

創作ノート

インタラクティブサウンドスケープ作品「アーバン・グリーン・シンフォニー」 The interactive sound installation "Urban Green Symphony"

顔 鑫

Xin YAN

東京工科大学デザイン研究科
Graduate School of Design
Tokyo University of Technology

松村 誠一郎

Seiichiro MATSUMURA

東京工科大学デザイン研究科
Graduate School of Design
Tokyo University of Technology

概要

インタラクティブサウンドスケープ作品「アーバン・グリーン・シンフォニー」は、鑑賞者がインタラクティブを通して都市の緑地のバランスについて理解することを促す作品である。作品には、Derivative 社 TouchDesigner と PureData (以下,Pd),Ultrahaptics 社 LeapMotion を用いている。鑑賞者は手の動きによって都市や緑地の選択、現地録音したサウンドスケープ音源の再生、エフェクトのかかり具合を操作することができる。これにより単なる音源の再生ではなく、鑑賞者がサウンドスケープを「指揮をするように奏でる」事ができる環境の創出を目指している。

The interactive soundscape work "Urban Green Symphony" encourages viewers to understand the balance of urban green spaces through active participation. The exhibition utilizes TouchDesigner, PureData, and LeapMotion. Viewers can select cities and green spaces, play back soundscapes recorded on-site, and manipulate sound effects through hand movements. Our goal is to create an environment where viewers can "conduct" the soundscape, rather than merely playing the sound sources.

1. 制作背景と目的

1.1. 都市の緑地と音環境

都市における緑地の植物は騒音を吸収し、騒音汚染の低減と住民の生活満足度の向上につながる利点がある(宮城, 2021)。緑地の植物の種類、高さ、組み合わせ方法によっても住民の心理的健康と生活の質の向上にも寄与することが示されている(Jefferson, 2022; Dzhambov, 2018)。そのメカニズムや具体的な効果を理解するために、伝統的な研究方法に加えて、直感的な体験が必要と考えられる。

1.2. 体験型インスタレーション

体験型インスタレーションは、鑑賞者が作品と直接的に対話し、音や映像などによって、感情や感知を経験することが可能であり、作品に込められた特定の概念を理解することを促進することが期待できる。これについては、作品の概念を積極的に理解し受け入れる能動性が向上し、コンテンツのエンターテインメント性も高まること示されている(吉田,2021)。また、体験型インスタレーション作品は、鑑賞者が体験を通じて以前とは異なる感情を味わったり、思い出を振り返ったりするような感動を意図していることが示唆されている(桐山, 佐藤, 2019)。

ゆえに視覚・聴覚情報を体験することを通して、都市の音環境と緑地のバランスについての問題に一般の人々が関心を持ち、直感的な理解を促すことができる場の創出は、有効な手段であると考えに至った。

本研究では、都市内の緑地で録音された音素材と撮影された画像素材を、鑑賞者自身が操り、「指揮をする動作」によって音の表現を行うことができる体験型サウンドインスタレーション作品をデザインした。

本作品の聴覚情報と表現については、展示と展示後の改良という2段階のフェーズを有している。

1. 緑地で現地録音された音源を再生する「サウンドスケープ」音の出力とそれらへのエフェクトの付加。
2. そのサウンドスケープからクラシック音楽の機械学習データを用いて変換する「音楽」出力。

1.については作品に実装し、展示を行なった(顔, 松村, 2024)。2.については、展示後のフィードバックに基づいて機能を改良した。

2. 作品の聴覚情報の構成要素と表現

2.1. サウンドスケープ

サウンドスケープは、「音」と「～の眺め/景」を組み合わせた概念で「音の風景」を意味する。「サウンドスケープ研究」においては、音を物理的存在としてとらえるだけでなく、特定の社会で生活する人々がどのような音を聞き取り、それらをいかに意味付け、価値づけているかについて取り扱う。サウンドスケープの構成音の種類として、「音楽」などの人為的な創作やデザインされたものばかりでなく、雨や風の音、動物や昆虫の鳴き声などの自然界の音から、人間の発する音、道具などの「物音」、機械の音など多種多様な音が含まれる(鳥越,1990)。

環境音の感情的な価値に関する研究では、特定の音や環境音が意識的に体験されるか潜在的に作用するかにかかわらず、関連する一連の記憶を引き起こすことが示されている(McCartney,2002)。すなわち音の想起により、人々は特定の場所や時間に再び繋がることのできる。

