SPATIAL INTERPOLATION OF ROOM IMPULSE RESPONSES USING COMPRESSED SENSING

Fabrice Katzberg, Radoslaw Mazur, Marco Maass, Martina Böhme, and Alfred Mertins

Institute for Signal Processing
University of Lübeck
Ratzeburger Allee 160, 23562 Lübeck, Germany

ABSTRACT

Measuring a large set of room impulse responses inside a volume of interest is time-consuming unless a large number of microphones is involved. However, increasing the number of microphones requires more hardware and raises effort, e.g., in calibration. Instead of measuring at any desired position, it is possible to spatially interpolate the sound field between sampled positions, in order to obtain estimates at unknown positions. Nevertheless, the Nyquist-Shannon sampling theorem should be met, which still demands a large number of spatial sampling points for large bandwidths. In this paper, we present a compressed-sensing approach that allows for stable and robust interpolation of room impulse responses using less measurements than required by the sampling theorem. Based on a small set of spatially subsampled room impulse responses, the proposed method is capable of providing an enlarged set allowing for aliasing-free reconstruction in space.

Index Terms— Room impulse responses, spatial interpolation, compressed sensing

1. INTRODUCTION

The knowledge of spatio-temporal room impulse responses (RIRs) provides useful information on acoustic scenes and is essential for algorithms focussing on, e.g., listening-room compensation [1, 2] and sound-field reproduction [3, 4].

For determining the RIR between a fixed emitter-receiver pair inside a room, common techniques use perfect sequences [5, 6], maximum-length sequences [7], and sine sweeps [8] for excitation. However, measuring multiple RIRs with stationary microphones over larger volumes involves either high effort in calibration and positioning of multiple microphones or a time-consuming measurement procedure when only one microphone is used. The use of moving microphones, which allow for accelerating the process, either requires specific demands on the microphone trajectory and speed [9] or needs real-time position tracking [10, 11].

Instead of measuring RIRs at all targeted positions, a smaller set of RIRs may be used for spatial interpolation in order to obtain estimates at positions where no microphone measurements were taken. Usually, to allow for spatial interpolation between measured locations, the Nyquist-Shannon sampling theorem should be met. For larger audio bandwidths, this still requires a high number of microphone locations. Considering conventional sound-field sampling with equidistant microphone arrays, spatial intervals of

$$\Delta < \frac{c_0}{2f_c} \tag{1}$$

are required for aliasing-free reconstruction [12], where c_0 is the speed of sound and f_c is the temporal cutoff frequency.

In order to relax the spatial sampling and interpolation problem for sound fields, the principle of compressed sensing (CS) may be used [13, 14]. By encapsulating the problems of sampling and compression to one joint object, CS allows for sampling a signal below the Nyquist rate, provided that the signal has a sparse representation and that incoherent measurements are available [15].

There exist various methods that exploit the compressed-sensing paradigm for RIR interpolation. In [16], a large set of RIRs is recovered from a smaller set of RIR samples by using convex optimization methods that enforce both sparsity of the early-reflection parts and the exponential decay of RIRs. Sparsity in time domain is also used in [17] for interpolating early reflections of RIRs within a volume of interest (VOI). Sparsity in frequency domain is exploited in [18], in order to reconstruct spatio-temporal RIRs at low frequencies via plane-wave approximations. In [19, 20], the sound field is spatially parametrized via spherical-harmonics solutions of the wave equation. The method in [19] extracts spherical-harmonics parameters, which may be used for sound-field reconstruction, from measurements of a spherical microphone array. A sparse approximation employing spherical harmonics is also applied in [20] for interpolating RIRs between stationary measurements.

