

テーマパークでの没入感向上を目指した自然音による BGM のサウンドゾーン制御

Sound zone control of background music with natural sounds to improve immersion in theme parks

長尾真侑 (Mayu Nagao)*

法政大学 情報科学部 デジタルメディア学科
mayu.nagao.4e@stu.hosei.ac.jp

Abstract

In theme parks, background music is played in each area according to a theme in order to provide an immersive experience. In order not to mix BGM from adjacent areas, sound zone control using various natural sounds in accordance with the theme of each area becomes effective. For this purpose, a model was created such that when natural sounds and BGM are input, they are modified to be suitable for sound control, and an experiment was conducted to evaluate the model. Experiments were conducted outdoors using 9 types of natural sounds and 7 types of BGM, and the sound control effectiveness of natural sounds was evaluated based on the obtained signal-to-noise ratio threshold values. The results of the experiment showed that the natural sounds that were effective were bird and frog calls, with signal-to-noise ratio thresholds of 14 dB and 7.2 dB, respectively. The most sound control effect was obtained when natural sounds were normalized by the largest component of the logarithmic amplitude spectrum as a feature.

1 はじめに

テーマパークが非日常感を与える上で重要視している要素の一つとして、没入感がある。アトラクションや建物などによって没入感が与えられるが、本研究は音による没入感、特にBGMに着目する。テーマパークで流れているBGMは、エリアごとにその雰囲気に合わせた音楽が採用されており、視覚から得られる情報とBGMを一体化させることで没入感を高めることに繋がる。しかし、エリアごとのBGMが混ざると視覚から得られる情報と聴覚から得られる情報に齟齬が生じてしまう。それを防ぐために様々な工夫がされている。例えば、東京ディズニーランドでは指向性スピーカーを使用することでの聞こえる範囲を限定する方法が行われているが、聞こえる範囲が狭いため歩いている途中で聞こえる範囲を出ると突然音が聞こえなくなるといった問題点がある。また、異なるBGMが流

れる2つのエリアの間に本物の滝を流すことで、2つのBGMを混ぜない工夫をしている。図1の滝に向かって歩いていると、滝に近づくにつれて、聞こえていたBGMが滝の音に変化する。また更に歩き続けると、また新たな別のエリアのBGMに変化する。



図1 ディズニーランドにある滝

その結果、エリアにいる際のBGMによる没入感がより高まっている。指向性スピーカーに比べて滝を用いた方法がテーマパークでの没入感向上の観点ではより適している。そのため、様々なエリアでBGMを違和感なく混ぜないために、エリアごとのテーマに沿った様々な自然音を用いたサウンドゾーン制御が有効になる。自然音は田村 [1] の研究で、1240人のアンケート結果から「好ましい」と感じる音だと評価されている。このことから自然音がテーマパークにおいて没入感を損なわずにBGMのサウンド制御をする点において適している。自然音の音量を大きくすればBGMは聞こえなくなるが、テーマパーク外での騒音問題に繋がってしまうため、適切な音量でのサウンド制御が求められる。

そこで、本研究ではBGMと自然音を入力すると、サウンド制御に適するように自然音を修正するようなモデルを作成する。修正とは、音の間隔を詰めたり、周波数を変えるなどである。サウンドゾーン制御の実験で得られた結果から自然音の特

* 指導教員：伊藤克亘 教授

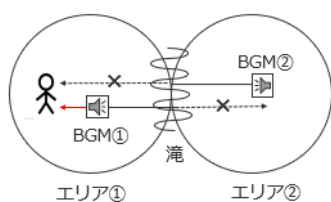


図2 ディズニーランドでのサウンドゾーン制御

微分析を行い、作成する。モデルを作成することで今後のテーマパークづくりに活かせる。例えば、新エリアや新しいテーマパークを造る際に、BGMのサウンドゾーン制御に適した滝のような自然音をエリア間に設置することで没入感を高めることが可能になる。本研究ではテーマパークを想定し、環境を合わせて実験することでこれらに応用できるようにする。

