Götz, Georg; Ananthabhotla, Ishwarya; Amengual Garí, Sebastià V.; Calamia, Paul T.

Autonomous Room Acoustic Measurements using Rapidly-Exploring Random Trees and Gaussian Processes

Published in:
Proceedings of the 10th Convention of the European Acoustics Association Forum Acusticum 2023

DOI:
10.61782/fa.2023.0796

Published: 01/01/2023

Document Version
Publisher's PDF, also known as Version of record

forum acusticum 2023

AUTONOMOUS ROOM ACOUSTIC MEASUREMENTS USING RAPIDLY-EXPLORING RANDOM TREES AND GAUSSIAN PROCESSES

Georg Götz[1†*] **Ishwarya Ananthabhotla**[2] **Sebastià V. Amengual Garí**[2] **Paul Calamia**[2]

[1] Aalto Acoustics Lab, Aalto University, Finland
[2] Reality Labs Research at Meta, USA
† This work was conducted during a research internship at Reality Labs Research.

ABSTRACT

Various robot systems have been proposed in the past to automate the tedious and time-consuming room acoustic measurement process. While small-scale measurements within a limited area can be realized with robotic arms, room-scale measurements require robots that can travel larger distances and ideally navigate through their environment autonomously. In this paper, we propose a new measurement strategy for large-scale, autonomous, room-acoustic measurement robots. The measurement strategy uses rapidly-exploring random trees to determine multiple candidate paths, from which it chooses the best path for exploring the unvisited parts of the environment and reconstructing a target acoustic metric. Gaussian process regression is used to incrementally merge new acoustic data into a global estimate. We evaluate the measurement strategy in a multi-room scenario, utilizing a late reverberation metric and a robot system consisting of a source and a receiver robot. We demonstrate that the measurement strategy can be used to map and reconstruct late reverberation characteristics over a large area.

Keywords: *room acoustic measurements, measurement robot, path planning, late reverberation*

*__Corresponding author__: *georg.gotz@aalto.fi.*

1. INTRODUCTION

The time-domain acoustical response between a sound source and receiver is called the room impulse response (RIR). Knowledge about this transfer function is crucial for many acoustic disciplines, but RIR measurements can be time-consuming and cumbersome.

Various RIR measurement robots have been developed in the past, ranging from robotic arms that measure within a limited spatial region [1] to larger-scale setups that can capture RIRs over extended environments [2,3]. More recently, a highly-portable autonomous robot system was proposed that captures large quantities of RIRs, without requiring prior geometric information about the measured environment, such as floor plans or 3D models [3]. Grid-based robots for large-scale RIR measurements also exist [4].

The present study proposes a new measurement strategy for room acoustic measurements with autonomous robots. It builds upon the system described in [3], extending it in two aspects. First, the proposed strategy leverages the potential of rapidly-exploring random trees [5] for exploring more general multi-room environments. Second, it uses Gaussian process regression to reconstruct sound field quantities, thus allowing it to deal with inaccuracies and uncertainties of the robot system (e.g., position tracking or measurement noise). We demonstrate the strategy by measuring a late reverberation quantity, the common-slope amplitude [6], for a multi-room environment.

The remainder of this paper is organized as follows. Sec. 2 provides background information on rapidly-exploring random trees, informative path planning, Gaussian process regression, and the common-slope model of late reverberation. Sec. 3 describes the proposed measurement strategy. Sec. 4 evaluates the strategy in a multi-room

environment using a new robot system. Sec. 5 briefly discusses the results and Sec. 6 concludes the paper.

2. BACKGROUND

2.1 Rapidly-exploring random trees

The proposed measurement strategy is based on rapidly-exploring random trees (RRTs). RRTs were introduced by LaValle [5] and have been used extensively in robotic motion planning and information gathering [7, 8]. RRTs are constructed in three steps. First, a random unobstructed position is sampled from the search space. Second, the tree node with the smallest Euclidean distance to the drawn position is determined. Last, the tree is expanded from the closest node toward the random position with a specified step size. In large environments, a deep expansion of the RRT using long branches can become computationally intractable. Therefore, the expansion is typically restricted to a finite horizon with limited branch length [9]. An overview of path planning algorithms, comparison of RRTs with other approaches, and analysis of their optimality can be found in [7].

2.2 Informative path planning

Autonomous robots are often tasked with exploring and mapping scalar fields within unknown environments. This endeavor is usually limited by resource constraints, such as finite fuel or battery life of the robots. The informative path planning (IPP) problem [8] can be stated as

$$\mathcal{P}^* = \underset{\mathcal{P}\in\Psi}{\operatorname{argmax}}\, I(\mathcal{P}) \text{ such that } c(\mathcal{P}) \leq B\,, \qquad (1)$$

where the best path $\mathcal{P}^*$ regarding the information quality metric (IQM) I is found from the collection of all paths Ψ under the constraint that its cost c does not exceed the resource budget B. Different strategies for choosing the IQM exist [8, 10], and two of them are described in the following. In the first strategy, $I(\mathcal{P})$ quantifies the uncertainty about yet unobserved measurement positions after measuring along path $\mathcal{P}$, i.e., the robots aim to gain information about all of the environment and avoid visiting the same position repeatedly. In the second strategy, $I(\mathcal{P})$ is based on the measured field itself, e.g., the robots aim to measure in locations where the analyzed field quantity is high.

