

環境音付与による音楽印象操作の効果検証

藤本 直樹^{1,a)} 山西 良典^{1,b)} 松村 耕平^{2,c)}

概要：本稿では、環境音に由来するサウンドスケープを利用した新たな音楽の楽しみ方の端緒として、楽曲を聴取する際の音楽印象に環境音が及ぼす影響を分析した。かつて楽曲を聴取する際には、周囲の自然音など楽曲以外の音に影響を受ける聴取環境であることが当然であった。しかし現在では、各種機器の高度化やノイズキャンセリングの普及から楽曲に含まれない音をノイズとみなし、楽曲単体のみを聴取することが一般的な聴取形態となっている。我々は、「何を楽しむか」のみではなく、「どう楽しむか」に目を向けた。空間コンテキストの認識に有効に働く環境音に注目し、RQ1「環境音の追加により音楽への印象は変化するのか？」と RQ2「どのような環境音が、音楽の印象をどのように変化させるのか？」の2種類の課題を検証した。その結果、環境音の付与により音楽印象が変化しうることを確認した。また、音楽印象は環境音と楽曲それぞれのコンテキストが相互作用することで形成されるものであり、組み合わせによって多様な変化が起こり得ることが示唆された。

1. はじめに

かつて音楽コンテンツを楽しむ際、周囲の自然音やハードウェアの駆動音などといった楽曲に含まれない環境音が同時に聞こえることは当然であった。しかし、録音および再生環境の改善やイヤフォンのノイズキャンセリング機能に伴って、楽曲音源を単体で聴取するという音楽の楽しみ方が一般化した。楽曲聴取の場において環境音はノイズと捉えられているといっても良いであろう。大きな騒音や雑踏の音がノイズとして捉えられることが多い一方で、環境音の存在により音楽の魅力を高める可能性も考えられる。例えば、野外フェスなどの環境音が耳に入る状況での音楽聴取は一般的に人気である。また、海辺で波打ち際の音と共に聴取する場合には、楽曲単体での聴取とは違った印象が感じられる。言い換えれば、音楽印象は聴取時の環境音との組み合わせによって形成されとも言える。

環境音は、空間コンテキストの認識に有効に働くことが知られている [1]。そのため、音楽聴取時に環境音を組み合わせることで、聴取者に異なる空間コンテキストでの音楽聴取体験を仮想体験させられる可能性がある。視覚情報を必要とするイラストや動画との組み合わせとは異なり、環境音の音楽との組み合わせは聴覚のみでの仮想体験を実

現することとなる。音による空間コンテキスト（風景）は国際標準化機構により“acoustic environment as perceived or experienced and/or understood by a person or people, in context”と定義されている。環境音や音楽など様々な音メディアが想起させる風景は“soundscape”（サウンドスケープ）[2]と呼ばれ、様々な研究 [3] で活用されている。

音楽聴取環境のサウンドスケープをデザインするためには、音楽聴取時の環境音が楽曲印象にどのような影響を及ぼすのかを明らかにする必要があるが、

RQ1: 環境音の追加により音楽への印象は変化するのか？
RQ2: どのような環境音が、音楽の印象をどのように変化させるのか？

といった知見は見当たらない。

本稿では、楽曲印象操作を目的とした環境音付与の効果を検証するために、上記の2種類のRQを実験的に確認する。環境音が付与されていない音楽と環境音が付与された音楽をそれぞれ聴取した際の音楽印象を比較する。

2. 関連研究

本研究では、環境音と音楽がもつ空間コンテキストとの関係性を分析する。そこで、環境音と音楽がもつ空間コンテキストそれぞれの先行研究から本研究を位置づける。

2.1 環境音に関する研究

環境音の分類に関する先行研究は数多く見られる。例えば、環境音の分類に視覚ドメインモデルを適用した Guzhov et al. の研究 [4] や CNN アーキテクチャのパフォーマンス

¹ 関西大学

Ryozenji 2-1-1, Takatsuki, Osaka 569-1095, Japan

² 立命館大学

Nojihigashi 1-1-1, Kustasu, Shiga 525-8577, Japan

a) k503587@kansai-u.ac.jp

b) ryama@kansai-u.ac.jp

c) matsumur@fc.ritsumei.ac.jp

向上に必要な学習モデルと拡張トレーニングセットを用いた Alamon et al. の研究 [5] が報告されている。環境音は主に音から受ける印象と風景を対応付け、風景毎に特徴的な音によるサウンドスケープとして整理される。