In this paper, we propose a CS based method that allows for recovering spatio-temporal RIRs at desired positions from spatially subsampled RIRs acquired at arbitrary points. In order to do this, a linear system of equations is build up, similar to the dynamic measurement approach presented in [10]. Unlike the measured points, the targeted positions are supposed to ensure aliasing-free reconstruction in space, thus, the problem is underdetermined and is solved exploiting sparsity in frequency domain. We focus on a model with targeted positions forming a virtual grid in space, keeping the computational effort small and allowing for straightforward analysis. Then, the system matrix only consists of interpolation coefficients in structured diagonal blocks and both the interpolation kernel and the sparsifying transform may become separable, which enables us to provide a fast coherence analysis for a given set of measured points. Nevertheless, also other arrangements than regular ones would be possible for target positions, e.g., designs following spherical sampling patterns where a sparse spherical-harmonics expansion may be used for setting up the linear system.

This work has been supported by the German Research Foundation under Grants No. ME 1170/10-1 and ME 1170/8-1.

978-1-5386-8151-0/18/$31.00

2. SPATIO-TEMPORAL RECONSTRUCTION OF RIRS

Assuming an echoic environment that constitutes a linear time-invariant system, the transmission of sound pressure between fixed emitter-receiver positions may be described as $x(t) = \int_{-\infty}^{\infty} h(\tau)s(t-\tau)d\tau$, where $t \in \mathbb{R}$ is the time variable, $s(t)$ and $x(t)$ are the emitted and received signals, respectively, and $h(t)$ is the RIR for the fixed setup. Considering a variable listener position $\boldsymbol{r} = [r_x, r_y, r_z]^T \in \mathbb{R}^3$, the resulting sound pressure field $p(\boldsymbol{r}, t) = \int_{-\infty}^{\infty} h(\boldsymbol{r}, \tau)s(t-\tau)d\tau$ is characterized by $h(\boldsymbol{r}, t)$ composing spatio-temporal RIRs that describe sound transmission from the source position to various points $\boldsymbol{r}$.

For the temporal domain, let $T = 1/f_s$ be the sampling interval of a microphone sampling with frequency $f_s > 2f_c$. Then, the measured RIR may be perfectly reconstructed from equidistant sampling points $t_n = nT$ ($n \in \mathbb{N}$) by using a sinc filter with infinite support. Since the amplitudes of RIRs decrease exponentially and vanish into the noise level beyond t_{L-1}, finite length interpolation kernels may obtain reasonable approximations in time domain.

Let us now consider the spatial interpolation of RIRs, i.e., the reconstruction of RIRs at arbitrary listener positions $\boldsymbol{r}$ from the sound-field data $h(\boldsymbol{r}_m, n)$ acquired at M microphone positions $\boldsymbol{r}_m$ ($m \in \{1, \dots, M\}$). For the conventional arrangement of points $\boldsymbol{r}_m$ forming a uniform grid in space, i.e., $\boldsymbol{r}_m \in \mathcal{G}$ with

$$\mathcal{G} = \left\{ \boldsymbol{r}_{\boldsymbol{g}} \,|\, \boldsymbol{r}_{\boldsymbol{g}} = \boldsymbol{r}_0 + [g_x\Delta_x, g_y\Delta_y, g_z\Delta_z]^T \right\}, \tag{2}$$

$\boldsymbol{r}_0$ being the grid origin, and $\boldsymbol{g} = [g_x, g_y, g_z]^T \in \mathbb{Z}^3$ representing the discrete grid variables, the Nyquist-Shannon sampling theorem requires a grid spacing according to (1), in order to allow for aliasing-free reconstruction. Since RIRs are acquired at a finite number of $D = XYZ$ grid coordinates $G = \{0, \dots, X-1\} \times \{0, \dots, Y-1\} \times \{0, \dots, Z-1\}$, only estimates

$$h(\boldsymbol{r}, n) \approx \sum_{\boldsymbol{g} \in G} \varphi_{\boldsymbol{r}}(\boldsymbol{g}) h(\boldsymbol{g}, n) \tag{3}$$

are available for RIRs at arbitrary positions $\boldsymbol{r}$, where $\varphi_{\boldsymbol{r}}(\boldsymbol{g})$ with $\boldsymbol{g} \in G$ denotes a realizable interpolation kernel that approximates the sinc function.