2.3 Gaussian process regression

When the robots are placed into a new and unknown environment, no information about the measured field is available yet. As the measurements proceed, more data is gathered, thus requiring incremental updates of the robots' world estimate. Gaussian process (GP) regression [11] is commonly used in robotics to model scalar fields and their spatial correlation in a probabilistic way and progressively fuse data into an environment estimate that can be used during path planning [8, 10]. GP regression also has applications in acoustics, e.g., in sound field reconstruction [12].

The following mathematical formulation of two-dimensional GP regression is based on the thorough elaboration in [11]. Let $\mathbf{X}_\mathrm{r} = \{\mathbf{x}_\mathrm{r}^{(i)} = (x_\mathrm{r}^{(i)}, y_\mathrm{r}^{(i)} \mid i = 1, \ldots, N\}$ denote the collection of N positions, for which the receiver robot has already conducted a measurement. The field quantity under investigation at these positions is given by the vector $\mathbf{y}$. In GP regression, we aim to estimate the field quantity $\mathbf{y}^*$ at M previously unobserved positions $\mathbf{X}_\mathrm{r}^*$ in the environment, for instance, given by a dense uniform grid. We assume that $\mathbf{y}$ and $\mathbf{y}^*$ are jointly Gaussian random vectors with zero mean [11]

$$\begin{bmatrix}\mathbf{y}^* \\ \mathbf{y}\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}\mathbf{K}(\mathbf{X}_\mathrm{r}^*, \mathbf{X}_\mathrm{r}^*) & \mathbf{K}(\mathbf{X}_\mathrm{r}^*, \mathbf{X}_\mathrm{r}) \\ \mathbf{K}(\mathbf{X}_\mathrm{r}, \mathbf{X}_\mathrm{r}^*) & \mathbf{K}(\mathbf{X}_\mathrm{r}, \mathbf{X}_\mathrm{r})\end{bmatrix}\right), \qquad (2)$$

where $\mathcal{N}(\cdot)$ denotes the normal distribution, $\mathbf{0}$ the zero vector, and $\mathbf{K}(\mathbf{X}_\mathrm{r}, \mathbf{X}_\mathrm{r}{}^*)$ and its permutations the covariance or kernel function matrix. For example, $\mathbf{K}(\mathbf{X}_\mathrm{r}, \mathbf{X}_\mathrm{r}{}^*)$ is an $N \times M$ matrix, evaluating the covariance or kernel function for all pairs of measured and estimated positions. The kernel function models the spatial correlations of the investigated field, i.e., it describes how strongly a measurement point influences the field estimate at an unobserved, remote position. Various kernel functions have been described in the literature [11, 12], and choosing a suitable one depends heavily on the field quantity under investigation. The covariance matrix can also be adapted to deal with noisy observations [11], thus making the GP regression conceptually similar to Kalman filtering [13].

By conditioning the joint distribution on the previous observations, we can estimate the mean and covariance of the conditional distribution as follows [11]:

$$\begin{aligned}\mathbf{y}^*|\mathbf{y} \quad &\sim \quad \mathcal{N}\big(\boldsymbol{\mu}_{\mathbf{X}_\mathrm{r}^*|\mathbf{X}_\mathrm{r}}, \mathbf{K}_{\mathbf{X}_\mathrm{r}^*|\mathbf{X}_\mathrm{r}}\big) \qquad (3)\\ \boldsymbol{\mu}_{\mathbf{X}_\mathrm{r}^*|\mathbf{X}_\mathrm{r}} &= \quad \mathbf{K}(\mathbf{X}_\mathrm{r}^*, \mathbf{X}_\mathrm{r})\mathbf{K}(\mathbf{X}_\mathrm{r}, \mathbf{X}_\mathrm{r})^{-1}\mathbf{y}\,,\\ \mathbf{K}_{\mathbf{X}_\mathrm{r}^*|\mathbf{X}_\mathrm{r}} &= \quad \mathbf{K}(\mathbf{X}_\mathrm{r}^*, \mathbf{X}_\mathrm{r}^*)\\ &\quad\quad -\mathbf{K}(\mathbf{X}_\mathrm{r}^*, \mathbf{X}_\mathrm{r})\mathbf{K}(\mathbf{X}_\mathrm{r}, \mathbf{X}_\mathrm{r})^{-1}\mathbf{K}(\mathbf{X}_\mathrm{r}, \mathbf{X}_\mathrm{r}^*)\,.\end{aligned}$$

The mean $\boldsymbol{\mu}_{\mathbf{X}_\mathrm{r}^*|\mathbf{X}_\mathrm{r}}$ is the field quantity estimate for the unobserved positions $\mathbf{X}_\mathrm{r}^*$. This environment estimate can finally be used for planning the next measurement paths.

