

携帯端末を用いた簡便な音場測定と可視化

Adam Manato Maeda Bin Amizan¹ 椋木雅之¹

概要：本研究では、携帯端末を使用した簡便な音場測定を目指す。これにより、空間中の様々な位置での音の情報が測定でき、音場設計に役立つ。具体的には、音源付近にチェッカーボードを配置し、動画を撮影しながら携帯端末を移動させて音源からの音を測定する。得られた動画の画像（フレーム）に対して、コンピュータビジョンの自己位置推定技術を適用し、同時に音源由来の音に周波数解析を施して音圧を数値化する。実験では音圧再現性と位置推定再現性の確認を行い、これらを組み合わせて音場を可視化した。音圧の再現性はやや低い、位置推定精度は高かった。

キーワード：音響分析、音場測定、自己位置推定、可視化

1. はじめに

世の中に駅等のアナウンスがうるさいと思う人がある。その駅を頻りに利用する常連客にはアナウンスが案内の役には立たず、逆にストレスを与えてしまう場合がある。また、駅コンコースなどでは、雑音アナウンスと同じくらいの音圧となり、アナウンスが聞き取りにくいという調査がある[1]。これらの問題を解決するために、指向性スピーカーを用いることが考えられる。指向性スピーカーを用いてアナウンスを特定の人にだけ発信することができれば、遠隔からパーソナライズした情報を周囲に影響を与えず伝えることができる。

このようなシステムを作るためには、スピーカーを設置した環境下での音場測定が必要となる。音場測定とは、音源と様々な相対位置で音圧(sound pressure)を測定することである。ここで、音圧とは媒質（大気）で満たされた空気中のある1点において、ある短期間に生じる波（音）による気圧変化の大きさである。音場測定により、音源と様々な相対位置で音圧の強弱がわかるようになる。

従来の音場測定法としては、サウンドメータ（騒音計）を使ったものが一般的である。騒音計は、測定機であるマイクを置いた地点における様々な音の大きさを全て考慮して音圧を測ることができる。しかし、マイクを置いた位置の情報を別途与えるか、測定する必要がある上、特定の音源の影響のみを測るには適していない。他に、レーザードプラ振動計（LDV）を用いて音源から発した超音波の計測と可視化を行う研究[2]もある。LDVを用いた音場測定は、精度は高いが、大がかりな機材が必要で、屋外等様々な場所での測定は行いにくい。

そこで本研究では、携帯端末を用いてより簡便に音場測定することを目指す。具体的には、音源付近に目印となるチェッカーボードを置き、音源から一定周波数の音を流した状態で、携帯端末で動画を撮影しながら移動する。これにより、空間中の様々な位置での音の情報が測定できる。得られた動画の画像（フレーム）に対して、コンピュータ

ビジョンの自己位置推定技術を利用して測定位置の推定を行う。同時に、その位置で測定した音に対して周波数解析（フーリエ変換）を適用することで、音源に由来した音の音圧を数値化する。これら2つのデータを結びつけて可視化することで、音場測定を行う。

以下、2章では音の基本原理について述べる。3章では、画像列からの自己位置推定について述べる。4章では、携帯端末により音場測定を行う提案システムの処理手順について述べる。5章では、実験の結果と評価について述べる。最後に、6章では、提案したシステムの実験結果に基づいて考察、将来の改善について述べる。

2. 音の基本原理

2.1 音波[3]

音は、音源から出た振動が媒質を介して伝わって耳に届いた刺激である。物理的には空気中での波動であることから、物理面を取り扱う場合には、音波と呼ぶことが多い。

音波とは、空気その他の媒質（medium）中に存在する波動である。このとき、ある瞬間に媒質中には、疎密の縞ができ、この疎密に従って圧力変化が生じる。空気が密なところでは圧力が大気圧よりも高くなり、疎のところでは低くなる。この媒質粒子の振動している状態が、時間とともに一定速度で媒質の中を移動するのが音波の伝搬である。音波の場合は、媒質粒子の振動方向と、振動の伝搬方向が一致している縦波（longitudinal wave）であることが特徴である。

