THE PLENACOUSTIC FUNCTION AND ITS SAMPLING

Thibaut Ajdler [*] *and Martin Vetterli* [*◇]

[*] Laboratory for Audio-Visual Communications (LCAV)
Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland

[◇] Department of EECS, University of California at Berkeley, Berkeley CA 94720, USA

Email: {thibaut.ajdler, martin.vetterli}@epfl.ch

ABSTRACT

In this paper, we study the spatialization of the sound field in rooms, in particular the evolution of the room impulse responses in function of their spatial positions. Thanks to this study, we are now able to predict the sound field in any position knowing the sound field in a certain number of positions in the room. If enough measurements are taken, we are able to completely characterize the sound field of the room at any arbitrary location. The existing techniques usually make use of models of the room in order to recreate the sound field present at some point in the room. The models take in consideration the different walls materials, obstacles present in the room, etc. Our technique simply starts from the measurements of the room impulse response in a finite number of positions and starting from this information the total sound field is then recreated. Further, we determine the number and the spacing between the microphones needed to reconstruct the sound field in a room up to a certain temporal frequency. We give a link between the temporal frequency of the sound and the spatial frequency of the sensors (microphones) in order to reconstruct the sound field.

1. INTRODUCTION

We introduce a function called the plenacoustic function. The plenacoustic function characterizes the sound field in space, e.g. inside a room. The idea of the plenacoustic is similar to the one introduced by the plenoptic function already studied in the image processing world [1]. There, the idea is to consider the light field in the neighborhood of a scene. The plenoptic function records the light intensity of all the rays of light in this neighborhood. The plenoptic function depends on a lot of different parameters:

$$P_o(x, y, z, \theta, \phi, \lambda, t)\,.$$

First, the plenoptic function depends on the point of view from where we look at the scene given by the 3 coordinates of the observer (x, y, z). It also depends on the direction of the light ray arriving at that point. The direction of the ray of light is characterized by the two angles θ and ϕ. If the scene is colored we also need to know the wavelength of each ray. Finally, if the scene is time variant, we still need to add the time information. We realize thus that the plenoptic function would require a huge amount of memory to store all this data. Sampling of the plenoptic function is then studied in [5, 10].

We study a similar idea in the area of audio. We focus on the plenacoustic function in rooms. Therefore, we are interested in the room impulse responses to be able to characterize what an ear would hear at any point in the room. Knowing this information, we are then able to recreate the sound that one would hear at any point in the room. Convolution of the sound produced by the source at some point in the room with the room impulse response from the source's position to the listener's position gives precisely what the sound is at the listener's position. Our plenacoustic function is thus parameterized by the following factors:

$$P_a(S\ and\ R\ position,\ characteristics\ of\ room,\ time)$$

with S and R the source's and receiver's positions and the characteristics of the room being for example the reflection factors of the walls of the room.

We then answer the following question: "How many microphones do we need to place in the room in order to completely reconstruct the sound field at any position in the room?"

The name of the plenacoustic function has been for the first time mentioned in [7]. We have developed a mathematical derivation of the plenacoustic function and a more advanced study of the sampling and reconstruction in [2].

The outline of this paper is as follows. Section 2 introduces the model used for the room impulse responses. We explain the construction of the plenacoustic function in Section 3. Further, sampling of the plenacoustic function is discussed in 3.2. We then briefly discuss in Section 3.3 how the reconstruction of the plenacoustic function can be realized. Section 4 shows the experimental results obtained by measuring the room impulse responses in a room. The conclusions are finally drawn in Section 5.

2. ROOM IMPULSE RESPONSE

In order to calculate the plenacoustic function, we need to find the room impulse responses in different points of the room. We use the image method discussed in [3, 6] for the room impulse responses in the case of rigid and non-rigid walls.