2.4 Common-slope model of late reverberation

In this paper, we intend to use the proposed robot system and measurement strategy for studying late reverberation phenomena. Late reverberation can be modeled as a linear combination of multiple decaying exponentials and a noise term [14]. In the most general scenario, late reverberation is inhomogeneous and anisotropic. Inhomogeneity refers to spatially varying late reverberation, i.e., variations due to different source or receiver positions. Anisotropy refers to directional variations, e.g., due to directional sources or receivers. Inhomogeneous and anisotropic late reverberation has been studied extensively [15–17], highlighting the requirement for a suitable late reverberation model.

The common-slope model was proposed as a general and compact representation of inhomogeneous and anisotropic late reverberation [6]. It is a model for the energy decay function (EDF), which is commonly determined via the Schroeder backwards integration procedure [18]. EDFs are typically processed in frequency bands (e.g. octave bands), but we drop the frequency dependency in the following to improve readability. For inhomogeneous and anisotropic late reverberation, the EDF $d(\mathbf{x}, t)$ has to be modeled for each source-receiver-configuration $\mathbf{x} = (\mathbf{x}_\mathrm{s}, \mathbf{x}_\mathrm{r}, \mathbf{\Omega}_\mathrm{s}, \mathbf{\Omega}_\mathrm{r})$, where $\mathbf{x}_\mathrm{s} = (x_\mathrm{s}, y_\mathrm{s}, z_\mathrm{s})$ and $\mathbf{x}_\mathrm{r} = (x_\mathrm{r}, y_\mathrm{r}, z_\mathrm{r})$ are the source and receiver positions, respectively. With azimuth and elevation angle denoted by ϕ and θ, respectively, the direction of departure from the source and the direction of arrival at the receiver are denoted by $\mathbf{\Omega}_\mathrm{s} = (\phi_\mathrm{s}, \theta_\mathrm{s})$ and $\mathbf{\Omega}_\mathrm{r} = (\phi_\mathrm{r}, \theta_\mathrm{r})$, respectively. The fundamental idea of the common-slope model is that the decay times of the exponentials are invariant across different source-receiver configurations, thus modeling inhomogeneity and anisotropy purely via decay amplitudes. The common-slope model is given by [6]

$$d_\kappa(\mathbf{x}, t) = N_{0,\mathbf{x}}\, \Psi_0(t) + \sum_{k=1}^{\kappa} A_{k,\mathbf{x}} \big[\Psi_k(t) - \Psi_k(L)\big]\,, \quad (4)$$

with the decay kernel

$$\Psi_k(t) = \begin{cases} L - t\,, & \text{if } k = 0 \\ \exp\!\left(\frac{-13.8\; t}{f_\mathrm{s}\; T_k}\right), & \text{if } k > 0 \end{cases}. \quad (5)$$

The model is characterized by the model order κ, specifying the number of exponentials used to model the EDF, where T_k and $A_{k,\mathbf{x}}$ are the decay times and amplitudes of the kth exponential, respectively. Furthermore, the model includes a noise term, whose amplitude is denoted by $N_{0,\mathbf{x}}$. The constant $-13.8 = \ln(10^{-6})$ ensures that the sound energy has decayed by 60 dB after T_k seconds. The constant term $\Psi_k(L)$ accounts for the finite upper limit of integration during the Schroeder backwards integration and can be neglected for large L [19].

Various approaches exist for determining the common-slope decay times T_k from measured data [6]. Once the T_k values are available, the common-slope amplitudes $A_{k,\mathbf{x}}$ can be determined via a least-squares fit to the measured EDFs [6]. The amplitude values can be visualized in common-slope amplitude maps of the environment, illustrating anisotropic and inhomogeneous late reverberation in a compact way. The resulting common-slope model can be used in all applications relying on late reverberation models, such as parametric spatial audio rendering, source separation, echo cancellation, and dereverberation.

Parameter	Value
Measurement horizon	$H = 14$
Nodes per iteration	$N_\mathrm{max} = 1200$
Step size	$s = 0.4\,\mathrm{m}$

Table 1: Parameters of our proposed measurement strategy.

3. MEASUREMENT STRATEGY

We propose a new measurement strategy for autonomous room acoustic measurements that combines RRTs (c.f. Sec. 2.1), IPP (c.f. Sec. 2.2), and GP regression (c.f. Sec. 2.3). In this paper, we use the strategy to measure common-slope amplitude maps (c.f. Sec. 2.4) of multi-room environments that potentially contain obstacles, and for which no floor plan or 3D model is available before the measurements.

In our measurement strategy, the receiver robot explores the scene while the source robot remains static [1]. The strategy involves three major steps. First, an RRT is expanded using the parameters summarized in Table 1. Second, the best path regarding a predefined information quality metric (IQM, c.f. Sec. 2.2 and Eq. (6)) is determined. Third, the receiver robot moves along the resulting path, and RIR measurements are conducted for every source-receiver configuration. After every measurement, the obstacle and common-slope amplitude map estimates are updated with the newly obtained data. Once all posi-

[1] We use a static source to facilitate the exposition and obtain illustrative common-slope amplitude maps, but our approach should generalize to the scenario when both robots move. In this more general case, which could also include directional sources and receivers, the GPR would model 10-dimensional amplitude maps (3 spatial axes and 2 angles for each robot) instead of the 2-dimensional maps depicted in this manuscript.

tions of the path have been measured, or if new obstacles are discovered on the way, a new RRT is expanded and the procedure is repeated.

