Interpolation of room impulse responses in 3d using compressed sensing

Rémi Mignot, Laurent Daudet, Francois Ollivier

Interpolation of room impulse responses in 3d using compressed sensing

R. Mignot[a], L. Daudet[a] and F. Ollivier[b]

[a]Institut Langevin " Ondes et Images "", 10 rue Vauquelin, 75005 Paris, France
[b]UPMC - Institut Jean Le Rond d'Alembert, 2 place de la gare de ceinture, 78210 Saint Cyr L'Ecole, France
remi.mignot@espci.fr

In a room, the acoustic transfer between a source and a receiver is described by the so-called "Room Impulse Response", which depends on both positions and the room characteristics. According to the sampling theorem, directly measuring the full set of acoustic impulse responses within a 3D-space domain would require an unreasonably large number of measurements. Nevertheless, considering that the acoustic wavefield is sparse in some dictionaries, the Compressed Sensing framework allows the recovery of the full wavefield with a reduced set of measurements (microphones), but raises challenging computational and memory issues. In this paper, we exhibit two sparsity assumptions of the wavefield and we derive two practical algorithms for the wavefield estimation. The first one takes advantage of the Modal Theory for the sampling of the Room Impulse Responses in low frequencies (sparsity in frequency), and the second one exploits the Image Source Method for the interpolation of the early reflections (sparsity in time). These two complementary approaches are validated both by numerical and experimental measurements using a 120-microphone 3D array, and results are given as a function of the number of microphones.

1 Introduction

The acoustic properties of a reverberating room can be given by analysing its *Room Impulse Responses* (RIRs). In [1], the concept of *Plenacoustic Function* (PAF) is introduced. This function gathers all RIRs of the room, and therefore it depends on time, on the source position, on the receiver position and on the room characteristics (geometry and wall properties).

On one hand, in some applications the effect of room reverberation is undesirable and acoustic echo cancelers are used to estimate the anechoic sound. On the other hand, reverberation plays an important role in auditory scene synthesis, in virtual reality framework for example. In both cases, knowing the whole set of RIRs in a given room could potentially be used to improve their performance.

Measuring the PAF is fundamentally a sampling problem: from a limited number of point measurements, the goal is to reconstruct (i.e. interpolate) the acoustic wavefield at any position in space and at any time.

1.1 Sampling of the Plenacoustic Function

Standard acquisition of signals relies on a regular sampling of space and time with respect to Shannon-Nyquist theory. Basically, to sample a band limited signal between 0 and f_c [Hz], the sampling rate F_s has to be chosen higher than $2f_c$. In the same way, the space sampling has to be dense enough to avoid aliasing in reconstruction.

Considering a fixe source and a 3-D regular microphone array, the Nyquist criterion is satisfied if (cf. [1]):

$$\delta_v < \frac{c_0}{2f_c}, \ \forall v \in \{x, y, z\}, \tag{1}$$

where δ_x, δ_y and δ_z are the sampling steps along the 3 axes. Physically, this criterion states that the sampling steps must be smaller that the half of the minimal wavelength $\lambda = c_0/f_c$.

Unfortunately, the measurement of a time varying 3-D image requires a too high number of microphones to be realized as such in practice. For example, in order to reconstruct the PAF inside a cube with side 2 m, and with $f_c = 6$ kHz, we need at least 340.10^3 point measurements (microphones).

Nevertheless, in [2] it is shown that the RIR part above a certain *Schroeder frequency* (cf. [3]) and after a certain *mixing time* (cf. [4]) behaves like a colored, modulated noise and can be modeled and simulated using a stochastic model (part C in fig. 1). Consequently, to reduce the number of required microphones, we propose to sample only the hatched part of fig. 1. For example, in [5] the low frequencies of the RIRs

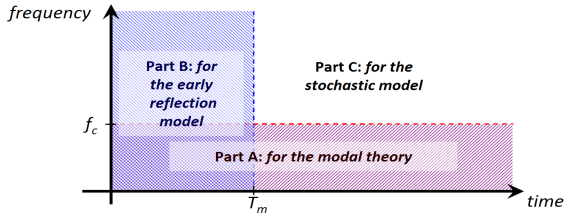

Figure 1: Validity parts of the models. Here f_c and T_m are associated to the Schroeder frequency and the mixing time respectively

(part A) are reconstructed using a model based on the modal theory, and in [6] a method based on Dynamic Time Warping is proposed for the interpolation of the early part of the RIRs (part B). But these 2 methods have been developped for the interpolation of the RIRs on a line, and here we want to sample the PAF in a 3-D domain.