2.1. Rigid walls

We consider a rigid wall as a wall where the air particles have zero normal velocity at its vicinity. Considering a source, the effect of the presence of a rigid wall can be replaced by placing an image of the source symmetrically on the far side of the wall as shown in Fig. 1. This can be generalized in the case of three rectangular adjacent walls. In this case, we will get 7 virtual image sources. In a case where two rigid walls are parallel, we get an infinite number of virtual image sources, each image being itself imaged. This is a similar effect as the one obtained when one looks at oneself between two parallel mirrors. Finally, in the case of a rectangular rigid-walls room of size (L_x, L_y, L_z), we get in total 8 sources (1 source and its seven virtual images) being each infinitely imaged by the other parallel walls. This leads to:

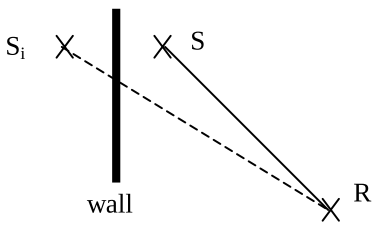

Figure 1: *Real and virtual sources (respectively S and S_i) modelling the presence of a rigid wall for the receiver R.*

$$p(t, S, R) = \sum_{p=0}^{7} \sum_{v=-\infty}^{\infty} \frac{\delta[t - |d_p + d_v|/c]}{4\pi |d_p + d_v|}\,, \qquad (1)$$

with (n, l, m) being an integer vector triplet, $d_p = (x_s \pm x_r, y_s \pm y_r, z_s \pm z_r)$ and $d_v = (2nL_x, 2lL_y, 2mL_z)$. The room impulse response is a function of time and is dependent on the source $S = (x_s, y_s, z_s)$ and the receiver's position $R = (x_r, y_r, z_r)$. The virtual sources are shown in a 2 dimensional case in Fig. 2. It can also be shown that the expression (1) gives us also the exact solution for the wave equation in a rectangular, rigid-wall room.

2.2. Non-rigid walls

The walls of a room are not always perfectly rigid. In [3], another expression is given for the room impulse response taking into account the different reflection factors of the walls. The reflection factors of the walls are considered as real. The expression is

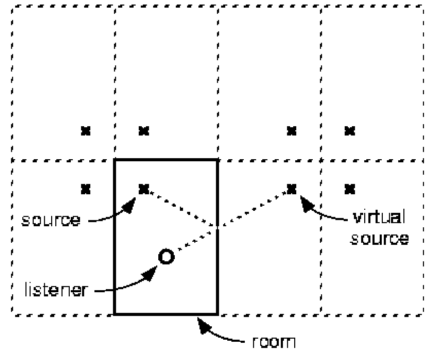

Figure 2: *Room with a listener, a real source and its virtual images.*

$$\begin{aligned} p(t, S, R) \;\; &= \;\; \sum_{p=0}^{1} \sum_{r=-\infty}^{\infty} \beta_{x1}^{|n-q|} \beta_{x2}^{|n|} \beta_{y1}^{|l-j|} \beta_{y2}^{|l|} \beta_{z1}^{|m-k|} \beta_{z2}^{|m|} \\ &\quad \times \frac{\delta[t - |d_p + d_r|/c]}{4\pi |d_p + d_r|}\,, \end{aligned} \qquad (2)$$

where $p = (q, j, k)$, $r = (n, l, m)$ and $d_r = (x_s - x_r + 2qx_r, y_s - yr + 2jy_r, z_s - z_r + 2kz_r)$ A more complete study of the room impulse response is given in [8].
In some application we can put a limit to the length of the impulse response (e.g. 2 seconds is generally enough for small rooms). In this case we can then calculate the distance travelled by a sound wave in this time. If we consider a sphere around the receiver's position with radius being the calculated distance, all the virtual sources located outside this sphere will not have any influence on the considered room impulse response. This transforms the infinite summations in (1) and (2) into finite summations.