3. RIR INTERPOLATION USING SPARSE PRIORS

The proposed method is based on solving a linear system of equations that is built up by representing spatially subsampled RIRs measured at arbitrary points $\boldsymbol{r}_m$ by means of targeted RIRs at desired positions $\tilde{\boldsymbol{r}}_d$ ($d \in \{1, \dots, D\}$). For targeted points forming an equidistant grid (2), this step is equivalent to solving the reverse interpolation problem according to (3), i.e., determining $h(\boldsymbol{g}, n)$ from measured data $h(\boldsymbol{r}, n)$. Unlike the measured points $\boldsymbol{r}_m$, the targeted positions $\tilde{\boldsymbol{r}}_d$ are supposed to satisfy the Nyquist-Shannon sampling theorem, thus, in general, the problem is underdetermined with $D > M$ and must be solved using CS. Finally, the recovered RIRs may be used for conventional interpolation in line with (3).

3.1. Modeling the Reverse Interpolation Problem

Mathematically, the linear system of equations is set up by projecting spatio-temporal RIRs on a uniform grid onto the measurement space spanned by RIRs acquired at arbitrary positions. Let us define the measurement vector

$$\boldsymbol{m} = \left[\boldsymbol{h}_1^T, \boldsymbol{h}_2^T, \dots, \boldsymbol{h}_M^T \right]^T \tag{4}$$

as concatenation of the measured RIRs

$$\boldsymbol{h}_m = [h(\boldsymbol{r}_m, 0), h(\boldsymbol{r}_m, 1), \dots, h(\boldsymbol{r}_m, L-1)]^T \tag{5}$$

at arbitrary positions $\boldsymbol{r}_m \in \mathbb{R}$, and the target vector

$$\boldsymbol{d} = \left[\tilde{\boldsymbol{h}}_1^T, \tilde{\boldsymbol{h}}_2^T, \dots, \tilde{\boldsymbol{h}}_D^T \right]^T \tag{6}$$

as concatenation of the desired RIRs

$$\tilde{\boldsymbol{h}}_d = [h(\boldsymbol{g}_d, 0), h(\boldsymbol{g}_d, 1), \dots, h(\boldsymbol{g}_d, L-1)]^T \tag{7}$$

at positions $\tilde{\boldsymbol{r}}_d \in \mathcal{G}$ on a uniform grid fulfilling (1). For reasons of clarity, we use the discrete variables $\boldsymbol{g}_d$ of the spatial grid. Further, let us define the vector

$$\boldsymbol{\varphi}_m = [\varphi_{\boldsymbol{r}_m}(\boldsymbol{g}_1), \varphi_{\boldsymbol{r}_m}(\boldsymbol{g}_2), \dots, \varphi_{\boldsymbol{r}_m}(\boldsymbol{g}_D)]^T \tag{8}$$

containing the kernel for interpolating the m-th measurement position relative to the spatial grid, and the $M \times D$ interpolation matrix

$$\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_M]^T \tag{9}$$

comprising any kernel for the M measured points. In the manner of the spatial interpolation given by (3), we can then formulate the overall interpolation problem in terms of the linear system of equations

$$\boldsymbol{m} = \boldsymbol{A}\boldsymbol{d} + \boldsymbol{\eta}, \tag{10}$$

where $\boldsymbol{\eta} \in \mathbb{R}^{LM}$ is a noise vector, incorporating both the measurement noise and the error of the bandlimited interpolation, and $\boldsymbol{A} \in \mathbb{R}^{LM \times LD}$ is the system matrix having the blockwise diagonal structure

$$\boldsymbol{A} = \boldsymbol{\Phi} \otimes \boldsymbol{I}_L, \tag{11}$$

with $\otimes$ denoting the Kronecker product and $\boldsymbol{I}_L$ being the identity matrix of size $L \times L$.