音波の物理的性質を表わす基本的な量は、その強さ（intensity）と周波数（frequency）である。媒質中の1点でのある瞬間の気圧（圧力）が、音のないときの大気圧より高くなった変化分を、音圧（sound pressure）という。一般に継続した音波の音圧は、時間的にも周期的に変化することが多い。音圧の瞬時値が時間の正弦関数として表わされる場合を基本として考える。このとき音圧の変化の1周期 T の間の2乗平均平方根の値を実効音圧という。音場は、ある

¹ 宮崎大学大学院工学研究科
The Graduate School of Engineering, University of Miyazaki

空間内での音圧の分布のことである。

2.2 音波の伝搬

音波は音源で発生した振動が伝搬したものである。音源は、理想的には、点音源 (point source)、線音源 (linear source)、および面音源 (plane source) に分類することができる。点音源は、大きさが 0 であるような音源、線音源は音源が一直線状をなしているもの、面音源は面積が十分広い面全体が振動して音源になっているものである[3]。

自由空間に置かれた点音源から出た音の強さは、音源からの距離 r の 2 乗に反比例する (逆 2 乗則)。

音源が点であれば音波は球面波となってどの方向にも同じ強さで伝搬する (無指向性)。一方、無限に広い板状音源のときは平面波が出ることになり、一方向にのみ音波が伝搬する (完全な単一指向性)。実際の音源はこの中間の性質を持ち、方向によって音波の伝搬する強さが異なる。これを指向性と呼ぶ。音源に指向性を持たせることで、目的の方向にのみ音を伝えることができる[4]。

音場測定を行うことで、音源から空間の様々な方向に対してどのように音が広がっているか把握することができる。

3. 画像列からの自己位置推定[5]

本研究では、音圧の測定位置を求めるために、コンピュータビジョンの自己位置推定技術の一種であるカメラキャリブレーションの外部パラメータ推定を用いる。本章では、その原理について述べる。

カメラキャリブレーションでは、3次元座標が即知の複数の点を複数のカメラで撮影し、得られた画像上での各点の2次元座標を利用する。ワールド座標系において既知の点の3次元座標を $\mathbf{s}_i = (X_i, Y_i, Z_i)$ 、 j 番目のカメラで撮影した画像上でのその点の2次元座標を $\mathbf{u}_{ji} = (u_{ji}, v_{ji})$ とすると、以下の関係式が成り立つ。

$$\mathbf{u}_{ji} = P[R_j \mathbf{s}_i + \mathbf{t}_j] = P \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}$$

ここで、 R_j, \mathbf{t}_j は j 番目のカメラの位置・姿勢を表す回転行列と平行移動ベクトルである。 \mathbf{u}_{ji} と \mathbf{s}_i の組が多数与えられれば、最小二乗法などの手法により、 R_j, \mathbf{t}_j を求めることができる。これにより、カメラの位置・姿勢の情報が得られる。

4. 携帯端末による音場測定

本研究では、携帯端末で動画を撮影しながら移動することで音場測定を行う。具体的には、以下の手順で測定する。

1. 音源の設置
2. 位置指標の設置
3. 動画の撮影
4. 動画の分割
5. 音声データに対する周波数解析による音圧測定
6. 動画のフレーム (画像) からの自己位置推定
7. 音場測定結果の可視化

以下、それぞれの処理について述べる。

4.1 音源の設定

音源としてスピーカーを置く。音源からの音圧を他の雑音と区別するために、音源から 1000Hz の正弦波 (純音) を発生させる。本研究では、ノートパソコンにスピーカーを接続し、純音を発生させる機能を提供している Web サービス[6]を利用して 1000Hz の純音を発生させる。

4.2 位置指標の設置

カメラキャリブレーションにおいて、既知の点を得るためにチェッカーボードを用いる (図 1)。チェッカーボードは、白黒の市松模様の平面である。市松模様の一辺の長さは 2cm で既知である。チェッカーボードの左上隅をワールド座標系の原点とする。市松模様の角の各点を既知の点とし、3次元座標を割り当てる。本研究では、 7×10 の市松模様のチェッカーボードを音源近くに置いたノートパソコンのディスプレイ上に表示して位置指標とする。

4.3 動画の撮影

携帯端末を用いて動画を撮影する。携帯端末のデフォルトの「カメラ」アプリケーションを用いて撮影する。撮影時には携帯端末のカメラを位置指標であるチェッカーボードの方に向けて、動画内にチェッカーボード全体が写るように撮影する。撮影位置において、およそ 5 秒間静止し、その後次の測定位置に移動し、再び 5 秒間静止して撮影することを繰り返す。

