

MeshRIR: 音場分析・合成のための 格子点上室内インパルス応答データセット*

○小山翔一, 西田智哉, 木村圭佑, 阿部拓実 (東京大学),
植野夏樹 (東京都立大学), Jesper Brunnström (東京大学)

1 はじめに

“MeshRIR”と呼ぶインパルス応答データセットを新たに公開した。既存のデータセットは、いくつかの音源位置からマイクロフォンアレイまでのインパルス応答を様々な室内環境で計測したものが多く、基本的には音声強調や音声認識の性能評価を目的に設計されている。一方、空間的な音場を分析/推定、あるいは合成/制御する手法が、ここ十年ほどで大きく発展した。しかしながら、現状のインパルス応答データセットは、測定点の空間解像度が極めて低く、音場分析・合成法の評価に用いることは難しい。MeshRIRは、ある領域内部を細かく離散化し、各点において測定したインパルス応答によるデータセットとなっており、音場分析・合成法の性能評価や結果の可視化などに適している。現在、2つのサブデータセットで構成されており、一方が単一音源から3次元の直方体領域内部のインパルス応答、もう一方が32音源のアレイから2次元の正方形領域内部のインパルス応答となっている。データセットはサンプルコードも含めて <https://sh01k.github.io/MeshRIR> からダウンロード可能である。

2 データセットの概要

インパルス応答は、音源から受音点までの音響伝播特性を線形時不変システムとして記述する。3次元空間の音響数値シミュレーションは膨大な計算コストを要することから、実環境で測定したインパルス応答は可聴化 (auralization) のような用途以外にも、音声強調、残響除去、音源定位・分離、音声認識など、様々な音響信号処理技術の性能を評価する上で有用である。これまでに多数のインパルス応答データセットが公開されているが、その多くは数個のスピーカ・マイクロフォン位置でのインパルス応答を、様々な残響環境で測定したものであった。これらの既存のデータセットに対し、我々が公開する“MeshRIR”は以下のような特徴を持つ。

- 3次元あるいは2次元の領域内部を細かく離散化した格子点上を測定点とする。測定点の間隔

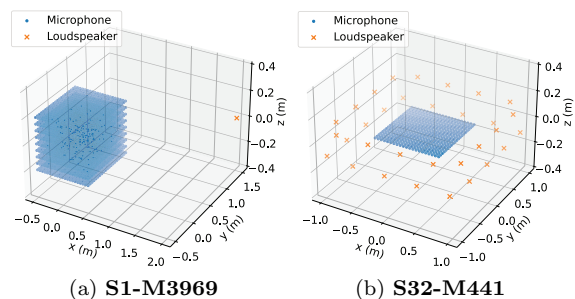


Fig. 1: インパルス応答測定時のスピーカ・マイクロフォン位置

は 0.05 m である。

- 現在、2種類のサブデータセットが利用可能であり、一方は単一音源から3次元直方体 (1.0 m × 1.0 m × 0.4 m) の測定領域、もう一方は32音源のアレイから2次元正方形 (1.0 m × 1.0 m) の測定領域のデータで構成されている。
- 複数のマイクロフォン信号から領域内の音圧分布を推定する音場再構成法や、スピーカアレイを用いて領域内に所望の音圧分布を合成する音場合成法のためのサンプルコードも同時に公開している。

上記2種類のサブデータセットは、3次元直方体測定領域のものを **S1-M3969**、2次元正方形測定領域のものを **S32-M441** と呼ぶ。これはそれぞれ音源数が1と32、測定点数が3969と441であることに由来する。**S1-M3969** および **S32-M441** の音源と測定点の位置を Fig. 1 に示す。

測定条件の詳細を Table 1 に示す。インパルス応答測定のためのスピーカ信号には、Linear swept-sine 信号 [1] を用いた。直交ロボットに無指向性マイクロフォン (Primo EM272J) を取り付けたものを使用し、その3次元位置を制御することで、各格子点におけるインパルス応答を測定した。音源用のスピーカとしては、**S1-M3969** では DIATONE DS-7、**S32-M441** では YAMAHA VXS1MLB を用いた。**S32-M441** の

*MeshRIR: A dataset of room impulse responses on meshed grid points for sound field analysis and synthesis. by KOYAMA, Shoichi, NISHIDA, Tomoya, KIMURA, Keisuke, ABE, Takumi (The University of Tokyo), UENO, Natsuki (Tokyo Metropolitan University), BRUNNSTRÖM, Jesper (The University of Tokyo)

