ECHO ROOM MAPPING

Digital Audio Systems, DESC9115, 2018
Graduate Program in Audio and Acoustics
Sydney School of Architecture, Design and Planning, The University of Sydney

INTRODUCTION

Signals originate in the real world as sensory data; seismic vibrations, visual images, sound waves, echoes, etc. Digital signal processing (DSP) is the mathematics and techniques used to manipulate those signals in digital form. DSP is used as a tool to effectively enhance multiple fields of science and engineering, including telecommunications, audio processing for speech and music, and echo location. (Smith, 1997)

Echolocation using DSP has been effectively demonstrated in areas like radar, sonar, seismology and medical imaging where the digital generation of the echo pulse is used for mapping those unique environments; aircraft movement, the ocean floor, the substructure of the earth, or the interior of the human body. But can this same process be applied more directly to room geometry? Wouldn't it be great to walk into a room, hit a button on your mobile device that creates an echo pulse and within moments understand the exact geometry of the room? This proposal is determined to effectively recreate the geometry of a room using acoustic echoes. Not only aimed at providing a useful tool for architects and engineers, the average person is given a chance to understand their own environment. This would provide a sense of security in an unknown space, which could be particularly beneficial for someone who was blind.

Section 1 describes the initial problem and the relevant factors associated with determining the geometry of a room using an echo based simulation. Section 2 introduces the appropriate mathematic equations and the geometric relationships that will be needed to recreate the room geometry. Section 3 describes the visual implementation of this process and identifies ways in which this product can enhance what is currently being developed. Section 4 establishes the procedures to effectively demonstrate the capabilities of this product.

1. PROBLEM DESCRIPTION

In attempting to model the geometry of a room from a single point, we can establish that both the sound source and the receiver are collocated and omnidirectional. By omitting an omnidirectional pulse within a room and collecting the responses at the same point, we are creating a set of delays that will be used to characterize the geometry. We can also assume that the speed of sound is constant (Dokmanic et al, 2011).

1.1. Image Source Modeling

The best approach to define the room acoustics is the image source model, described by Borish (as cited in Peng et al, 2015) as one which converts a sound source inside a room into multiple ones outside the boundaries, each corresponding to an image of the original. We are defining the number of reflective paths, which increase in time. Figure 1 shows an example of image source model with collocated source and receiver.

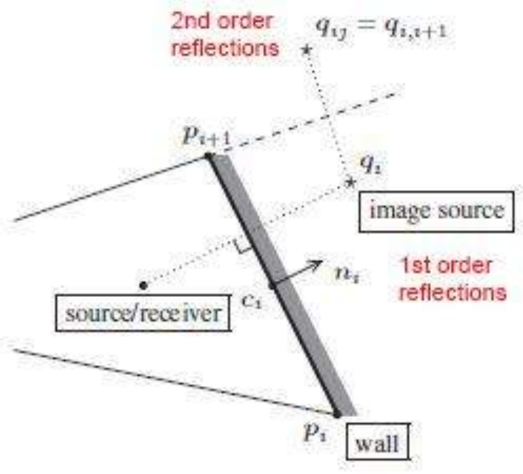


Figure 1. *Image source setup w/ collocated source/ receiver (*Dokmanić et al, 2011)

These represent the first and second order reflections, which is ultimately what we use to identify the room geometry. The timing of a reflection is characterized by the following:

Time = Distance/ Speed of Sound

1.2. Room Impulse Response Model

The sound propagation from source to receiver is described by the room impulse response (RIR). The RIR encompasses both the line of sight (direct path) and the reflective path. (Wang et al, 2016). The RIR models a channel between the fixed source and receiver. It is basically a train of pulses, each corresponding to an echo. The receiver hears the convolution of the emitted sound with a corresponding RIR. By measuring the impulse responses, we obtain the propagation times, which can be linked to the room geometry through the image source model. (Allen as cited in Dokmanić et al, 2013). We are replacing the reflections with virtual sources, which as we have discussed are mirrored images of the original source across the corresponding reflecting walls (Dokmanić et al, 2013).

2. SPECIFICATION

We have established the basic model structure to be used for the processing of these echo pulses. There are numerous processes and corrective measures that should be noted to achieve the desired final product. These include accounting for attenuation sue to absorption, compensating for irregular room geometry, influence of noise and effective assignment of echoes.

