a filter causal and physically realizable. Putting the two filters in series, only the (known) delay remains. Since the filters are linear, and if we assume the same of the system to be measured, we can put the system in between the two filters and obtain its impulse response using the filtered signal instead of the Dirac pulse (fig.1A).

Since we can regard any signal as the output of an FIR filter having the signal's sample values as its coefficients, we could in theory use any signal we want as long as the inverse filter exists and we can find some way to compute and implement it. For some classes of signals this can be done relatively easily, and that is the basis of the two methods discussed in the next sections. In practice the theoretical model of fig.1A is realized by generating the signal directly instead of filtering a Dirac pulse, and the inverse filtering is usually done by (de)convolution rather than a by real filter (fig.1B).

2.1 Maximum length binary sequences

Pseudo random binary sequences can be generated by a shift register with exclusive-or feedback from selected taps. Provided the correct feedback terms are used, a generator using N stages will produce a maximum length sequence (MLS) of length $L = 2^N - 1$. A sampled audio signal derived from such a sequence has exactly the same power spectrum as a Dirac pulse repeated every L samples, but it has L times more power for the same amplitude. For example, using an $L = 1023$ sequence will improve the signal to noise ratio by 30 dB.

The inverse filtering for such a signal can be done efficiently by using the Hadamard transform. Like for the Fourier transform, a 'fast' version of this transform exists (and it is even simpler than the FFT). This is the way the MLSSA software mentioned before (and many other systems) operate.

For acoustical measurements, the signal can be filtered further to obtain a 'pink' spectrum instead of 'white' noise, again improving the S/N ratio at low frequencies where it is usually the most problematic.

A more elaborate discussion of MLS based techniques can be found in the references (Vanderkooy, 1994).

The main difficulty with the MLS method is its sensitivity to non-linear behaviour. Most loudspeakers produce substantial amounts of distortion, and this will interfere with the measurement and show up as spurious signals in the impulse response. The method discussed in the next section, while more complex to implement, does not have this problem.

2.2 Swept sine techniques

A second class of signals for which the inverse filter can be computed easily are linear and logarithmic frequency sweeps. For a linear sweep, the inverse is just the time reversal of the original signal. Such a signal has a 'white' spectrum, and for acoustical measurements a logarithmic sweep, having a 'pink' power spectrum is often preferred. In that case, provided the sweep is not too fast, the inverse filter is again the time-reversed original, but modified by a +6 dB per octave gain factor (6 dB, and not 3, since a +3 dB per octave correction applied to each filter separately would make both of them, and their product, 'white'). In both cases the inverse filter can be realized efficiently by using FFT-based convolution.

The advantage of using a sweep is that at any time we produce only a single frequency, and any distortion introduced will consist of the harmonics of that frequency only. If we use a *rising* frequency sweep, the harmonics will be generated ahead of the same frequencies appearing in the signal. So after deconvolution, any distortion will appear as spurious peaks in *negative* time in the impulse response, and most of it can then be edited out easily.

Another interesting feature of this method is that it does not depend on exact synchronisation of the playback and capture sample clocks. Any frequency error between these will result in a 'smearing' of the impulse response, in the sense that a Dirac pulse becomes itself a very short sweep. It is even possible to correct for this after the deconvolution.

The sweep method was pioneered by Angelo Farina (Farina, 2000), and it is the one used in ALIKI.

3 Measurements derived from the impulse response

This section provides a quick overview of some acoustical measures that can be calculated from a captured impulse response.

If IR measurements are performed for use in a convolution reverb system, then the choice of the transducers used is largely a matter of common sense combined with aesthetic preferences. The same is to some extent true if the object is room correction.

In contrast, in order to derive the measures described below, the IR measurement must be done according to a standardised procedure, and by using the correct equipment. In practice this means the use of true omnidirectional speakers (purpose built), and in some cases of a microphone calibrated for diffuse-field measurements (i.e. having a flat response integrated over all directions rather than on-axis).

The two ISO documents mentioned in the References section provide a good introduction to what is involved in such measurements.

All these values can be calculated for the full frequency range signal, for an A-weighted version, or octave or sub-octave bands.

3.1 The Schroeder integral

Most of the values described in the following sections can be obtained by computing the *Schroeder integral* of the IR, defined as follows. Let $p(t)$ be the impulse response, with $t = 0$ corresponding to the arrival of the direct sound. Then the Schroeder integral of $p(t)$ is the function

$$S(t) = \int_t^\infty p^2(t)dt \tag{1}$$

In other words, $S(t)$ corresponds to the energy still remaining in the IR at time t. When plotted in dB relative to the maximum value at $t = 0$, $S(t)$ will be same as the level decay curve obtained after switching off a steady signal.

