



Course material

Course:

Digital Signal Processing II

Video:

2.1.6 SOTD Can one hear the shape of a room

Concepts (extracted from automatically generated subtitles):

Room acoustics. Impulse response. First sound. Shape of a room. Important problem. Second virtual source. Concept of a room impulse response. Impulse signal. Today's signal. Response of the system. Virtual sources of the original source. Complex room. Room impulse response. Part of the sound. Different room impulse responses.



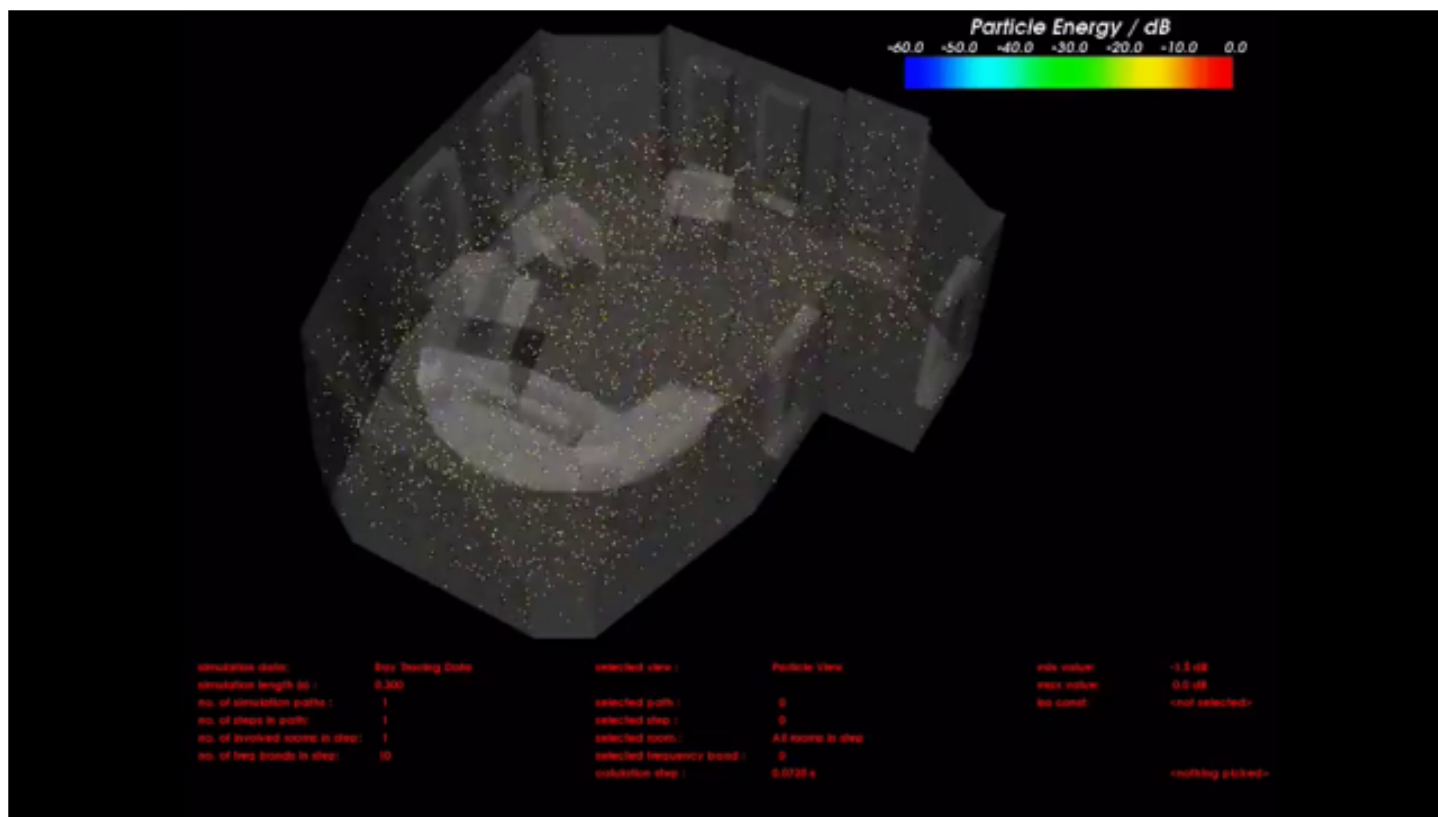
[to video sequence search](#)
(within Digital Signal Processing II.)



