Determining Room Shape from Acoustic Echoes

Team Mentor: Dr. Lalitha Vadlamani

GROUP

Sathya Sravya

Ruchita

Mehthab

Meghna

Abstract

- We show how to compute the shape of a “convex polyhedral room” from its response to a known sound, recorded by a few microphones. Geometric relationships between the distances derived from the arrival times of echoes enable us to “blindfoldedly” estimate the room geometry.
- We aim at solving the problem using the room impulse response (RIR) and reflection times from the walls.

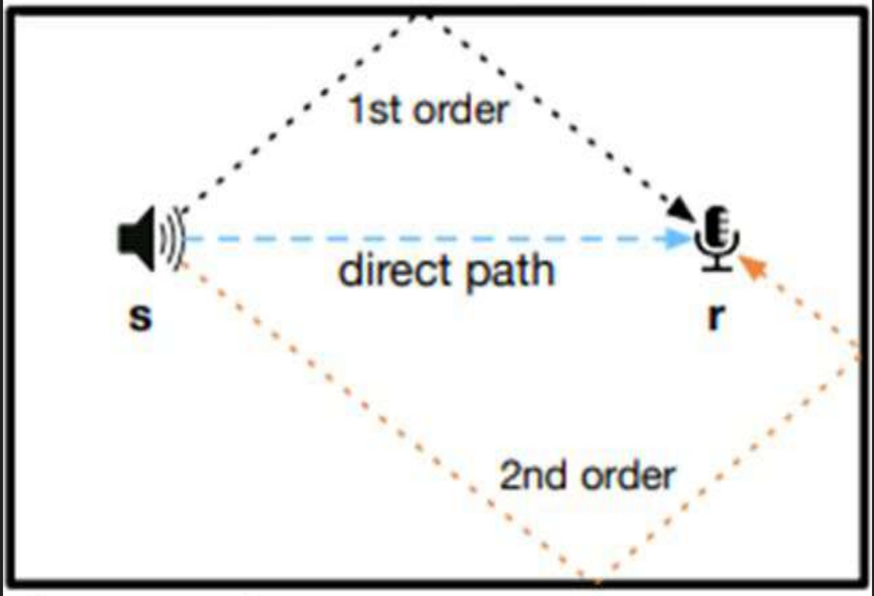
Modelling the Room

Let s be the audio source position and r be the receiver location.

- Assume that ‘s’ produces sound at time ‘To’ inside a room, part of the sound will travel directly to the receiver ‘r’ and will arrive at time ‘$t = T_o + \tau_o$.’

- Other parts of the sound will first reflect off surfaces in the room before reaching ‘*r*’ at ‘$t_i = T_o + \tau_i$’ with ‘$\tau_i > \tau_o$’ and the sound that was reflected first is called the First order reflection

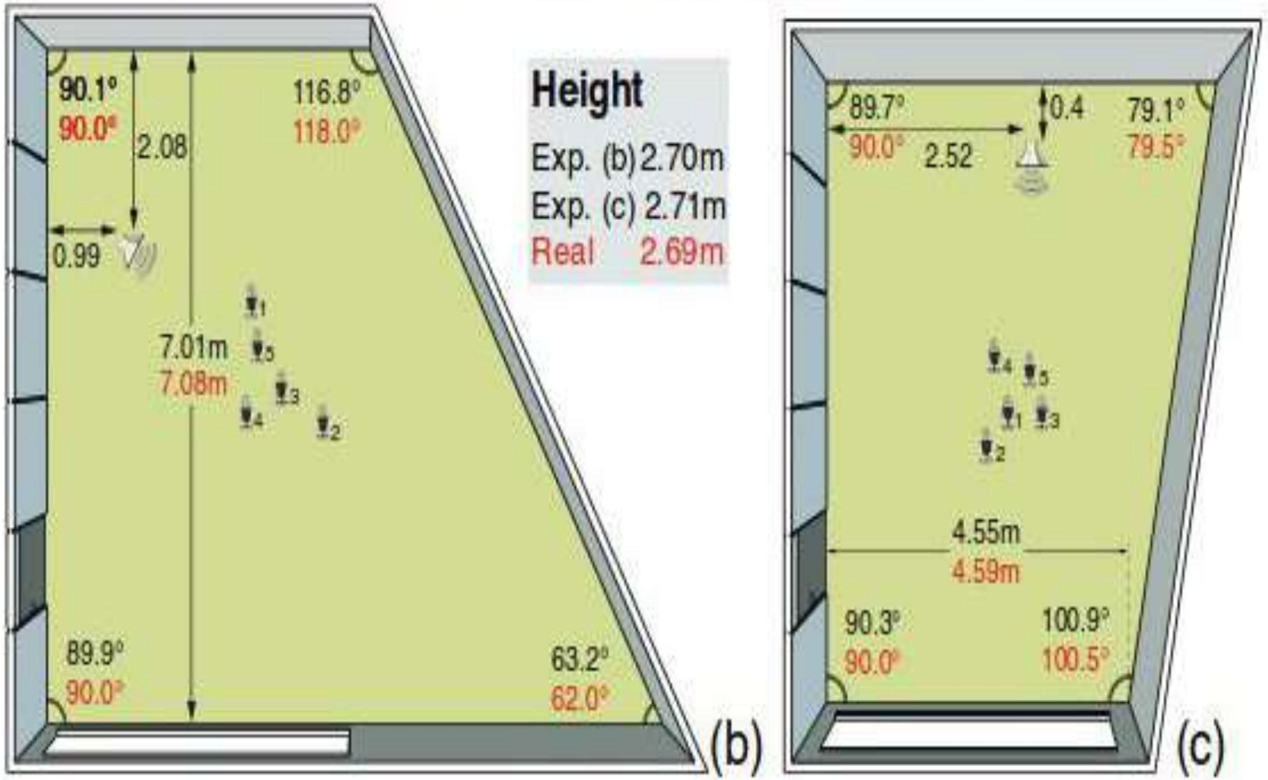
- These echoes reaching receiver can be assumed to be received from an imaginary source kept behind a wall(s’i).



Room topview

Why not concave polyhedral room ?

Our method requires the sound waves from source (input) to reflect of a wall at least once for its detection - In a concave polyhedron it **may not** reach some wall , also if it does reflect the image source obtained may be inside the room giving a wrong wall.