3.2. Compressed-Sensing Formulation

Since $D > M$, the linear system (10) provides an infinite number of least-squares solutions for $\boldsymbol{d}$. Nevertheless, the principle of CS allows for finding a stable and robust solution also in the underdetermined case. The spectrum of sound fields is ideally confined to a hypercone along the temporal frequency axis [12], thus, we can represent the equidistantly sampled sound field $h(\boldsymbol{g}, n)$ collected in $\boldsymbol{d}$ by a sparse coefficient vector

$$\boldsymbol{c} = \boldsymbol{\Psi}\boldsymbol{d},$$

where $\boldsymbol{\Psi} = \boldsymbol{T}_Z \otimes \boldsymbol{T}_Y \otimes \boldsymbol{T}_X \otimes \boldsymbol{T}_L$ performs a 4D frequency transform that is supposed to be separable into multiple 1D transforms given by unitary matrices $\boldsymbol{T}_U \in \mathbb{C}^{U \times U}$. The vector $\boldsymbol{c} \in \mathbb{C}^{LXYZ}$ is assumed to be K-sparse, i.e., at most K coefficients are non-zero. Accordingly, the least-squares problem may be regularized as

$$\underset{\boldsymbol{c} \in \mathbb{C}^U}{\operatorname{argmin}} \; \|\boldsymbol{m} - \boldsymbol{\mathcal{A}}\boldsymbol{c}\|_{\ell_2}^2 \quad \text{s.t.} \quad \|\boldsymbol{c}\|_{\ell_0} \leq K, \tag{12}$$

with the CS matrix $\boldsymbol{\mathcal{A}} = \boldsymbol{A}\boldsymbol{\Psi}^H$ and the pseudonorm $\|\boldsymbol{c}\|_{\ell_0}$ counting the number of non-zero elements in $\boldsymbol{c}$. The problem (12) is NP-hard [21] and is solved in practice by using greedy algorithms

[22, 23, 24] or applying convex optimization tools to corresponding ℓ_1-minimization problems [25, 26, 27, 28]. For stable and robust recovery of any (approximately) K-sparse signal, any set of K columns in $\boldsymbol{\mathcal{A}}$ must form a nearly orthogonal system. A useful property to evaluate the sampling process according to that is the coherence

$$\mu(\boldsymbol{\mathcal{A}}) = \max_{1 \le u \neq v \le LD} \frac{|\langle \boldsymbol{a}_u^c, \boldsymbol{a}_v^c \rangle|}{\|\boldsymbol{a}_u^c\|_{\ell_2} \|\boldsymbol{a}_v^c\|_{\ell_2}}, \tag{13}$$

with $\boldsymbol{a}_u^c$ being the u-th column in $\boldsymbol{\mathcal{A}}$ [15]. The error bounds of the CS based recovery improve for a smaller coherence [29].

3.3. Fast Coherence Analysis

Determining the coherence naively by calculating any scalar product of any two columns in $\boldsymbol{\mathcal{A}}$ poses a problem in $\mathcal{O}(L^2D^2)$. However, similar to the dynamic approach in [11], this effort may be easily reduced to complexity of $\mathcal{O}(D)$ for $\boldsymbol{\Psi}$ performing discrete Fourier transforms (DFTs).

To keep the description simple, let us first consider the measurement space to be a line in x-direction only, i.e., $\boldsymbol{\Psi} = \boldsymbol{T}_X \otimes \boldsymbol{T}_L$ achieves the 2D DFT involving the sampled frequency variables $k_x \in \{-\frac{X-1}{2}, \dots, \frac{X-1}{2}\}$ and $l \in \{-\frac{L-1}{2}, \dots, \frac{L-1}{2}\}$. Then, assuming perfect interpolation kernels $\varphi_{r_m}(g_x) = \operatorname{sinc}(g_x - \tau_x^m)$, with

$$\tau_x^m = \frac{r_m - r_0}{\Delta_x} \tag{14}$$

being the spatial-grid delay to measured point r_m, the columns building $\boldsymbol{\mathcal{A}}$ can be described by $\boldsymbol{a}^c_{(k_x,l)} = \left[\boldsymbol{\rho}_1^T, \dots, \boldsymbol{\rho}_M^T\right]^T$ using the structured phase vectors