In this work, we propose a method to sample and to interpolate in 3-D the parts A and B separately (cf. fig. 1), using the same measurements. In order to reduce the number of microphone in the array, an interesting approach is the *Compressed Sensing* framework which is briefly described below.

1.2 Compressed Sensing

The problem consists in the reconstruction of a signal $y \in \mathbb{R}^N$ from M observations x_m, linked by the linear system $x = \Phi y$. *Compressed Sensing* (CS) deals with the underdetermined case, for which there are more unknowns than equations ($N > M$), cf. e.g. [7, 8]. As such a problem cannot be solved without additional hypothesis, the underlying idea is that if y lives in a subspace of dimension K and with basis ψ, for $K < M$, we can solve $y = \psi a$ writing $x = \Phi y = \Phi\psi a = \theta a$. But, in general we do not know ψ.

Then, we define L vectors ψ_l, forming the matrix Ψ with $L \gg K$, and we look for a basis which explains y. In other words, we look for a vector $\alpha \in \mathbb{R}^L$ K-sparse (where no more than K coefficients are non-zero), such that $y = \Psi\alpha$. Unfortunately this problem is not convex and difficult to solve. However, we can change it into a convex problem by considering the following *Basis Pursuit Denoising* approach:

$$\min_{\alpha \in \mathbb{R}^L} \|\alpha\|_{\ell_1} \quad \text{subject to} \quad \|x - \Phi\Psi\alpha\|_{\ell_2} \leq \varepsilon, \tag{2}$$

where the norm ℓ_n is given by $\|y\|_{\ell_n} = (\sum_i |y_i|^n)^{1/n}$, and ε is a fidelity parameter. A high ε allows a stronger sparsity of α, and a small ε improves the reconstruction of y.

Some theoretical results (cf. e.g. [9, 10]) give a sufficient condition for reconstructing y in the case of sparse signals,

by the so-called Restricted Isometry Property (RIP). It quantifies how Φ and Ψ are mutually incoherent with respect to their use on sparse signals. In practice, the RIP is difficult to compute, but it is verified with high probability for some random sampling matrices. This encourages the use of randomly selected observation points in practice, which are here the microphone positions in the 3-D space.

In this work we will consider the *Modal Theory* and the *Image Source Method* in order to justify the use of CS for the parts A and B respectively (cf. fig. 1). We will see that these 2 deterministic models give 2 sparse representations of the PAF in their respective part: sparsity in frequency for part A, and sparsity in time for part B.

1.3 Outline

The outline of this paper is as follows. In section 2, considering two acoustic models consecutively, we show that the two parts of the PAF (low frequencies and early reflections) can have a sparse representation in some respective dictionaries, and so that the Compressed Sensing framework can be used. For these two models, we briefly present the derived algorithms of decomposition. Then, in section 3, we present some experimental results and some comparisons. Note that this paper focuses on the experimental results of the proposed approaches, for more details about the algorithms, see [11, 12]. Finally, in section 4, we conclude this paper.

2 Modelling and CS algorithms

2.1 Modal Theory

Here, we proposed two methods based on the modal theory for the interpolation of the PAF in low frequencies.