Table 1: **S1-M3969** および **S32-M441** の測定条件の詳細

	S1-M3969	S32-M441
サンプリング周波数	48000 Hz	
インパルス応答長	32768 samples	
室の寸法	7.0 m $\times$ 6.4 m $\times$ 2.7 m	
音源数	1	32
測定領域	3次元直方体: 1.0 m $\times$ 1.0 m $\times$ 0.4 m	2次元正方形: 1.0 m $\times$ 1.0 m
測定位置間隔	0.05 m	
測定点数	21 $\times$ 21 $\times$ 9 (= 3969) points	21 $\times$ 21 (= 441) points
残響時間 (T_{60})	0.38 s	0.19 s
平均室温	26.3°C	17.1°C

測定時には吸音パネルをいくつか設置したため、残響時間が **S1-M3969** と比べて短くなっている。信号の入出力には Dante インタフェースを備えた PC を利用した。インパルス応答データは MATLAB 用の MAT フォーマット、および NUMPY 用の NPY フォーマットが利用可能であり、それ以外の付加的なデータは JSON ファイルとして提供している。より詳細なデータセットの情報は、ウェブサイト (<https://sh01k.github.io/MeshRIR>) あるいは文献 [2] を参照されたい。

また、データセットと合わせて、複数のマイクロフォン信号から領域内の音圧分布を推定する音場再構成法、スピーカアレイを用いて領域内に所望の音圧分布を合成する音場合成法のためのサンプルコードも公開している。音場再構成法としては、無限次元調和解析に基づく文献 [3] の手法を用いた。音場合成法では、多点音圧制御法 [4,5] の他、各スピーカの伝達関数の球波動関数展開係数を文献 [3] の方法で推定し、重み付きモードマッチング法 [6] を用いる手法のコードを公開している。

3 測定音場の可視化

S1-M3969 を用いて、音源位置の z 座標と同じ $z = 0$ の xy 平面上の音圧分布を可視化した結果を示す。このような可視化は、測定点の空間解像度が高い MeshRIR によって可能となる利用例の一つである。 $z = 0$ の測定面のインパルス応答データに対して、カットオフ周波数 600 Hz の低域通過フィルタを適用後、40 sample ごとの xy 平面上の瞬時音圧分布を Fig. 2 に図示した。 xy 平面上を直接波および初期反射が通過する様子が見て取れる。

4 おわりに

音場分析・合成の技術はここ十年ほどで大きく発展した一方、その評価は数値シミュレーションによるものに限られている場合が多く、実環境での性能評価

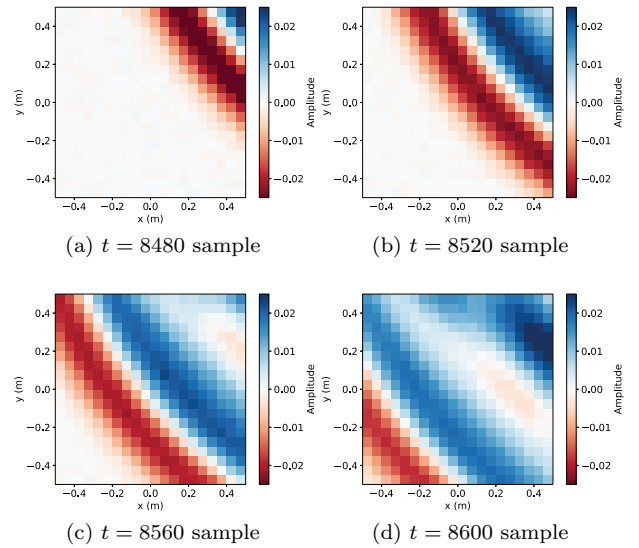


Fig. 2: **S1-M3869** のインパルス応答に対して低域通過フィルタを適用し、瞬時音圧分布を可視化した例。

は十分になされてこなかった。本データセットの公開が、今後、音場分析・合成技術の発展の一助となれば幸いである。

謝辞 本研究は、JST さきがけ JPMJPR18J4 の助成を受けたものです。

参考文献

- [1] Suzuki, et al. In *JASA*, 97(2), 1995.
- [2] Koyama, et al. In *arXiv preprint arXiv: 2106.10801*, 2021.
- [3] Ueno, et al. In *IEEE Signal Process. Lett.*, 25(1), 2018.
- [4] Kirkeby, et al. In *JASA*, 94(5), 1993.
- [5] Koyama, et al. In *IEEE/ACM Trans. ASLP*, 28, 2020.
- [6] Ueno, et al. In *IEEE/ACM Trans. ASLP*, 27(12), 2019.