

ENVIRONMENT SIMULATOR FOR AUDIO SIGNALS

Antonio Satué-Villar¹, Juan Fernández-Rubio²

¹Escuela Universitaria Politecnica de Mataro (EUPMT), Universidad Politécnica de Cataluña, Av. Puig i Cadafalch 101-111, 08303 Mataró, Spain, phone +34937574404, fax +34937570524, e-mail satue@eupmt.es

²Escuela Técnica Superior de Ingenieros de Telecomunicación de Barcelona (ETSETB), Universidad Politécnica de Cataluña, C. Jordi Girona 1-3, 08034 Barcelona, Spain, phone +34934016431, fax +34934016447, e-mail juan@gps.tsc.upc.es

ABSTRACT

This paper describes an application developed in Matlab to achieve, from different signals, the signal reaching the different microphones of a linear cluster with equally-spaced elements. These signals can be used to perform simulations and listen to them to see the room effects.

The application permits to obtain the signals in an anechoical environment (open space) and in a rectangular room, using the image method. In this last case, in addition to the cluster signals, the impulse response between source and receiver (selectable) can be viewed. Response viewing and simultaneous hearing of signals enable students of introductory modules in signal processing to understand the concept of impulse response and its meaning in room acoustics.

1. INTRODUCTION

In introductory modules of signal processing in Engineering studies, one of the first concepts to be assimilated by students is "impulse response". Most of text books and related literature ([1],[2]) dedicate quite many pages to this topic. This concept is basic, and the student does not always understand it clearly. Because of that the student often knows how to carry out the pertinent mathematical operations but he/she does not interpret the results correctly. This paper presents a simulator that intends to improve the understanding of what impulse response means. Therefore we decided to work with voice signals, since they are easy to obtain and besides they can be reproduced in almost all current computers. To simulate the reverberations in the room (with the restriction of this being a parallelepiped), image method is programmed [3].

2. IMAGE METHOD

Image method is a theory that allows us to explain the behavior of the acoustic field in a room. A sound wave can be altered in its way by reflection, refraction or diffraction

phenomena. In our case, we will only take into account reflections, as the refraction phenomena are produced when there is a change of means of propagation and those of diffraction when there are obstacles of size comparable to wave length.

Image method consists of finding some "virtual image sources", which come from reflecting the signal source in the different walls of the room. This method can be iterated, so that each of the walls generates a first order image of the original source; each of the walls will also generate an image of each of these first order images (second order images), etc... Summarizing it, n-order images are obtained from those of n-1 order and the virtual walls of n-1 order. The coordinates of these images are obtained through geometrical calculations. The method will not take into account all these sources, as some of them can then be really meaningless and for high orders the necessary order increase does not compensate the images contribution to the result.

The overall effect is that between two points in a room there is an impulse response $h(t)$ that depends on the position of the points and the geometry of the room. Thus, in an anechoic room, this $h(t)$ has only one delta due to the space that runs the sound wave from the transmitter (mouth) up to the receiver (microphone). However, if the room is very reverberative, these $h(t)$ answers can be even of several seconds, with the consequent processing delays. These differences are further noticed with its hearing.

Rabinkin proved that it is possible to truncate the number of reflections to deal with, thus reducing the computational time. Truncating idea is very intuitive, for normally the strongest reflections usually happen soon, due to wave attenuation with distance. The conclusions obtained were that we can truncate soon (minor calculation complexity) if the array has many sensors, the room is small or the absorption factor of walls is high [4].

We observe that if the working space were always the same one (the same room and the same position of speaker and microphone), at every microphone output we could apply an inverse filter in such a way that the convolution of the $h(t)$ with this inverse filter $h^{-1}(t)$ would generate the delta of the direct signal. In general, it is not like this.

Image method has been chosen because our interest is to obtain a transference function between two points in the room. To calculate it with more exact models it would be required to calculate all the modes in the margin of interest frequencies (typically, between 100 Hz and 4 KHz). However, image method only includes those images that contribute to the impulse response. What matters is to know that every image contributes with an impulse of known level and delay, whereas in the real case it is a decreasing exponential one.