2.1.1 Structured Sparsity

Considering linear acoustic propagation away from the sources, the acoustic pressure $p(t,\vec{X})$ is governed by the wave equation $\mathrm{c_0}^2\,\Delta p(t,\vec{X}) - \partial_t^2 p(t,\vec{X}) = 0$, where $\Delta = \nabla^2$ is the laplacian operator and ∂_t is the time derivative. Assuming a modal behavior (at low frequencies) for closed rooms with ideally rigid walls, the solution can be decomposed as a discrete sum of complex harmonic signals with the angular frequencies ω_q:

$$p(t,\vec{X}) = \sum_{q\in\mathbb{Z}^\star} A_q\,\phi_q(\vec{X})\,g_q(t), \tag{3}$$

where $g_q(t) = \mathrm{e}^{j\omega_q t}$, ϕ_q is the modal shape of the mode q and A_q is a related complex amplitude. With the wavenumber $k_q = \omega_q/\mathrm{c_0}$, we get the Helmholtz equation for every mode: $\Delta\phi_q + k_q^2\phi_q = 0$.

In the Helmholtz equation, ϕ_q is the eigenmode of the laplacian operator with eigenvalue $-k_q^2$. If the room is star-shaped, previous studies [13, 14] have shown that an eigenmode of the laplacian with a negative eigenvalue can be approximated by a finite sum of plane waves incoming from various directions, and sharing the same wavenumber k_q. Then

$$\phi_q(\vec{X}) \approx \sum_{r=1}^{R} a_{q,r}\,\mathrm{e}^{j\vec{k}_{q,r}\vec{X}} \tag{4}$$

is the R-order approximation of ϕ_q, with $\vec{k}_{q,r}$ the 3-D wavevector r of the mode q, such that $\|\vec{k}_{q,r}\|_2 = |k_q|$.

In the case of non rigid walls, the modes are damped in time, k_q now has an imaginary part: $k_q = (\omega_q - j\xi_q)/\mathrm{c_0}$, where $\xi_q < 0$ is the damping coefficient. Therefore, $g_q(t)$ of eq. (3) becomes: $g_q(t) = \mathrm{e}^{jk_q\,\mathrm{c_0}\,t} = \mathrm{e}^{\xi_q t}\,\mathrm{e}^{j\omega_q t}$. In theory, these losses modify ϕ_q, nevertheless we assume that the approximation (4) remains valid, at least for $\vec{X}$ far from the walls.

Consequently, considering a finite frequency range $[0,\omega_c]$ containing Q real modes, or equivalently $2Q$ complex modes, and considering R-order approximations of the ϕ_q's, the PAF $p(t,\vec{X})$ can be approximated by a sum of $2QR$ damped harmonic plane waves, $\exp(j(k_q\,\mathrm{c_0}\,t + \vec{k}_{q,r}\vec{X}))$.

Now, taking advantage of this *Structured Sparsity*, and starting from the sampled signals $p(t_n,\vec{X}_m)$ of a array of M microphones at $\vec{X}_m$ covering a 3-D domain of interest Ω (a finite convex domain within the room), we present an algorithm previously proposed for the near-field acoustic holography of plates [15]:

(a) The shared wavenumbers k_q are estimated using a joint estimation of damped sinusoidal components (cf. [16]).

(b) The operational shapes ϕ_q of (3) are estimated using the projection of the measured signal onto a basis formed by the damped exponentials $\mathrm{e}^{jk_q\,\mathrm{c_0}\,t}$.

(c) Every ϕ_q is approximated using a finite sum of plane waves sharing the same wavenumber k_q (cf. eq. (4)).

(d) Finally, the PAF can be interpolated $\forall t \in [0, N/Fs]$ and $\forall\vec{X}\in\Omega$ using $\tilde{p}(t,\vec{X}) = \sum_q \tilde{\alpha}_{q,r}\,\mathrm{e}^{j(k_q\,\mathrm{c_0}\,t+\vec{k}_{q,r}\vec{X})}$, where $\tilde{\alpha}_{q,r}$ are the coefficients optimised in stage (c).

2.1.2 Modal analysis in a rectangular room

In this section, we study the solutions of the wave equation in the simple case of a rectangular room. From this study, we exhibit a stronger property of sparsity which justifies the use of the Compressed Sensing framework (CS).