[to video](#)

Center for Digital Education. More educational support material here:

<https://exoset.epfl.ch/booc/>



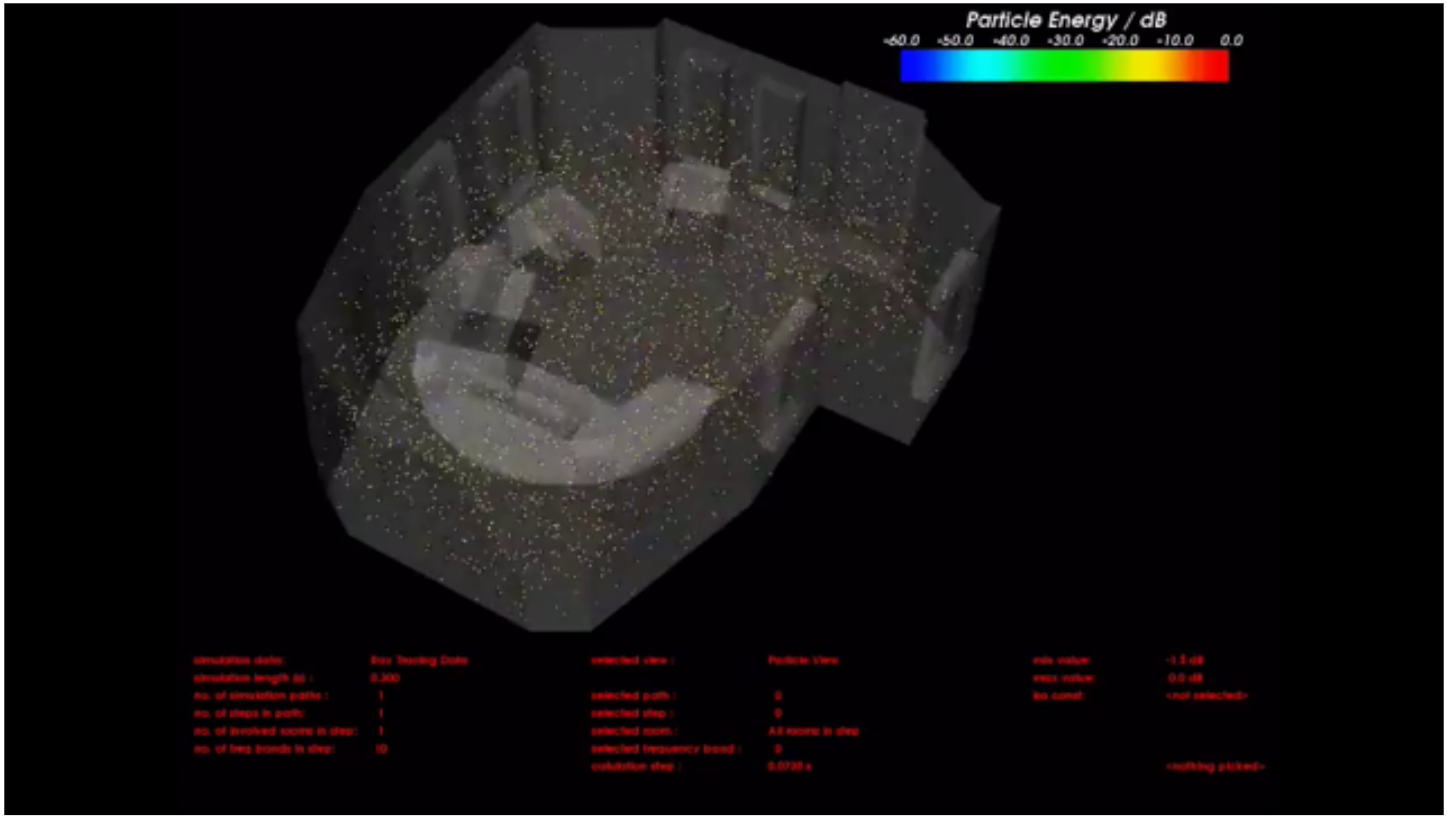
In today's Signal of the Day we have a look at an important problem in room acoustics, entitled Can one hear the shape of a room? Let us first investigate how sound propagates in a room between a sender or a loudspeaker and the receiver or a microphone R, which is illustrated in the following figure. The first sound to reach the microphone corresponds to the direct transmission path depicted in orange. But together with this direct transmission, echoes from the walls also arrive at the microphone, depicted in purple here. The walls act essentially as mirrors. Every time a sound hits a wall, part of the sound is reflected in such a way that the incidence angle equals the reflection angle. These angles are defined with respect to the normal of the walls, n_I and n_J respectively. Part of the signal is also absorbed by the wall, which depends on the type of material the wall is made of. The energy absorbed by the wall is transformed into heat, which increases the wall's temperature very slightly. This phenomenon is generally referred to as diffuse reflection. These echoes act as virtual sources of the original source. For example, here is a virtual source which corresponds to the reflection with respect to this wall, and here is a second virtual source corresponding to the reflection given here. Echoes generate themselves new echoes from other walls that we call higher generation echoes. And this goes on and on until the energy of the signal is fully absorbed. The line in green represents, for example, a second generation echo. First echo here, this point, second here,