90.1°
90.0°
2.08
116.8°
118.0°
0.99
7.01m
7.08m
89.9°
90.0°
63.2°
62.0°
(b)
Height
Exp. (b) 2.70m
Exp. (c) 2.71m
Real 2.69m
89.7°
90.0°
2.52
0.4
79.1°
79.5°
4.55m
4.59m
90.3°
90.0°
100.9°
100.5°
(c)

1. Let us consider a convex polyhedral room with K planar walls.
2. The acoustic channel between the source and the receiver is modeled by the room impulse response as : $h(t)=\sum$ ai □(t-ti)
3. The Time Of Arrival can be calculated as ti=||si-r||/c where c is the speed of sound ,s is the source position and ri is the the location of the microphones (a minimum of 4 are required)(from proof1).
4. The source position can be calculated by the direct method using TOA.
5. Image sources are located by extracting and labeling the echoes corresponding to it from the RIR of the microphones. $s_i=s+2 <p_i-s,n_i> n_i$
6. The image sources are the mirror images of the true sources.Therefore finding an image source will give us a wall.

It is subfficient to prove for case $M=4$ cases when $M>4$ follow by considering any subset of 4 microphones.

Consider 4 microphone locations that are randomly feasible.

$$y_{k,m} \approx \|\tilde{s}_k - r_m\|^2$$

$$\approx \|\tilde{s}_k\|^2 + \|r_m\|^2 - 2 (\tilde{s}_k^T r_m)$$

$$\tilde{y}_{k,m} \approx \frac{1}{2} (y_{k,m} - \|r_m\|^2)$$

$$= \tilde{s}_k^T r_m - \frac{1}{2} \|\tilde{s}_k\|^2$$

$\tilde{s}_k$ is the location of image source w.r.t wall k

$$\begin{bmatrix} \tilde{y}_{k,1} \\ \tilde{y}_{k,2} \\ \tilde{y}_{k,3} \\ \tilde{y}_{k,4} \end{bmatrix} = \begin{bmatrix} r_1^T & -1/2 \\ r_2^T & -1/2 \\ r_3^T & -1/2 \\ r_4^T & -1/2 \end{bmatrix} \begin{bmatrix} \tilde{s}_k \\ \|\tilde{s}_k\|^2 \end{bmatrix}$$

$$\tilde{y}_k = R \tilde{u}_k$$

$$1^T \tilde{y}_k = -\frac{M}{2} \cdot \|\tilde{z}_k\|^2$$

$$\|\tilde{z}_k\|^2 = \sum_{m=1}^{M} \frac{2}{M} \tilde{y}_{k,m}$$

The image source is found as

$$\tilde{z}_k = S\tilde{y}_k$$

$$SR = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

S satisfies this

$$\Rightarrow \quad \frac{2}{4} 1^T \tilde{y}_k + \|S \tilde{y}_k\|^2 = 0$$

where $\tilde{y}_k$ corresponds to k^{th} wall.

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} R^+$$

R^+ is the pseudo inverse of R.

with this choice, any column submatrix of S with $n \leq 3$ column is rank n with probabality 1.

Procedure for Identification of Wall

(i.e image source)

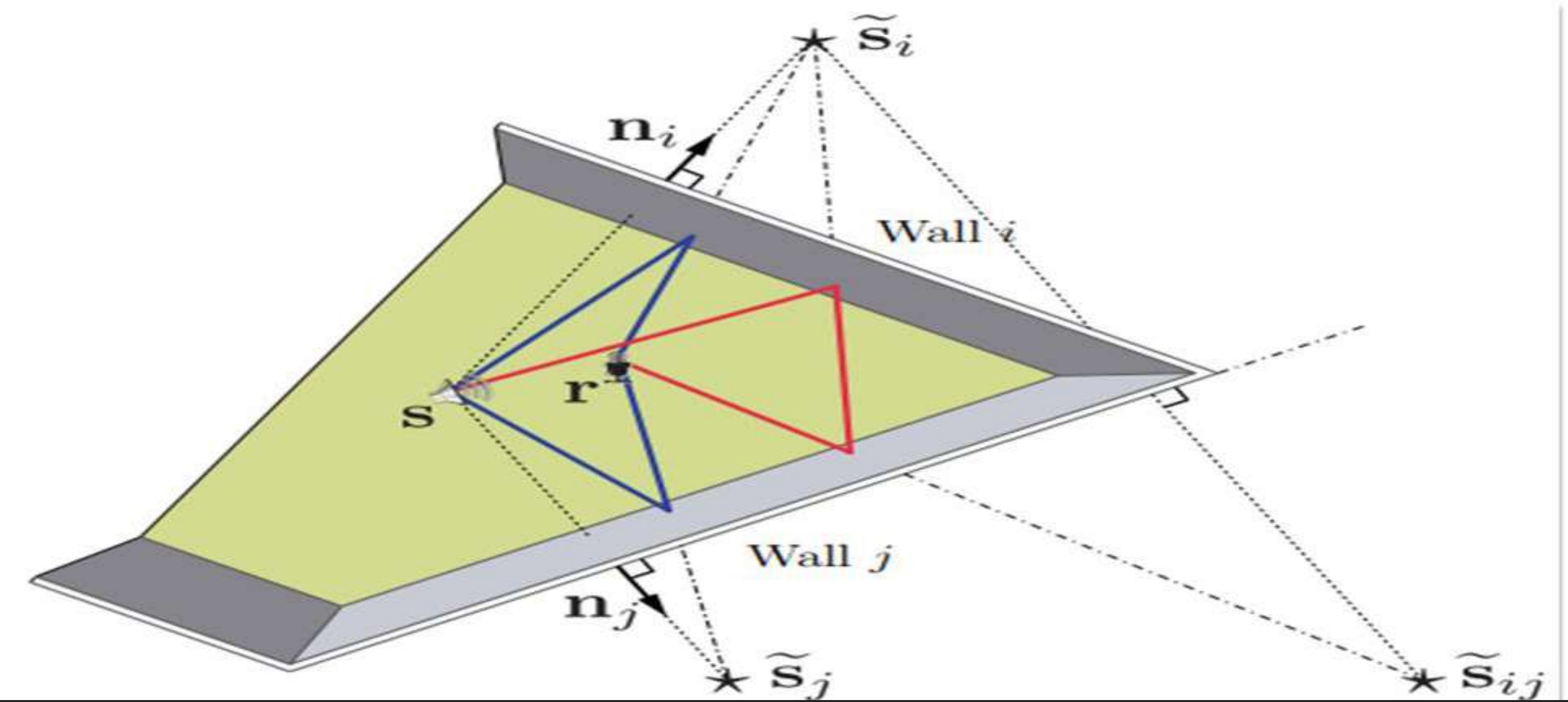
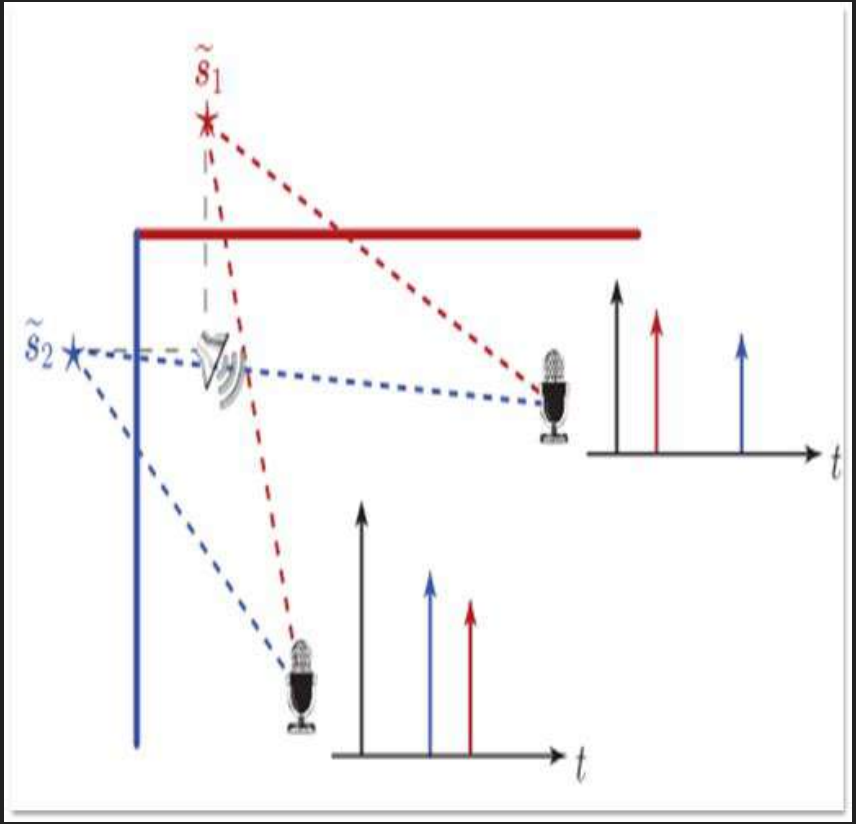


