

# 深層学習を用いた三次元室内音場における壁面吸音率の推定

丸山 和旺<sup>†</sup> 新井 大斗<sup>†</sup> 佐藤 元<sup>†</sup> 池田 雄介<sup>†</sup>

<sup>†</sup>東京電機大

## 1 はじめに

室内音響の把握には、吸音率を始めとする境界条件の情報が重要である。しかし、コンサートホールなどにおいて空間全体の境界条件を現場で測定する事は容易ではない。このため、近年、局所に配置された複数のマイクロホンで計測された室内インパルス応答からデータ駆動型の吸音率推定手法が提案されている[1][2][3]。しかし、三次元音場に関しては、複雑な部屋形状に対する検討はあまり行われていない。

そこで、本研究では深層学習を用いて、少数の室内インパルス応答から、水平方向を任意の四辺形、高さを一定とした部屋形状において、壁面と天井、床面の6面に対して吸音率推定を行う手法を提案する。

## 2 提案手法

提案手法の概要を図1に示す。天井と床面が平行な六面体の形状の三次元室内音場を考える。音源の位置は固定とする。部屋の中央付近の領域に配置された少数のマイクロホンから、深層学習を用いて天井、床、4壁面のそれぞれの吸音率を推定する。

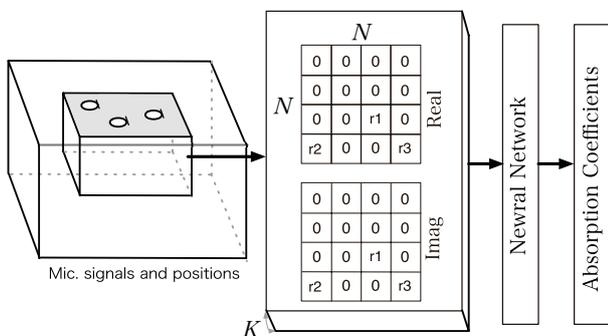


図1: 提案手法の概念図

### 2.1 データセットの作成

部屋の中央にマイクロホンの配置を想定する領域として、格子点 ( $N \times N \times K$ ) を想定する。この格子点の中からランダムに選択された少数点における複素音圧を入力信号とする。得られた複素音圧を実部と虚部に分け、音圧データの無い部分を零とした実数の

Deep learning based estimation of sound absorption coefficients in 3D room acoustics

<sup>†</sup> Waoh Maruyama (23fmi28@ms.dendai.ac.jp)

<sup>†</sup> Hiroto Arai (23fmi03@ms.dendai.ac.jp)

<sup>†</sup> Gen Sato (22fmi21@ms.dendai.ac.jp)

<sup>†</sup> Yusuke Ikeda (yusuke.ikeda@mail.dendai.ac.jp)

Tokyo Denki University (<sup>†</sup>)

行列を入力データセット (幅, 高さ, チャンネル) = ( $N \times N \times (K \times 2)$ ) として作成する。教師データとして6壁面 (4壁面, 天井, 床) の吸音率 (0.0 - 1.0) を一次元配列に格納する。

部屋の形状は、水平面上でランダムに選ばれた4点を頂点とする四辺形を、高さ方向に押し出して作成された6面体とする。このとき、マイクロホンと音源の位置は必ず部屋の内側に含む。また、床面と天井は平行で高さは固定とする。

データはすべて虚像法を用いたシミュレーションによって作成され、本稿では、一次反射音までを考慮することとする。

### 2.2 深層学習による推定

シミュレーションによって作成した音圧のデータセットを用いて、壁面の吸音率を教師データとした回帰学習を行う。図2に学習に用いた畳み込みニューラルネットワーク (CNN) のアーキテクチャを示す。カーネルサイズは (2, 2) とし、すべての Dropout の比率は 0.25 とした。

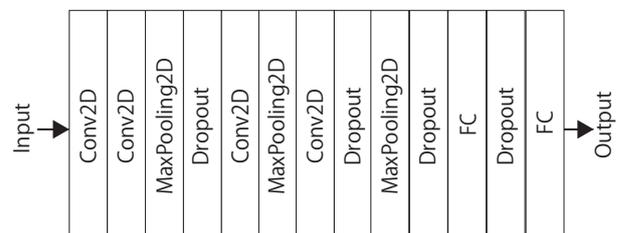


図2: ニューラルネットワークのアーキテクチャ

## 3 シミュレーション実験

### 3.1 実験条件

表1: データセットのシミュレーション条件

音速 [m/s]	340
周波数 [Hz]	500
格子数	21 × 21 × 5 (間隔 $x-y:0.05$ m, $z:0.5$ m)
データ数	10000
吸音率	0-1(一様ランダム)
マイクロホン数	15(1つの水平面につき3本)