This model considers a speaker in a room as a point in a rectangular cavity. A single point in the open space emits a pressure wave of the form

$$P(\omega, X, X') = \frac{e^{i\omega(R/c-t)}}{4\pi R}$$

where P = pressure, $\omega = 2\pi f$, f = frequency t = time, R = |X-X'|, X = emitter position vector (x,y,z), X' = receiver position vector (x',y',z') y c = sound speed.

In a rigid wall (zero normal speed), bound condition is satisfied by placing a symmetrical image in the far side of the wall. So,

$$P(\omega, X, X') = \left[\frac{e^{i(\omega/c)R_+}}{4\pi R_+} + \frac{e^{i(\omega/c)R_-}}{4\pi R_-} \right] e^{-i\omega t}$$

where two distances are defined from the microphone to the origin (R₋) and to the image (R₊):

$$R_-^2 = (x - x')^2 + (y - y')^2 + (z - z')^2$$

$$R_+^2 = (x + x')^2 + (y - y')^2 + (z - z')^2$$

The wall is placed in x=0.

In a general case of six walls, the situation is more complicated because every image has other images. Pressure can be written as

$$P(\omega, X, X') = \sum_{p=1}^8 \sum_{r=-\infty}^{\infty} \frac{e^{i(\omega/c)|R_p+R_r|}}{4\pi|R_p+R_r|} e^{-i\omega t} \quad (1)$$

where R_p represents the eight vectors of the 8 permutations over ± of:

$$R_p = (x \pm x'), (y \pm y'), (z \pm z')$$

r is a three-integer vector (n,l,m), and:

$$R_r = 2(nL_x, lL_y, mL_z)$$

where (L_x, L_y, L_z) are the room dimensions. (1) equation is the frequency response of the pressure assuming rigid walls for an origin point in X=(x,y,z) and a receiver point in X' = (x',y',z'). If we apply the inverse Fourier transform at (1) equation we obtain the room impulse response equation.

$$p(t, X, X') = \sum_{p=1}^8 \sum_{r=-\infty}^{\infty} \frac{\delta\left[\left|R_p + R_r\right|/c\right]}{4\pi\left|R_p + R_r\right|} \quad (2)$$

If the walls of the room are not rigid, the solution as image points cannot be exact. A precise detail of the effects of some walls with finite impedance is impossible since even the effects on a simple image are quite complicated. Therefore, the image model of a point has been kept, besides meaning a wall reflection coefficient independent on the angle. In practice, in the range of 0.1-4 KHz, with reflection coefficients higher than 0.7, rectangular geometry and emitter-receiver not very near the wall, these assumptions are true.

Adding the reflection coefficient to (2) equation, this remains modified in the following way

$$p(t, X, X') = \sum_{p=0}^1 \sum_{r=-\infty}^{\infty} \beta_{x1}^{n-q} \beta_{x2}^{n-l} \beta_{y1}^{l-j} \beta_{y2}^{l-k} \beta_{z1}^{m-k} \beta_{z2}^{m-l} \times A$$

where

$$A = \frac{\delta\left[t - \left(\left|R_p + R_r\right|/c\right)\right]}{4\pi\left|R_p + R_r\right|}$$

and where R_p is expressed in function of p = (q,j,k) vector as:

$$R_p = (x - x' + 2qx', y - y' + 2jy', z - z' + 2kz')$$

Impulse response is constructed based on a histogram of received image pulses with different time delays. Bin amplitude of histogram equals sampling period.

3. THE APPLICATION

The application carried out allows, from different signals arriving from different directions, obtaining the signal arriving from different microphones of a linear cluster with equal-spaced elements [5]. These signals can be used to carry out further simulations or to listen to them to see the room effects. The application allows obtaining the signals in an anechoic environment (open space) and in a rectangular room, using the image method. In this last case, apart from having the signals in the grouping, we can view the impulse response between source and receiver (selectable)

3.1. Main data

'Main data' window (figure 1) is used to insert in the program the data common to the two different simulations: receiver number, noise variance, signal-noise ratios of signals, and the names of signal files.