Illustration of receiver receiving first and second order echoes from imaginary sources corresponding to different walls.

- For locating the image source, it’s corresponding echoes must be extracted from the RIRs of the microphones.The challenge is that these echoes are unlabelled in each RIR as the order of arrival of echoes maynot be the same for all microphones.
- The echoes corresponding to different image sources do not reach the two microphones in the same order. Thus, the main problem is the identification of right combination of echoes that correspond to the same image source.



Room Geometry Estimation Using EDMs

Properties of EDM Matrices

- They are matrices whose entries are the squared distance between the receivers,i.e the microphones(M). D ϵ R MXM   Where as D[i,j]=|| ri-rj |^2 ri , rj ∈ R.
- They are symmetric matrices (dij=dji) with diagonal elements = 0 ,as dii is always zero.
- If we use the 5 microphones r1 to r5 as our point set X, we can construct D
- We can define a function that outputs D ∈ EDM from point set X as

- D(X) = diag(XX^t) 1 T + 1 diag((XX^t))T − 2XX^t,

where diag (·) denotes a column vector containing all diagonal entries of the input matrix and 1 is a column vector of ones in R .

- This implies that every element of EDM must satisfy the following properties:

Triangle inequality ,non-negativity,self-distance and symmetry.

- To identify the echoes coming from same wall we use EDM based technique.
- One echo from each microphone's RIR is selected. The aim is to find the collection of echoes that come from the same wall.
- It exploits the properties of Euclidean matrices.(here rank)
- For the removal of ghost echoes detected while considering peaks from the RIR.
- We assume that we know the distances between the microphones.

Echo labeling :EDM-based Approach

1. Each microphone receives direct sound and K first order echoes from the K walls.
2. The arrival times are directly proportional to the distance between the virtual sources and the receivers $t_i=||s_i-r||/c$.
3. With selection of each set of distances obtained from TOAs of echoes corresponding to all receivers, we augment the EDM "D",obtain Daug.
4. The maximum value of rank of obtained Daug will be 5 ,if all selected echoes are from the same wall, else the rank will be greater than 5.(proof 2 since we chose 3D room).

The row rank of a matrix is the maximum number of rows, thought of as vectors, which are linearly independent. Similarly, the column rank is the maximum number of columns which are linearly independent. It is an important result, not too hard to show that the row and
column ranks of a matrix are equal to each other. Thus one simply speaks of the rank of a matrix.

The augmented matrix is constructed as,

$$D_{aug}\left(d_{(i_1,i_2,\cdots i_M)}\right) = \begin{bmatrix} D & d_{(i_1,\cdots i_M)} \\ d^T_{(i_1,\cdots i_M)} & 0 \end{bmatrix}$$

Echoes in RIRs corresponding to different walls….

$$
\begin{matrix}
r_1 : [e_? \;\; e_? \;\; e_? \;\; e_? \;\; e_?] \\
r_2 : [e_? \;\; e_? \;\; e_? \;\; e_? \;\; e_?] \\
r_3 : [e_? \;\; e_? \;\; e_? \;\; e_? \;\; e_?] \\
r_4 : [e_? \;\; e_? \;\; e_? \;\; e_? \;\; e_?]
\end{matrix}
\Rightarrow
\begin{matrix}
r_1 : [e_0 \;\; e_2 \;\; e_1 \;\; e_3 \;\; e_4] \\
r_2 : [e_0 \;\; e_1 \;\; e_2 \;\; e_3 \;\; e_4] \\
r_3 : [e_0 \;\; e_2 \;\; e_3 \;\; e_1 \;\; e_4] \\
r_4 : [e_0 \;\; e_3 \;\; e_1 \;\; e_2 \;\; e_4]
\end{matrix}
$$

- If rank(Daug) ≤ 5 or more specifically Daug verifies the EDM property, then the selected combination of echoes corresponds to an image source, or equivalently to a wall.This approach usually requires testing all possible combinations of echoes taken from all receivers(mics).
- In practical cases the number of combinations is small enough that this does not pose a problem .
- Also,It is not necessary to test all echo combinations. An echo from a fixed wall will arrive at all the microphones within the time given by the largest inter-microphone distance.Therefore, it suffices to combine echoes within a temporal window corresponding to the array diameter. This substantially reduces the running time of the algorithm. Consequently, we can be less conservative in the peak-picking stage.