サウンドスケープには様々なシーンが想定されるが、都市 (urban) に関する環境音を用いたサウンドスケープの分類、認識の試みは特に多い。InJo et al. の研究 [6] で音風景のための音源の適切性が視覚情報と相互作用した事が報告されている。他にも、敵対的生成ネットワークを用いた環境音分類のためのアノテーション付きデータセット拡張の試み [7] や DCASE ^{*1} による音源から音響空間を評価する研究 [8] などによってサウンドスケープをアノテーションした信頼性の高い環境音データセットが構築・拡張されてきた。

2.2 音楽が持つコンテキストに関する研究

音楽が持つコンテキストを音楽体験の向上のために活用する試みは広範に行われている。例えば、感性語を用いた音楽印象に基づくマッピング及び検索 [9] や、意外性のある音楽推薦のための分岐型人力音楽推薦手法 [10] などが報告されている。特にコンテキスト認識を基にした推薦システム “Context-Aware Recommender Systems” (CARS) は、音楽推薦の精度向上や多様化を目的とした多くの研究で課題として取り組まれている。例えば, Khalid et al. による音楽コンテキストの抽出やフィルタリング [11] や、ソーシャルメディアから獲得した情報と音楽のコンテキスト情報からの音楽推薦 [12] が提案されている。

音楽情報科学研究において、環境音やサウンドスケープを音楽の生成や分類に利用する試みは、武井らによるモザイク音楽生成 [13] に見られるものの多くはない。既存の音楽が持つ音楽印象の操作や分類において、環境音を用いたアプローチは少ない。これは、環境音と音楽コンテキストを対応付けたデータセットの不足や、CARS においてサウンドスケープを活用するためのフレームワーク構築が十分ではない事に起因すると考えられる。

本研究では、環境音を音楽印象の操作を最終目的として、環境音が持つサウンドスケープが音楽印象に影響を及ぼすかを検証する。この仮説検証では、音楽印象の主観評価を含む研究 [14], [15], [16] における実験設定を課題条件と統制条件に参照した。

3. 仮説検証実験

楽曲と環境音はそれぞれ音声のみで様々なコンテキストの表現が可能であることが知られている [17]。また、環境音が持つサウンドスケープが VR 空間の状況把握に有効である [18] ことから、環境音は空間コンテキストの把握に影

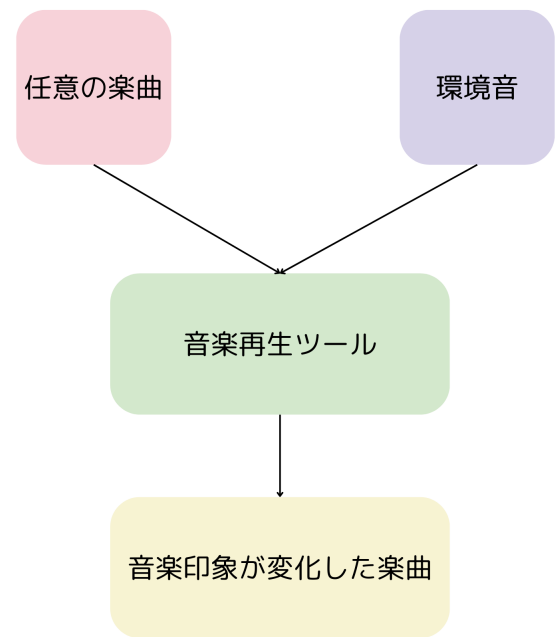


図 1 環境音を用いたインタラクションのイメージ

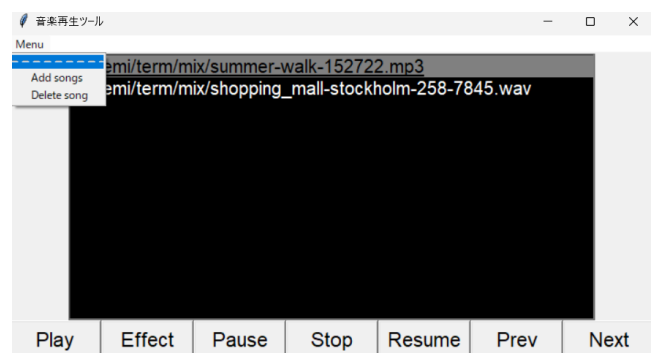


図 2 音楽再生ツールの GUI

響を及ぼすことが予測される。そこで、RQ1: 環境音の追加により音楽への印象は変化するのか？に対して、「環境音の有無が音楽印象に影響を与えない」を帰無仮説として仮説検証する。また、この実験の結果から RQ2: どのような環境音が、音楽の印象をどのように変化させるのか？を考証する。ここで本実験の独立変数は楽曲に付与される環境音、従属変数は楽曲から得られた音楽印象である。