3. PLENACOUSTIC FUNCTION

3.1. Construction of the plenacoustic function

We want to know the room impulse responses in every point of a room. Let us consider (without loss of generality) that we study the plenacoustic function along a line in the room. We can construct a 3-dimensional graph formed by all the room impulse responses placed one next to the other along one axis. In Fig. 3, the two base axes are space and time. Space corresponds to the listener's position, while the time axis corresponds to the duration of the impulse responses. The third axis is the amplitude (in dB) of the room impulse responses. Studying the plenacoustic function along a line in the room would require to store the room impulse responses for every single point along this line. Furthermore,

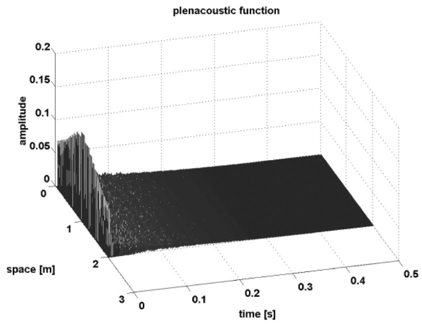

Figure 3: *Plenacoustic function along a line in a room.*

the time continuous impulse responses should be known. This data is impossible to store, therefore we need to sample the plenacoustic function.

3.2. Plenacoustic sampling

In order to sample the plenacoustic function, we need to sample the room impulse responses at a certain temporal sampling rate depending on the audio bandwidth ($44,1$ kHz for CD quality recordings). Further, by taking an evenly spaced finite number of impulse responses, we uniformly sample the plenacoustic function in space. When the plenacoustic function is sampled, repetitions of the spectrum occur [9]. The 2-dimensional spectrum will here be repeated along the temporal and the spatial frequency axes.
By taking the 2-dimensional Fourier transform (2D-FT) of this obtained sampled plenacoustic function, we get Fig. 4. In this figure, the two base axes represent the correspond-

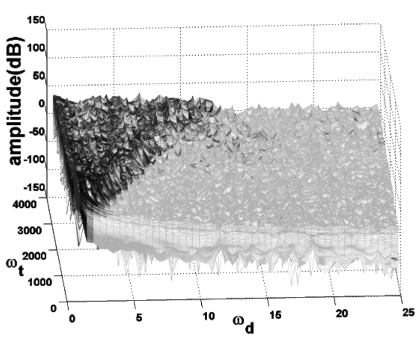

Figure 4: *Two dimensional Fourier transform of the plenacoustic function.*

ing axes of time and space, being a temporal and a spatial frequency. Temporal and spatial pulsations are defined respectively as

$$\omega_t = \frac{2\pi}{\Delta t} \tag{3}$$

$$\omega_d = \frac{2\pi}{\Delta d} \tag{4}$$

with Δt the sampling period of the impulse response and Δd the sampling interval between the different positions of the measured impulse responses. The third axis is here the amplitude (in dB) of the 2D-FT. We clearly see a triangular shape in the 2D-FT of the plenacoustic function. We can furthermore observe that the part outside of the triangle is of much lower amplitude than the part inside the triangle. The contents of the 2D-FT of the plenacoustic function can be schematized as in Fig. 5, where ω_{tmax} corresponds to half the temporal sampling frequency. In this figure, we can

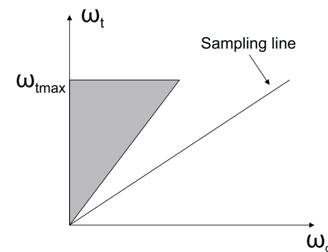

Figure 5: *Schematic view of the plenacoustic function in frequency domain.*

see that the spatial frequency support grows for increasing temporal frequencies. Spatial sampling of the plenacoustic function at some spatial sampling frequency will lead to aliasing for all the temporal frequency above some corresponding temporal frequency. Perfect reconstruction becomes impossible above this temporal frequency. Similarly, the sampling line at the right of the triangle in Fig. 5 shows for any temporal frequency, the minimal spatial sampling frequency to avoid aliasing. It can be shown that the equation of this line is:

$$\omega_d = \frac{2\omega_t}{c}\,, \tag{5}$$

with c the sound propagation speed in the air.

3.3. Reconstruction

Knowing the sound field in every point of the sampling grid, we apply the usual interpolation techniques [9] in order to reconstruct the sound field at any location. First, we need to upsample our time domain signal accordingly to the desired location. We then filter the upsampled plenacoustic function with an appropriate 2-dimensional filter. The value at the location of interest is then obtained by interpolation.