携帯端末は Google Pixel 6a を使用した。動画の各フレーム (画像) の大きさは 1920×1080 画素である。また、音声のサンプリングレートは、48000Hz である。取得した動画は、MP4 形式で保存される。

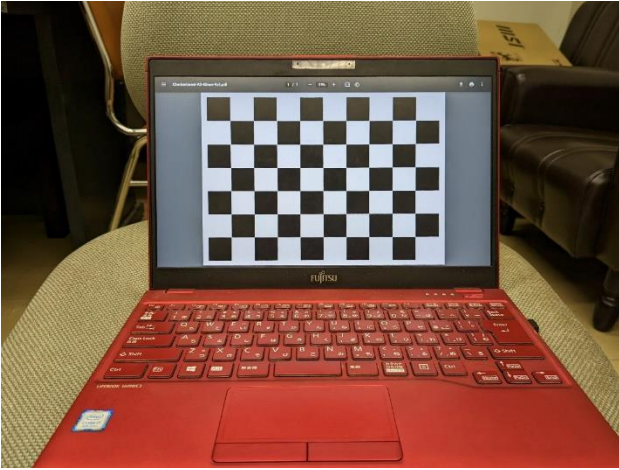


図 1 パソコンに表示されたチェッカーボード

4.4 動画の分割

動画撮影時に5秒間静止した位置での音圧測定を行うために、撮影した動画を5秒間ずつの動画に分割する。動画は、FFmpegのsegment_timeオプションで分割し、5秒間ずつの動画が出力される。

4.5 音声データに対する周波数解析による音圧測定

4.4節で分割した動画から音声データを抽出する。携帯端末では、マイクが2つあるので、音声データがステレオになっている。これをモノラルにするためにFFmpegのac1オプションを使う。さらに、位置の移動時に発生する音を考慮し、分割した5秒間の動画の内、最後の3秒を切り取って音声データとして残す。FFmpegのsscifオプションに3秒間を指定することで、最後の3秒間だけの音声データを出力する。これらの処理を行うことにより3秒間のモノラル音声データが得られる。

得られた音声データに対して周波数解析（フーリエ変換）を適用することで、音源に由来した音の音圧を数値化する。フーリエ変換は、音声のような時間的に変動する信号を、種々の周波数を持つ正弦波の和に分解する手法である。サンプリングレート $1/N$ で観測された N 個の標本点 $\{x = 0, 1, \dots, N-1\}$ における音声の観測値を $f(x)$ とすると、その離散フーリエ変換 $F(t)$ は以下で定義される。

$$F(t) = \sum_{x=0}^{N-1} f(x) \exp\left(-i \frac{2\pi tx}{N}\right) \quad (4.1)$$

ここで、 i は虚数単位 $i^2 = -1$ である。 $F(t)$ を使って元の観測値 $f(x)$ は次式で表現される。

$$f(x) = \frac{1}{N} \sum_{t=0}^{N-1} F(t) \exp\left(i \frac{2\pi tx}{N}\right) \quad (4.2)$$

すなわち、 $F(t)$ は元の観測値 $f(x)$ を種々の周波数を持つ正弦波の和として表した時の、周波数 t の正弦波の強さを示している。

本実験では、音源から1000Hzの音を発している。1000Hzの音の測定位置における強さを求めるために、4.4節で得られた音声データから観測値を1秒分取り出し、高速フーリエ変換する。高速フーリエ変換は、離散フーリエ変換の高速な計算法である。高速フーリエ変換後の1000Hzに対応する正弦波の成分の大きさ $|F(1000)|$ を、測定位置での音圧とする。