notes

summary

0m 0s





and this corresponds to a second order image source until it reaches the microphone. We can now look at the simulation of how sound propagates inside a complex room. This cool video was produced by one of our lab members, Dirk Schroeder, and illustrates how a pulse emitted at a loudspeaker propagates in a room until all the pulse energy is dissipated.

notes

summary

- ▶ Linear model
 - Sound level in room acoustics is low
 - Linear model is good approximation
 - Entirely characterized by impulse response, i.e. response to Dirac impulse
- ▶ Room impulse response
 - Describe audio channel between sender S and receiver R
 - Sum up effect of direct path transmission and subsequent attenuated echoes

3

From our explanation, it seems that sound propagation is a fairly complex phenomenon, involving the superposition of a signal and its different echoes,

notes

summary

2m 25s



- ▶ Linear model
 - Sound level in room acoustics is low
 - Linear model is good approximation
 - Entirely characterized by impulse response, i.e. response to Dirac impulse
- ▶ Room impulse response
 - Describe audio channel between sender S and receiver R
 - Sum up effect of direct path transmission and subsequent attenuated echoes

3

such that the derivation of a model might not be very simple.

notes

summary

2m 37s



- ▶ Linear model
 - Sound level in room acoustics is low
 - Linear model is good approximation
 - Entirely characterized by impulse response, i.e. response to Dirac impulse
- ▶ Room impulse response
 - Describe audio channel between sender S and receiver R
 - Sum up effect of direct path transmission and subsequent attenuated echoes

3

Nevertheless, in room acoustics, sound levels are low, such that a linear model is a good approximation.

notes

summary

2m 42s



- ▶ Linear model
 - Sound level in room acoustics is low
 - Linear model is good approximation
 - Entirely characterized by impulse response, i.e. response to Dirac impulse
- ▶ Room impulse response
 - Describe audio channel between sender S and receiver R
 - Sum up effect of direct path transmission and subsequent attenuated echoes

3

From what we have just learned in class, a linear time invariant system is entirely characterized by its impulse response, by definition, the response of the system to an impulse signal. We introduce by analogy the concept of a room impulse response that describes the audio channel between a sender S and a receiver R. The room impulse response simply sums up the direct pass transmission and the subsequent attenuated echoes. Let us now look at different room impulse responses.