We expand RRTs as follows:

1. The expansion starts with $i = 0$ from a random initial position $\mathbf{x}_\mathrm{r}^{(0)}$, which also serves as the root of the first RRT. We save $\mathbf{x}_\mathrm{r}^{(0)}$ to our list of measurement positions $\mathbf{X}_\mathrm{r}$, yielding $N = 1$.
2. If $N > N_\mathrm{max}$, the expansion is finished.
3. Sample a random position $\widetilde{\mathbf{x}}_\mathrm{r}$ from the unobstructed area $\mathcal{S}$ of the scene. The exact layout of $\mathcal{S}$ is typically unknown before the measurement, and therefore we just chose a large enough square that will certainly fit the entire scene.
4. Determine $\mathbf{x}_\mathrm{r}^{(\mathrm{c})}$, the measurement position in $\mathbf{X}_\mathrm{r}$ that is closest to $\widetilde{\mathbf{x}}_\mathrm{r}$.
5. Determine the path length from $\mathbf{x}_\mathrm{r}^{(0)}$ to $\mathbf{x}_\mathrm{r}^{(\mathrm{c})}$. If it is smaller than the measurement horizon H, continue with step 6. Otherwise the branch should not be expanded further and the algorithm returns to step 2.
6. Calculate a candidate for the next measurement position in the tree by making a step from $\mathbf{x}_\mathrm{r}^{(\mathrm{c})}$ into the direction of $\widetilde{\mathbf{x}}_\mathrm{r}$:
$$\widehat{\mathbf{x}}_\mathrm{r}^{(i+1)} = \mathbf{x}_\mathrm{r}^{(\mathrm{c})} + s * \frac{\widetilde{\mathbf{x}}_\mathrm{r} - \mathbf{x}_\mathrm{r}^{(\mathrm{c})}}{\|\widetilde{\mathbf{x}}_\mathrm{r} - \mathbf{x}_\mathrm{r}^{(\mathrm{c})}\|},$$
where $\|\cdot\|$ denotes the Euclidean distance.
7. Check whether $\widehat{\mathbf{x}}_\mathrm{r}^{(i+1)} \in \mathbf{X}_\mathrm{r}$. If it is, the position is already part of the tree and the algorithm returns to step 2.
8. Check from the current obstacle map estimate whether $\widehat{\mathbf{x}}_\mathrm{r}^{(i+1)}$ is reachable for the robot, i.e., no obstacles block the way. If the position cannot be reached, the algorithm returns to step 2.
9. If both checks passed, set $\mathbf{x}_\mathrm{r}^{(i+1)} = \widehat{\mathbf{x}}_\mathrm{r}^{(i+1)}$, add $\mathbf{x}_\mathrm{r}^{(i+1)}$ to $\mathbf{X}_\mathrm{r}$, start the next iteration with $i = i + 1$, and go back to step 2.

Once the RRT is expanded, we determine the best path for the next measurements. As we start the measurement strategy in a yet unknown environment, we use an IQM that emphasizes exploration. It is calculated by counting the number of unvisited cells in the neighborhoods around the path and penalizing the path if its cells have already been visited multiple times. The cells refer to a dense $10\,\mathrm{cm}$ grid of the entire environment. Throughout this paper, we calculate the IQM as follows:

$$\begin{aligned}
I(P) &= N_\mathrm{unvisited}(P) - N_\mathrm{multi}(P)\,, \qquad (6)\\
N_\mathrm{unvisited}(P) &= \sum_{j=0}^{\|P\|}\Big(\|\mathcal{X}_j\| - \sum_{\mathbf{x}_\mathrm{r}\in\mathcal{X}_j} v(\mathbf{x}_\mathrm{r})\Big)\,,\\
N_\mathrm{multi}(P) &= \sum_{j=0}^{\|P\|}\sum_{\mathbf{x}_\mathrm{r}\in\mathcal{X}_j} w(\mathbf{x}_\mathrm{r})\,,\\
\mathcal{X}_j &= \Big\{\mathbf{x}_\mathrm{r}\,,\ \text{such that } \|\mathbf{x}_\mathrm{r} - \mathbf{x}_\mathrm{r}^{(j)}\| \le 0.5\,\mathrm{m}\Big\}\,,\\
v(\mathbf{x}) &= \begin{cases} 1\,, & \text{if } \mathbf{x} \text{ was visited before}\\ 0\,, & \text{otherwise}\end{cases}\,,\\
w(\mathbf{x}) &= \text{Number of visits at } \mathbf{x}\,,
\end{aligned}$$

where $\|P\|$ and $\|\mathcal{X}_j\|$ denote the number of cells in path P or neighborhood $\mathcal{X}_j$, respectively. Alternatively, the IQM could also be based on the variance reduction in the Gaussian process achieved by the path [8], inherently tying the metric to the measured field quantity. The IQM could also be designed to consider properties of the measured field quantity. For example, one could preferably measure paths where the common-slope amplitudes (c.f. Sec. 2.4) vary considerably, maximizing the gradient magnitudes of the common-slope values along the path.