4.6 動画のフレーム（画像）からの自己位置推定

音圧を測定した位置を求めるために、4.4節で分割した5秒の動画の4秒目のフレーム（画像）を取得する。この画像内に写っているチェッカーボードに対して、市松模様の角の点を抽出する。点の抽出には、OpenCVのfindChessboardCorners関数を用いる。得られた点をワールド座標系における既知の3次元座標と対応づける。分割した5秒間の動画それぞれに対してこの処理を行うことで、それぞれにおいて既知の3次元座標が画像内に撮影された2次元座標が得られる。これらに対して3章で述べた画像列からの自己位置推定処理を適用することで、それぞれの画像を撮影した際のカメラの位置・姿勢情報 R_i, t_i が得られる。ワールド座標系におけるカメラの位置は、各カメラのカメラ座標における原点の位置なので、下式によりカメラ位置を求める。これにより、音圧測定時の携帯端末の位置が得られる。

$$\begin{aligned} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} &= -R_i^{-1} t_i \\ &= -R_i^T t_i \\ &= -\begin{pmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} \end{aligned} \quad (4.3)$$

4.7 音場測定結果の可視化

4.5節で得られた音圧と4.6節で得られたその時のカメラ位置を対応づけることで、測定した空間のどの位置でどの程度の強さの音圧があるかわかる。これらを組み合わせてプロットすることで、音場測定の結果を可視化することができる。

5. 実験

5.1 実験環境

実験は、宮崎大学工学部の研究室で行う。防音や無響設備等は備えられていない。実験環境を図2に示す。4章で

述べたように、ノートパソコンに音源になるモノラルスピーカーを繋げて 1000Hz の音を発信する。ノートパソコンの画面にはチェッカーボードの画像を表示する。ノートパソコンから 1メートル離れた位置には、ノートパソコンに対してほぼ垂直に正対するようにパネルを設置する。パネル上の5カ所に印をつける。各場所を P1~P5 とする。P1 から P5 はパネル上でほぼ水平に、ノートパソコンに向かって右から左に約 16cm の等間隔に並んでいる。このパネル上の点 P1~P5 に携帯端末を移動させながら動画を撮影する。撮影した動画は、常にノートパソコン上のチェッカーボード全体が写っている。撮影した動画の一部を図 3 に示す。これらの動画を用いて、提案した音場測定法の確認を行う。

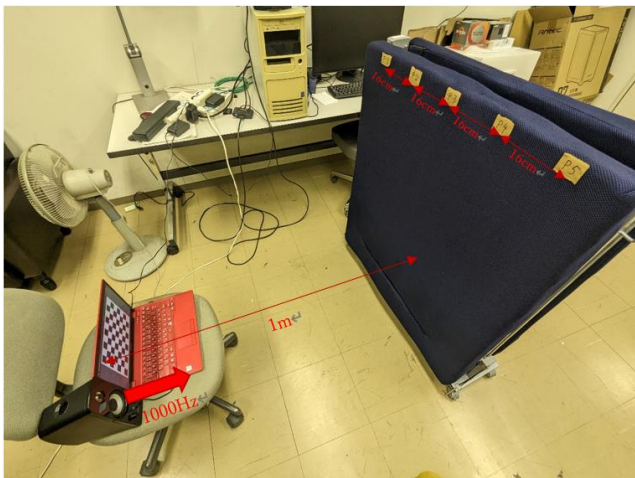


図 2 音場測定の手法

5.2 音圧の再現性の確認

本実験では、音圧測定の再現性を確認する。まず、P1~P5 の順に3回移動しながら動画を撮影する。次に、それぞれの動画から、P1~P5 での音圧を求める。得られた結果を図 4 に示す。

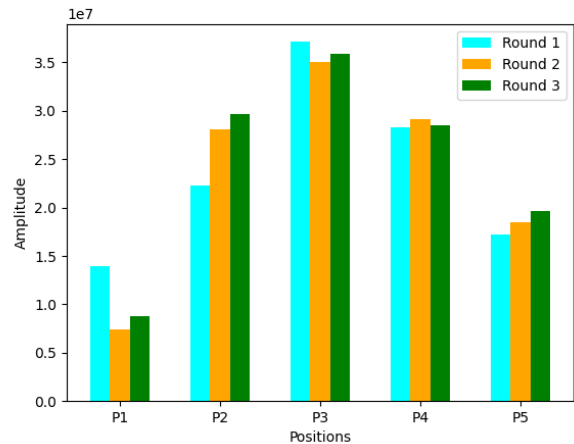


図 4 音圧の再現性の結果

図 4 より、多少の変動はあるが、全体として各点において 3 回の測定とも同じ程度の値が得られている。特に、P3、P4 はほぼ同じ値が再現できている。一方、P1 は誤差が大きい。P1 はスピーカーから最も離れており、測定された音圧が低い。その際には、少しの反射の影響も大きく受ける。撮影の際には、撮影者が移動するため、その反射波の影響を受けた可能性がある。