2 従来研究

サウンドゾーン制御に関する先行研究として、マスキング、サウンドスケープに関する論文がある。

2.1 マスキング

マスキングとは、音の知覚を他の音で妨害することである。マスキングに関する研究では、多様な騒音に対する自然音でのマスキングにおいて、どの自然音が適しているかを実験によって検証されている。Karlら[2]は、風力発電のWT音による睡眠障害などの騒音問題軽減のために、自然音(海の波、針葉樹、落葉樹)によるマスキングを行った。3種類の自然音のマスキング効果を信号対雑音比(SN比)の閾値検出実験から求めた閾値によって比較し、海の波の音が最も高いマスキング効果が見られるという結果を出した。SN比はSをマスキングした音、つまり本研究のBGMやWT音、Nを自然音としたときのラウドネスのパワー比率である。また、自然音でのマスキングに関する既存のモデルを使用し、そのモデル予測値と経験的結果の比較を行った。

2.2 サウンドスケープ

他に、環境音など、その場で流れている全ての音の一つの音楽としてとらえる、サウンドスケープという考え方もある。Kaiserら[3]はアメリカのテーマパークでのサウンドスケープについて研究していた。テーマパーク全体の騒音レベルの測定をし、音響設計にサウンドスケープ分析を用いることを議論していた。また、サウンドスケープ設計について手順が示されている。サウンドスケープを考慮することでテーマパークの没入感向上に繋がる。

2.3 BGMと没入感

BGMが没入感に与える影響についての論文もある。[4]では、映画などの映像作品の映像と音楽の雰囲気的一致が視聴者の没入体験に与える影響について検討した。映像と音楽の雰囲気が一致している場合、視聴者の没入感が高まることが示されていた。

2.4 発展

自身の卒業研究と従来研究を比較して、従来研究では「同じ種類の2つの音を消すことはできるのか」、「片方の音楽を消して、もう片方は残す」、「滝以外の他の自然音でBGMを消す

ことはできるのか」、「屋内ではない屋外でのサウンドゾーン制御」が不足している。どの先行研究においても、音楽と音楽のような同じ種類の音から1つの音を消す方法ではなく、音声と騒音のような2つの違う種類の音から特定の音を消す方法について研究されている。本研究は、音楽と音楽を混ぜないためのマスキングであるため、同じ種類の音から特定の音を消す研究を行う必要がある。また、テーマパークでは自身がいない反対側のエリアのBGMは消さなければいけないが、自身がいる側のエリアは聞こえていなければいけない。そのため、2つの音楽のうち片方だけを消す必要がある。他にも、テーマパークの環境に合わせた様々な自然音でのサウンド制御や屋外でのサウンドゾーン制御が求められる。テーマパークで過ごすうえで必要な音、不必要な音を取捨選択しながら、サウンドスケープを組み立てることが必要となる。

3 提案手法

自然音を入力とし、BGMをマスクするように自然音を変換するモデルを作成する。マスキングは、2つの音の対数振幅スペクトルの振幅を比較した際に小さい方がその周波数帯ではマスクされるとされている。その条件を満たすために、まず自然音とBGMをそれぞれ比較する。人間の聴覚特性を考慮して、周波数軸を対数にしたBGMの対数振幅スペクトル(図3)で比較する。ラウドネスを表すdBが対数で表されるので、より人の音の聞こえ方に似た状態でスペクトルを比較できる。BGMの方が対数振幅スペクトルが大きい場合は、自然音とBGMの対数振幅スペクトルをグラフに起こし、2つの対数振幅スペクトルで囲まれた部分の面積を求める。BGMが自然音の対数振幅スペクトルを下回る場合は、統計量を用いて、自然音を正規化する。特徴量としては、標準偏差、最大成分、平均値を用いる。自然音とBGMの対数振幅スペクトルの標準偏差、最大成分、平均値をとる。自然音とBGMで特徴量が一致するよう自然音を調整する。

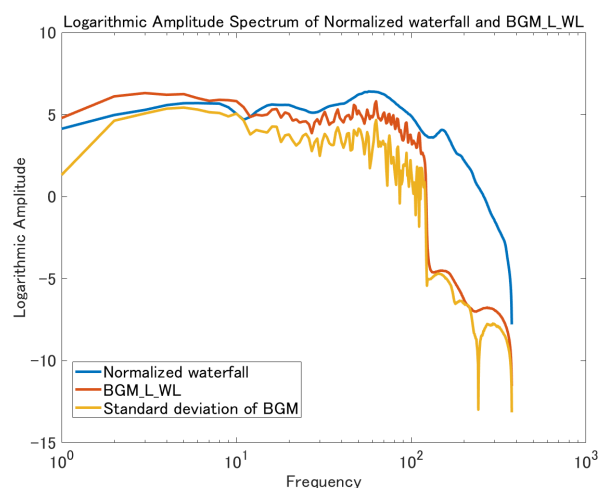


図3 周波数軸が対数の対数振幅スペクトル

4 評価・分析方法

本研究は SN 比による客観評価とアンケートによる主観評価を行う。BGM が自然音によって消されるかどうかは人の感覚によって異なるため、その主観評価を実験で行う。さらに、[2] で用いられている SN 比を計測する客観評価を行う。主観と客観の両面からモデルを評価する。