3.1 実験環境

実験用の刺激音楽は、自作した任意の音楽と効果音を合成し再生するツールによって準備した。図 1 に、本ツールを用いた環境音による音楽印象操作のイメージを示す。本ツールを利用することで、音楽の再生中に任意のタイミングで環境音を付与可能であり、環境音が持つ様々なサウン

^{*1} <https://dcase.community>

表 1 音楽再生ツールの基本的な機能

操作	機能
add song	楽曲, 効果音を登録する
delete song	楽曲, 効果音を削除する
play	選択中の楽曲を再生する
effect	選択中の環境音を再生する
pause	音源を一時停止する
stop	音源の再生を終了する
music list	登録した音源を選択する

ドスケープを用いた音楽体験を作成可能である。本ツールは図 2 に示す GUI を実装しており、各種ボタンを用いて簡単に操作可能である。また、表 1 に示すように音楽と環境音を個別に制御して再生できる他 mp3 及び wav 形式のファイルを読み込む事で選択リストに加える機能を備えている。

環境音について、音と空間コンテキストを対応付けたデータセットはいくつか公開されている。実験で用いる環境音には、DCASE が公開している自然言語による環境状況のアノテーションが付いた環境音データセットを使用する。このデータセットは DCASE challenge という音環境理解のための取り組みにより、各音源に対応する音環境が自然言語でアノテーションされている。

楽曲について、著作権フリーの音楽素材などを提供する投稿サイト“pixabay^{*2}”にて公開されている楽曲を使用する。本サイトでは、投稿されている楽曲に付与されている音楽ジャンルや情景に関するタグから楽曲の特徴を把握できる。本稿では、環境音の付与によって変化する楽曲から想起される感情および情景を分析する。そこで、楽曲の選定においては 1) 異なる音楽ジャンルのタグを持つこと、2) 場所や感情に関するタグを持たないこと、の 2 点を条件とした。実験では、表 2 に示すようなジャンル、場所・感情タグが付与された 2 曲を用意した。

3.2 実験手順

実験では、10 代後半から 20 代前半の男女合計 30 名を被験者として用意した。本実験では、実験環境結果への影響を防ぐために、実施環境について以下 2 点の制限を設けた。

- (1) ヘッドホンを装着した状態で音源を再生した時、音源以外の音が聞こえないこと
- (2) 室温や照明などの環境について不快に感じないこと

実験の基本的な手順は次の通りである。はじめに、被験者を A 群、B 群および C 群の 3 群（10 名ずつ）に分割する。次に、表 3 に示すように群ごとに異なる環境音あるいは無音が合成された楽曲 2 曲をヘッドホンを用いて聴取させる。このとき、想起される情景および感情に自覚的な状態で聴取するように指示する。その後、得られた音楽印象

^{*2} <https://pixabay.com/ja/music/> (2023 年 11 月 28 日確認)

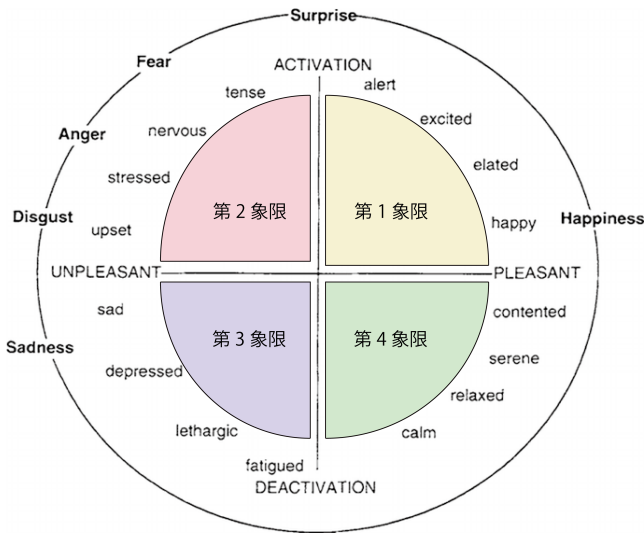


図 3 本稿で参照するラッセルの感情円環モデルを基にした感情空間。文献 [19] により引用したものに、説明に必要な象限名の付記と色付けをおこなった。

に関する設問に回答させ、回答された結果から環境音が音楽印象に与えた影響について考証する。設問への回答形式は設問内容に応じて自由記述またはリッカート尺度を用いた 1~5 の選択式のいずれかを指定する。選択式の設問では被験者にあるトピックを提示し、それに最も同意できる状態を 5、同意しない場合を 1 として 1~5 の 5 段階でトピックを評価させる。表 4 と表 5 に、それぞれ自由記述形式の設問と 5 段階評価形式の設問を示す。