In the case of a rectangular room with rigid walls, we can make the variable separation in cartesian coordinates (x,y,z). Then, each modal shape is written as the product of 3 functions of one variable. With $\vec{X} = [x,y,z]^T$, the PAF becomes:

$$p(t,\vec{X}) = \sum_{q\in\mathbb{Z}^\star} A_q\,F_{xq}(x)\,F_{yq}(y)\,F_{zq}(z)\;\mathrm{e}^{jk_q\,\mathrm{c_0}\,t}. \tag{5}$$

For each mode, these functions verify the 1-D Helmholtz equation $\partial_v^2 F_v + k_v^2 F_v = 0$ for $v\in\{x,y,z\}$. With rigid walls, the k_v's are real constants such that $k_x^2+k_y^2+k_z^2 = k^2$ (cf. [17]). According to the Helmholtz equation, for each cartesian coordinate v the F_v's are the sum of 2 solutions: $F_v(v) = A_v^+\,\mathrm{e}^{jk_v v} + A_v^-\,\mathrm{e}^{-jk_v v}$. Then, expanding $F_xF_yF_z$, the modal shape $\phi_q(\vec{X})$ is written as the sum of 8 plane waves $\mathrm{e}^{\pm jk_x x\pm jk_y y\pm jk_z z} = \mathrm{e}^{j\vec{k}\vec{X}}$, with $\vec{k} = [\pm k_x, \pm k_y, \pm k_z]^T$.

In the case of non rigid walls, as the wavenumber k is complex: $k = (\omega - j\xi)/\mathrm{c_0}$, the k_v's are complex. This implies a slight decrease of the F_v's near the walls. Nevertheless, for $\vec{X}$ far from the walls, we assume that the imaginary part of k_v is negligible, and that $k_x^2 + k_y^2 + k_z^2 = \mathcal{R}_\mathrm{e}(k)^2 = \omega^2/\mathrm{c_0}^2$.

Consequently, in a bandwidth containing $2Q$ complex modes, the PAF can be written as the sum of $16Q$ harmonic plane waves in the case of rectangular rooms. Note that in

the previous section, each modal shape was approximated by R fixed plane waves sampling uniformly the sphere of radius $|\omega|/c_0$, whereas here, with the assumption of rectangular room, only 8 plane waves are required by complex mode.

In this work, the proposed method is based on the *Matching Pursuit* algorithm (cf. e.g. [18]). It consists in iteratively substracting from the signal the group of 8 harmonic planes waves $\exp(j(kt + \vec{k}\vec{X}))$ that best approximates the measured PAF. These plane waves are associated to one mode and share the same wavenumber k.

Although this model doesn't hold for arbitrary geometries (cylindrical rooms for example), it can nevertheless be extended to non rectangular rooms. Indeed, if all walls are plane, we can assume that the modal shapes are still sparse on a dictionary of plane waves. The corresponding wavevectors remain on the sphere of radius $|k|$.

2.2 Image Source Method

The principle of the *Image Source Method* is that an acoustic path involving reflections in an enclosed space can be represented by a straight line path connecting the listener to a corresponding *Virtual Source* (VS) emitting in free space. For a fixed real source at position $\vec{Y}_0$ and a receiver at position $\vec{X}$, the RIR have the general form (cf. e.g. [19]):

$$p(t, \vec{X}) = \sum_{s \in \mathcal{J}} \beta_s \frac{\delta(t - \|\vec{X} - \vec{Y}_s\|_2 / c_0)}{4\pi \|\vec{X} - \vec{Y}_s\|_2}, \qquad (6)$$

where the $\vec{Y}_s$ are the VSs positions, the β_s coefficients take into account the damping caused by the multiple reflections against the walls, and $\mathcal{J}$ is a set indexing the virtual sources.