After determining the best path, the robots conduct measurements in the corresponding positions. The obstacle and common-slope maps are consecutively updated with the gathered data. We use the rational quadratic kernel in the Gaussian process regression, given by [11]

$$\mathbf{K}_\mathrm{RQ}(\mathbf{x}_\mathrm{r}, \mathbf{x}'_\mathrm{r}) = \Big(1 + \frac{\|\mathbf{x}_\mathrm{r} - \mathbf{x}'_\mathrm{r}\|^2}{2\alpha l^2}\Big)^{-\alpha}\,, \qquad (7)$$

where $\|\cdot\|$ denotes the Euclidean distance, α is the scale mixture parameter, and l is the kernel length scale. The rational quadratic kernel can describe smooth spatial functions that vary across many length scales. In other words, it can represent patterns that vary with different spatial frequencies. The hyperparameters α and l are learned while fitting the kernel to the measured data by maximizing the log-marginal likelihood [11]. During the fitting process, we assume noisy observations, thus adding a uniform noise term to the kernel matrices in Eqs. (2) and (3) as outlined in [11, 13]. Uncertainties regarding the position tracking or common-slope analysis can then be included in the model, such that the GPR does not necessarily fit the measurements exactly but allows small deviations.

Figure 1: Robot system used in the evaluation. It contains two Boston Dynamic Spot robots, equipped with a loudspeaker and a microphone array, respectively.

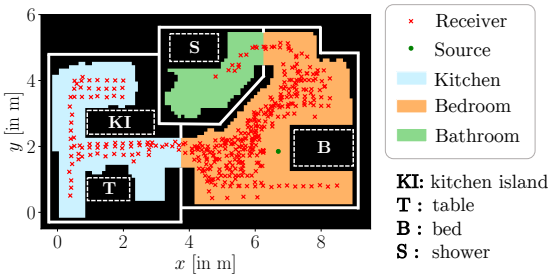


Figure 2: Flat-like environment consisting of a kitchen, bedroom, and bathroom. A robot system was used to measure RIRs from 420 different source-receiver configurations. The source remained static for all measurements. The walls are depicted as thick white lines. Some obstacles and their limits are indicated with dashed white lines. Black areas are obstructed for the robots.

4. EVALUATION

We evaluate the proposed measurement strategy using a new robot system in a multi-room environment with obstacles, for which no geometric information (floor plan, 3D model) is available before the measurements. The evaluation has two parts. In the first part (c.f. Sec. 4.4), the RRT-based strategy is compared against other path planning approaches while the interpolation method is fixed. In the second part (c.f. Sec. 4.5), the reconstruction of common-slope maps via GP regression is compared against other interpolation methods while the path planning method is fixed. In both cases, we measure success by the ability of the system to approximate the ground truth data, which is a common-slope amplitude map obtained via GP regression on a dense measurement grid.

4.1 Robot system

We evaluate the proposed measurement strategy using the robot system shown in Fig. 1, which is similar to the one outlined in [3]. It contains two Boston Dynamics Spot robots, one acting as a source and one as a receiver. The source robot is equipped with an MSE Audio Soundsphere loudspeaker, and the receiver robot uses an open SDM array with Earthworks M30 as the omnidirectional channel and six DPA 4060s as the orthogonal pairs. Both robots are equipped with the Spot Core processor, and the audio signals are played back and recorded with a Focusrite Scarlett Solo and Zoom F8 audio interface, respectively.

The robots, their processors, and the measurement laptop communicate via a wireless TCP network socket. The measurement script for moving the robots, planning paths, and conducting measurements is implemented in Python. The Boston Dynamics robots provide position tracking and obstacle detection capabilities via cameras at each side. The obstacle detection returns occupancy maps of approximately $1.5\,\mathrm{m} \times 1.5\,\mathrm{m}$ around the robots. We use exponential sine sweeps to measure RIRs [20] and turn off the robot fans during the measurements.

4.2 Dataset under investigation

We used the above robot system to measure RIRs in a flat-like environment consisting of three connected rooms (bedroom, kitchen, and bathroom). The measured dataset contains RIRs from 420 different source-receiver-configurations, for which the source remained static and the receiver robot measured the positions as illustrated in Fig. 2. All rooms feature various obstacles, such as a bed, a table with chairs, a kitchen island, a shower, or a toilet.

4.3 RRT expansion in the investigated environment

Fig. 3 shows an example of the RRT expansion in the above multi-room environment. In Fig. 3a, the receiver robot only sees its immediate environment (approx. $1.5\,\mathrm{m} \times 1.5\,\mathrm{m}$) and plans its path accordingly. After a few steps, it discovers new obstacles and determines that the path cannot be measured until the end. A new RRT expansion brings the robot into the bathroom (c.f. Fig. 3b). Once the robot has explored the bathroom and made its way back to the bedroom, the RRT directs the robot into the kitchen (c.f. Fig. 3c). After 120 steps, the robot has explored all three rooms (c.f. Fig. 3d).