$$\boldsymbol{\rho}_m = \xi_m \left[e^{-2\pi j 0 \frac{l}{L}}, e^{-2\pi j 1 \frac{l}{L}}, \dots, e^{-2\pi j (L-1) \frac{l}{L}}\right]^T, \tag{15}$$

with $\xi_m = e^{-2\pi j \tau_x^1 \frac{k_x}{X}} e^{-2\pi j (\tau_x^m - \tau_x^1) \frac{k_x}{X}}$ giving the phase shift corresponding to position r_m relative to r_1. Using column representations through (15) and defining differences between two discrete frequencies as in [11], $\Delta k_x = k_x' - k_x''$, $\Delta l = l' - l''$, the coherence according to (13) may be found by determining the maximum over tupels $(\Delta k_x, \Delta l) \neq (0,0)$ as

$$\mu(\boldsymbol{\mathcal{A}}) = \max_{(\Delta k_x, \Delta l)} \frac{1}{LM} \left| \sum_{m=0}^{M-1} \sum_{n=0}^{L-1} e^{-2\pi j \tau_x^m \frac{\Delta k_x}{X}} e^{-2\pi j n \frac{\Delta l}{L}} \right| \tag{16}$$

$$= \frac{1}{LM} \max_{\Delta k_x} \left| \sum_{m=0}^{M-1} e^{-2\pi j \tau_x^m \frac{\Delta k_x}{X}} \right| \max_{\Delta l} \left| \sum_{n=0}^{L-1} e^{-2\pi j n \frac{\Delta l}{L}} \right| \tag{17}$$

$$= \frac{1}{M} \max_{\Delta k_x \neq 0} \left| \sum_{m=0}^{M-1} e^{-2\pi j \tau_x(m) \frac{\Delta k_x}{X}} \right|, \tag{18}$$

where (17) results from the Cauchy product formula and (18) is obtained since the maximum over Δl must be L at $\Delta l = 0$.

For measurements in 3D space, we can exploit the separability of the uniform-grid dimensions and obtain

$$\mu(\boldsymbol{\mathcal{A}}) = \frac{1}{M} \max_{\Delta \boldsymbol{k}} \left| \sum_{m=0}^{M-1} e^{-2\pi j \left(\tau_x^m \frac{\Delta k_x}{X} + \tau_y^m \frac{\Delta k_y}{X} + \tau_z^m \frac{\Delta k_z}{Z}\right)} \right|, \tag{19}$$

with $\Delta \boldsymbol{k} = (\Delta k_x, \Delta k_y, \Delta k_z) \neq (0,0,0)$. For a grid design with suffiently small spacing, the expression (19) may be used for finding microphone positions being optimal for the CS problem, also in practical applications where realizable interpolation filters are involved (cf. [11]).

4. EXPERIMENTS AND RESULTS

Experiments have been carried out on simulations using the image source method [30]. A box-shaped room with dimensions of size $5.8\,\mathrm{m} \times 4.15\,\mathrm{m} \times 2.55\,\mathrm{m}$ and reverberation time of $RT_{60} = 0.3\,\mathrm{s}$ was considered. An omnidirectional sound source was placed at $[1.4, 1.6, 1.0]^T\,\mathrm{m}$. The sampling frequency $f_s = 8\,\mathrm{kHz}$ was used. The simulated RIR were corrupted by additive white Gaussian noise with a signal-to-noise ratio of $40\,\mathrm{dB}$ and their length was set to $L = 500$. Interpolation results are presented for a VOI of size $(0.1\,\mathrm{m})^3$ having the origin according to (2) at $\boldsymbol{r}_0 = [2.75, 1.4, 0.8]^T\,\mathrm{m}$.