3.2 Reverb Time and Early Decay Time

The conventional definition of the *Reverb Time* is the time required for the sound level to decay to -60 dB relative to the original level, after a steady signal (normally white or filtered noise) is switched off. This time is normally denoted RT_{60} if the S/N ratio permits a reliable measurement down to that level, or RT_{30} if it is extrapolated from the -30 dB time.

For a measurement derived from an IR, it can be read directly from the Schroeder integral. The ISO standard prescribes that RT_{30} should be derived from the times the integral reaches respectively -5 dB and -35 dB, obtained by least-squares fitting, and extrapolated to the

60 dB range. The RT_{20} value is computed form the -5 dB and -25 dB times in the same way.

The *Early Decay Time EDT* is similar but derived from the -10 dB point of the integral, again by least-squares fitting.

3.3 Clarity, Definition and Central Time

The *Clarity* measure describes the ratio of the energies (in dB) before and after a given time referred to the arrival of the direct sound. This value provides a good indication of how 'clear' or 'transparent' the sound heard by a listener at the measurement position is. For speech, it is measured at 50 ms, while for music 80 ms is used. The two values are denoted C_{50} and C_{80}. The definition of C_{50} is

$$C_{50} = 10log_{10}\frac{\int_0^{0.050} p^2(t)dt}{\int_{0.050}^{\infty} p^2(t)dt} \quad (2)$$

$$= 10log_{10}\frac{S(0) - S(0.050)}{S(0.050)} \quad (3)$$

and similar for the 80 ms value.

The *Definition* is similar to Clarity, but is the simple ratio (not in dB) of the early sound energy to the *total* energy. In practice it's not necessary to compute both C and D, as they can easily be derived from each other.

The *Central Time* is the 'centre of gravity' of the energy in the IR, defined as

$$T_S = \frac{\int_0^{\infty} tp^2(t)dt}{\int_0^{\infty} p^2(t)dt} \quad (4)$$

3.4 Early Lateral Energy

To compute the values introduced in this section, two simultaneously captured IRs are required, one using an omnidirectional free-field microphone, and the second using a figure-of-eight (velocity) microphone pointing sideways (i.e. with the null pointing at the 'centre of stage'). These values provide some measure of the 'width' and 'spaciousness' of the sound. Extreme care and high quality equipment is required in order to obtain meaningful results from these computations.

Let $p_L(t)$ be the lateral IR, then

$$LF = \frac{\int_{0.005}^{0.080} p_L^2(t)dt}{\int_0^{0.080} p^2(t)dt} \quad (5)$$

and

$$LFC = \frac{\int_{0.005}^{0.080} |p_L(t)p(t)|dt}{\int_0^{0.080} p^2(t)dt} \quad (6)$$

LFC is said to correspond closer to subjective observation.

4 ALIKI program structure

ALIKI is written as an integrated package controlled by a graphical user interface[5]. Technically speaking it consists of two separate executables (the audio interface part is an independent process), but this is hidden from the user who just sees a single interface. Figure 2 shows the main modules and files used.

ALIKI can interface via JACK, or use ALSA devices directly, or it can be run without any audio hardware for off-line processing of stored data. It uses its own sound file format, but facilities to import or export WAV-format or raw sample files are included. The special format keeps all data for multichannel impulses conveniently together, facilitating handling (in particular when you have many files, all with similar and confusing names). It also allows to include specific metadata, for example parameters to be used by a convolution reverb. The metadata will become even more important when the program is extended to allow automated multiple measurements, e.g. to obtain polar diagrams.

4.1 The capture module

Functions of this module include

- input measurement parameters (frequency range, sweep time, channel names, etc.),
- generate and store the sweep and inverse filter waveforms,
- provide test signals and metering,
- perform the actual measurements and store the results.

ALIKI will handle up to 8 audio inputs. So it is possible to record e.g. an Ambisonics B-format, a stereo pair and a binaural format in one operation. Capturing the IRs can be also be done without using ALIKI, e.g. by using Ardour to play the sweep file and record the microphone signals (ALIKI will read Ardour's Broadcast Wave files).

4.2 The deconvolution module

This module reads the recorded waveforms and the inverse filter file, and calculates the actual impulse responses. It uses a fast FFT-based convolution algorithm. An optional correction

[5] At the time of writing, the GUI is still in full development, therefore no screenshots are yet available.

