On room impulse response measurements with sine sweeps

Carmine-Emanuele Cella

February 25, 2017

Abstract

This short note will describe a common approach to compute the impulse response of a room using sine sweeps, based on Wiener deconvolution.

1 Introduction

Under specific assumptions, a room can be considered as a linear and time invariant system (LTI) characterized by an impulse response $h(t)$. The accurate measurement of the impulse response can be useful to derive many acoustical parameters of the room. Unfortunately, such measurement is difficult to obtain using directly an impulse as excitation signal and it is often easier to use sine sweeps. A sine sweep is a better realizable signal compared to an ideal impulse (which is only approximated up to the physical limitations of the system) and has an excellent signal-to-noise ratio. The usual method is as follows:

- a sine sweep is played in a room, using a sufficient energy;
- at the same time of the playing, the recording of the room is effectuated;
- finally, a deconvolution of the original sweep from the recorded sound is performed, in order to estimate the real impulse response.

The following section will describe a common method to perform the deconvolution step, pointing out inherent problems.

2 Wiener deconvolution

A convolution in time domain can be tought as a complex multiplication in frequency domain:

$$y(t) = x(t) * h(t) = X(k)H(k) \tag{1}$$

where X and H are the Fourier transforms of the respective time signals x and h. As such, knowing the input x and the output y of a system, it is possible to calculate its transfer function h performing a complex division in frequency domain:

$$H(k) = \frac{Y(k)}{X(k)} \tag{2}$$

and then computing the inverse Fourier transform of H.

While this method is mathematically sound, it suffers from extreme sensitivity to noise and it is impractical to be used in reality. A better way to retrive h is given by a process called *Wiener deconvolution,* that consists in regularizing the complex division by some quantities depending on the known signals:

$$W(k) = \frac{Y(k)\overline{X(k)}}{|X(k)|^2 + \sigma^2} \tag{3}$$

where $\overline{X}$ denotes the complex conjuage and $\sigma = \lambda \max(|X(k)|)$ (λ is a scaling parameter). Here h will be the inverse Fourier transform of W.

In the specific case of estimating the impulse response of a room with sine sweeps, Y will be the recording of the room, X will be the excitation sweep and W the estimated impulse response in frequency domain. The effect of the parameter λ is as follows: if it is too small, the regularization will be not enough and the impulse response will be noisy; on the other hand, if λ is too big the deconvolution will not be effective and the impulse response will still contain components of Y. The following code shows an implementation of the discussed method in Matlab:

```
y = audioread ('room_dump.wav');
x = audioread ('sweep.aif');
Y = fft (y);
X = fft(x, size(Y, 1));
lambda = 0.005;
sigma = max(abs(X)) * lambda;
W = (Y.* conj(X)./(abs (X).^2 + sigma^2));
h = real (ifft (W));
sig_len = (size(y, 1) - size (x, 1));
h = h(1:sig_len);
audiowrite ('ir_rebuild.wav', h, 44100);
```

3 Known problems

The described method is only efficient in specific cases, among which the one discussed in this note. A simple way to improve this method would be to make σ frequency dependent. For more general approaches for deconvolution, other methods can be used such as least squares and independent component analysis.