Practical Algorithm

- ❖ The above stated rank test mostly fails in practical scenarios as the Daug is not EDM as there is error in measurements (precision) made prior to this (impulse times,speed of sound , distances between microphones ,) and also due to sampling resolution
- ❖ So instead of applying the strict rank rule we modify in that we check how close Daug is to being an EDM.
- ❖ The cost function s-stress is used .
- ❖ Let ~Daug be the noisy matrix then s-stress (~Daug) is the score of the matrix which says how close it is to being an EDM
- ❖ Lower score - closer to EDM
- ❖ Sort scores and corresponding echoes in ascending order

Algorithm :

Input: Microphone EDM D, and em, $m \in \{1, \cdots, M\}$.

Output:Shape of the room.

1.For every d{i1,...,iM),score[d(i1,...,iM)]$\leftarrow$s$-$stress(Daug).

2.Sort the scores collected in score.

3.Compute the image source locations.

4.Remove image sources that do not correspond to walls (higher-order by their geometric dependencies to first order echoes and the “ghost” sources by heuristics).
5.Reconstruct the room

Filtering out Higher Order Image Sources

- The location of a first order image source location can be computed as

$$s' = s + 2\langle p - s, n \rangle n$$

Where i corresponds with wall i, n is the unit normal and p is any point on the wall. In words this is twice the distance from the true source through the wall along the normal unit vector.

- In the same way the location of second order image sources can be expressed in terms of the first order image sources.
- S''_{ij} - second order image source ′ is a combination of s'_i and s'_j - two first order sources . When image source s'_i is

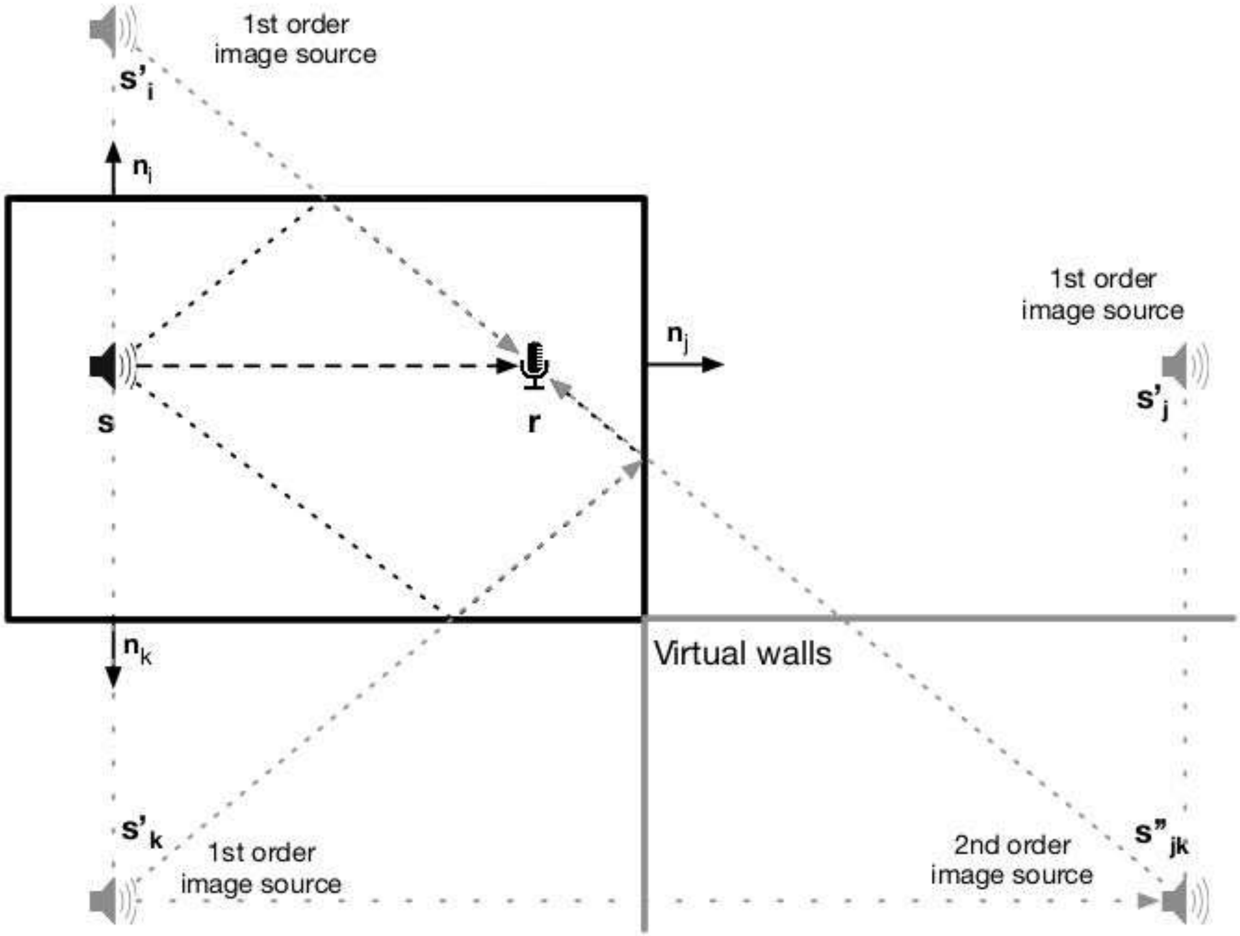
mirrored in the (extended virtual) wall corresponding to image source s'_j , the result is second order image source s''_{ij} .

$$s'' = s' + 2\langle p - s' , n \rangle n$$

We check if image source s''_{ij} ,is close to the result of mirroring s'_i in the wall of s'_j as

$$||s'_i + 2\langle p_j - s'_i \, n_j \rangle n_j - s''_{ij} || < \varepsilon$$

with ε some small number, we know whether s'' is a higher order image source.