4.1 SN 比による客観評価

実験から求めた BGM と自然音の SN 比から、自然音のサウンドゾーン制御効果を評価する。サウンド制御効果があるのは SN 比が -10dB 以下の時 [2] だとされているので、その値を基準とする。同時に録音する側の BGM でも同様に実験し、聞こえるかを評価する。ここでは、テーマパークで計測した SN 比を基準とする。

東京ディズニーランド内の実際にサウンドゾーン制御が行われている場所で、BGM と滝の録音、騒音の大きさ (ラウドネス) の計測を行った。録音はバイノーラルマイク RM-ATZ19R を用いた。計測した滝の騒音の大きさと録音位置までの距離から、滝と BGM の SN 比も求めた。結果、滝のラウドネスは約 75dB 、BGM のラウドネスは約 65dB 、滝と BGM の SN 比は 18dB であった。この値を評価基準とする。ただし、ディズニーランドの該当箇所のスピーカーは、建物の 2 階部分にスピーカーが設置されているため、スピーカーから出ている正確な騒音の大きさは測ることができなかった。該当箇所ではないスピーカーから出ている音のラウドネスは約 80dB であった。

4.2 アンケートによる主観評価

リスニング実験では、反対側の BGM が気にならないかのアンケートを取る。本研究は先行研究とは違い、リスニングする側のスピーカーから流れている BGM が、自然音で消されてはいけない。そのため、自身がいるエリアの BGM を、自然音が邪魔していないかを被験者に主観評価してもらう。5 段階でアンケートを取り平均を取る、MOS 評価を取り入れる。

2 つの評価結果を合わせてテーマパークでの BGM のサウンド制御として適切な自然音を生成するモデルを評価する。

5 実験

サウンド制御に適した自然音のモデルを評価するため、スピーカーを用いて BGM と自然音を流す実験を行う。4 の 2 つの評価を行うための実験を行う。

[5] のサウンドゾーン実験では、64 個のスピーカーに対して 2688 点でサウンドゾーンを評価していた。本実験でも事前に何点かでリスニングし、効果的でありそうな点を選択してリスニング実験を行う。

本研究はテーマパークでの実装を想定しているため、屋内ではなく屋外で実験を行う必要がある。屋内で行うと壁や天井からの反射でラウドネスが変化し、テーマパークでの環境とは異なってしまうためである。実験時の風速は 0.1 m/s だった。

5.1 実験方法

スピーカー 3 つを、図 4 のように 5m ずつ離して設置する。端の 2 つから別エリアの BGM を流し、中央の 1 つから自然音を流す。この際、全てのスピーカーの向きを同じにする。リ

スニングは、BGM を流しているスピーカー片方の近くから行う。BGM を流すスピーカーは BEHRINGER EUROPORT PPA2000BT、自然音を流すスピーカーは YAMAHA Stagepas 400i を使用した。DAW として Cakewalk by BandLab、オーディオインターフェースは Roland Rubix24 を使用した。

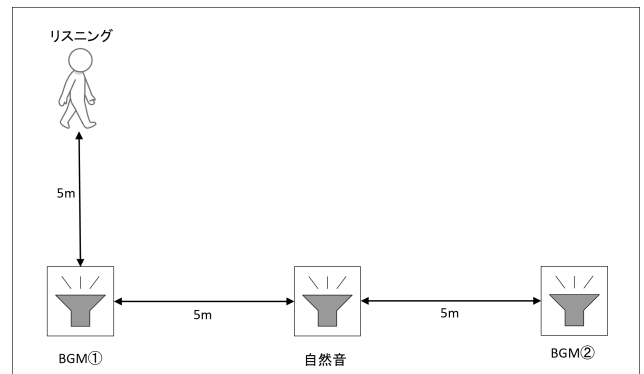


図 4 本実験模式図

男性 3 名、女性 2 名の計 5 名の被験者に 4.2 節の主観評価を行ってもらった。スピーカーから 5.2.4 で決定した 11 通りの音を流す。ディズニーランドで計測したラウドネスを基に、BGM を 70dB で流し、自然音は各組み合わせの正規化後の大きさに流す。予備実験から決めたリスニング場所から、4.1 節の SN 比による客観評価も行う。複数点計測した結果から、リスニングの場所をスピーカーから 5m 離れた場所とした。