これらの研究から、サウンドスケープは単なる環境音や環境雑音の否定的な定義を超えて、生活の品質の一部を構成する要素としてみなされることわかる(Schulte,2006;Deng,2009;Schulte,2013)。

2.2. クラシック音楽の効用

本作品の展示後の改良として、出力する音楽にクラシック音楽の「型」を利用することを検討した。その理由は、西洋のクラシック音楽は、その楽曲構造や音色の特性により音楽療法分野においても広く活用され、感情的な安定やストレス軽減、集中力の向上、感情の表出や処理、リラクゼーション、そして身体的なリハビリテーションに効果的であることが知られているためである。例えば、そのゆったりとしたリズムやメロディは、自然な呼吸や心拍数を促進し、身体的な緊張が緩和され、心身のリラクセスが促進されて、人々のストレスを軽減する有効性について報告されている(高橋,1995; 森下2003)。

2.3. 機械学習を用いた音色変換

次に、そのクラシック音楽の型を残したまま、本作品の音の表現を作るために、機械学習モデルの導入を検討した。

AIによる機械学習は、データの種類や状況などに応じて「教師あり学習」「教師なし学習」「強化学習」という3つに分けられる。例えば、SVM(Support Vector Machine)は、「教師あり学習(教師データを学習させてから、認識・予測を行う手法)」である。

RAVE(Realtime Audio Variational Autoencoder)は,Cailonらによって提案された機械学習モデルであり、リアルタイムで音色変換を生成することが可能である。同様の機械学習モデルには、So-vits-svc(SoftVC VITS Singing Voice Conversion),DDSP(Differentiable Digital Signal Processing),BERT(Bidirectional Encoder Representations from Transformers)などがある。

上記の機械学習モデルを活用した音楽作品も多く制作されている。例えば、事前に水の音を学習させたモデルに、バイオリンの音を入力変換している事例(公文,小坂2022)、SVM(Support Vector Machine)を利用して、リアルタイムの脳波データによって楽曲生成プログラムを作成する事例(井上2022)等、機械学習モデルによる様々な音響合成が試みられている。

上記を鑑みて、本研究では、クラシック音楽の要素と都市の音とサウンドスケープを融合した音楽表現のために、音楽の構造と特徴を分析した機械学習モデルを使用した。作品展示後の改良として、クラシック音楽の構造を維持しながらも、次節で述べるRAVEを用いて都市の音と緑地のサウンドスケープを音色に変換して用いる楽曲の生成システムを試みている。

2.4. RAVEの概要

RAVE(Realtime Audio Variational autoEncoder)は、高速かつ高品質な音声合成を可能にするために特別に構築された変分オートエンコーダーモデルである。RAVEは、特定の二段階トレーニングプロセスを採用している。まず、モデルは通常のVAE(Variational AutoEncoders)としてトレーニングされ、表現学習に使用される。その後、高品質な音声合成を実現するために、対抗的な生成目標でモデルを微調整する。このモデルは、元の波形のマルチバンド分解と,Engelらのインスピレーションを受けた古典的な合成ブロックを組み合わせており、計算複雑性を大幅に増加させることなく、48kHzのサンプリングレートまで高品質の音声合成を実現する。また、トレーニング後の潜在空間の分析により、再構築の忠実度と表現のコンパクト性との直接的な関係を制御できるようになった。これにより、潜在的な軌跡の探索と操作が容易になる。RAVEモデルはパラメータ数が少なく、複雑なデータセットでも収束し、標準のノートパソコンCPU上でリアルタイムに使用できる。非監督学習の音声モデリングにおいて、RAVEモデルは優れた性能を発揮し、信号の圧縮や音色スタイルの変換などの分野にも適用できる(Antoine, 2021)。

2.5. Pdでの「nn～」の利用

「nn～」オブジェクトは、Pdと深層学習のlibtorch C++インタフェースとの間に介在し、事前学習(Pre-training)

した RAVE を読み込むことができる変換レイヤーである。これにより, Pd で直接深層学習の機能を使用できるようになる。今回は予め開発者である Antoine Caillon によって RAVE に 80 時間分のクラシック音楽を学習させたモデルを利用した。