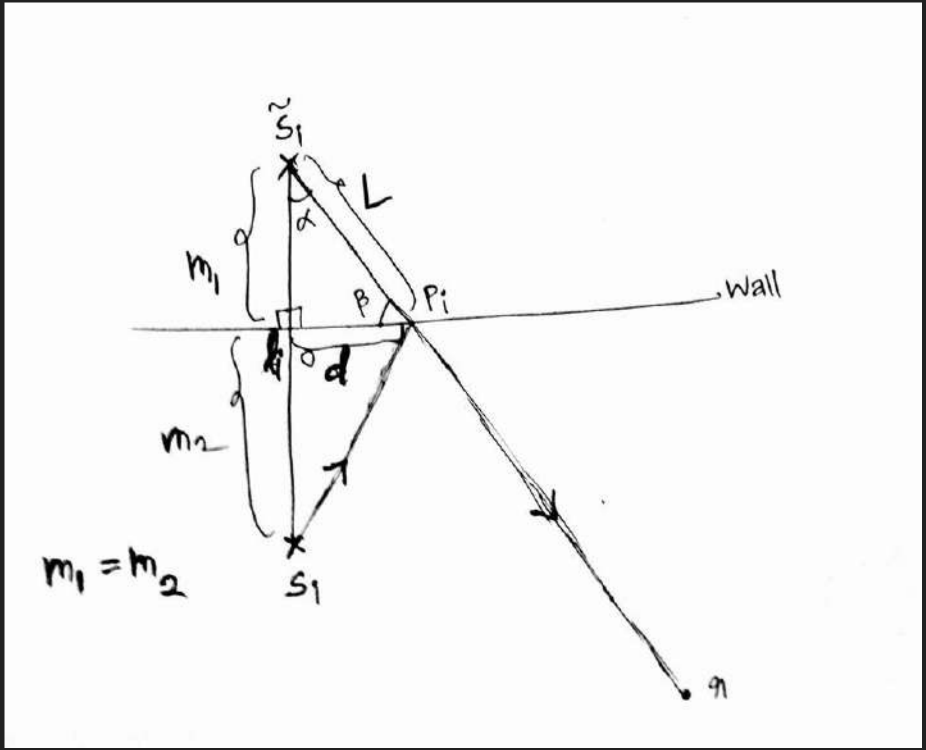

1st order
image source
s'ᵢ
n_i
1st order
image source
n_j
s
r
s'ⱼ
n_k
Virtual walls
s'ₖ
1st order
image source
2nd order
image source
s"ⱼₖ

Room Reconstruction

- ❖ After running the above algorithm we obtain K arrays of M elements each
- ❖ Where , K is number of walls - so by the number of sets of echoes we obtain by echo labeling gives us the number of walls we can find - and m is the number of microphones.
- ❖ Each element in an array is the image source's distance from one receiver .
- ❖ We need to find the coordinates of the image source.
- ❖ Using three receivers (as there are at least four - proof in Uniqueness), we find intersection of three spheres - intersection of two spheres is a point or a circle,
- ❖ And intersection of a sphere and a circle is 1 point , 2 points or infinite(circle)
- ❖ For circle of intersection two of the the receivers must be at the same point which is not the case,
- ❖ So we obtain 1 or 2 points ,we then use distances from remaining receivers to choose from these.

Room Reconstruction (Contd.)

- ❖ Now, we know three points - receiver r, S and S'i - thus we have two
- ❖ Once we have the image source , we find the midpoint of real source and this image source to obtain a point on the wall corresponding to this image source
- ❖ li = (S'i + S)/2
- ❖ We also obtain the perpendicular to the wall at this point as - (S - li)



Using angle $(\vec{S} - \vec{\tilde{S}}_i)$ & $(\vec{r} - \vec{\tilde{S}}_i)$ we get angle 'α',

We Know $m_1 = \dfrac{\| S_1 - \tilde{S}_1 \|}{2}$

$$\cos\alpha = \frac{m_1}{L} \Rightarrow L = \frac{m_1}{\cos\alpha}$$

Now P_i,

$$\frac{\| \vec{\tilde{S}}_i - \vec{P}_i \|}{\| \vec{P}_i - \vec{r} \|} = \frac{L}{D - L}$$ here $D = \| \vec{\tilde{S}}_1 - \vec{r} \|$

$\rightarrow$ (section formula in 3D)

P_i divides line joining $\tilde{S}_i$ and r internally,

P_i, l_i = points on plane/wall.

$(\vec{S}_i - \vec{l}_i)$ is the perpendicular to plane/wall, so we can write the equation of wall.

Also $l_i = \frac{\tilde{s}_i + s_i}{2}$ (co-ordinates)

We know p_i, l_i points on plane

$$\boxed{\left(\vec{\tilde{s}}_i - \vec{l}_i\right) \cdot \left(\vec{z} - \vec{l}_i\right) = 0 .}$$

Hence desired equation of Wall.

Angles between walls are given by angles between their normals $\left(\vec{\tilde{s}}_i - \vec{s}\right)$s.

Hence we will be able to determine shape of Room.

$$\boxed{\theta_{ij} = \frac{\vec{s}_i - \vec{s}_j}{\|\vec{s} - \vec{s}_i\| \, \|\vec{s} - \vec{s}_j\|}}$$

Uniqueness

- Only one room corresponds to a selected set of first order microphones if the room-microphone array is “feasible” -
- Given a room R and a loudspeaker position s, we say that the point $x \in R$ is feasible if a microphone placed at x can receive all the first-order echoes of a pulse emitted from s.
- Since our algorithm needs the knowledge of first order echoes, we need to restrict ourselves to the cases in which we can hear them.
- We need at least four possible unique receivers to accomplish this
- Feasible - in the feasible region
- Unique - no two of these four receivers are at the same point/location.

Proof 2 (rank of EDM <= 3+2)

We need,

EDM for $X = \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \vdots \\ \bar{r}_n \end{bmatrix}_{n\times 1}$. Let 1 be $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

We use $XX^T = \begin{bmatrix} \bar{r}_1^2 & \bar{r}_1\bar{r}_2 & \cdots & \bar{r}_1\bar{r}_n \\ \bar{r}_1\bar{r}_2 & \bar{r}_2^2 & \cdots & \bar{r}_2\bar{r}_n \\ \vdots & & & \\ \bar{r}_n\bar{r}_1 & \cdots & \cdots & \bar{r}_n^2 \end{bmatrix}$

We want to get $EDM(X) = \begin{bmatrix} (\bar{r}_1-\bar{r}_1)^2 & (\bar{r}_1-\bar{r}_2)^2 & \cdots \\ & & \end{bmatrix}$

$$= \begin{bmatrix} 0 & \|\bar{r}_1\|^2 + \|\bar{r}_2\|^2 - 2\bar{r}_1\cdot\bar{r}_2 & \cdots \\ & 0 & \\ & & \end{bmatrix}$$

$$= \begin{bmatrix} |\bar{r}_i|^2 + |\bar{r}_j|^2 - 2|\bar{r}_i|.|\bar{r}_j| \end{bmatrix}_{n\times n}$$

$$= \left[|\bar{r}_i|^2\right] + \left[|\bar{r}_j|^2\right] - 2\left[|\bar{r}_i|.|\bar{r}_j|\right]$$

$$= \begin{bmatrix} r_1^2 & r_1^2 & \cdots \\ \vdots & & \\ r_n^2 & r_n^2 & \cdots \end{bmatrix} + \begin{bmatrix} r_1^2 & r_2^2 & \cdots \\ r_1^2 & r_2^2 & \cdots \\ \vdots & & \end{bmatrix} - 2\begin{bmatrix} r_1^2 & \bar{r}_1\bar{r}_2 & \cdots \\ \bar{r}_2\bar{r}_1 & r_2^2 & \cdots \\ & & \end{bmatrix}$$