4. EXPERIMENTS

After observing this phenomenon by simulating different room impulse responses with different reflection factors using (2), we made some experiments in order to check if the triangle shape of the 2D-FT of the plenacoustic function was also observed with measured room impulse responses.

4.1. MLS theory

We used the Maximum Length Sequences (MLS) [4] technique in order to calculate the room impulse responses. A MLS is a periodic pseudo-random binary sequence. Consider a system with an input $x(k)$, a linear system under test characterized by an impulse response $h(k)$ and the output of the system $y(k)$. The cross-correlation between in- and output is given by:

$$R_{xy}(k) = R_{xx}(k) * h(k)\,. \tag{6}$$

A MLS sequence has the property of having its autocorrelation being essentially an impulse. This impulse is represented by a Dirac function (the approximation gets more valid for longer sequences). Therefore we obtain that

$$R_{xy}(k) \approx h(k)\,. \tag{7}$$

We can thus measure the impulse response of a linear system by calculating the cross-correlation between the MLS and the output signal. The MLS technique is highly immune to noise.

4.2. Measurements

The measures were made in a sound insulated room (LCAV audio room at EPFL). 40 measures were taken with 1.8 cm spatial spacing. As seen in (5), we can easily calculate that for our spatial spacing we should be able to reconstruct the signal up to a frequency around 9.2 kHz. With these measurements, the obtained 2D-FT of the plenacoustic function is shown in Fig. 6. In this figure, we show the whole spatial axis until the sampling frequency (and not until the Nyquist frequency) to see clearer until where we get the triangle shape and when the aliasing begins. We can see in Fig. 6 that the aliasing begins around 9 kHz. We get thus a good

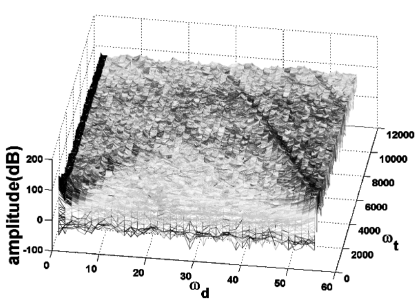

Figure 6: *Measured plenacoustic function in frequency domain.*

correspondence between theoretical and practical results. Remark that perfect precision cannot be achieved in the measurements. The uniform sampling cannot be perfectly applied at 1.8 cm but it is really interesting to still observe the triangular shape in the Fourier transform of the measured plenacoustic function.

5. CONCLUSION

In this article we have introduced the plenacoustic function. The plenacoustic function characterizes the sound field at any place in a room. We then studied the sampling of this function and explained how to reconstruct the sound field between the microphones using interpolation techniques. We are now able to predict the number of microphones to place in a room to perfectly reconstruct the sound field at any point in the room up to a certain temporal frequency. We also showed the relation between the critical spacing of the microphones and the corresponding maximal frequency before aliasing occurs. We presented finally experimental results and compare theoretical results with measurements.

6. REFERENCES

[1] E.H. Adelson and J.R. Bergen. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, pages 3–20. MIT Press, 1991.

[2] T. Ajdler and M. Vetterli. The plenacoustic function, sampling and reconstruction. *IEEE ICASSP, submitted*, 2003.

[3] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *JASA*, 65, 1979.

[4] J. ”Borish and J.B. Angell. An efficient algorithm for measuring the impulse response using pseudorandom noise. *J. Audio Eng. Soc*, 31(7), 1983.

[5] J.X. Chai, X Tong, S.C. Chan, and H.Y. Shum. Plenoptic sampling. In *Proceedings of the conference on Computer graphics*, pages 307–318, 2000.

[6] W.G. Gardner. The virtual acoustic room. Master's thesis, MIT, 1992.

[7] M. Kubovy and D. Van Valkenburg. Auditory and visual objects. *Cognition*, 80:97–126, 2001.

[8] P.M. Morse and K.U Ingard. *Theoretical Acoustics*. McGraw-Hill, 1968.

[9] M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Signal Processing. Prentice Hall, A.V. oppenheim edition, 1995.

[10] C. Zhang and T. Chen. Generalized plenoptic sampling. Technical Report Amp 01-06, Carneggie Mellon University, September 2001.