For the early part $t \in [0, T]$ of the RIRs, we only have to take into account the VSs within a ball of radius $r = T\, c_0$ and center $\vec{X}$. As a consequence, since the RIRs are represented as linear combinations of contributions of a few VSs (cf. eq. (6)), the early part of the 4 dimensional function $p(t, \vec{X})$ has a sparse representation on a dictionary of spherical waves, where the centers are the VSs. Then, to reconstruct $p(t, \vec{X})$ in a finite time interval $[0, T]$ and in a domain Ω of the space, we propose to exploit this sparsity using the CS approach.

As in section 2.1, the present method is based on the *Matching Pursuit* algorithm (cf. e.g. [18]). Here, it iteratively substracts from the measured PAF the atom that best approximates it. This atom is chosen among a dictionary Ψ of monopoles within a virtual space which is a ball of radius $r = Tc_0$ and of center the microphone array. Then, every atom ψ_s of Ψ is given by

$$\psi_s[n, m] = \frac{\delta(t_n - \|\vec{X}_m - \vec{Y}_s\|_2 / c_0)}{4\pi \|\vec{X}_m - \vec{Y}_s\|_2}. \qquad (7)$$

The underlying assumptions of this model are as follows: the walls have to be plane, only specular reflections are considered (no diffuse reflections), diffraction is not taken into account, the domain Ω of analysis and reconstruction has to be small enough in order to avoid a problem of visibility of the VSs according to $\vec{X}$ (cf. [19]), and the imaginary part of the wall reflections (in the frequency domain) have to be sufficiently small. Note that for convex rooms, diffraction can be neglected, and we consider complex wall reflections for a more realistic modeling (cf. [12] for more details).

3 Experiments and results

We have designed a real 3-D array with 120 electret microphones (cf. fig. 2), randomly positioned within a cube of size 2m with a statistical distribution close to uniform - up to mechanical constraints. The room has dimensions (3.8, 8.15, 3.6)m, it was emptied but still had features that made it non-ideal: a doorway, two windows, a cornice, concrete walls, wood panels, etc.

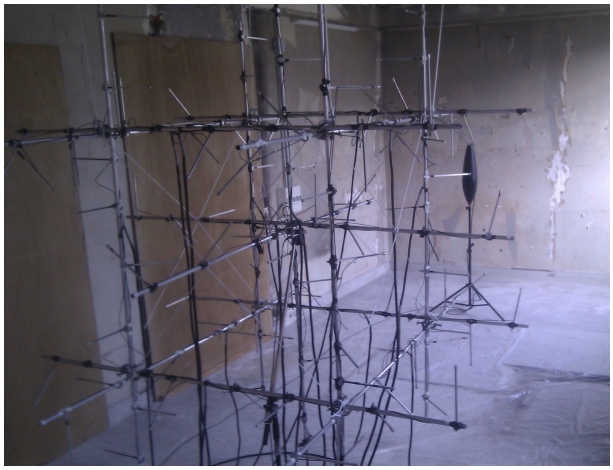

Figure 2: Photograph of the experimental microphone array

Here, we present some results of the proposed algorithms, using the same experiment. They will be named: methods *CS-A1* and *CS-A2* for the reconstruction in low frequencies (cf. sec. 2.1.1 and 2.1.2 respectively), and method *CS-B* for the reconstruction of the early part (cf. sec. 2.2). Interpolation CS-A1 and CS-A2 are compared using the estimation of their Signal-to-Noise Ratio (SNR) [dB], and method CS-B is evaluated using its SNR and its Pearson correlation coefficient c [%]. With s the $(N \times 1)$ vector of the target RIR, such that $s_{[n]} = p(t_n, \vec{X})$, and $\widetilde{s}$ its interpolation: $\text{SNR} = 20 \log(\|s\|_2 \,/\, \|s - \widetilde{s}\|_2)$ and $c = 100\, |\langle s, \widetilde{s} \rangle| \,/\, (\|s\|_2 . \|\widetilde{s}\|_2)$.