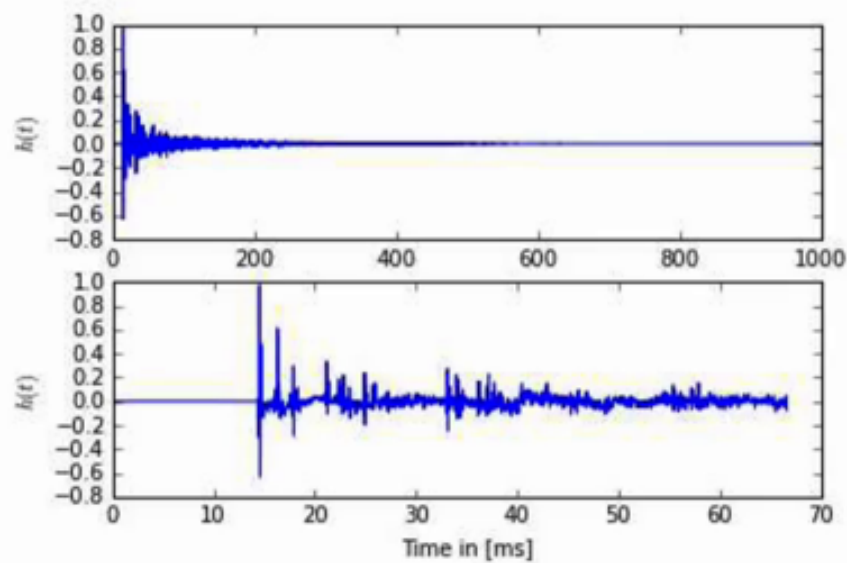
notes

summary

2m 52s



Example of RIR, Conference Room



play stop

4

For example, the one of our lab conference room. Here we plot on the top the first second of the room impulse response,

notes

summary

3m 25s





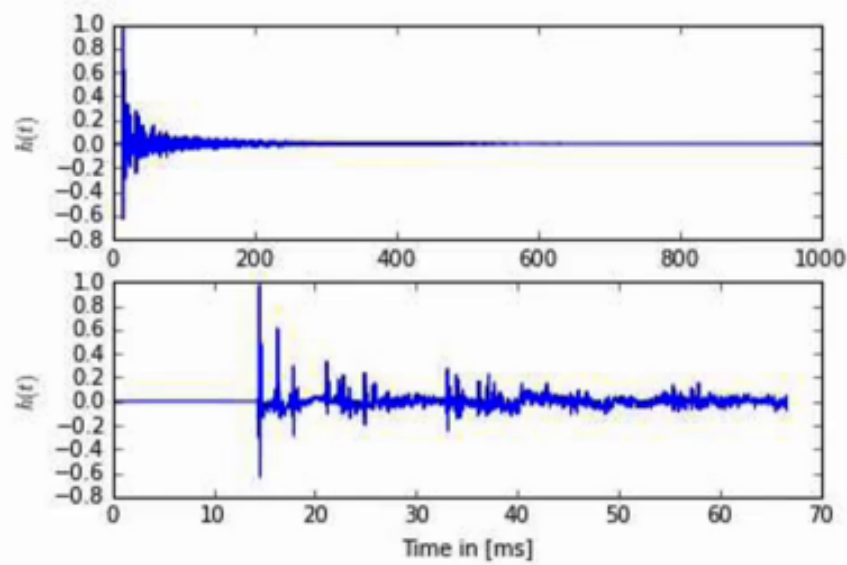
notes

summary

3m 37s



Example of RIR, Conference Room



play stop

4

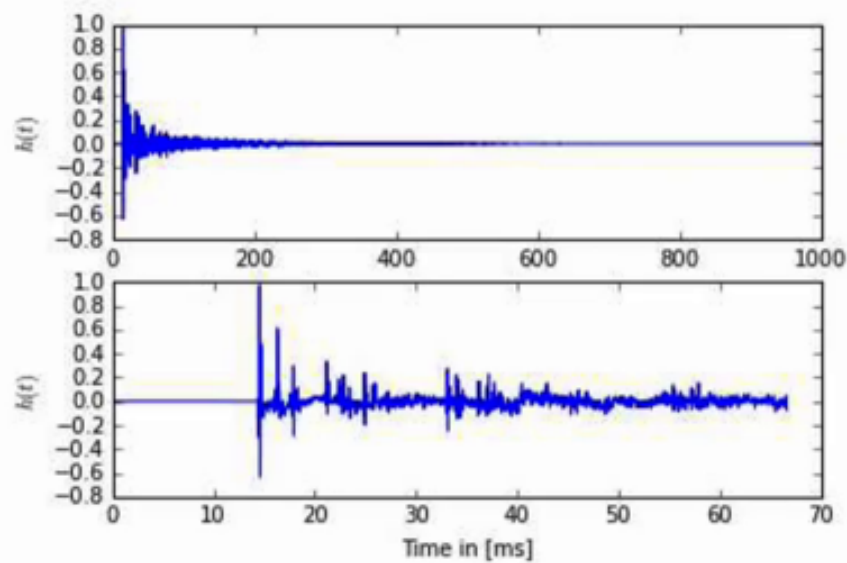
But non-flat fading, diffusivity of reflections, and the diffraction all unite