4.4 Comparison against other path planning methods

We compare the RRT-based path planning against two other approaches. Each approach yields a sequence of steps, and we use the closest receiver position from the above dataset

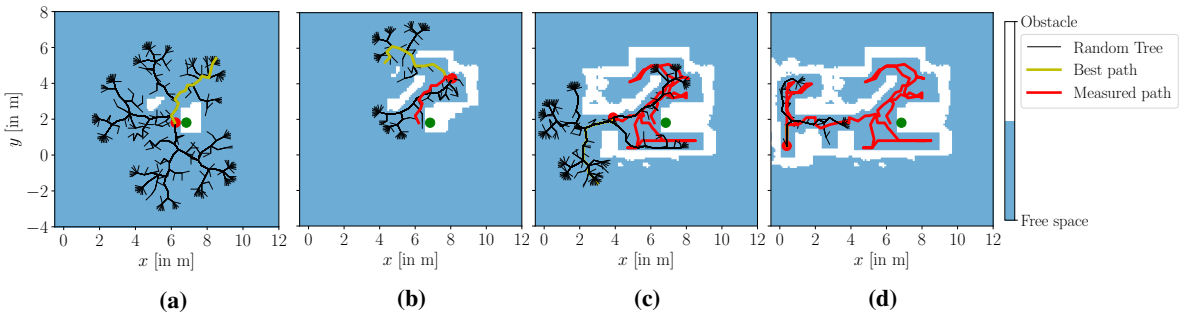


Figure 3: Example of the RRT-based measurement strategy in the investigated multi-room environment. The figure illustrates the measurement strategy at various time instances. For all of the snapshots, the source and receiver robot positions are indicated by the green and red dot, respectively. The source robot is modeled as a static obstacle, around which the receiver robot plans its paths. The RRT is shown in black and the best path in each iteration is depicted in yellow. The path measured by the receiver robot is indicated in red.

for calculating the evaluation metrics and plots. The first baseline approach, *"Random Neighbor" (RandN)*, extends the pure random walk approach by allowing the robots to plan ahead until a specified measurement horizon. The paths are determined by making a random step within a $0.8\,\mathrm{m} \times 0.8\,\mathrm{m}$ neighborhood centered around the previous step. During each planning iteration, 86 paths with $H = 14$ steps are planned, resulting in $86 \times 14 = 1204$ planned steps in total. RandN uses the same IQM as the proposed RRT-based approach, c.f. Eq. (6).

The second baseline approach, *"Random Walk" (RandW)*, implements the random walk proposed in [3] without a measurement horizon. The receiver robot chooses a random position within a $3.2\,\mathrm{m} \times 3.2\,\mathrm{m}$ neighborhood around the previous step and tries to reach it. If an obstacle is detected along the way, the robot stops in front of it and continues with the next step.

The goal of the proposed measurement strategy is to reconstruct late reverberation characteristics from common-slope amplitude maps. Therefore, the ground truth for our evaluation is the GP regression using all 420 measurements from the environment (c.f. Fig. 2). We obtain the ground truth via GP regression [c.f. Eqs. (2) and (3)] by setting

$$\begin{aligned} \mathbf{X}_{\mathrm{r,true}} &:= \text{420 receiver positions from Fig. 2}\,, \\ \mathbf{A}_{k,\mathrm{true}} &:= A_{k,\mathbf{x}} \text{ values at } \mathbf{X}_{\mathrm{r,true}}\,, \\ \mathbf{X}_{\mathrm{r}}^{*} &:= \text{uniform grid with } 10\,\mathrm{cm} \text{ resolution}\,. \end{aligned}$$

Fig. 4 shows the ground truth common-slope maps. Our analyses use the common-slope model [c.f. Eqs. (4) and (5)], thus inherently modeling background noise as a separate term. The two decay processes are localized in the bedroom/kitchen and bathroom, respectively, exhibiting a cross-fade between the bedroom and bathroom.

In the comparison, the $\mathbf{X}_{\mathrm{r}}$ and $\mathbf{A}_k$ values are set according to the receiver positions and common-slope amplitude values from the RRT, RandN, and RandW approach. The GP regression is then recalculated on the corresponding subsets to determine the common-slope maps for the different approaches. The dB-MSE ε between the ground truth and the common-slope amplitude maps obtained from a certain approach is calculated by

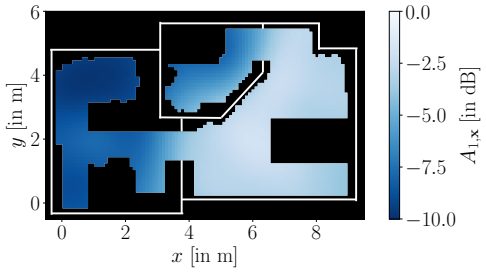
$$\varepsilon = \frac{1}{\kappa \|\mathbf{X}_{\mathrm{r}}^{*}\|} \sum_{k=1}^{\kappa} \sum_{\mathbf{x}_{\mathrm{r}} \in \mathbf{X}_{\mathrm{r}}^{*}} \left[A_{k,\mathbf{x}_{\mathrm{r}},\mathrm{true}}^{(\mathrm{dB})} - A_{k,\mathbf{x}_{\mathrm{r}},\mathrm{appr}}^{(\mathrm{dB})} \right]^2 , \quad (8)$$

where both amplitude values are represented on a logarithmic scale in dB, and $\|\mathbf{X}_{\mathrm{r}}^{*}\|$ denotes the number of positions in $\mathbf{X}_{\mathrm{r}}^{*}$. A dB-MSE value of 0 dB corresponds to a perfect reconstruction of the ground truth.