4. 結果と考察

表 6 に示す 6 条件でそれぞれ得られた結果を比較・分析する。考察の観点としては感情と情景を設定し、それぞれの観点から得られた結果を考察する。

感情に関する結果の分析では、設問 1.4 で被験者から得られた感情をラッセルの感情円環モデル [20] を参考に 4 つのカテゴリに分類した。このモデルでは感情を“arousal-sleepiness”と“pleasure-displeasure”の 2 軸を用いて 2 次元上のベクトルとして表現している。本稿では、上記の 2 軸により成立する平面を感情空間と呼称し、実験で得られた感情をいずれかの象限に分類して考察する。図 3 に、本稿で扱う感情空間を示す。この感情空間は、文献 [19] により引用したものに、説明に必要な象限名の付記と色付けを行ったものである。

情景に関する考察では、時間帯、場所、聴取環境についてを詳細に考察した。楽曲から想起される時間帯の分析において、一日を“朝”“昼”“夕方”“夜”に大別し、時間帯についての回答を分類した。後述する実験で得られた結果を俯瞰すると、回答の大部分が上記の 4 種のいずれかに分類可能であり、他の“深夜”や“未明”といった特徴的な時間帯を示す回答は非常に少数であった。そこで、気象庁が公

表 2 実験に用いる楽曲がもつ各カテゴリごとのタグ

曲名	Genre	Movement	Mood	Theme
Summer Walk	“Acoustic Group”, “Beautiful Plays”, “Main Title”, “Solo Guitar”, “Small Emotions”, “Indie Pop”, “Folk”	“Smooth”, “Floating”, “Elegant”, “Medium”	“Relaxing”, “Peaceful”, “Laid Back”, “Dreamy”, “Bright”, “Romantic”, “Hopeful”	“Music For Videos”, “Music For YouTube Videos”, “Vlog Music”, “Background Music”, “Film Music”
Trap	“Trap”, “Beats”, “Electronic”, “Mainstream Hip Hop”, “Alternative Hip Hop”, “Action”	“Chasing”, “Heavy & Ponderous”, “Medium”	“Restless”, “Laid Back”, “Energetic”, “Suspense”, “Quirky”	“Music For Videos”, “Music For YouTube Videos”, “Vlog Music”, “Background Music”

表 3 各群が扱う環境音

A 群	B 群	C 群
shopping mall	urban park	無し

表 4 自由記述形式の設問

設問番号	内容
1.1	音楽をいつ聴いているように感じましたか
1.2	音楽をどこで聴いているように感じましたか
1.3	どのような聴取環境であると感じましたか
1.4	どのような感情を感じ取りましたか

表 5 5段階評価形式の設問（最大値は 5）

設問番号	内容
2.1	感情が想起される
2.2	情景が想起される

表 6 実験で使用した条件

条件	楽曲	環境音
S-N	Summer Walk	無し
S-S	Summer Walk	shopping mall
S-P	Summer Walk	urban park
T-N	Trap	無し
T-S	Trap	shopping mall
T-P	Trap	urban park

開する一日の時間細分図^{*3}を参考に“夜”の時間帯を定め、“日中”と“夜”のどちらにも属さない時間帯を“朝”、“日中”の中でも特に回答の多かった“夕方”を除く時間帯を“昼”として考察する。楽曲から想起される場所の分析では、出現頻度の高い単語や特徴的な表現を抽出する。これらの単語や表現を条件間で比較することで、音楽を単独で聴取した場合から環境音を付与したことで変化した音楽に想起する場所の違いを考察する。楽曲から想起される聴取環境については、音源と自身の距離（超近距離、近距離、遠距離）と再生デバイスの観点から考察する。ここで、超近距離は再生デバイスを身に着けている状態、近距離は再