3.2. Simulation selector

In this window there are two buttons only: "Free simulation" and "Simulation in room", which are for simulation in open space or in a room respectively. The first one is for the simulation in an imaginary room of infinite dimensions and the second one in a room where only the effect of ceiling, ground and walls is taken into account.

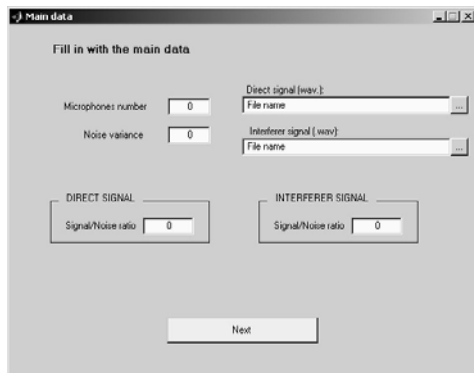


Figure 1: Main data

3.3. Angles

If the simulation in open space has been selected, a new window is open, which is used to select the incident angles of signals. These angles are selected through slider bars. Once the bar is moved there appears a number in red color indicating its value.

3.4. Results

This is the last window of the free simulation, which is clearly divided up into two parts: one to listen to the sound and another to save it (figure 2). The former allows listening to the direct source, the interfering source or the signal in a determined receiver, indicating the number of receiver that is wanted to listen to.

The latter allows carrying out two options. One is to save the sound of a specific receiver (writing the number), and afterwards selecting the name of the file. The other option saves the sounds of all the receivers, with the "Room" name followed by the number of receiver. With this last option the user does not have to save the names 3, 4, 5,... times (depending on how many receivers there are). In this last option the name cannot be chosen but we can choose the directory name.

3.5. Room dimensions

If the simulation in room has been chosen, the following window that is shown is that of figure 3 in which length, width and altitude in meters of the room are written. The absorption factors of the ceiling, ground, and walls of the room, which are indicated by a nearby letter, are also written on the right of the picture. The window shows a drawing with letters in the ceiling, ground and walls, so that they relate to the letters of reference used for the absorption factors.

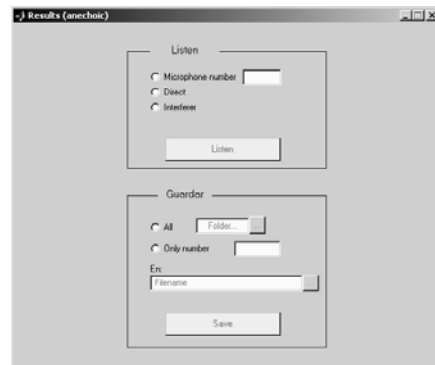


Figure 2: Results (open space)

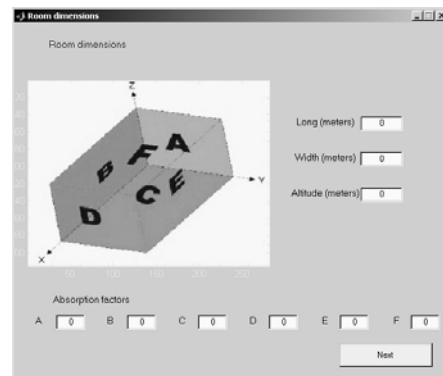


Figure 3: Room size

3.6. Positions in the room

This is the window (figure 4) where the positions of the sources and receivers (the reference receiver is introduced) are introduced and afterwards the distance between receivers is indicated (regarding to the reference position) in the room. For that purpose there are three cells with a letter each, and a drawing is shown to know which axis runs every letter and, this way, be able to see how the point remains. After this window, the final window is loaded, which takes several seconds as the processing of sound vector matrix is to be done. A progressive bar is used to see that the program is loading up.