3. 作品のシステム構成

3.1. ハードウェアとソフトウェア

本作品では, LeapMotion ジェスチャートラッカーや Zoom 社 H3-VR 全方位録音デバイスを用いた。展示作品の実装には Pd や TouchDesigner, のソフトウェアツールを使用した。Zoom H3-VR は 360 度全方位の音声を記録でき, 今回は緑地のサウンドスケープ素材の現地録音に用いた。LeapMotion は, 赤外線カメラによって手や指の位置を追跡するデバイスである。TouchDesigner は地図、手の動きを表すポインタと視覚エフェクト等の視覚情報の提示, Pd は都市や緑地の環境音の聴覚情報の出力を制御するためのプログラミング環境である。Pd では録音した音源の再生に加えて, 手の動きによってディレイ、リバーブ、音量のモジュレーションの度合いを変化させるエフェクト機能も制御している (図 1)。

そして, 展示後の改良に機械学習を取り入れた際の RAVE モデルは, 音楽生成と音色変換に使用する機械学習モデルである。



図 1: 作品全景

3.2. 作品の体験・操作方法

以下の手順で, ディスプレイの前で鑑賞者が手を空中で動かして本作品を操作してインタラクションを行う。ディスプレイに映る視覚情報, スピーカーから出力される聴覚情報を視聴しながら鑑賞者は操作を行う。手による操作方法は, シーンによって変わる。

1. 都市の選択：左手を握るたびに, 都市が切り替わる。右手を上下すると, 緑化率によって量の違うエフェクトのかかった音が鳴る。(図.2)

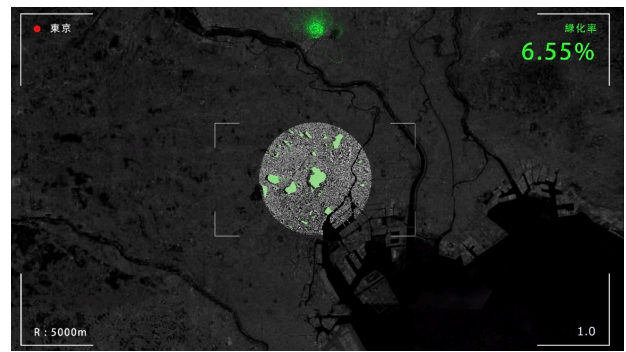


図 2: 都市の緑化率に応じた音が鳴る

2. 緑地の探索：右手の移動によって, 画面のポインタを動かし, 緑地に指して選択する。(図.3) 左手を上下すると緑地のサウンドスケープの音にモジュレーションなどのエフェクトがかかる。(図.4, 5)



図 3: ポインタで緑地を選択する



図 4: 左手が低い時に環境音の原音再生

3.3. 作品のシステムの動作の流れ

以下の流れでシステムは動作する。(図.6)

1. LeapMotion を使用して, 鑑賞者の手の位置 (空間座標データ) をキャプチャし, PC に送信する。



図 5: 左手が高い時に視聴覚出力のエフェクトがかかる

2. TouchDesigner を介して読み取られた、手の位置データに基づいて、画面上のターゲットの位置を制御する。
3. 同時に手の座標データは、OSC (Open Sound Control) を使用して Pd に送信される。
4. Pd は、手の上下によってサウンドスケープ音の音量やディレイやモジュレーションなどのかけるエフェクトの量を変化させる。

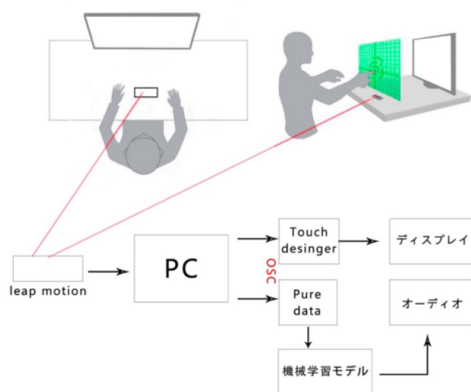


図 6: システム図

4. 展示の観察と発展

4.1. 展示概要

2024年02月23日から2024年02月24日にかけて愛知県立芸術大学にて開催された「インターカレッジ・ソニックアーツ・フェスティバル 2023」において、体験型インスタレーション「アーバン・グリーン・シンフォニー」を展示した(顔, 松村, 2024)。インターカレッジ登録校の教員、学生と一般来場者が作品を体験した様子について観察を行ない、対話によるフィードバックを得た。