2.1. Ideal RIR

An ideal RIR for a collocated device is shown below:

$$h_i(t) = \sum_{n=0}^{\tilde{N}} a_{i,n} \cdot \delta(t - t_{i,n}),$$

where we are effectively desribing the attentuation of the direct path (refelctive path from source to image source) based on the number of walls and echoes as a function of time. t_i is the time of arrival, which is needed in order to map the room geometry. (Peng et al, 2015) As you move away from the original source, the density of the virtual source increases. In a regular polygonal shaped room, one can distinguish between sources corresponding to early reflections and diffuse reverberation. The same does not hold for irregular room geometry (Dokmanić et al, 2011). Peng et al, 2015 noted that first order reflections alone are inadequate for reconstruction of all types of room shapes but are quite effective for simple rooms. Dokmanic et al (2011) measured both first order and second order reflections (two distinct boundaries) and concluded that noise makes it difficult to collect entire sets due to the increase in pulses as time (delay) increases, but the room reconstruction was quite effective for a wide range of polygonal shapes.

2.2. Matrix Analysis

These studies show the viability of the process and demonstrate room for improvement. However, how do we deal with these issues? One effective application is matrix analysis described by Dokmanic et al (2011), referencing Figure 1.

$$q_i = 2\langle c_i, n_i \rangle n_i,$$

$$\begin{aligned} q_{ij} &= q_i + 2\langle c_j - q_i, n_j \rangle n_j \\ &= q_i + q_j - 2\langle q_i, n_j \rangle n_j. \end{aligned}$$

To summarize, second-generation image sources are defined in terms of the first-generation images instead of the original source, giving a simple expression that links the room geometry with the delay times in the impulse response. However, there is still an issue referred to as echo labelling.

2.3. Echo Labeling

Time measurements are imperfect. RIRs contain peaks introduced by various sources of noise that do not correspond to any image source or knowing which echo corresponds to which wall. Echo labelling attempts to solve the problem of matching echoes to corresponding walls. A possible strategy is to extract a large number of echo candidates and then group them using combinatorial search based on established criteria. (Krekovic et al, 2016).

3. IMPLEMENTATION

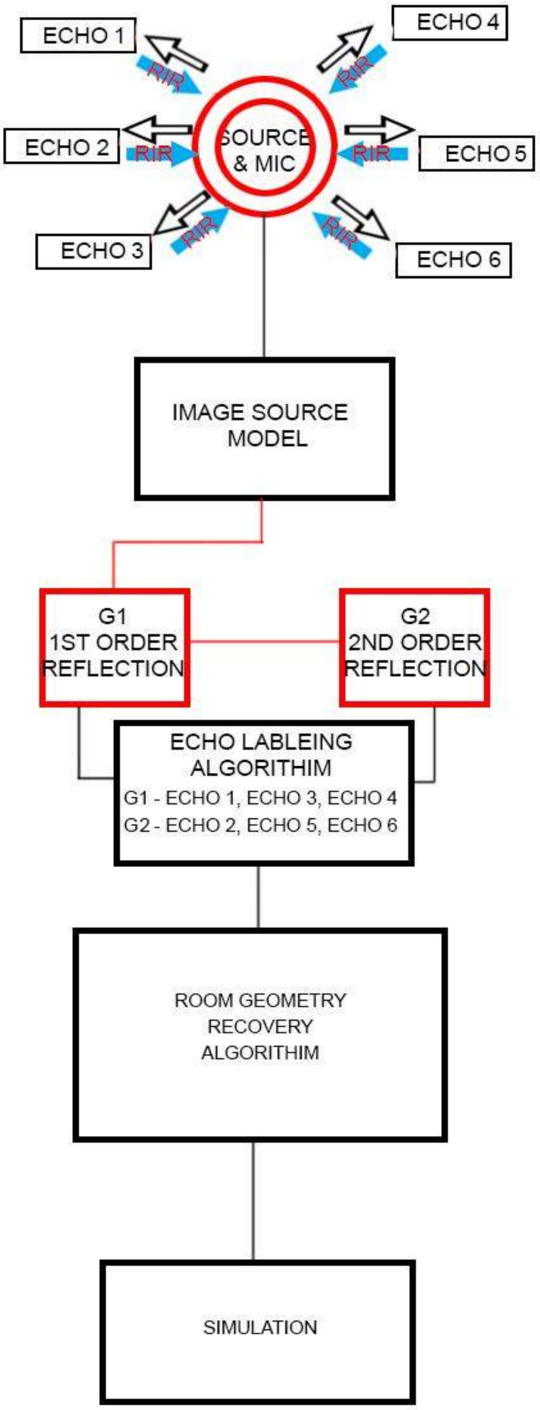


Figure 2: *Flow diagram describing the process of obtaining the room geometry from RIR measurements, which includes simulation as a form of evaluation.*

For the average user, the process will be developed as an app on a mobile device where the user can identify several parameters, including the type of materials, estimated volume, whether the room is polygonal or irregular and even add noise.