notes

summary

3m 49s





play stop

4

in the latter part of the room impulse response to create noise-like exponential decay. Of course, an equally important contribution is from the fact that echoes get denser and denser as time passes. The rate of the decay varies wildly between different rooms and is linked to the so-called reverberation time, or T_{60} , which is defined as a time for reflections of a direct signal to decay by 60 dB. It varies from around a second in a speech-only lecture hall to about two seconds for a general purpose auditorium, to, for example, more than eight seconds in the Notre-Dame Cathedral in Paris. It is a very good idea to listen to room impulse responses. By doing that, we can really appreciate the difference between various acoustic spaces.

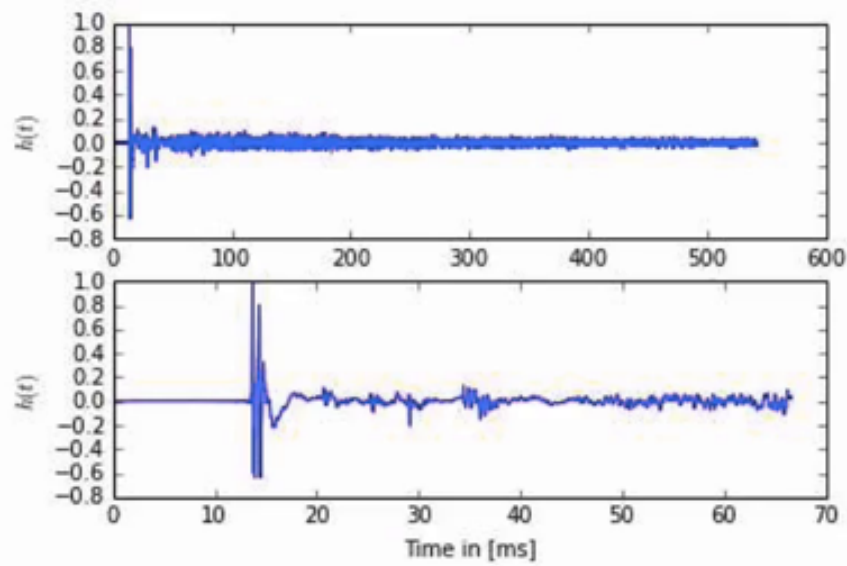
notes

summary

3m 56s



Example of RIR, Lausanne Cathedral



play stop

5

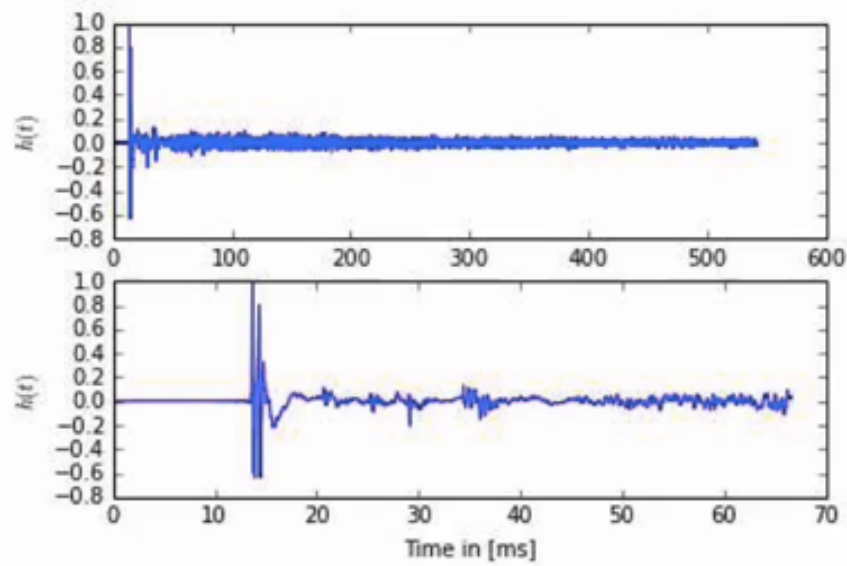
Let us look at another room impulse response, that of the Lausanne Cathedral.