5.2 音源

様々な音に対応できるよう、多くの種類の音を使用する。BGM として、東京ディズニーリゾートで実際に BGM として使用されている音楽を使用する。自然音は、Karl ら [2] の研究で使用されていた波の音を使用する。また、葉の音より波の音の方がサウンド制御効果があったこと、テーマパークに川などの水が多いことから、他の水に関連する波と川の音、テーマパークで流れていても不自然ではない虫や鳥の鳴き声での実験も行う。ディズニーランドでサウンド制御効果のために使用される、滝の音でも実験を行う。滝と似たスペクトルを持つピンクノイズも使用する。

5.2.1 自然音の録音

自然音は自ら録音した音を使用する。Karl らの録音方法を参考に、バイノーラルマイク 3DIO FS ProII、ショットガンマイク ZOOM SGH-6 を用いて 1 分間録音する。本研究は、滝、蝉、波、夜の虫、川の音を録音した。滝と波、川の音は高さ 1.4m 、 0.3m の 2 段階で録音し、蝉、夜の虫の音は高さ 1.4m で録音した。滝と蝉のラウドネスはともに約 80dB で、波のラウドネスは約 70dB であった。録音した音をスピーカーで流し、どのマイクで録音したものが実験に適した音を選定する。滝、波、川に関しては録音の高さでも比較し適した音を選ぶ。

5.2.2 録音していない自然音

動物の鳴き声は録音環境を管理することが困難であるため、配布されている音を使用する。[6] では、アメリカ全国の国立公園で録音された自然音が掲載されている。[6] から鳥の鳴き声と、予備実験でサウンド制御効果の見られた蛙の鳴き声をダ

ウンロードした。

5.2.3 BGM

実験で使用する BGM は、東京ディズニーリゾートで実際に使用されている 7 曲である。これらは、録音した自然音が違和感なく使用できるようなエリアの BGM を選択した。録音の有無に関わらず、全ての自然音、BGM を 30 秒、サンプリング周波数 48000Hz にし、ステレオであったものは R と L をそれぞれ 1/2 にして足すことでモノラルとした。

5.2.4 実験で使用する音の組み合わせの決定方法

対数振幅スペクトルから音量を変更する際に、周波数軸を対数になるようにプロットした。その BGM 7 曲と自然音 9 種類のスペクトルの差を積分する。3 節で変更する前後のグラフ面積の比較を行う。ディズニーリゾートで使用されている自然音と BGM の組み合わせの対数振幅スペクトルのグラフ面積を求める。その面積と誤差が 1 以内となるその他の組み合わせから BGM、自然音、対数振幅スペクトルの調整方法を全て選べるような自然音と BGM の組み合わせ、全 10 通りを決めた。

5.3 結果

本実験を行った結果、10 通りの自然音と BGM の組み合わせのサウンド制御効果の有無、SN 比はそれぞれ表 1 となった。MOS1 は反対側の BGM が聞こえないか、MOS2 は手前の BGM が自然音で消されていないかを MOS で評価した値である。SN 比 1 は BGM1（リスニング側）と自然音の SN 比、SN 比 2 は自然音と BGM2（反対側）の SN 比である。

また、自然音と BGM の対数振幅スペクトルの平均と標準偏差をグラフにしたものが図 5 と図 6 である。

本研究では、BGM1 側の SN 比の基準を 18dB、BGM2 側の基準を -10dB としていた。SN 比で条件を満たす組み合わせはなかった。主観評価では、鳥、蛙の音にサウンド制御効果があった。特徴量として最大成分で自然音を正規化する方法が、最も効果が得られた。MOS の結果と SN 比を散布図にしたものが図 7 と図 8 である。これらの図からわかる通り、MOS と SN 比の結果は比例していなかった。