4.2. 体験型インスタレーションの鑑賞者からのフィードバック

鑑賞者の自由回答と実際に在廊していた際の鑑賞者との対話から得られたフィードバックは以下であった。

- このインタラクティブシステムを通じて、都市の音環境と緑地の面積が正の相関関係にあることが良く理解できるという感想が共通していた。
- システムの視覚的な誘導により、抽象的な音を感知することが可能であり、都市の音環境と緑地の関係をより深く理解できる。
- 初めて操作する鑑賞者にとって、両手を同時に使って操作するのはやや困難である。
- 都市の音を音楽リソースとして活用する方法について、さらに追加する余地があるとの意見があった。
- 弱視の方が展示を訪れたが、彼らは視覚的な効果を十分に体験できないと感じていた。

4.3. システムの改良と発展

以下の操作方法の改良と機械学習による音楽生成機能の追加を行なった。

1. 手の操作の変更：両手の操作から右手の操作とキーボードのキー入力（スペースキー）の併用に切り替えた。
2. スペースキーによる6つのマップとモードの切り替え：押す度に東京の都市環境音、東京の異なる緑地環境音、横浜の都市環境音、横浜の異なる緑地環境音、さいたま市の都市環境音、さいたま市の異なる緑地環境音に切り替わる。
3. 都市の緑地の探索：右手を動かして画面内のターゲットを赤い枠内に移動させ、その後、右手を握って選択を確定する。
4. 音の加工とミックスの調整：握った状態で右手を上下に動かすと、画面が揺れる効果が現れる。揺れが強いほど、緑地音と都市音の対比効果が大きくなる。揺れがない平穏な状態では、生音が再現される。
5. 都市の緑化率の認識：右手を動かして画面内の緑色の粒子円の動きを制御する。緑色の粒子円が画面の上部にある場合、画面の右上隅に都市の緑化面積率が表示する。

6. RAVE 機械学習モデルによる音色変換：右手を握ると、都市の音と緑地の環境音が RAVE モデルによってクラシック音楽の音色に変換される。具体的には Pd で「nn~」オブジェクトを使用し、RAVE 機械学習モデルに接続して実現している (図 7)。

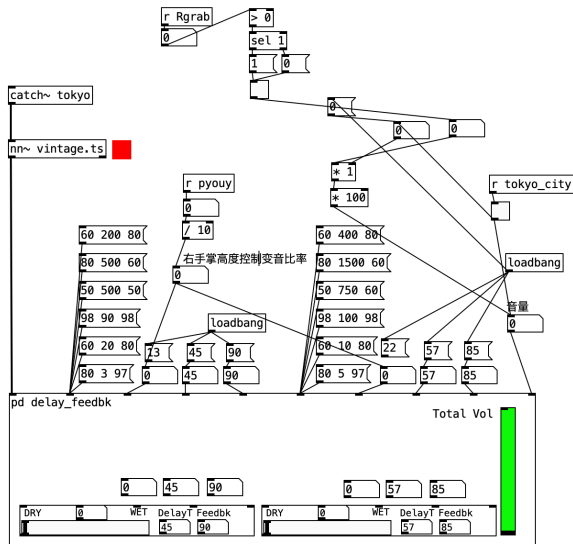


図 7: PureData パッチ

5. まとめと今後の展開

展示期間中、本作品の体験によって、鑑賞者が都市の緑地と音環境に対して興味を示す場面が見受けられた。特に、緑地環境音モードでは、鑑賞者が手のジェスチャーで緑地の環境音を操作する体験によって、サウンドスケープを「奏でる」のを楽しむ様子も見受けられた。今後は前述の改良を施した後の作品の展示を行い、新たに鑑賞者からのフィードバックを得る予定である。

一方で、より包括的な(インクルーシブ)体験の検討などの課題も浮き彫りになった。近年、情感触覚の研究は著しい進歩を遂げている。そのため、従来の視覚・聴覚情報に加えて触覚体験を導入することを通じて、より多様で包括的なインクルーシブ体験へと発展させることも視野に入れている。

6. 参考文献

- TouchDesigner . Derivative. 2024 年 03 月 28 日アクセス (<https://derivative.ca/>)
- PureData . Miller Puckette. 2024 年 03 月 28 日アクセス (<https://puredata.info/>)

LeapMotion . Ultraleap. 2024 年 03 月 28 日アクセス (<https://www.ultraleap.com/>)

宮城昭博氏ら.2021.「森林空間がもたらす騒音に対する音の物理的減衰効果」『森林計画学会誌』54 (2) ,pp.93-103.

Jefferson Dias de Oliveira,Daniela Biondi,Allan Rodrigo Nunho dos Reis.2022.“The role of urban green areas in noise pollution attenuation”*Dyna* ,89(220),pp.210-215.