As error measure for the interpolated RIRs, the normalized system misalignment [10]

$$\mathrm{NSM} = \frac{\|\boldsymbol{h}_{\boldsymbol{r}} - \hat{\boldsymbol{h}}_{\boldsymbol{r}}\|_{\ell_2}^2}{\|\boldsymbol{h}_{\boldsymbol{r}}\|_{\ell_2}^2} \tag{20}$$

is used, where $\boldsymbol{h}_{\boldsymbol{r}} \in \mathbb{R}^L$ contains the true RIR at position $\boldsymbol{r}$ and $\hat{\boldsymbol{h}}_{\boldsymbol{r}} \in \mathbb{R}^L$ contains the corresponding interpolation result.

In order to allow for a baseline of the interpolation quality, we first sampled the VOI conventionally using regular microphone arrays with spacing $\Delta \in \{0.033\,\mathrm{m}, 0.02\,\mathrm{m}\}$ for each dimension. Thus, we obtained two data sets of equidistantly sampled RIRs for baseline: spatio-temporal RIRs at positions on a coarse grid close to the Nyquist rate and RIRs on a finer grid that leads to twofold spatial oversampling. The coarse grid involves $D = 64$ grid points, the fine grid leads to $D = 216$. Based on these sets of data, the interpolation errors over the centered xy-plane of the VOI ($z = 0.85\,\mathrm{m}$) are presented in Figs. 1(a) and 1(e), respectively. The errors result from separable Lagrange interpolators of order three for each dimension (cf. [10]). The smallest errors are, of course, obtained close to the regular points. The largest interpolation errors occur at intermediate positions on the left-hand and right-hand border. At these sides, the direct sound paths and their reflection paths initially enter the sampled volume, which obviously decreases the performance of the interpolation kernels that had to be truncated at grid edges.

In the following, we compare the above mentioned results for regular sampling with the outcomes of our proposed CS based strategy. Therefore, we modeled the linear system (10) for each of the two targeted grids. For measurements, we acquired RIRs at random positions $\boldsymbol{r}_m \in \mathbb{R}^3$ inside the VOI. The linear systems were solved for several cases with full and partial measurement data provided ($M = \alpha D$ with $\alpha \in \{1, 0.8, 0.6\}$), where for $\alpha = 1$ the number of measured positions is equal to the number of points on the grid to be recovered, and for $\alpha < 1$, fewer measurements are taken and an underdetermined system is obtained. For the CS based recovery, we used the iterative hard-thresholding algorithm (IHT) [24] with step size $\mu = 5 \cdot 10^{-3}$ and 3000 iterations. The 4D DFT was applied for sparse frequency representations. The iterative recovery was started with the initial estimate being the zero vector. Beginning with a small K, $K_0 = \alpha 500$ for $\Delta = 0.033\,\mathrm{m}$ and $K_0 = \alpha 1500$ for $\Delta = 0.02\,\mathrm{m}$, the sparsity constraint was successively relaxed every 50 iterations by $K_{i+1} = K_i + K_0$ (cf. [11]).

We observed that the extension of the modeled grid beyond the measured VOI improves the CS recovery for an unchanged number of measurements, although this results in linear systems with even more unknowns. Following this strategy, a truncation of the interpolation kernels may be avoided. Interpolation errors based on RIRs