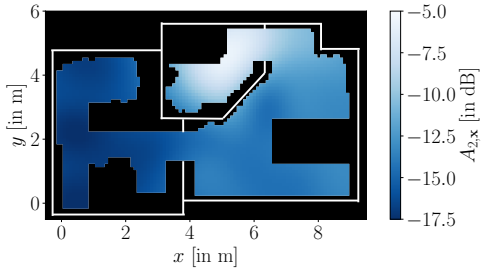
Fig. 5a illustrates how many steps the path planning approaches require to achieve a small dB-MSE. The proposed RRT-based approach achieves a dB-MSE below 1 dB already after approximately 100 steps. At this time, each room was visited at least once. In contrast, the RandN and RandW approaches require more than 200 or 1050 steps, respectively, to achieve a comparably low reconstruction error. Both approaches spend considerably more time in each room, repeatedly measuring configurations that are similar or identical. Especially the RandW approach seems to get easily stuck in the individual rooms, having difficulties to explore the entire environment.

4.5 Comparison against other interpolation methods

We compare the GP regression (c.f. Sec. 2.3) against nearest neighbor, linear, and cubic interpolation regarding their reconstruction performance of common-slope amplitude



(a) $T_1 = 0.53\,\mathrm{s}$

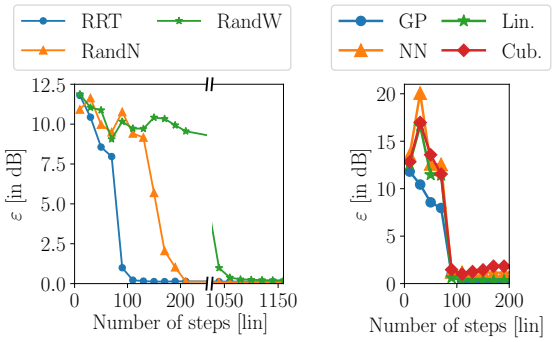


(b) $T_2 = 1.02\,\mathrm{s}$

Figure 4: Ground truth common-slope amplitude maps of the investigated multi-room environment, obtained via GP regression on all 420 receiver positions with static source. The second-order common-slope model [c.f. Eq. (4)] is used to describe the inhomogeneous late reverberation of the environment. (a) The decay process with the smaller decay time is strongest in the bedroom and kitchen. (b) The decay process with the larger decay time is strongest in the bathroom. The two decay processes cross-fade during the transition from bedroom to bathroom.

maps. We use the receiver positions obtained from the RRT-based path planning approach and reconstruct the $\mathbf{A}_k^*$ values for the uniform grid $\mathbf{X}_\mathrm{r}^*$. For linear and cubic interpolation, we only interpolate within the convex hull spanned by the data points, and determine outside positions by using their nearest neighbors.

Fig. 5b illustrates the reconstruction performance of the different approaches regarding the dB-MSE ε [c.f. Eq. (8)]. Generally, the approaches perform similarly with low dB-MSE values after approximately 100 steps, but GP regression and linear interpolation achieve the smallest errors. Nearest neighbor, linear, and cubic interpolation exhibit an increased peak after approximately 30 steps because the nearest neighbor interpolation for receiver positions outside the convex hull extends the amplitude values across room boundaries. This phenomenon is not as pronounced for the GP regression due to the finite kernel length scale, which models the uncertainty for positions far away from the measured data.



(a) Path planning approaches

(b) Interpolation approaches

Figure 5: Comparison of (a) different path planning approaches (RRT: Rapidly-exploring Random Tree, RandN: Random Neighbor, RandW: Random Walk) and (b) different interpolation approaches (GP: Gaussian Process regression, NN: Nearest Neighbor, Lin.: Linear interpolation, Cub.: Cubic interpolation). The plot shows the dB-MSE [c.f. Eq. (8)] between the reconstruction achieved by the respective strategies and the ground truth obtained from GP regression on a dense grid (c.f. Fig. 2). A dB-MSE value of 0 dB corresponds to a perfect reconstruction.

The error increases again for the cubic interpolation after approximately 100 steps due to overshooting.

5. DISCUSSION

The evaluation demonstrated the proposed measurement strategy and indicated that RRT-based path planning and GP regression outperformed other approaches in the investigated example. Path planning using RRTs quickly explores the investigated multi-room environment, whereas the RandN and RandW approach had difficulties transitioning between the rooms. Although the RandW approach has reportedly yielded uniform coverage in environments with a single room [3], it seems to fail in more general multi-room environments.

The comparison of different interpolation approaches did not show marked differences in the investigated scenario. Standard interpolation approaches reconstruct the common-slope maps similarly well as the GP regression. However, we assume that the GP regression will leverage its full potential when reconstructing other field quantities that exhibit more complicated spatial interactions, or when reconstructing the entire sound field [12]. Furthermore, GP regression has a significant advantage in the robotic mapping scenario because it allows dealing with uncertainty and measurement noise [11, 13].