(1)

$$EDM(X) = \underbrace{\operatorname{diag}(XX^T)}_{n \times 1} \cdot \underbrace{1^T}_{1 \times n} + 1\left(\operatorname{diag}(XX^T)\right)^T - 2(XX^T)$$

So $D(X) = EDM(X)$

$$= \operatorname{diag}(XX^T) \cdot 1^T + 1 \cdot \left(\operatorname{diag}(XX^T)\right)^T - 2(XX^T)$$

$$\operatorname{Rank}(EDM(X))$$

$$\leq \cancel{\text{rank}}$$

$$1 + 1 + r$$

$\to$ min. rank(X) for it to form an EDM.

Experiment

Sampling frequency : Minimum 16kHZ ; 96kHz adequate [1]. Higher the better.

Number of samples - 4096

Time window for sampling Tw = Number of sample * 1/fs

Therefore farthest wall that can be heard - is at a distance < c*Tw

Where c is speed of sound.

Peak-picking :

We can consider the reading at every sample to be an impulse and remove these in later echo labeling but this has much higher complexity so we pick peaks when the input value at a certain receiver is greater than certain threshold .

Place an additional receiver r_dummy ,other than those being used ,at a distance q from source, q<<di ,where di is distance from source for receiver i.
The largest peak of this r_dummy is directly from source,
Let its magnitude be m, for each receiver -
Threshold = y*((m - Mi)/di)*c*Tw
Where y - reflection coefficient of wall (assumed equal for all walls) Mi is maximum value of input at receive i.

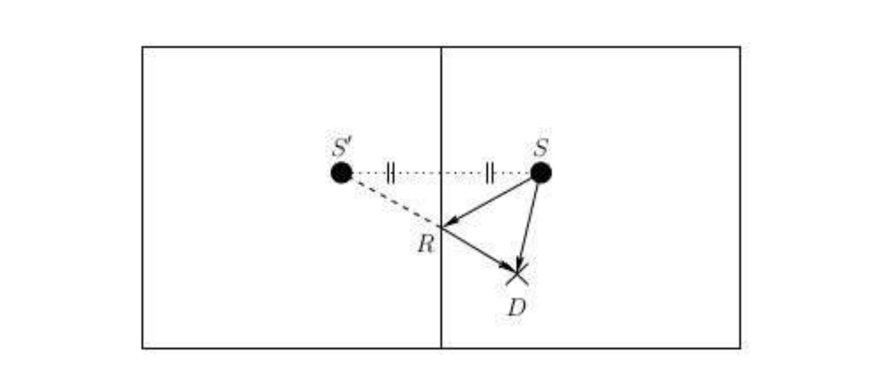
Number of microphone = 5
Number of real sources = 1

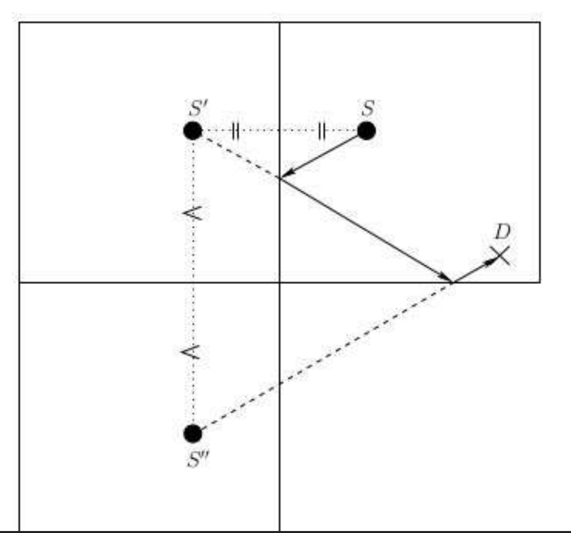
RIR Generation
Allen & Berkley’s Image Method

We can control

- Order of reflection
- Room dimension
- Mic directivity of impulse produced.

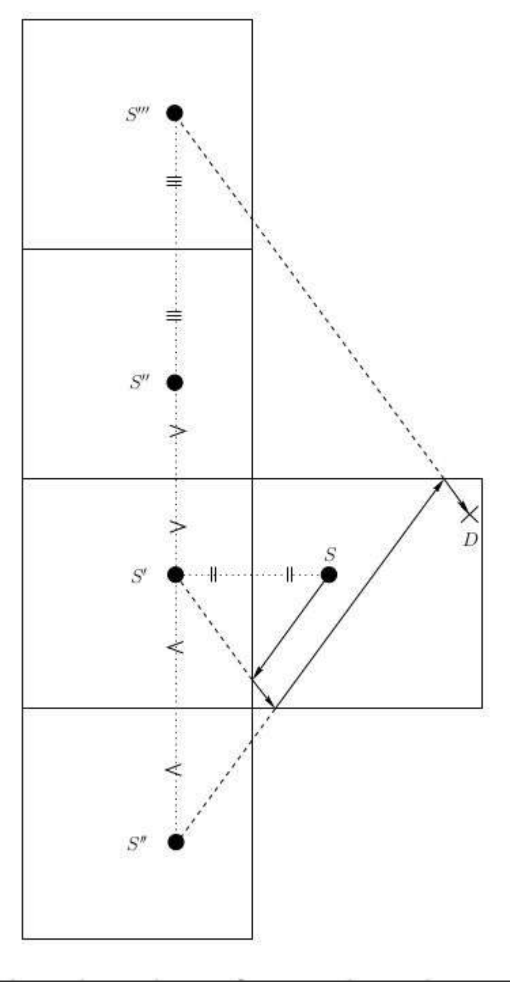
At destination D two signals arrive, one from the direct path and a second one from the reflection. The path length of the direct path can be directly calculated from the known locations of the source and the destination. Also shown is an image of the source, S’, located behind the wall at a distance equal to the distance of the source from the wall. Because of symmetry, the triangle SRS’ is isosceles and therefore the path length SR + RD is the same as S’D. Hence, to compute the path length of the reflected path, we can construct an image of the source and compute the distance between destination and image. Also, the fact that we are computing the distance using one image means that there was one reflection in the path.