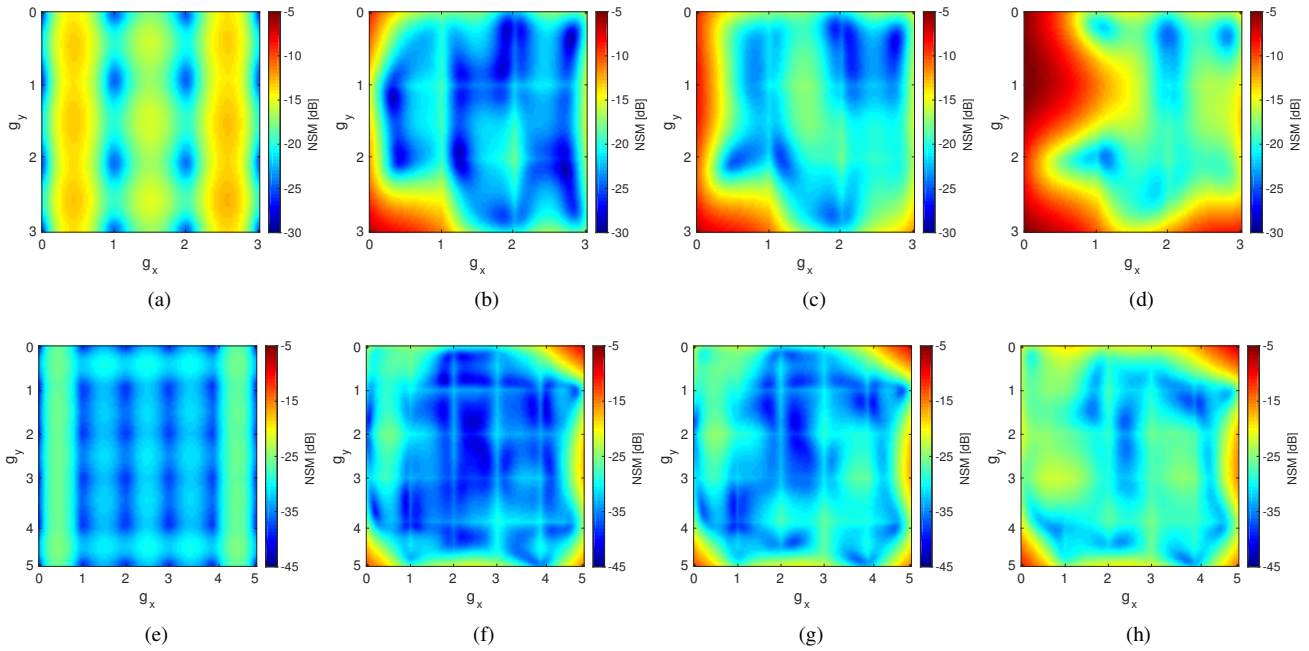

Fig. 1. Errors of RIR interpolation over the centered xy-plane of a VOI of size $(0.1\,\mathrm{m})^3$ using sets of RIRs on a virtual uniform grid with spacing $\Delta = 0.033\,\mathrm{m}$ (first row) and $\Delta = 0.02\,\mathrm{m}$ (second row). (a), (e) Baseline with M grid RIRs conventionally sampled. (b), (f) Grid RIRs recovered by the proposed CS method using M RIRs, (c), (g) $0.8M$ RIRs, and (d), (h) $0.6M$ RIRs measured at random positions.

recovered in this way are presented in Figs. 1(b)-(d) and Figs. 1(f)-(h), regarding $\Delta = 0.033\,\mathrm{m}$ and $\Delta = 0.02\,\mathrm{m}$, respectively, for different numbers of measurements. For setting up the linear systems and subsequent interpolation of positions inside the VOI, also Lagrange interpolators of order three were used.

Since, in general, the random measurements are located at intermediate positions, the proposed CS method obtains the smallest interpolation errors between grid points. In Figs. 1(b) and 1(f), for example, the modeled grid may be reproduced by tracking lines of slightly higher errors among areas where the interpolation error is very small. Compared to the RIR interpolation by using measurements on a uniform grid, the CS approach obtains a smaller interpolation error on average for an equal number of random measurements (Figs. 1(b) and 1(f)). However, a small number of outliers with high error appear at some border positions where, obviously, not enough random measurement points were available. Providing only 80% and even 60% of the measured data still yields robust interpolation results for the fine-grid scenario (Figs. 1(g) and 1(h)). For the coarse-grid model, we also obtain adequate results when solving the linear system with 80% of the measured data (Figs. 1(c)). However, for the case using 60% of data, larger regions with high interpolation errors are present at the grid border Fig. 1(d).