6. CONCLUSION

This paper proposed a new measurement strategy for autonomous robot-based room acoustic measurements in environments, for which no geometric information is available before the measurements. The strategy uses rapidly-exploring random trees (RRTs) for path planning and continuously estimates the investigated acoustic field quantity using Gaussian process (GP) regression. We demonstrate the strategy in a multi-room environment, for which we measure and reconstruct a late reverberation quantity, the common-slope amplitude. Our evaluation shows that the RRT-based path planning quickly explores the environment, while other approaches have difficulties transitioning between the rooms. The GP regression is suitable for noisy measurements and achieves slightly improved reconstruction performance compared to other interpolation methods. Autonomous robot systems using the proposed strategy will speed up room acoustic data collection significantly.

7. ACKNOWLEDGMENTS

The authors would like to thank Scott Colburn and David Kalamen for designing the robot system, building the loudspeaker and microphone array fixtures, and attaching them on the robots.

8. REFERENCES

[1] S. A. Verburg and E. Fernandez-Grande, "Reconstruction of the sound field in a room using compressive sensing," *J. Acoust. Soc. Am.*, vol. 143, no. 6, pp. 3770–3779, 2018.

[2] I. B. Witew, M. Vorländer, and N. Xiang, "Sampling the sound field in auditoria using large natural-scale array measurements," *J. Acoust. Soc. Am.*, vol. 141, no. 3, pp. EL300–EL306, 2017.

[3] G. Götz, A. M. Ornelas, S. J. Schlecht, and V. Pulkki, "Autonomous Robot Twin System for Room Acoustic Measurements," *J. Audio Eng. Soc.*, vol. 69, no. 4, pp. 261–272, 2021.

[4] G. Stolz, S. Werner, F. Klein, L. Treybig, A. Bley, and C. Martin, "Autonomous Robotic Platform To Measure Spatial Room Impulse Responses," in *Jahrestagung für Akustik (DAGA)*, pp. 59–61, 2023.

[5] S. M. LaValle, "Rapidly-Exploring Random Trees: A New Tool for Path Planning," tech. rep., Iowa State University, Computer Science Department, 1998.

[6] G. Götz, S. J. Schlecht, and V. Pulkki, "Common-slope modeling of late reverberation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, 2023.

[7] S. Karaman and E. Frazzoli, "Sampling-based Algorithms for Optimal Motion Planning," *Int. J. Robot. Res.*, vol. 30, no. 7, pp. 846–894, 2011.

[8] G. A. Hollinger and G. S. Sukhatme, "Sampling-based robotic information gathering algorithms," *Int. J. Robot. Res.*, vol. 33, no. 9, pp. 1271–1287, 2014.

[9] G. Hollinger and S. Singh, "Proofs and Experiments in Scalable, Near-Optimal Search by Multiple Robots," in *Robotics: Science and Systems IV* (J. Trinkle, F. Ramos, and O. Brock, eds.), pp. 206–213, Cambridge, MA, USA: MIT Press, 1st ed., 2009.

[10] G. Hitz, E. Galceran, M. Garneau, F. Pomerleau, and R. Siegwart, "Adaptive continuous-space informative path planning for online environmental monitoring," *J. Field Robot.*, vol. 34, no. 8, pp. 1427–1449, 2017.

[11] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: The MIT Press, 2006.

[12] D. Caviedes-Nozal, N. A. B. Riis, F. M. Heuchel, J. Brunskog, P. Gerstoft, and E. Fernandez-Grande, "Gaussian processes for sound field reconstruction," *J. Acoust. Soc. Am.*, vol. 149, no. 2, pp. 1107–1119, 2021.

[13] S. Reece and S. Roberts, "An introduction to Gaussian processes for the Kalman filter expert," *2010 13th Int. Conf. Inf. Fusion*, pp. 1–9, 2010.

[14] N. Xiang and P. M. Goggans, "Evaluation of decay times in coupled spaces: Bayesian parameter estimation," *J. Acoust. Soc. Am.*, vol. 110, no. 3, pp. 1415–1424, 2001.

[15] T. Akama, H. Suzuki, and A. Omoto, "Distribution of selected monaural acoustical parameters in concert halls," *Appl. Acoust.*, vol. 71, no. 6, pp. 564–577, 2010.

[16] N. Xiang, J. Escolano, J. M. Navarro, and Y. Jing, "Investigation on the effect of aperture sizes and receiver positions in coupled rooms," *J. Acoust. Soc. Am.*, vol. 133, no. 6, pp. 3975–3985, 2013.

[17] M. Berzborn and M. Vorländer, "Directional sound field decay analysis in performance spaces," *Building Acoustics*, pp. 1–15, 2021.

[18] M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, no. 3, pp. 409–412, 1965.

[19] N. Xiang, P. M. Goggans, T. Jasa, and M. Kleiner, "Evaluation of decay times in coupled spaces: Reliability analysis of Bayeisan decay time estimation," *J. Acoust. Soc. Am.*, vol. 117, no. 6, pp. 3707–3715, 2005.

[20] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique," in *108th Conv. AES*, pp. 1–23, 2000.