生デバイスを認知可能な距離、遠距離はそれ以上の距離とした。

4.1 感情に関する結果と考察

表 8 に、設問番号 2.1 “感情が想起される” に対する、各条件別での回答の平均値を示す。回収した結果を楽曲別で環境音毎に回答をグループ化し分散分析をおこなったところ、S-N 条件と S-S 条件および S-P 条件では p 値が 0.84、T-N 条件と T-S 条件および T-P 条件では 0.51 であり楽曲に関わらず環境音の違いによる有意差は検出されなかった。以上の結果より、感情が想起されるか否かには環境音付与の有無や種類が影響するとは言えないことが示された。

環境音が付与された楽曲の聴取から想起された感情を図 3 に示した感情空間を用いて分類し、楽曲別で考察をおこなった。楽曲が持つ “Mood” カテゴリのタグに注目すると “Summer Walk” では感情空間上にて第 1 および第 4 象限に位置するものが多く、“Trap” では第 2 象限に位置するものが多くなっている。S-N 条件と S-S 条件および S-P 条件において感情空間上にて第 1 または第 2 象限に位置する感情に注目すると、S-N 条件では 0 件、id2 では 1 件、id3 では 4 件であり環境音 “urban park” が付与されている時により多く回答された。また、T-N 条件と T-S 条件および T-P 条件において感情空間上にて第 1 または第 4 象限に位置する感情に注目すると、環境音無しでは 1 件、環境音 s では 6 件、環境音 p では 3 件であり環境音が付与された際により多く回答された。どちらの結果においても環境音を付与することにより楽曲単体からは得られた感情とは異なる象限に位置する感情が回答されていることが確認された。環境音を付与することで、楽曲印象に含まれない感情の想起を促すことが有効である可能性が示唆された。

4.2 情景に関する結果と考察

表 9 に、設問番号 2.2 “情景が想起される” に対する、各条件別での回答の平均値を示す。回収した結果を楽曲別で環境音毎に回答をグループ化し分散分析をおこなったところ、S-N 条件と S-S 条件および S-P 条件では p 値が

^{*3} https://www.jma.go.jp/jma/kishou/known/yougo_hp/saibun.html (2023 年 11 月 28 日確認)

表 7 設問番号別での実験結果の一例

条件	設問 1.1	設問 1.2	設問 1.3	設問 1.4	設問 2.1	設問 2.2
S-N	朝方	リゾートの海	遠くから聞こえてくる	くつろぎ、癒し、安らぎ	3	5
S-N	朝方	港	道端で弾き語りしている	穏やか	3	5
S-N	朝方 (午前 8 時くらい)	海辺のガラス張りのレストラン	天井に付いている店内のスピーカーから聞こえてくる	穏やか、安らぎ	4	5
S-S	昼間	駅の中のカフェの店内	遠くから聞こえてくる	怒り (うるさい)	5	5
S-S	夕方頃	街中	近くで弾いている	落ち着き	4	5
S-S	夕方	駅の中の喫茶店	人の喧騒は遠くから聞こえるが、音楽をイヤホンで聴いてる印象	平和、日常、優雅、余裕	2	5
S-P	夕方	放課後の屋外	帰りながら聞こえてくる	疲労	4	4
S-P	夕暮れ	海辺	どこかから聞こえてくる感じ	リラックス、穏やか	4	5
S-P	夕暮れ	沖縄の海沿いお土産屋さん	スピーカーと店外からの音	おだやか	5	5
T-N	夜	屋敷	辺り一帯から響いてくる	憤り	3	5
T-N	深夜	ライブ会場	スピーカー。	怒り	1	4
T-N	夜	クラブ	スピーカーから、地面に響くように	復讐心、威嚇など	4	2
T-S	昼間	教室	天井側に設置されたスピーカーから聞こえてくる音のように感じる	怒り	2	1
T-S	夜	クラブ、USJ とか	広い建物内でスピーカーで流れている、大きめの音	楽しい、興奮、高揚	2	2
T-S	深夜	路上	近くから聞こえてくる	興奮	2	4
T-P	朝方	自室	外から鳥の囀りが聞こえてくる部屋	興奮	3	4
T-P	夕方	都会の港	スピーカー (自分と音源の距離は近い)	気分が沈んでいる感じ	3	4
T-P	夜	映画館	スピーカー	疑心	4	1

表 8 設問 2.1 “感情が想起される” への回答の平均値 (最大値は 5)

条件	平均値
S-N	3.2
S-S	3.3
S-P	3.5
T-N	3.1
T-S	2.9
T-P	3.6

表 9 設問 2.2 “情景が想起される” への回答の平均値 (最大値は 5)