5.3 位置推定の再現性確認

本実験では、自己位置推定の再現性を確認する。5.2 節の実験で撮影した 3 回分の動画を利用する。それぞれの動画から点 P1~P5 での自己位置を推定した結果を図 5 に示す。図 5 の図中の赤枠がチェッカーボード、四角錐がカメラでその頂点が焦点位置になっている。チェッカーボードの左上隅が原点、ノートパソコンに向かって右向きが x 軸、鉛直下向きが y 軸、ノートパソコンより奥向きが z 軸である。



図 3 P1~P5 から撮影した動画

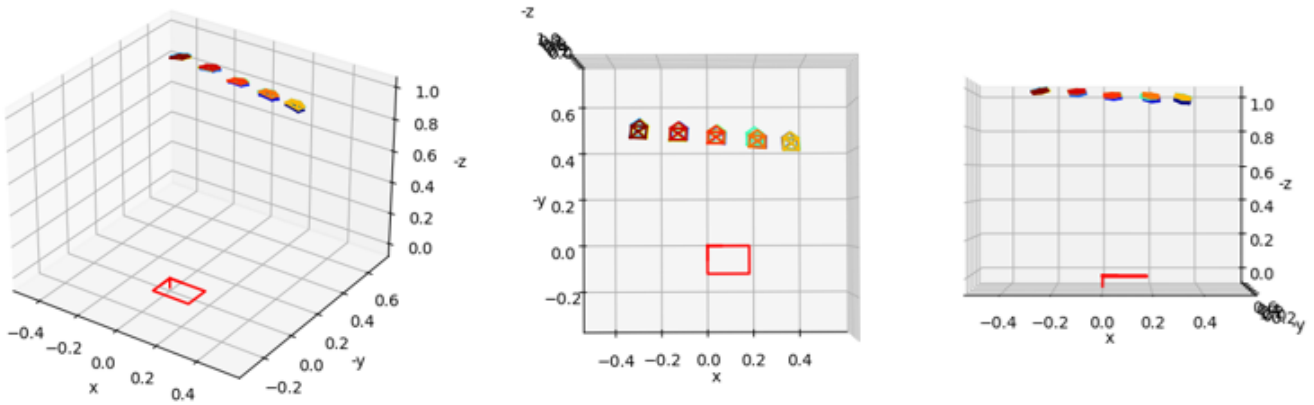


図 5 位置測定の再現性の結果

各軸の単位はメートルである。P1~P5 の 5 点での撮影を 3 回行っているが、推定された自己位置は 3 回でほぼ同じ位置となっている。P1~P5 はチェッカーボードを見下ろす位置関係になっているが、推定されたカメラ位置の y 座標は 0.5 メートル程度で、目測との矛盾はない。X 軸上で P1~P5 が右から左にほぼ等間隔に並んでいることも、配置に矛盾はない。また、カメラ位置の z 座標はほぼ -1メートルとなっている。これらの結果から、自己位置推定の再現性は高いと言える。

5.4 音圧測定と自己位置推定による音場の可視化

音圧測定と自己位置推定を同時に行うことにより音場測定が可能になる。5.2 節で得られた音圧と 5.3 節で得られたその時のカメラ位置を対応づけて、以下の図 6 に結果を示す。図 6 の x, y 軸は、5.3 節の図 5 と同じカメラ位置の x, y 軸となっている。点 P1~P5 は音源に正対した平面上を移動しているため、図 6 では、カメラ位置の z 座標の代わりに、そのカメラ位置における音圧 (Amplitude) を軸に取り棒グラフで表している。音源位置は、ノートパソコンに向かって左側 (x 軸上の正方向) に置かれている。音場の可視化結果から音源に最も近い位置 (P3) で音圧が最も高くなっており、そこから離れるにつれて音圧が低下していること