Path is involving two reflections. The length of this path can be obtained from the length of S’’D.

The length of a path involving three reflections is obtained from the length of S’’’D.



- In general the path lengths (and thus the delays) of reflections can be obtained by computing the distance between the source images and the destination.
- The strength of the reflection can be obtained from the path length and the number of reflections involved in the path.
- The number of reflections involved in the path is equal to the level of images that was used to compute the path.

Image Method

- Consider a rectangular room with length, width and height given by L_x, L_y and L_z. Let the sound source be at a location represented by the vector $r_s = [x_s, y_s, z_s]$ and let the microphone be at a location represented by the vector $r = [x, y, z]$
- The relative positions of the images measured with respect to the receiver position and obtained using the walls at $x = 0$, $y = 0$ and $z = 0$ can be written as

$$\mathbf{R_p} = [(1-2q)x_s - x, (1-2j)y_s - y, (1-2k)z_s - z]$$

Each of the elements in the triple $p = (q, j, k)$ can take on values 0 or 1, resulting in eight different combinations that specify a set P, i.e., $P = \{(q, j, k) : q, j, k \in \{0, 1\}\}$. When the value of p is 1 in any dimension, then an image of the source in that direction is considered.

It should be noted that some of these images correspond to higher order reflections. To consider all images, we add the vector Rm to Rp where

$$\mathbf{R_m} = [2m_x L_x,\ 2m_y L_y,\ 2m_z L_z]$$

where m x , m y , and m z are integer values. Each of the elements of the triple m = (m x , m y , m z) takes on values from −N to +N .

The distance between any source image and the microphone can be written as $d = \|\mathbf{R_p} + \mathbf{R_m}\|$

The time delay of arrival of the reflected sound ray corresponding to any source image can be expressed as

$$\tau = \frac{d}{c} = \frac{\|\mathbf{R_p} + \mathbf{R_m}\|}{c}$$

Now the impulse response is

$$h(\mathbf{r}, \mathbf{r}_s, t) = \sum_{\mathbf{p} \in \mathcal{P}} \sum_{\mathbf{m} \in \mathcal{M}} \beta_{x1}^{|m_x - q|} \beta_{x2}^{|m_x|} \beta_{y1}^{|m_y - j|} \beta_{y2}^{|m_y|} \beta_{z1}^{|m_z - k|} \beta_{z2}^{|m_z|} \frac{\delta(t - \tau)}{4\pi d}$$

βx1, βx2, βy1, βy2, βz1 and βz2 are the reflection coefficient of the walls

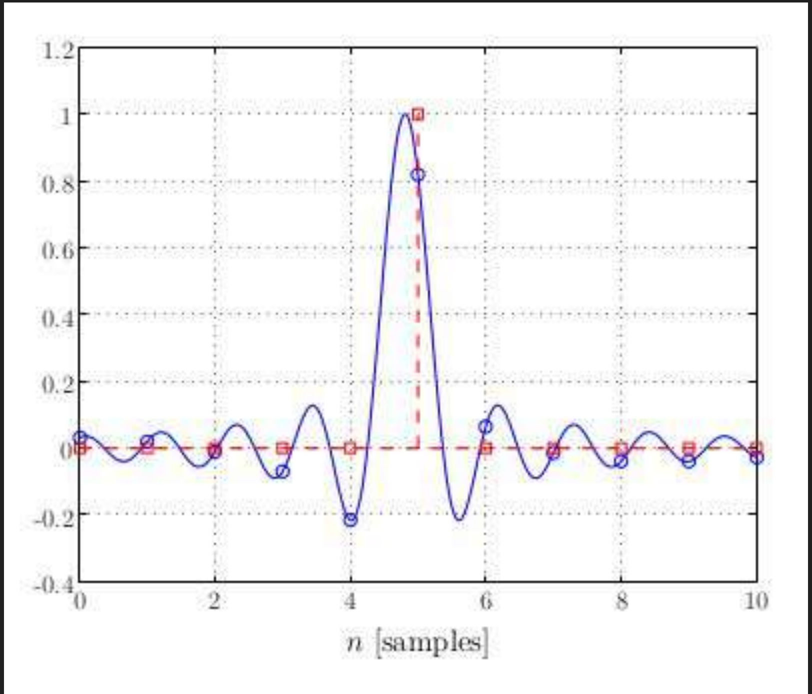
The elements of the triple p are 0 or 1, which means that there are 8 different combinations, (0,0,0) to (1,1,1). The elements of the triple m range from −N to +N , which means that there are (2N + 1)^3 combinations. Therefore for a given N, we get 8(2N+1)^3 different paths.

An important consideration while simulating the discrete version of this impulse response using a computer is that the delays given by do not always fall at sampling instants. Ideally, the discrete version of is given by

$$h(\mathbf{r}, \mathbf{r}_s, n) = \sum_{\mathbf{p} \in \mathcal{P}} \sum_{\mathbf{m} \in \mathcal{M}} \beta_{x_1}^{|m_x - q|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y - j|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z - k|} \beta_{z_2}^{|m_z|} \frac{\mathrm{LPF}\{\delta(n - \tau f_s)\}}{4\pi d}$$

where fs is the sampling frequency and LPF{·} denotes a theoretically perfect Low-Pass Filter with cut-off frequency fs /2. We also make the following approximation

$$\mathrm{LPF}\{\delta(n - \tau f_s)\} \approx \delta(n - \mathrm{round}\{\tau f_s\})$$