The goal is to provide the user with the ability to create the most realistic room geometry possible. For example, an architect may want to utilize the app to understand the effect of the roughness of the room materials or the addition of noise which could inform how they characterize the feeling of the space.

In addition to the parameters, we will aim to develop a device that can be attached to your mobile phone that contains an omnidirectional microphone and loudspeaker, to ensure the accuracy and dynamic capability of the individual user. Typical microphones on mobile devices are power limited and non-omnidirectional.

4. EVALUATION

To evaluate the accuracy and efficiency of the product, we can implement a few techniques that will allow us to judge its success. One such example is noted by Wang et al, 2016 where a laptop was used as a microphone and a mobile phone was used as a loudspeaker. A chirp signal was then emitted by the cell phone (directed to individual walls) and the resulting echoed signals were measured. This was then compared to actual tape measurements of the space. It was found that by directing the output of the cell phone, it was easier to identify peaks that corresponded to walls versus those from noise or higher order reflections.

Another method identified by Dokmanić et al, 2011 attempts to simulate the uncertainties in the timing equation by adding Gaussian noise to the simulated delay lines and feeds them into the proposed algorithms. They used a geometric estimation of a quadrilateral room to compare the estimated room with noiseless conditions and varying signal to noise ratio dB levels to that of the actual room.

This proposal demonstrates the effectiveness and achievability of this technique. This product has numerous potential applications for audio forensics, virtual reality and architectural acoustics. For instance, being able to design a space with specific acoustic qualities or to change the entire acoustic perception of the space? It's an opportunity that one would not want to pass up.

5. REFERENCES

Dokmanić, I., Lu, Y. M., & Vetterli, M. (2011). Can one hear the shape of a room: The 2-D polygonal case. Paper presented at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011 p 321-324.

Dokmanić, I., Parhizkar, R., Walther, A., Lu, Y. M., & Vetterli, M. (2013). Acoustic echoes reveal room shape. Proceedings of the National Academy of Sciences of the United States of America, 110 (30), 12186-12191.

Krekovic, M., Dokmanic, I., & Vetterli, M. (2016). EchoSLAM: Simultaneous localization and mapping with acoustic echoes. Paper presented at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016 p 11-15.

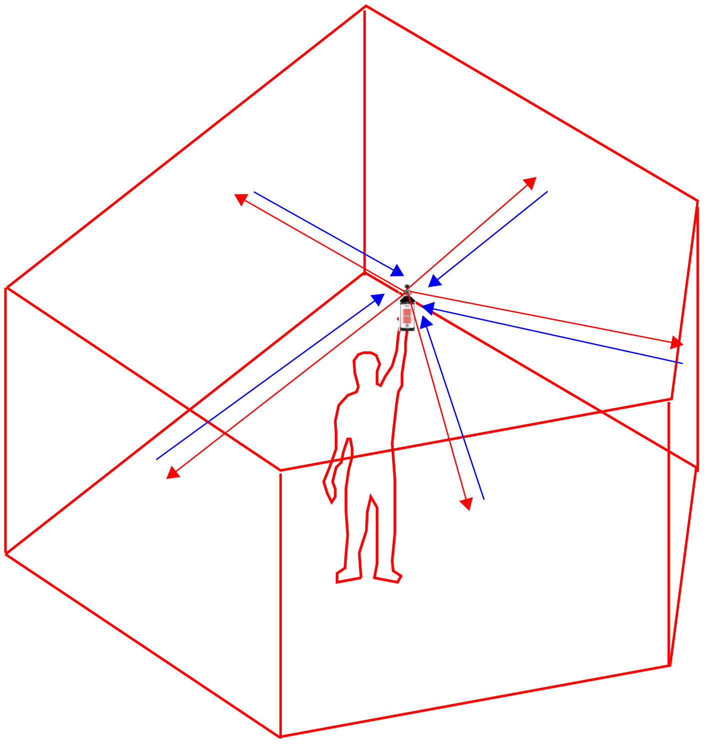
Peng, F., Wang, T., & Chen, B. (2015). Room shape reconstruction with a single mobile acoustic sensor. Paper presented at the IEEE Global Conference on Signal and Information Processing (GlobalSIP) pp 1116-1120.

Smith, Steven W. (1997). The scientist and engineer’s guide to digital signal processing. San Diego, CA: California Technical Publishing

Wang, T., Peng, F & Chen, B. (2016) First order echo based room shape recovery using a single mobile device, 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, pp. 21-25.