条件	平均値
S-N	4.7
S-S	4.7
S-P	4.4
T-N	3.4
T-S	3.8
T-P	3.0

0.40、T-N 条件と T-S 条件および T-P 条件では 0.48 であり環境音の違いによる有意差は検出されなかった。以上の結果より、情景が想起されるか否かには環境音付与の有無や種類が影響するとは言えないことが示された。

4.2.1 時間帯に関する結果と考察

設問番号 1.1 “音楽をいつ聴いているように感じましたか” に対する回答を実験条件および時間帯ごとに分類し整

理した。表 10 に、各条件での回答件数を示す。

はじめに、S-N 条件と S-S 条件および S-P 条件に注目する。S-N 条件では、6 件の回答が “朝” に分類されている一方で “昼” に分類された回答は 0 件であった。しかし S-S 条件では、回答の半数が “昼” に分類され、また S-P 条件では 6 件の回答が “夕方” に分類された。次に、T-N 条件と T-S 条件および T-P 条件を注目すると、T-N 条件では

表 10 設問 1.1 “音楽をいつ聴いているように感じましたか” に対する時間区分別での回答数

条件	朝	昼	夕方	夜	分類不可
S-N	6	0	4	0	0
S-S	1	5	4	0	0
S-P	3	1	6	0	0
T-N	0	0	0	10	0
T-S	0	3	0	6	1
T-P	2	3	1	4	0

表 11 設問 1.3 “どのような聴取環境であると感じましたか” に対する距離別での回答数

条件	超近距離	近距離	遠距離	分類不可
S-N	0	7	1	2
S-S	2	4	4	0
S-P	0	4	3	3
T-N	0	1	2	7
T-S	0	4	4	2
T-P	1	4	1	4

全ての回答が“夜”に分類された一方で T-S 条件では 3 件が“昼”に分類されており、T-P 条件では過半数が“夜”以外に分類されている。これらの結果から、音楽印象に対しそれぞれの環境音固有の影響が及んだことが示唆された。

4.2.2 場所に関する結果と考察

S-N 条件の場合では“港”や“リゾートの海”など海辺を指す表現が 4 件みられた。また回答の半数が“レストラン”や“部屋の中”など何らかの施設を場所として挙げていた。

S-S 条件の場合では“カフェ”または“喫茶店”という単語が 6 件の回答に含まれており、また“駅”という単語が 5 件の回答に含まれていることが特徴的であった。同じ楽曲を用いた際の“カフェ”の出現回数は S-N 条件で 1 件、S-P 条件では 0 件であり環境音“Shopping mall”に強く影響を受けた音楽印象であると考えられる。

S-P 条件の場合では海辺に関係する表現が 6 件に出現している。S-N 条件の場合にも海辺を指す表現が頻出していたことを踏まえると、環境音“urban park”が“shopping mall”と比較して自然環境に近い環境音であったことから、環境音が楽曲本来の音楽印象を増長させた結果であると考えられる。

T-N 条件と T-S 条件および T-P 条件の場合では環境音の違いによる表現の差が確認できなかった。S-N 条件と S-S 条件および S-P 条件のとき環境音毎に特徴的であった“カフェ”や“港”といった表現は T-N 条件と T-S 条件および T-P 条件においては出現しないことから、環境音のみが持つ印象ではなく楽曲由来の音楽印象と環境音が持つ情報が相互に作用することで新たな音楽印象が形成されることが示唆された。

表 12 設問 1.3 “どのような聴取環境であると感じましたか” に対する回答における再生デバイスへの言及

条件	再生デバイスに言及した回答数	特徴的な表現
S-N	7	演奏
S-S	5	演奏、イヤホン
S-P	4	演奏
T-N	7	スピーカー
T-S	6	スピーカー
T-P	7	スピーカー

4.2.3 聴取環境に関する結果と考察

表 11 に、音源と自身の距離について条件ごとに何件の回答が得られたかを示す。S-N 条件の場合において、10 件の回答のうち 7 件が近距離に音源があることを示すものであった。一方で、S-S 条件および S-P 条件では音源との距離が近距離であることを示す回答は各条件で 4 件ずつとなっており、環境音の付与によって自身と音源との距離に関する認識が変化したことが伺える。また、S-S 条件では距離の特定が不可能な回答が 0 件であった。これに加え、他の楽曲を扱った T-N 条件と T-S 条件および T-P 条件においても環境音付与によって距離の分類が不可能であるとされた回答が減少していることから、環境音付与が音楽印象が持つ空間コンテキストをより詳細化したことが示唆された。