notes

summary

4m 56s





play stop

5

Again, we plot the first second of the room impulse response on the top, and we also zoom in the first 70 milliseconds at the bottom. Again, let's listen to it. Measured or synthetic room impulse responses can be used to recreate

notes

summary

5m 1s





- ▶ Anechoic sound [play](#) [stop](#)
- ▶ Conference room [play](#) [stop](#)
- ▶ Lausanne cathedral [play](#) [stop](#)

the acoustic impression of the corresponding space. If we convolve a room impulse response of a cathedral with a sound sample measured in a room without echo, the resulting signal will sound as if it was recorded in the cathedral. This process is called *oralization*. It is used, for example, in virtual reality

notes

summary

5m 21s





- ▶ Anechoic sound [play](#) [stop](#)
- ▶ Conference room [play](#) [stop](#)
- ▶ Lausanne cathedral [play](#) [stop](#)

to create realistically sounding virtual spaces. Another application example is, of course, in music production.

notes

summary

5m 46s





- ▶ Anechoic sound [play](#) [stop](#)
- ▶ Conference room [play](#) [stop](#)
- ▶ Lausanne cathedral [play](#) [stop](#)

As you will hear, even this simple convolution can evoke correct mental imagery about the space we are oralizing. Consider the following sound sample. It does sound dry. Now let's oralize it. First, let's hear what happens when we convolve this dry sound with the impulse response of our lab conference room. Now let's do the oralization with the cathedral's room impulse response.

notes

summary

5m 50s





- ▶ Anechoic sound [play](#) [stop](#)
- ▶ Conference room [play](#) [stop](#)
- ▶ Lausanne cathedral [play](#) [stop](#)

1

An important practical question is how to measure room impulse responses.

notes

summary

6m 25s





As they are response to an impulse, the first idea is to simply use an impulsive source of sound and record the room response with a microphone. But there are numerous drawbacks to this simple idea. Room impulse responses is the response to an idealized pulse, a Dirac delta function. Realistic impulsive sources are only approximations of the Dirac, and usually not very good ones. Even if we might think that popping a balloon or firing a starter pistol