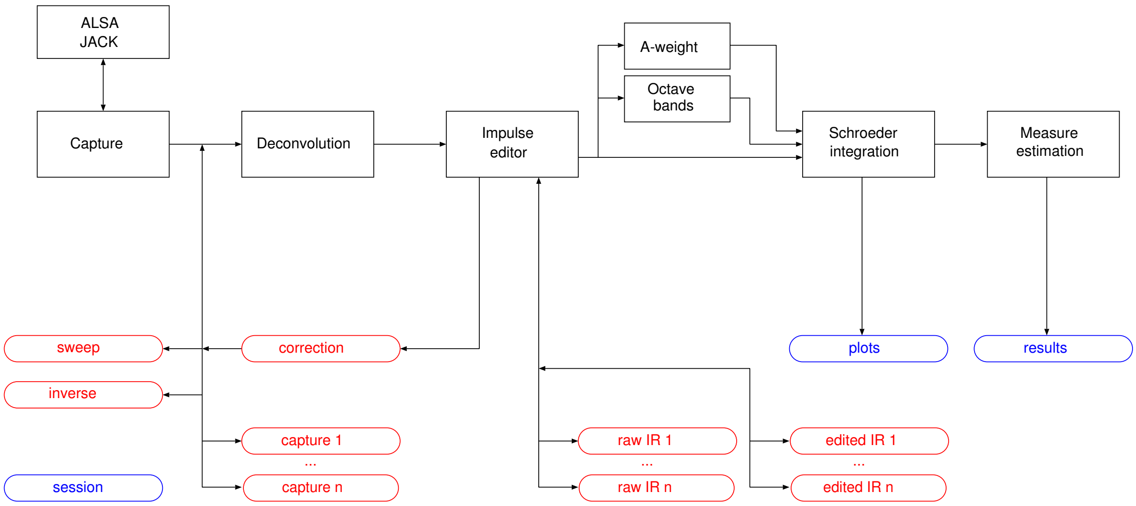

Figure 2: ALIKI program structure and files

filter compensating for the response of the loudspeaker and/or microphone can be used.

The results are saved to the raw impulse response files, and transferred to the editor module.

4.3 The editor module

This provides visualisation and basic editing of impulse responses. It is used to

- normalise the IR to a standard level,
- calibrate the time axis (i.e. put the direct sound at $t = 0$),
- trim the end of the IR to remove noise,
- remove the direct sound if required.

The editor will operate on groups of impulse responses (e.g. a stereo pair or B-format) preserving relative timing and levels. Edited IRs can be saved for later use. Unless the user really wants it, the editor will never overwrite the original waveforms.

This module has one additional function: the first few milliseconds of an IR can be used to compute an inverse FIR filter (up to 4096 taps) that will be used by the deconvolution engine to compensate for the response of the speaker and microphone. It uses a simple FFT-based inversion method, and an interactive procedure steered by the user in order to avoid major errors that could result from a simple automated calculation.

4.4 The filter, integration and measure modules

The remaining modules in fig.2 are closely integrated from the user's point of view.

The Schroeder integral can be computed and visualised for the filtered or full-range IRs. A-weighing and octave band filtering is again performed using FFT-base convolution, and is combined with the backwards integration. The integrals are stored with a resolution of about 1 millisecond. They can be exported in a format readable by applications such as Gnuplot, which can convert them to a number of graphical formats.

The first release of ALIKI computes EDT, RT_{20}, RT_{30}, RT_{user}, T_S, $C_{50,80,user}$, $D_{50,80,user}$, LF and LFC. Others (such as IACC) maybe added in future versions.

All measured values can be exported as text, CSV, or Latex table format for use in spreadsheets or reports.

5 Acknowledgements

The IR measurement method used by ALIKI is based on the work of Prof. Angelo Farina (Dipartimento Ingegneria Industriale, University of Parma, Italy). His many papers (freely available via his website [6]) are required reading for anyone exploring the subject of acoustical IR measurement, in particular in the context of surround sound.

Also the work of Anders Torger (author of BruteFir [7]) and Denis Sbragion (author of

[6] http://pcfarina.eng.unipr.it
[7] http://www.ludd.luth.se/ torger/brutefir.html

DRC) has been very inspiring.

References

Angelo Farina. 2000. Simultanuous measurement of impulse response and distortion with a swept-sine technique. *Audio Engineering Society Preprint 5093*.

ISO TC43/SC2. 2003. Acoustics — Measurement of the reverberation time — Part 1: Performance spaces. ISO CD 3382-1.

ISO TC43/SC2. 2004. Acoustics — Application of new measurement methods in building acoustics. ISO WD 18233.

John Vanderkooy. 1994. Aspects of MLS measuring systems. *Journal of the Audio Engineering Society*, 42(4):219–231.