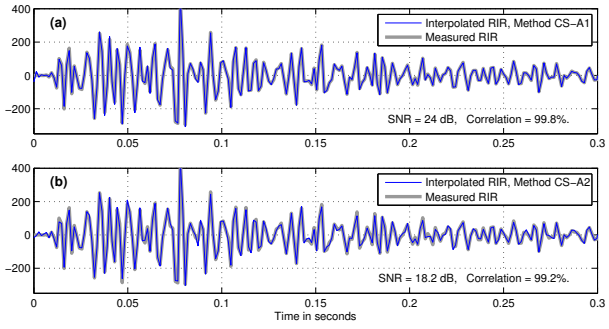

Figure 3: Methods CS-Ax. Example of two interpolated responses in low frequencies: $f_c = 300$ Hz

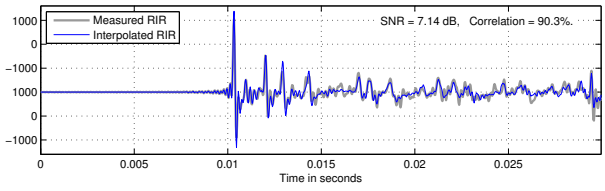

Figure 4: Method CS-B. Example of the interpolation of the early part

Figures 3 and 4 present some examples of the interpolated RIRs, togehter with the measured ones. The analysis is performed using 119 microphones of the array, and the interpolations are evaluated using the measured RIR of the 120th microphone. Remark 1: as opposed to fig. 4, in fig. 3 the contributions of the reflections are not distinguishable because of the low-pass filtering at 300Hz. Remark 2: the RIR of fig. 4 takes into account the non-ideal responses of the microphones and the loudspeaker. Remark 3: whereas the SNRs of methods CS-A1 and CS-A2 are quite good (24dB and 18.2dB respectively), the SNR of method CS-B doesn't seem as good (7.14dB). But we can remark that the correlation remains quite good (90.3%). This point will be discussed in conclusion.

In figure 5, the performances are evaluated according to the number of microphones for the analysis. Here, we have randomly selected 15 microphones for the interpolation and the SNR averages, and the analyses are performed using the 105 other microphones. As a general trend, performance decreases with M. It is however interesting to notice that with $M > 40$ method CS-A1 outperfoms method CS-A2 which suddenly falls for $M < 16$. Nevertheless, between 19 and 28 microphones method CS-A2 is better with a stable SNR (approximately 10dB); this observation has been noticed with other tests.

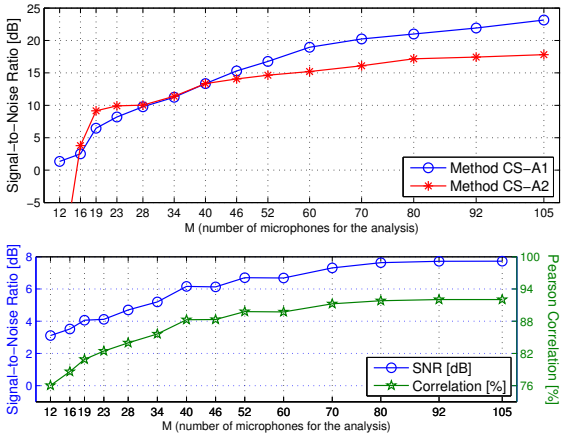

Figure 5: Effect of the number of microphones for the analysis. Top: Methods CS-Ax, bottom: Method CS-B

Figure 6 presents the performances according to the distance of the interpolation position from the center of the array. In this figure, the vertical lines represent the standard deviation of the SNRs per interval. As expected, the performances decrease when the position moves away from the center, although it can be noticed that whereas method CS-A1 is better around the center of the array, the SNR of method CS-A2 decreases slower and is better at the edge.