推定する部屋の高さは3mで固定とした。部屋の最小体積は2m × 2m × 3m, 最大体積は10m × 10

m × 3 m となり、六面体形状の室内の三次元音場を虚像法 (pyroomacoustics [4]) を用いてシミュレーションした。シミュレーション条件を表 1 に示す。

壁面の水平方向の頂点の座標はそれぞれ 4 m × 4 m の範囲でランダムに決定した。入力データとなる測定点は、水平方向において、部屋の中心における 1 m × 1 m の範囲を 21 × 21 の格子状に分割し、その中からランダムに選ばれた 3 つの格子点とした。これを高さ方向においても同様に、床面を基準とした 0.5 m から 2.5 m まで 0.5 m ごとに 5 つの平面それぞれで設定をする。したがって、合計 15 点の音圧を入力信号に用いた。音源は高さ 1.5 m で、計測領域に対して水平方向に 1 m 離れた位置に配置した。

6 壁面を 1 組とした吸音率設定を合計 8000 パターン用意し、それぞれ音場シミュレーションを行い、入力信号を作成した。作成したデータのうち 70% を学習データ、30% を検証データとした。また、テスト用に新たに生成した 2000 パターンの吸音率設定からシミュレーションした音場データを用いて、推定精度の検証を行った。学習は Tensorflow-gpu (Ver. 2.9.1) を用いて行った。

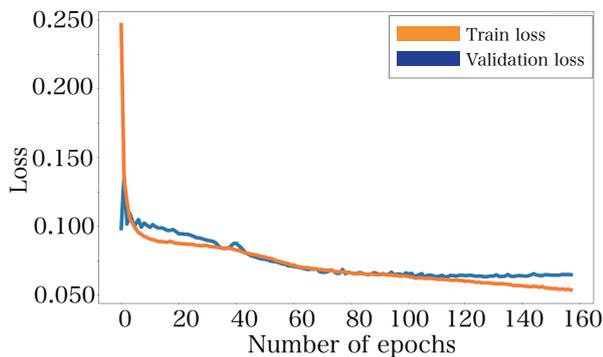


図 3: 学習損失と検証損失

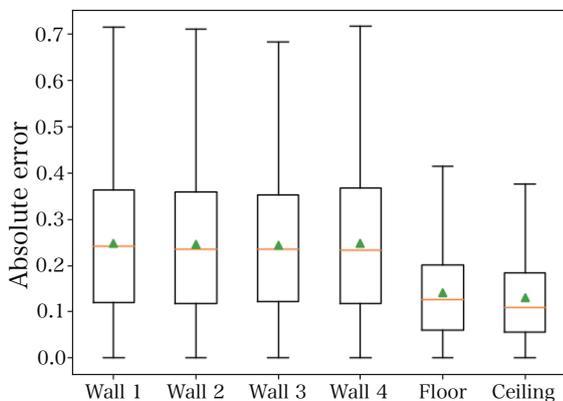


図 4: 壁面・床・天井に対する吸音率推定精度の絶対誤差の箱ひげ図

### 3.2 結果

図 3 に訓練時と検証時の損失を示す。訓練損失と検証損失は学習が進むにつれ共に減少し、おおよそ収束した。また、100 Epoch 付近から訓練損失と検証損失の差が開き始め、過学習が生じ始めているが、早期終了の適用により過学習が過度に進む前に停止されている。

図 4 に推定吸音率の絶対誤差の箱ひげ図を示す。4 つの壁面においては中央値で約 0.25 程度の誤差があり、推定精度は十分とは言えない。一方で、天井と床面における絶対誤差の中央値は約 0.1 程度と、壁面と比べて推定精度が得られた。これは水平面の壁面は様々な形に変化するが、床面、天井面は平行で同じ距離に保たれていた条件であることやマイクロホンの高さも固定されているため、水平面と比べてデータに偏りが大きく推定が容易になったと考えられる。

### 4 おわりに

本稿では、高さ固定の任意の 6 面体の部屋形状に対して、データ駆動型の吸音率推定手法を提案した。実験から推定精度に課題が見られたため、今後は、特に水平方向の壁面の推定精度の向上を目指して手法の検討を行う。

**謝辞** 本研究は東京電機大学総合研究所研究 Q22J-02 の助成を受けたものです。

### 参考文献

- [1] Cédric Foy, *et al.*, “Mean absorption estimation from room impulse responses using virtually supervised learning,” *Journal of the Acoustical Society of America*, 150 (2), pp.1286–1299, 2021.
- [2] Stéphane Dilungana, *et al.*, “Learning-based estimation of individual absorption profiles from a single room impulse response with known positions of source, sensor and surfaces,” *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, pp 5623–5630, 2021.
- [3] Hiroto Arai, *et al.*, “Data-driven estimation of sound absorption coefficients considering the positions of microphones,” *InrerNoise23*, 2023.
- [4] pyroomacoustics, <https://github.com/LCAV/pyroomacoustics> (最終閲覧日 2024-1-17)