5.4 考察

5.4.1 SN 比の客観評価

対数振幅スペクトルでマスキング効果が得られるように調整しても、SN 比評価が良くなかった理由として、距離と音の減衰の関係がある。音は距離が離れるほど聞こえにくくなる。周波数が高い音ほど早く減衰する。図 5、図 6 より、自然音のほうが高周波数成分を多く持つ音が多いため、スピーカーから離れると自然音のラウドネスが小さくなる。対数振幅スペクトルを BGM と自然音で比較したとき、自然音の方が高周波数成分を多く持つ組み合わせが、SN 比が小さくなる傾向にあった。

5.4.2 MOS による主観評価

人間が最もよく聞こえる周波数は 3000~4000Hz である。この周波数帯の自然音と BGM の対数振幅スペクトルの差を特徴量ごとで比較した。しかし、効果のあった最大成分と残りの 2 つで大きな差は見られなかった。最大成分で正規化することが最も効果があるという、本実験の研究結果を確固たるものにするには、今後 3000~4000Hz での帯域での正規化の分析が必

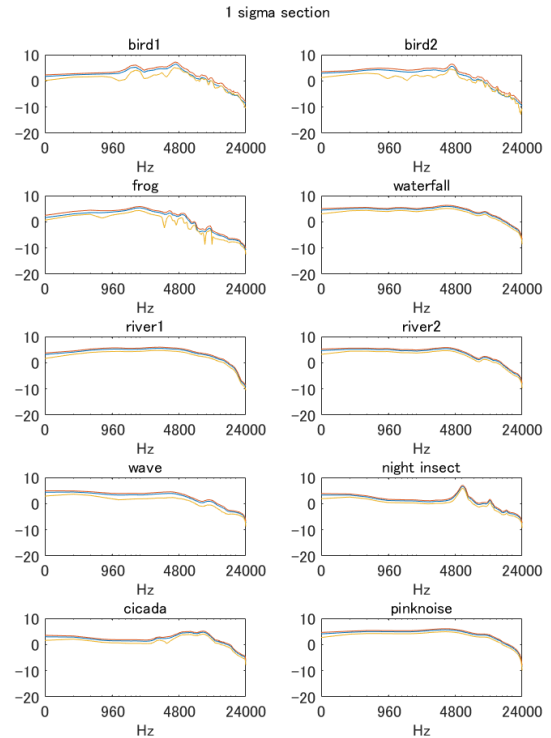


図 5 本実験で用いた自然音のスペクトル平均、標準偏差

要である。

MOS と BGM の短時間エネルギーの分散の相関係数は -0.8271 であった。短時間エネルギーの分散が大きいということは、BGM が大きく聞こえる瞬間が多くあるということだ。その BGM が聞こえた一瞬のインパクトによって、評価が下がってしまう組み合わせがあった。

自然音に、聴覚フィルタと、人間が最もよく聞こえる周波数の 3000~4000Hz のバンドパスフィルタをかけた。その波形を見ると、鳥 (bird1、bird2)、波、蛙が音の間隔が開いていた。同時に、短時間エネルギーの分散を求めると、bird1 が最も値が大きかった。そのため、bird2、蛙、波に比べて bird1 が MOS 評価が低かった。この問題を解消するには、他の自然音と合成する必要がある。滝と bird1 を合成してリスニングした結果、BGM が聞こえにくくなっていた。

5.4.3 SN 比と MOS 値の関係

リスニング実験の結果と比較すると、SN 比の結果とは比例していない。MOS は、BGM の短時間エネルギーの分散に関係していた。主観評価では、BGM の音量が一瞬大きくなって自然音以上に聞こえることにより、サウンド制御されていないと評価されてしまった。しかし、SN 比は全体のラウドネスの平均を取っているため、一瞬のラウドネスの大きさはあまり影響を及ぼさない。これらが原因で主観評価と客観評価の結果に相関がなかった。このことから、SN 比を考慮するだけではサ