An interesting experiment is the comparison of the methods according to the array configuration. In figure 7, we have numerically simulated and tested 6 different array configurations with 125 microphones: 3 random arrays (with a uniform distribution within the cube with side 2m), the experimental array (which is approximately random, cf. fig. 2), a spherical array with radius 1.24m (for which the volume is $8m^3$ like the cube), and a regular array (where the receivers are uniformly positioned within the cube with side 2m). The noticeable result is that: whereas methods CS-A1 and CS-A2 (for low frequencies) are better with random arrays, as expected, the best result for method CS-B (early part) is ob-

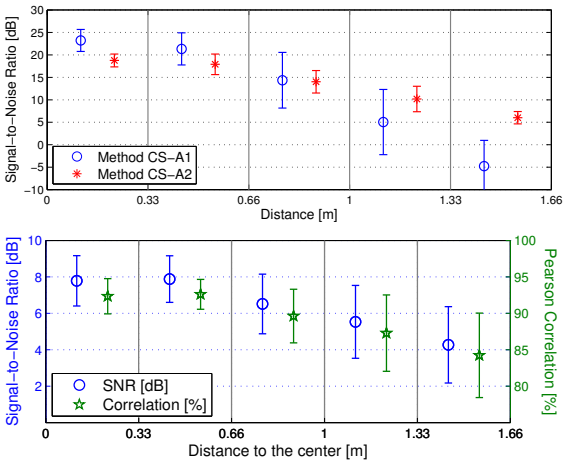

Figure 6: Effect of the distance from the array center. Top: Methods CS-Ax, bottom: Method CS-B

tained with the regular array. The first observation can be explained because for Compressed Sensing, the measured signals have to be decorrelated from the decomposition dictionary (plane waves), consequently random arrays are preferable (cf. the RIP property in sec. 1.2) . Moreover with random arrays, the probability that all microphones lie on the zeroes of a modal shape is smaller than with regular arrays. For method CS-B (for early reflections), the observation can be explained because the diffences of time of arrival have to be maximized between the microphones of the array, and this is done with a regular array where the microphones are uniformly spaced. Nevertheless, we have realised a random array because method CS-B remains acceptable, and because method CS-A1 cannot be used with regular array.

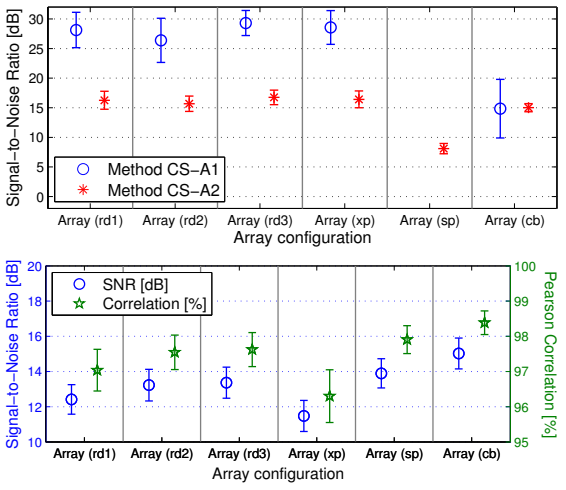

Figure 7: Six different microphone arrays. Top: Methods CS-Ax, bottom: Method CS-B

We have checked the robustness of both methods to the geometry of the room, in particular when the measured room gets further away from the "ideal" rectangular room. This has been made by opening the windows and the door, and by placing a chair and a wood panel. Moreover, in the first set of experiments the baffled loudspeaker, which has a complex directivity pattern, was oriented towards the array ; in this second set of experiments the loudspeaker was turned so as to not face the array. The results of these new measurements reveal that methods CS-A1 and CS-A2 (based on the modal

theory) are stable with obstacles and with several orientations of the loudspeaker, while the SNR for method CS-B (based the image source method) decreases.

For methods CS-A1 and CS-A2, since the modal density strongly increases with the frequency, the sparsity assumption is less and less valid in high frequencies. Indeed, the performances of method CS-A1 decreases when the cutoff frequency f_c increases, but method CS-A2 seems very stable when f_c varies, at least up to 400Hz.

In the same line for method CS-B, since the echo density strongly increases with time, the sparsity assumption is less and less valid. Indeed, when the RIRs duration increases for the analysis, the global performance decreases. Nevertheless, we observe that the interpolation quality doesn't change at the begining of the reconstructed response.