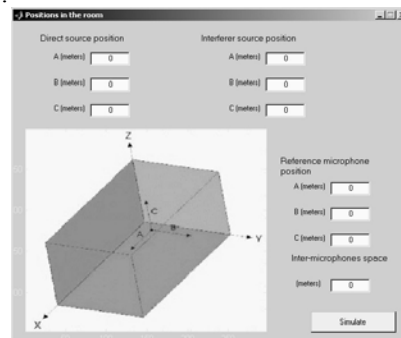


Figure 4: Sources/Receivers positions

3.7. Results in the room

In this window (figure 5), we can do the same as in the anechoic environment window. But, in addition, we are able to see the impulse response between a source and a determined receiver, indicating direct or interfering source and the number of receiver. Then, a new window appears with the graph of the impulse response

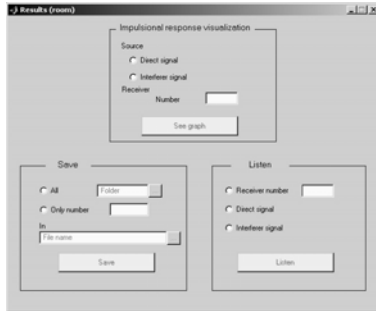


Figure 5: Results (room)

4. EXPERIMENTS

Different tests have been carried out to verify the correct functioning of the application. As an example, two impulse responses are shown in two extreme cases. They both have been carried out with 4 receivers, 40 dB of signal-noise ratio in the direct and interfering signal, and a noise variance of 0.1 in a room of typical dimensions (6 meters long, 4 width and 3 high), where the direct source is in the (1,1,1) coordinate, the interfering one in (3,1,1) and the receiver of reference in (4,2,1) with a distance of 0.05 meters (4 receivers) between them. In the first simulation (figure 6) the absorption factors are of 0.01 for the walls and 0.1 for ceiling and ground. That is, the absorption is almost null and they reflect practically all the signals reaching them. The signal obtained is listened to with many overlaps, due to the impulse response complexity.

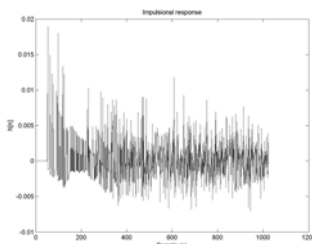


Figure 6: Impulse response. Example (little absorption)

In the second simulation (fig. 7) the data are the same except for the absorption factors, which are of 0.9 for the walls and of 0.99 for ceiling and ground. In this case there are hardly reverberations and it is appreciated that there is a delta for every room surface, of minor amplitude in the first case.

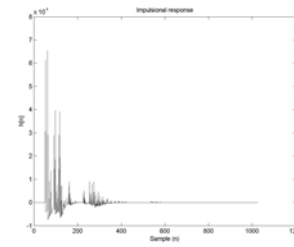


Figure 7: Impulse response. Example (great absorption)

5. CONCLUSIONS

The present application complies with the indicated premises and has a friendly user interface that allows its use without wasting time to understand it. Preliminary files are available at [6].

A possibility to improve the application performance would be to turn it into C language. As it come to specifications, the program could have the option of including more sources and even including some objects in the room. An inconvenience of the application is that image method is indicated for small/medium sized rooms, losing accuracy in big rooms.

We think that the application is a good educational tool to help understand the effects of reverberations, either by means of the files of sound that generates or through the impulse responses that calculates. This application has been used as support to theory in a university module on "Signal Processing". Academic results have raised 8% in 2004 compared to 2003 when the application was not used. Speech signals are complex signals to model but we want to emphasize the motivating effect for the student to be able to listen to the effects of a system (in our case, a room) over voice signals.

6. ACKNOWLEDGEMENTS

This work has been partially supported by the European Commission (FEDER) and Spanish/Catalan Government under projects TIC2003-08382-C05-02, TIC2003-05482, TEC2004-04526 and 2001SGR-00268.

7. REFERENCES

- [1] Haykin, Van Veen, Signals & systems, Wiley & Sons, 2001
- [2] Oppenheim, Schafer, Discrete-time signal processing, Prentice Hall, 1999
- [3] Allen, Berkley, "Image method for efficiently simulating small-room acoustics", *Journal of the Acoustical Society of America* vol 65 núm. 4 pp.943-950, april 1979
- [4] Rabinkin, Renomeron, Flanagan, Macomber, *Optimal truncation time for matched filter array processing*, ICASSP 98
- [5] Brandstein, Ward, Microphone arrays, SpringerVerlag, 2001
- [6] <http://www.eupmt.es/~satue/index3>