DESC9115 - ECHO ROOM MAPPING
GUI PROTOTYPE

echo mapz



SIMULATION

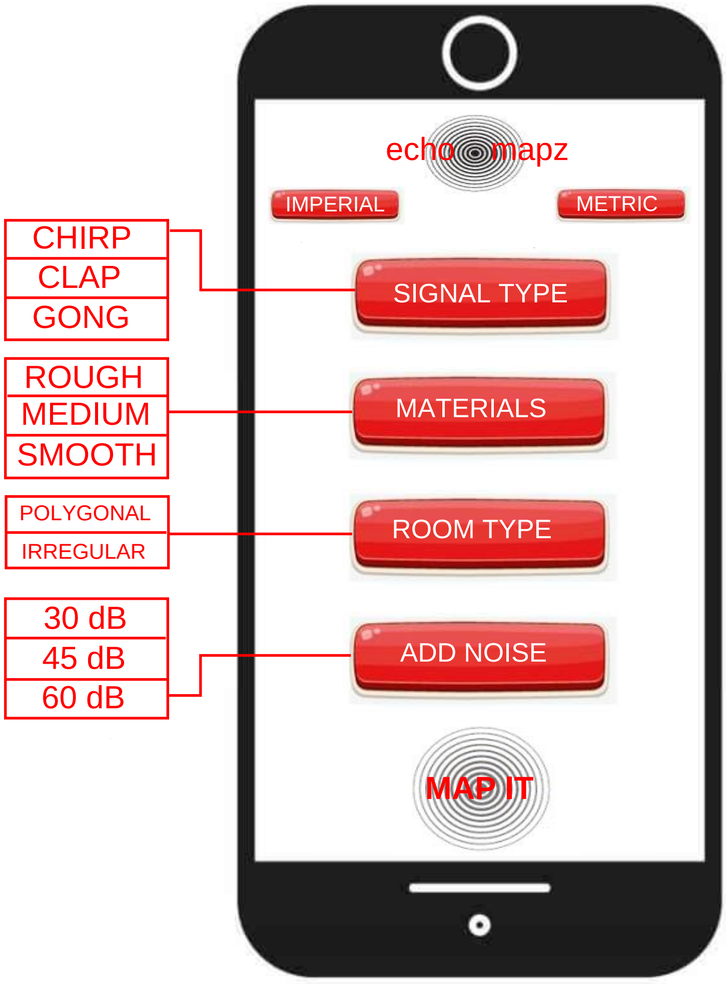
OMNIDIRECTIONAL MIC
(READING DELAYED TO
ALLOW FOR INITIAL SIGNAL)

OMNIDIRECTIONAL
SPEAKER

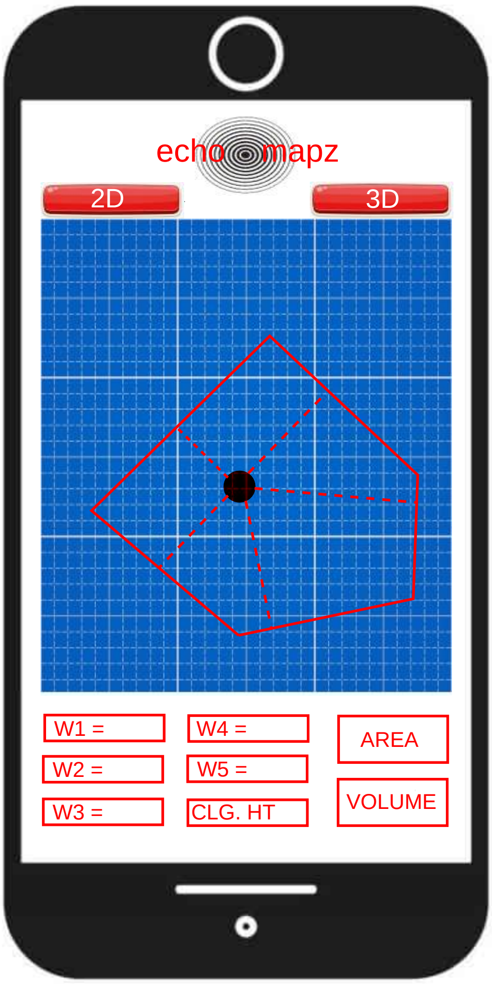
ATTACHMENT DEVICE
(ADJUSTABLE VOLUME
CONTROLS)

INTERFACE INVERTED TO
ALLOW FOR ATTACHMENT
AT BOTTOM OF PHONE

GUI INTERFACE



INPUT SCREEN



OUTPUT SCREEN