再生デバイスについて、S-N 条件と S-S 条件および S-P 条件においては“演奏”や“弾き語り”などの表現が多くみられた一方で、T-N 条件と T-S 条件および T-P 条件では 30 件中 19 件の回答に“スピーカー”という単語が出現した。表 12 に、各条件において再生デバイスについて言及している回答数と特に出現回数の多い表現を示す。“弾き語り”や“弾いている”などの楽器を演奏していることを示す表現は“演奏”として扱った。再生デバイスに言及した回答数および特徴的な表現は S-N 条件と S-S 条件および S-P 条件および T-N 条件と T-S 条件および T-P 条件の比較では確認されなかった。前述した音源と自身の距離についての変化と合わせた考察から、再生デバイスについては楽曲自体が持つ音楽印象への依存度が高い一方で、空間情報の認識においては環境音が影響を及ぼしたことが示唆された。

4.3 情景に関する複数観点からの議論

音楽印象から想起される情景について、時間帯、場所および聴取環境の観点から複合的に考察した。S-N 条件と S-S 条件および S-P 条件で示された条件を比較すると、S-N 条件のときには朝の海辺で演奏に耳を傾ける情景が想起されたと言える。一方で、S-S 条件では昼のカフェや駅といった施設が想起され、S-P 条件では夕方や朝の海辺が多く想

起された。T-N 条件と T-S 条件および T-P 条件においては共通してクラブやイベント会場などでスピーカーから流れる楽曲を聴取している情景が想起された。一方で、想起される時間帯について T-N 条件の場合には夜を指す表現のみが見られたが、T-S 条件、T-P 条件においては他の時間帯を示す回答が半数程度確認できた。各条件で付与された環境音の性質をもとに、これらの結果を考察する。S-S 条件および T-N 条件で使用されている環境音 “Shopping mall” は、人の話し声が含まれており、「人出の多い時間帯である昼」かつ「賑わいのある場所」の情景を想起させる可能性がある。また、S-P 条件および T-P 条件で使用した環境音 “Urban park” には鳥の鳴き声などが含まれており、「野鳥の活動時間である朝および昼」かつ「屋外の音が聞こえる場所」が想起されたと考えられる。S-N 条件と S-S 条件および S-P 条件および T-N 条件と T-S 条件および T-P 条件から得られた結果を比較すると、これらの環境音の性質による影響が及んでいると考えられる回答が多いものの、同一の環境音がどのような楽曲に対しても同一の情景を思い起こさせたわけではないことが確認できる。このことから、楽曲自体がもつ印象と組み合わせられることで様々な情景を呼び起こすきっかけとして作用することが示唆された。

実験で使用した 2 曲の楽曲と 2 種の環境音の全ての組み合わせで、環境音付与による音楽印象の変化を確認した。一方で、付与した環境音が同じであった場合にも特徴的な表現として抽出された単語や最も多く想起された時間帯は楽曲によって異なった。また、4.2 章で示したように、情景が想起されるか否かについては、条件の変化による有意差が確認されなかった。これらのことから、

- 楽曲に環境音を付与した場合には楽曲と環境音の組み合わせによって、組み合わせ固有の音楽印象が形成される
- 環境音付与によって形成された音楽印象は楽曲自体が本来持つ印象と同程度に強く想起される

といった考察が得られた。

5. おわりに

本稿では、楽曲に環境音を付与することで音楽印象がどのように変化するかを自由記述および段階評価を用いて調査し、音楽印象を操作する手段としての環境音付与の効果を検証した。実験の結果から、環境音の付与によって楽曲聴取から想起される感情や情景の認識に変化が確認された。このことから、環境音に由来する様々なサウンドスケープを利用した音楽印象操作の可能性が示唆された。一方で、同じ環境音であっても異なる楽曲には異なる感情や情景を想起させている様子も確認でき、音楽印象操作は楽曲と環境音の組み合わせに応じて異なることが確認された。

音楽印象を環境音によって操作するためには、環境音と

音楽の組み合わせの多様な条件を分析し、音響特徴量同士や環境音が持つサウンドスケープが与える意味的な効果を取り入れたモデルを構築する必要がある。今後は、本稿で示された知見と明確になった問題点を参照し、実験条件のアップデートを行った上で被験者人数を充実させた実験を行う。また、環境音と楽曲の組み合わせを活用して特定のサウンドスケープをデザインするような音楽鑑賞の新しいインタラクションを目指していきたい。