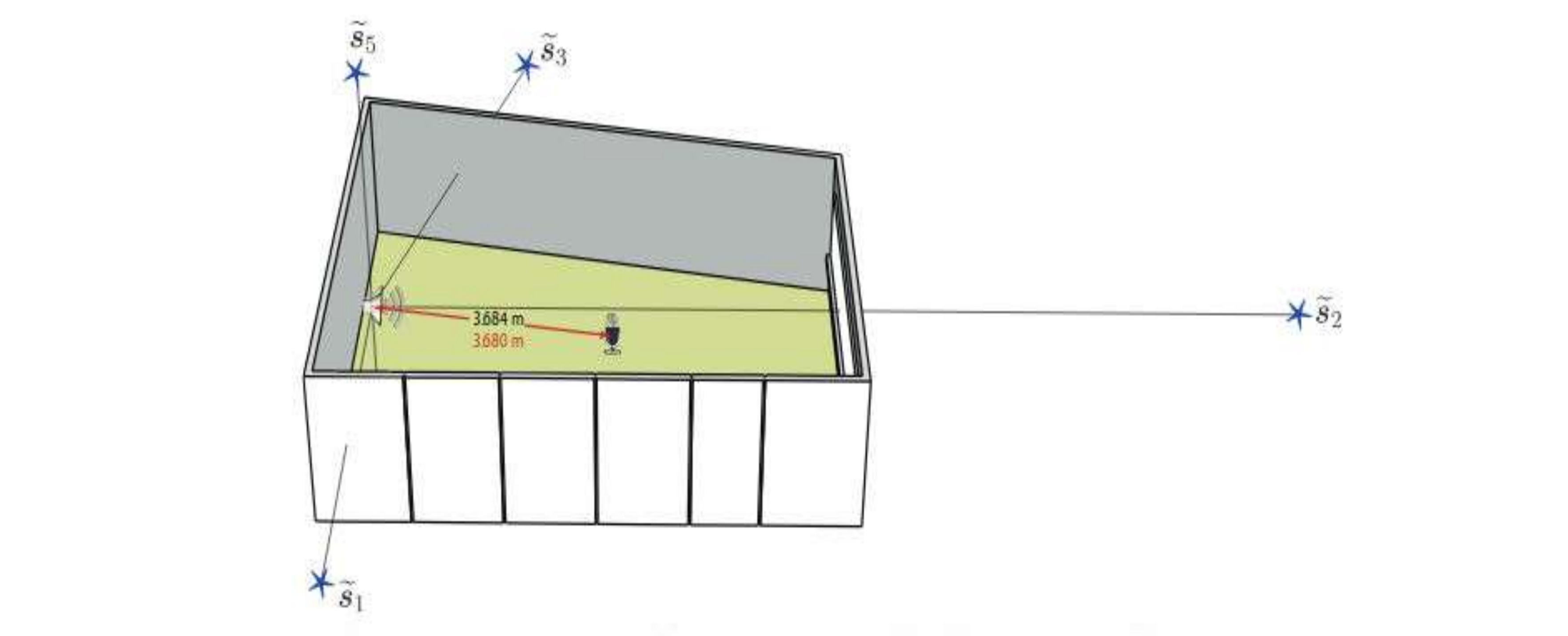


Comparison of the shifted and low-pass impulse method.

Peterson suggested another modification to the image method. In this approach, each impulse is replaced by the impulse response of a Hanning-windowed ideal low-pass filter of the form

$$\delta_{\mathrm{LPF}}(t) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{2\pi t}{T_w}\right)\right) \operatorname{sinc}(2\pi f_c t) & \text{for } -\frac{T_w}{2} < t < \frac{T_w}{2} \\ 0 & \text{otherwise} \end{cases}$$

Where Tw is the width (in time) of the impulse response and f c is the cut-off frequency of the low-pass filter



The virtual sources of the loudspeaker are represented with stars. The image source of the floor (~s4) is not shown for better visualization. The distance of the loudspeaker and the microphone is shown in red while the estimated distance is in black.

Code-MATLAB

Functions Used

Function: mdscale

Example: [P,sstress] = mdscale(K,3,’criterion’,’metricsstress’);

Function: findpeaks

Example: [pks5,locs5] = findpeaks(h5);

Function: pdist

Example: D = pdist(r,'euclidean');

D_Matrix = squareform(D);

Input parameters:

Parameter	Description
c	sound velocity in m/s.
fs	sampling frequency in Hz.
r	M x 3 matrix specifying the (x,y,z) coordinates of the receiver(s) in m.
s	1 x 3 vector specifying the (x,y,z) coordinates of the source in m.
L	1 x 3 vector specifying the room dimensions (x,y,z) in m.
beta	1 x 6 vector specifying the reflection coefficients $[\beta_{x_1}\ \beta_{x_2}\ \beta_{y_1}\ \beta_{y_2}\ \beta_{z_1}\ \beta_{z_2}]$ or beta = Reverberation Time (RT_{60}) in seconds.

Optional input parameters:

Parameter	Description	Default value
nsample	number of samples to calculate.	$RT_{60}\, f_s$
mtype	type of microphone that is used ['omnidirectional', 'subcardioid', 'cardioid', 'hypercardioid', 'bidirectional'].	'omnidirectional'
order	maximum reflection order.	-1
dim	room dimension (2 or 3).	3
orientation	direction in which the microphone is pointed, specified using azimuth and elevation angles in radians.	[0 0]
hp_filter	use 'false' to disable high-pass filter.	'true'

Output parameters:

Parameter	Description
h	M x nsample matrix containing the calculated room impulse response(s).
beta_hat	In case a reverberation time is specified as an input parameter the corresponding reflection coefficient is returned.

Intuitive methods

1. Based on advantageous placing(positioning) of receivers.(weak signals from farthest walls)
2. Also we can take advantage of reverberation time of echoes produced.

Reverberation

- Reverberation time is the length of an echo you hear in a room.As sound vibrations trail off you hear a "decay." In a more strict definition, reverberation time measures the length of time for sound to decay.
- When sound bounces off a surface it produces reflections which are absorbed or bounced to other surfaces depending on the surface composition. Reverberation time is expressed as RT60, which equals the number of seconds needed for reflections to decay by 60 decibels below the direct sound level.
- Room dimensions and shape are major factors in determining room reverberation time, which can be measured in all architectural structures. Many of the factors affecting sound quality involve how sound waves bounce around the room. The more walls, the greater chance of longer reverberation.

References

1.J.C. Gower, “Properties of Euclidean and Non-Euclidean Distance Matrices”, Harpenden, Hertfordshire AL5 2.Q United Kingdom.
http://www.convexoptimization.com/TOOLS/Gower1.pdf

2.Reza Parhizkar ,”Euclidean Distance Matrices: Properties, Algorithms and Applications”,
3.Ingmar Jager,”Room Shape Estimation from Acoustic Echoes using Graph-based Echo Labeling”.
4.Alex Cosper, “How Does the Shape of a Room Affect Reverberation Time?"
5.Ivan Dokmanić, Reza Parhizkar, Andreas Walther, Yue M. Lu, and Martin Vetterli,
“Acoustic echoes reveal room shape”,
http://www.pnas.org/content/110/30/12186
6.INTERNATIONAL AUDIO LABORATORIES ERLANGEN , RIR Generator,
https://www.audiolabs-erlangen.de/fau/professor/habets/software/rir-generator

Thank you !!!