notes

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

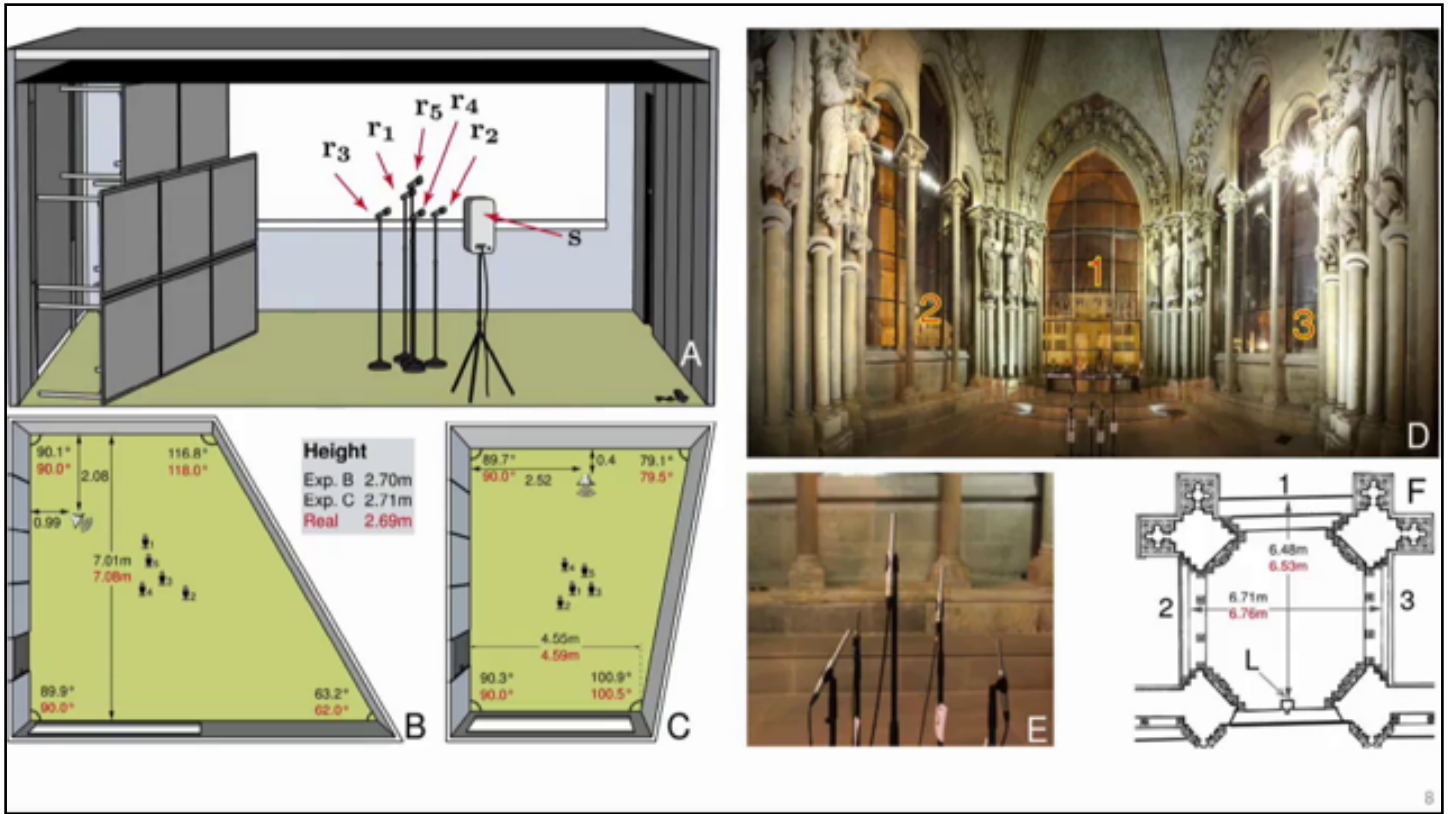
summary

.....

.....