5. CONCLUSIONS

In this paper, we presented a CS based method for RIR interpolation. By using a small set of spatially subsampled RIRs measured at arbitrary positions inside the VOI, a linear system of equations was set up that leads to the estimation of RIRs at targeted positions allowing for aliasing-free interpolation in space. In general, the linear system is underdetermined. In order to obtain a stable and robust solution, principles of CS were used exploiting the sparsity of spatio-temporal RIRs in frequency domain. For Fourier representations and targeted positions on a virtual grid in space, the structure of the CS matrix and its dependency on measuring positions have been shown. Further, a simple position-dependent expression for the coherence of measurements has been derived, allowing for the efficient evaluation of microphone arrangements in terms of CS reconstruction. Future works will be directed toward using the fast coherence computation to optimize the microphone positions and to include sparsity measures for the early-reflection part of RIRs.

6. REFERENCES

[1] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Audio, Speech, Lang. Process*, vol. 36, no. 2, pp. 145–152, Feb. 1988.

[2] J. N. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.*, vol. 42, no. 11, pp. 884–900, Nov. 1994.

[3] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[4] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Trans. Audio, Speech, Lang. Process*, vol. 17, no. 1, pp. 107–116, Sept. 2009.

[5] H.-D. Lüke, "Sequences and arrays with perfect periodic correlation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 3, pp. 287–294, May 1988.

[6] V. P. Ipatov, "Ternary sequences with ideal periodic autocorrelation properties," *Radio Eng. Electron. Phys.*, vol. 24, pp. 75–99, Oct. 1979.

[7] D. D. Rife and J. Vanderkooy, "Transfer-function measurement with maximum-length sequences," *J. Audio Eng. Soc.*, vol. 37, no. 6, pp. 419–444, June 1989.

[8] A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Proc. 122nd Audio Engineering Society Convention*, May 2007, pp. 1–21.

[9] T. Ajdler, L. Sbaiz, and M. Vetterli, "Dynamic measurement of room impulse responses using a moving microphone," *J. Acoust. Soc. Am.*, vol. 122, no. 3, pp. 1636–1645, July 2007.

[10] F. Katzberg, R. Mazur, M. Maass, P. Koch, and A. Mertins, "Sound-field measurement with moving microphones," *J. Acoust. Soc. Am.*, vol. 141, no. 5, pp. 3220–3235, May 2017.

[11] F. Katzberg, R. Mazur, M. Maass, P. Koch, and A. Mertins, "Compressive sampling of sound fields using moving microphones," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 2018, pp. 181–185.

[12] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3790–3804, Sept. 2006.

[13] E. Candès and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Febr. 2006.

[14] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[15] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[16] A. Benichoux, L. Simon, E. Vincent, and R. Gribonval, "Convex regularizations for the simultaneous recording of room impulse responses," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 1976–1986, April 2014.

[17] R. Mignot, L. Daudet, and F. Ollivier, "Room reverberation reconstruction: interpolation of the early part using compressed sensing," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 11, pp. 2301–2312, July 2013.

[18] R. Mignot, G. Chardon, and L. Daudet, "Low frequency interpolation of room impulse responses using compressed sensing," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 1, pp. 205–216, Jan. 2014.

[19] E. Fernandez-Grande and A. Xenaki, "Compressive sensing with a spherical microphone array," *J. Acoust. Soc. Am.*, vol. 139, no. 2, pp. EL45–EL49, Feb. 2016.

[20] N. Antonello, E. De Sena, M. Moonen, P. A. Naylor, and T. van Waterschoot, "Room impulse response interpolation using a sparse spatio-temporal representation of the sound field," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 10, pp. 1929–1941, Oct. 2017.

[21] B. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, April 1995.

[22] J. A Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[23] D. Needell and J. A. Tropp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.

[24] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 629–654, Dec. 2008.

[25] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, Aug. 1998.

[26] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, March 2006.

[27] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. R. Stat. Soc., Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[28] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *Ann. Statist.*, vol. 35, no. 6, pp. 2313–2351, 2007.

[29] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," in *Proc. Natl. Acad. Sci.*, March 2003, pp. 2197–2202.

[30] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, 1979.