Dzhambov A M, Markevych I, Tilov B, et al. 2018.“Lower noise annoyance associated with GIS-derived greenspace: pathways through perceived greenspace and residential noise”*International journal of environmental research and public health*, 15(7),pp.1533.

吉田匠吾, 森川純樹, 小川泰輝, 等.2021.「体験型プロジェクトマップがもたらすインタラクティブ性の効果に関する考察」『映像情報メディア学会技術報告』41(12),pp.5-8.

桐山孝司, 佐藤雅彦.2019.「インタラクティブな展示の設計」『科学基礎論研究』46(2),pp.65-70.

公文太, 一小坂直敏.2022.「RAVE によるリアルタイム音色変換を用いたライブエレクトロニクス作品の制作」『先端芸術音楽創作学会会報』14(3),pp.6-10.

井上英章.2022.「機械学習を利用した脳波データによるリアルタイム楽曲生成」『先端芸術音楽創作学会会報』14(3),pp.24-30.

鳥越けい子.1997.「サウンドスケープ-その思想と実践」『日本音響学会誌』53(12),pp.964-971.

鳥越けい子.1990.「サウンドスケープとはなにか」『環境技術』19(7),pp.409-411.

McCartney, A. 2002. “Circumscribed journeys through soundscape composition” *Organised Sound*, 7(1),pp.1-3.

Schulte-Fortkamp, B.,Dubois, D.2006.“Recent advances in soundscape research-Preface”*Acta Acustica united with Acustica*, 92(6),pp.5-8.

Deng, Z., Wu, W, Shi, D.2009.“Two case studies on the soundscape in historical area and its subjective assessment from the local people”In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* ,Vol. 2009, No. 4, pp.2672-2680.

Schulte-Fortkamp, B.2013.“Soundscape-A matter of human resources”In *INTER-NOISE and*

NOISE-CON Congress and Conference Proceedings ,Vol. 247, No. 3, pp.4738-4742.

高橋滋, 本多正喜, 宇部弘子, 等.1995.「クラシック音楽の情動的効果について」 [D]. Gunma University.

森下晶代.2003.「安静臥床を強いられた健康な女子学生のストレス反応と音楽の効果」『日本看護科学会誌』 23(1),pp.36-45.

Antoine Caillon, Philippe Esling .RAVE (Real-time Audio Variational autoEncoder) , IR-CAM - Sorbonne Universite ,On-line PDF,2021. (<https://arxiv.org/pdf/2111.05011.pdf>)

Ircam . nn-tilde (2023) 「github-actions」 2024年 03 月 28 日アクセス (<https://acids-ircam.github.io/nn-tilde/>)

顔 鑫, 松村誠一郎.2024.「Urban Green Symphony」作品展示, インターカレッジ・ソニックアーツ・フェスティバル 2023, 愛知県立芸術大学, 2024年 03 月 28 日アクセス (<https://ic.jssa.info/>) .

7. 著者プロフィール

顔 鑫 (Xin YAN)

東京工科大学デザイン研究科2年生. 松村研究室に所属. 中国湖北美術学院大学にて版画を学び, 卒業後に来日. 現在はアートとデジタルテクノロジーの融合を追求して, 都市の音環境と緑地のバランスに焦点を当てたマルチモーダル (聴覚・視覚・触覚) 体験型インスタレーションの制作をしている.

松村 誠一郎 (Seiichiro MATSUMURA)

東京工科大学デザイン研究科教授. 博士 (学際情報学) . 株式会社セガ・エンタープライゼス, 東京大学大学院学際情報学府, オランダ・デン・ハーグ王立音楽院ソノロジー研究科ソノロジーコース, 東京藝術大学芸術情報センターを経て, 2010年より東京工科大学デザイン学部准教授, 2018年より教授, 2019年より現職. 2019年より明治大学総合数理学部兼任講師. 専門はサウンド・デザイン, コンピュータ音楽, サウンドを軸としたインタラクティブインスタレーションの制作.



この作品は、クリエイティブ・コモンズの表示 - 非営利 - 改変禁止 4.0 国際 ライセンスで提供されています。ライセンスの写しをご覧になるには、<http://creativecommons.org/licenses/by-nc-nd/4.0/> をご覧頂くか、Creative Commons, PO Box 1866, Mountain View, CA 94042, USA までお手紙をお送りください。