6m 37s





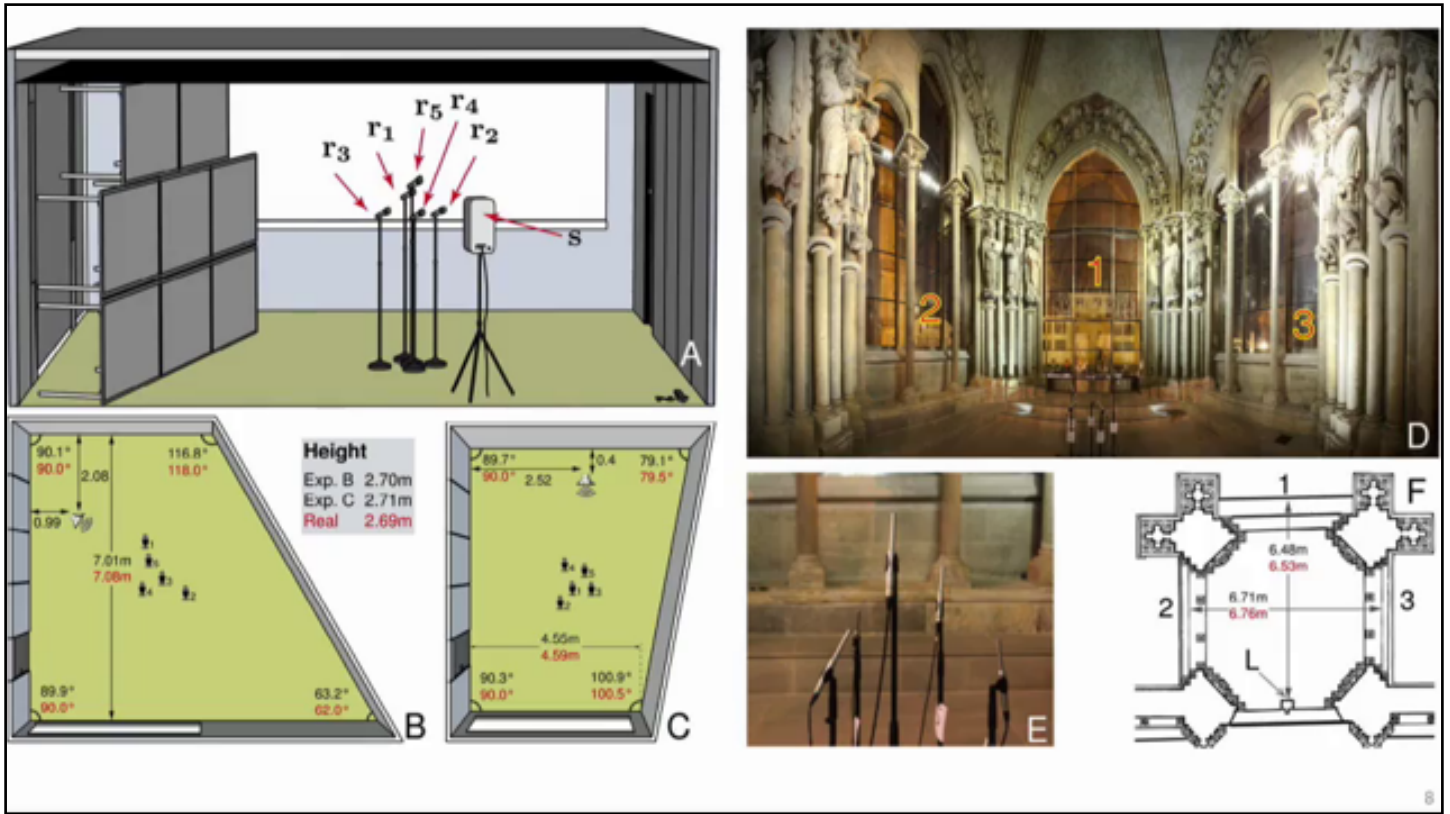
generates a very good impulse, in reality this is not true. Repatibility is also an issue. Popping different balloons generates different sounds. Another way of measuring a room impulse response consists of playing a known sound through the loudspeaker and then deconvolve it in the spectral domain. In the picture you see how we measured the room impulse response of the Lausanne cathedral. We placed a special loudspeaker, a so-called omnidirectional loudspeaker, in the center of the picture, and we played actually what is known as a sine sweep, which is a sinusoid of increasing frequency. It resembles a little bit the sounds that bats emit when they actually do the process of echolocation. But let's not get too much into the details here and come back to our main question of interest. Can one hear the shape of a room? The question is whether we can reconstruct the shape of a room using only echoes. Consider the following experimental setup. We put some microphones into a room. We would like to reconstruct the room's shape from what the microphones record when we clap our hands, that is, from the recorded impulse responses. Room impulse responses contain peaks that correspond to echoes, echoes are modeled by image sources. But if the room is not too complicated, locations of image sources tell us how the room looks like. So we can rephrase the room reconstruction problem into a problem of localizing a certain number of image sources. At the next level of abstraction, we forget that they are image sources, and we just attack a problem of localizing multiple sources emitting impulses.

notes

summary

7m 14s





In our lab, we developed the theory for reconstructing the shape of a room from impulse responses, and we ran some experiments to test it. The first round of experiments was run in our conference room, mentioned already multiple times. This is shown in the left part of the image, and you see the model of the room at the bottom. We built a wall out of tables so we could move around the room and change its geometry. We then added some microphones and loudspeakers, and we were ready to go. Here are the results for two different room configurations. In both experiments, all dimensions of the room were estimated to high accuracy using five microphones. For example, the height of the ceiling is estimated within an error of only one or two centimeters. Encouraged by the success of this experiment, we went straight for the major league. We did experiments in the Lausanne Cathedral. Of course, we didn't really hope to reconstruct the detailed geometry of the cathedral. That's because cathedrals, with their curved surfaces, domes, decorations, and columns, are much more complicated than the rooms we developed our algorithm for. Nevertheless, we were able to reconstruct the main dimensions of an alcove at the same time, and we were able to build a cathedral with quite good accuracy.

notes

summary