がわかる。音源を点音源とみなすと、音圧は音源からの距離が離れるほど低下する。音場の可視化結果から、その様子が読み取れる。

6. おわりに

本研究では、携帯端末を用いた簡便な音場測定システムを構築した。音源からの音圧を周囲の雑音と区別できるように、1000Hz の純音をノートパソコンに接続したスピーカーから発信した。また、ノートパソコンのディスプレイ上に位置測定のための目印となる指標 (チェッカーボード) を表示した。この環境下で、移動しながら携帯端末で動画を撮影することで、音源と様々な相対位置での音源からの音を記録した。得られた動画中の音から、1000Hz の音の強さをフーリエ変換で求めることで、その位置での音圧を得た。また、動画のフレーム (画像) から、コンピュータビジョンの自己位置推定技術を使って、音圧の測定位置を求めた。得られた音圧とその測定位置を組み合わせることで、音場の可視化を行った。

実験では、音圧の再現性の確認、位置推定の再現性の確認、両者を組み合わせた音場の可視化を行った。これらの実験により、携帯端末を用いて簡便に音場測定を可視化ができることを示した。

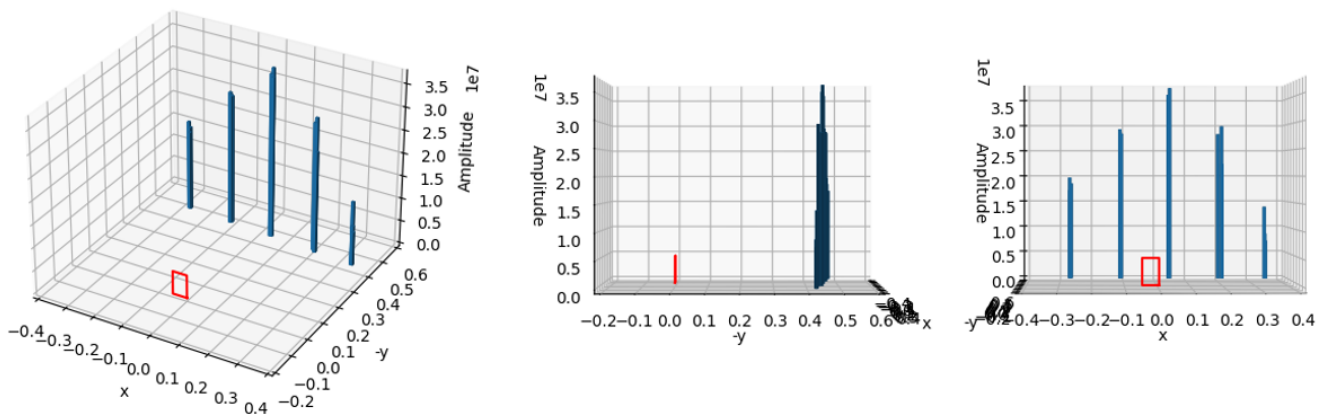


図 6 音場の可視化の結果

今後の課題としては、提案システムでの音場測定の精度評価が挙げられる。実験では、音圧の測定値の正解値が不明であった。実験の場所は無響室ではなかったため、あらゆる反射音が発生して音波が打ち消しあう現象が起こる。これによって耳で直に音を聞いていても、音圧が急に減少したり、増加したりすることがあった。よって、実験場所を無響室に変えることによって、音場測定の条件を整え、音量測定装置等による精密な測定結果と提案システムの測定結果を比較し、精度評価する必要がある。

参考文献

- [1] 李孝珍, 坂本慎一, 菅原彬子, 池田佳樹, “駅コンコースにおける音環境評価のための実測調査および聴感評価実験”, 日本建築学会環境系論文集第 84 巻第 765 号, 983-991, (2019)
- [2] Yasuhiro Oikawa, Makoto Goto, Yusuke Ikeda, Toshikazu Takizawa and Yoshio Yamasaki, “Sound Field Measurements Based On Reconstruction From Laser Projections”, ICASSP, 661-664, (2008)
- [3] 古井貞熙, “音響・音声工学”, 電子・情報工学入門シリーズ 2, 近代科学社, 1-11, (1992)
- [4] 牛来辰巳, “騒音と振動の計測”, 株式会社コロナ社, 17-18, (1986)
- [5] 石田剛史, “画像列からの 3 次元形状復元における非線形最小 2 乗法の高速化と安定化に関する研究”, 広島市立大学大学院情報科学研究科博士論文, 7-9, (2004)
- [6] Online Tone Generator: <https://onlinetonegenerator.com/>