参考文献

- [1] 井本桂右: DCASE Challenge: 環境音分析・理解のための統合的コンペティション, 日本音響学会誌, Vol. 79, No. 9, pp. 470–476 (2023).
- [2] Philip, D.: ISO/TS 12913-2: 2018- Soundscape-Part 2: Data collection and reporting requirements-what’s it all about, *Acoustics Bulletin*, pp. 55–57 (2018).
- [3] A.L.Brown, Kang, J. and Gjestland, T.: Towards standardization in soundscape preference assessment, *Applied Acoustics*, Vol. 72, No. 6, pp. 387–392 (2011).
- [4] Guzhov, A., Raue, F., Hees, J. and Dengel, A.: ES-ResNet: Environmental Sound Classification Based on Visual Domain Models, *2020 25th International Conference on Pattern Recognition*, pp. 4933–4940 (2021).
- [5] Justin, S. and Pablo, B. J.: Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification, *IEEE Signal Processing Letters*, Vol. 24, No. 3, pp. 279–283 (2017).
- [6] Jo, H. I. and Jeon, J. Y.: Effect of the appropriateness of sound environment on urban soundscape assessment, *Building and Environment*, Vol. 179, p. 106975 (2020).
- [7] Madhu, A. and Kumaraswamy, S.: Data Augmentation Using Generative Adversarial Network for Environmental Sound Classification, pp. 1–5 (2019).
- [8] Politis, A., Mesaros, A., Adavanne, S., Heittola, T. and Virtanen, T.: Overview and Evaluation of Sound Event Localization and Detection in DCASE 2019, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 29, pp. 684–698 (2021).
- [9] 池添 剛, 梶川嘉延, 野村康雄: 音楽感性空間を用いた感性語による音楽データベース検索システム, 情報処理学会論文誌, Vol. 42, No. 12, pp. 3201–3212 (2001).
- [10] 野中滉介, 中村聡史: 未知の音楽に誘導することを目的とした分岐型人力音楽推薦手法の提案, 研究報告ヒューマンコンピュータインタラクション, Vol. 2020-HCI-187, No. 15, pp. 1–7 (2020).
- [11] Haruna, K., Ismail, M. A., Suhendroyono, S., Damiasih, D., Pierewan, A. C., Chiroma, H. and Herawan, T.: Context-Aware Recommender System: A Review of Recent Developmental Process and Future Research Direction, *Applied Sciences*, Vol. 7, No. 12, p. 1211 (2017).
- [12] Tan, S., Bu, J., Chen, C., Xu, B., Wang, C. and He, X.: Using Rich Social Media Information for Music Recommendation via Hypergraph Model, *ACM Trans. Multi-media Comput. Commun. Appl.*, Vol. 7S, No. 1, p. 22 (2011).
- [13] 武井 祥, 徳永幸生, 杉山 精, 阿部匡伸: 環境音を用いたモザイク音楽の生成条件の検討, 第 73 回全国大会講演論文集, Vol. 2011, No. 1, pp. 251–252 (2011).
- [14] 佐藤敬子, 満倉靖恵: 音楽が画像の印象に及ぼす影響と印象と物理的特性との関連に関する一考察, 電気学会論文誌 C, Vol. 131, No. 8, pp. 1451–1458 (2011).

- [15] 菅野禎盛, 岩宮眞一郎: 映像と音楽の情緒的印象に対する同期要因と速度対応要因の効果, 日本音響学会誌, Vol. 56, No. 10, pp. 695–704 (2000).
- [16] 石田理代, 佐藤敬子, 満倉靖恵: 音楽が画像印象に与える影響と音楽特徴量との関連性, 電気学会論文誌C, Vol. 136, No. 9, pp. 1376–1385 (2016).
- [17] Fuda, S., Aletta, F., Kang, J. and Astolfi, A.: Sound Perception of Different Materials for the Footpaths of Urban Parks, *Energy Procedia*, Vol. 78, pp. 13–18 (2015).
- [18] Hong, J. Y., Lam, B., Ong, Z.-T., Ooi, K., Gan, W.-S., Kang, J., Feng, J. and Tan, S.-T.: Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications, *Building and Environment*, Vol. 149, pp. 1–14 (2019).
- [19] Posner, J., Russell, J. A., Gerber, A., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlou, A., Zhu, H. and Peterson, B. S.: The neurophysiological bases of emotion: An fMRI study of the affective circumplex using emotion-denoting words, *Human brain mapping*, Vol. 30, No. 3, pp. 883–895 (2009).
- [20] Russell, J. A.: A circumplex model of affect., *Journal of personality and social psychology*, Vol. 39, No. 6, p. 1161 (1980).