表 1 リスニング実験結果と SN 比

BGM1	BGM2	自然音	特徴量	MOS1	MOS2	SN 比 1(dB)	SN 比 2(dB)
L_AL	L_WL	waterfall	なし	2.2	3.0	-5.0	6.1
L_AL	L_CW	bird1	標準偏差	1.0	2.8	-1.5	6.5
L_CW	L_AL	bird2	平均値	3.8	4.4	-0.7	14.0
L_CW	S_PC	wave	標準偏差	3.8	2.6	-2.4	11.7
L_WL	L_CW	river2	標準偏差	2.6	2.0	-4.1	4.6
L_WL	L_NFL	river1	最大成分	4.2	3.0	-10.5	1.7
L_WL	L_NFL	cicada	なし	2.0	4.4	-0.1	12.1
L_NFL	L_CW	waterfall	最大成分	3.4	3.0	-4.2	4.7
S_PD	S_AW	night insect	標準偏差	2.2	3.2	-2.6	14.7
S_PC	S_AW	frog	最大成分	4.2	4.8	-2.4	7.2
L_AL	L_WL	pinknoise	なし	2.0	3.0	-3.6	7.5

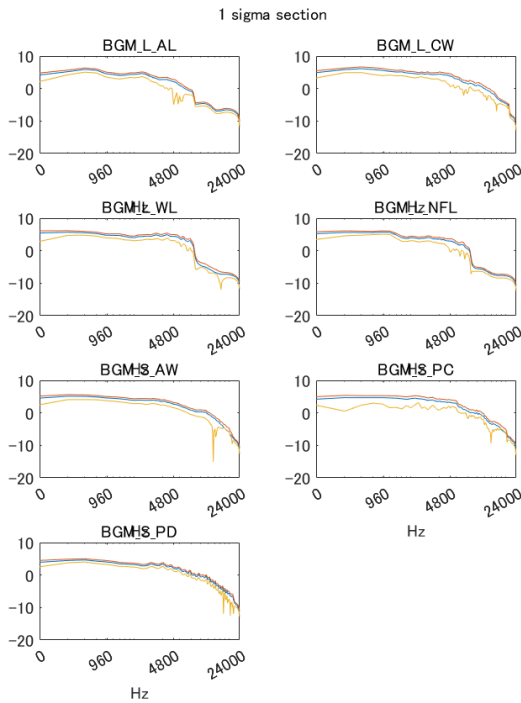
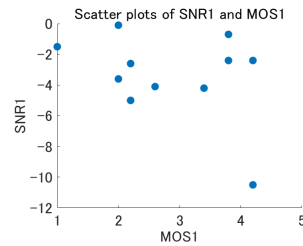
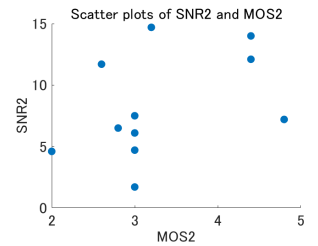


図 6 本実験で用いた BGM のスペクトル平均、標準偏差

ウンド制御効果は得られないことがわかる。

6 結論

BGM をマスクするように自然音を変換するモデルを作成した。BGM の対数振幅スペクトルの特徴量を用いて、自然音を正規化した。屋外でリスニング実験を行い、モデルを評価した。SN 比ではなく、最大成分のような特徴量を使用することでサウンド制御の性能が上がるということがわかった。作成したモデルは入力として BGM を 1 曲のみ受け入れるものであった。しかし、これは「片方の音楽を消して、もう片

図 7 SN 比 1 と MOS1 の
散布図図 8 SN 比 2 と MOS2 の
散布図

方は残す」という本来の目的には不足している。2 曲を入力させ、この条件に合うよう自然音を修正し出力するモデルを作成することが今後の課題である。

参考文献

- [1] Akihiro Tamura. Recognition of sounds in residential areas. In *Journal of Asian Architecture and Building Engineering*, Vol. 1, pp. 41–48, 2002.
- [2] Shafiquzzaman Khan Karl Bolin, Mats E. Nilsson. The potential of natural sounds to mask wind turbine noise. In *Acta Acustica united with Acustica*, Vol. 96, pp. 131–137, 2010.
- [3] Fabio Kaiser and Thorsten Rohde. Orlando theme park acoustics - a soundscape analysis. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, pp. 4312–4321.
- [4] 張馨月, 浅野倫子. 映像と音楽の感情の調和性が没入感に及ぼす影響. *認知科学*, Vol. 31, No. 4, pp. 552–572, 2024.
- [5] Takuma Okamoto and Atsushi Sakaguchi. Experimental validation of spatial fourier transform-based multiple sound zone generation with a linear loudspeaker array. *The Journal of the Acoustical Society of America*, Vol. 141 3, p. 1769, 2017.
- [6] National Park Service. Sound gallery, 2022. <https://www.nps.gov/subjects/sound/gallery.htm>.