4 Conclusion

Starting from a random array with a small number of microphones, and exploiting two sparsity assumptions of the acoustic wavefield in reverberating room (sparsity in frequency and sparsity in time), we have shown that it is possible to use the Compressed Sensing framework to interpolate in space the Room Impulse Responses. Consequently, we can seperately sample the low frequencies and the early reflections of the Plenacoustic Function (respectively part A and part B of fig. 1), but using the same measures. Then, part C can be synthesized using a stochastic model (cf. [2]).

Note that, the performances are evaluated using the SNR and the correlation coefficient which are probably not significant. For specific applications, more relevant criteria should be designed. For example human perception should be taken into account in the case of a digital reverberation.

Acknowledgments

The authors want to thank Dominique Busquet and Christian Ollivon, for their precious help in the realisation of the microphone array and the acquisition of the RIRs.

References

[1] T. Ajdler, L. Sbaiz, M. Vetterli, "The Plenacoustic Function and its Sampling", *IEEE Transaction on Signal Processing*, p. 3790-3804 (2006)

[2] J.-M. Jot, L. Cerveau, O. Warusfel, "Analysis and Synthesis of Room Reverberation Based on a Statistical Time-Frequency Model", *AES 103rd Conv.*, (1997)

[3] M. Schroeder, "Statistical parameters of the frequency response curves of large rooms", *Journal of the AES*, p. 299-306 (1987).

[4] J.-D. Polack, "Playing billard in the concert hall: the mathematical foundations of geometrical room acoustics", *Applied Acoustics*, **38** p. 235-244 (1993).

[5] Y. Haneda, Y. Kaneda, N. Kitawaki, "Common-Acoustical-Pole and Residue Model and Its Application to Spatial Interpolation andExtrapolation of a Room Transfer Function", *IEEE Trans. on Speech and Audio Proc.*, p. 709-717 (1999)

[6] C. Masterson, G. Kearney, F. Boland, "Acoustic Impulse Response Interpolation for Multichannel Systems using Dynamic Time Warping", *AES 35th Conf.* (2009)

[7] E. Candès, M. Wakin, "An introduction to compressive sampling", *IEEE Sig. Proc. Mag.*, p. 21-30 (2008)

[8] R. Baraniuk, "Compressive sensing", *IEEE Sig. Proc. Mag.*, p. 118-121 (2007)

[9] E. Candès, "Compressive sampling", *International Congress of Mathematicians*, (2006)

[10] E. Candès, T. Tao, "Decoding by linear programming", *IEEE Trans. on Inform. Theory*, p. 4206-4215 (2005)

[11] R. Mignot, G. Chardon, L. Daudet, "Compressively sampling the plenacoustic function", *SPIE Conf. Wavelets and Sparsity XIV*, **vol. 8138 08**, (2011).

[12] R. Mignot, L. Daudet, F. Ollivier, "Compressed sensing for acoustic response reconstruction: interpolation of the early part", *IEEE WASPAA'11*, p. 225-228 (2011)

[13] E. Perrey-Debain, "Plane wave decomposition in the unit disc: Convergence estimates and computational aspects", *J. Comp. App. Math.*, p. 140-156 (2006)

[14] A. Moiola, R. Hiptmair, I. Perugia, "Approximation by plane waves", *Research Report (ETH Zürich)*, (2009)

[15] G. Chardon, A. Leblanc, L. Daudet, "Plate impulse response spatial interpolation with sub-Nyquist sampling", *JSV*, p. 5678-5689 (2011)

[16] G. Chardon, L. Daudet, "Optimal Subsampling of Multichannel Damped Sinusoids", *6th IEEE SAM*, (2010)

[17] H. Kuttruff, "Room Acoustics", *Spon press* (2000)

[18] M. Elad "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing", *Springer*, (2010)

[19] J. Borish, "Extension of the image model to arbitrary polyhedra", *JASA*, p. 1827-1836 (1984)