



# 雰囲気制御のための環境音変換手法の基礎的検討

## Basic Study of Environmental Sound Conversion Method for Ambience Control

平田祐登<sup>1)</sup>, 竹内俊貴<sup>1)</sup>, 谷川智洋<sup>1)</sup>, 鳴海拓志<sup>1)</sup>, 廣瀬通孝<sup>1)</sup>

Yuto HIRATA, Toshiki TAKEUCHI, Tomohiro TANIKAWA, Takuji NARUMI, and Michitaka HIROSE

1) 東京大学大学院 情報理工学系研究科 (〒 113-0033 東京都文京区本郷 7-3-1)

**概要:** 環境音は空間の印象を特徴づける重要な要素であり, 環境音によって雰囲気を変えることができる. 実空間の文脈を保持したまま環境音を補正することで, 違和感のない雰囲気の制御が可能になると考えられる. 本研究では, 実際の環境音に対し音響パラメータを操作し, 印象評価を変えることができるのかを調査するため, 環境音の変換手法を検討し, 知覚される情動の変化との関係を評価する検証実験を行った. その結果, 高音増幅により覚醒度が高いと知覚される傾向や, 低音増幅により高級感が高いと知覚される傾向, 快さの知覚には音源依存性があることが示唆された.

**キーワード:** 環境音, 雰囲気制御, 聴覚, 主観評価実験

### 1. 研究の背景と目的

環境音は, 日常生活の中で無意識に耳にし, その空間の印象を特徴づける重要な要素であり, また環境音によって感じる雰囲気を変えることができる. 例えば BGM にはマスキング効果やイメージ誘導効果があるとされ [1], 例として喫茶店において家具や照明とともに空間を演出するために用いられている. しかし, 環境音の評価構造は個人・社会・文化的要因により多様であり [2], 個人の嗜好や心理状態に適合した雰囲気の制御は難しい.

また記憶・計算作業時, 外来雑音を遮断し集中力を向上するために, 密閉型ヘッドホンやノイズキャンセリングヘッドホンが用いられる. ノイズキャンセリングヘッドホンは, ヘッドホンの外側から伝わる雑音と逆位相の音波を生成し, 打ち消すことで, ユーザはヘッドホンから再生する音声のみを鮮明に聞くことを可能にする. しかし, これらのヘッドホンは, ユーザにとって空間の文脈を崩し, 生活に必要な音までも遮断してしまう.

一方で, 地域や場における様々な音を一つの「風景」として捉えた概念を「サウンドスケープ」と呼び [3], 環境音の中から不必要な音を排除するだけでなく, 必要な音を積極的に活用し, 音を配慮した環境改善の動きも見られる.

このように, 従来環境音は個人にとってそのまま受け容れる (図 1 上) か遮断する (図 1 下) かであり, 個人の嗜好や心理状態に適合させることは困難だった. そこで本研究では, 図 1 中央に示すように, 個人の嗜好や心理状態と空間の文脈の両方を考慮した雰囲気の制御を目的とする. その基礎検討として, 実空間において環境音を遮断せず, 空間の文脈を保持したまま環境音を補正することで, 違和感なく環境音に対する印象を制御する手法を検討した.

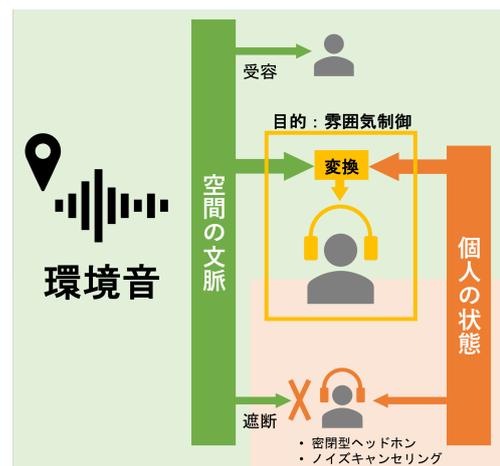


図 1: 本研究が目指す環境音変換の概要

### 2. 関連研究

Weiyi Ma & William Forde Thompson は, 人間の行動, 動物の音, 機械の騒音, 自然の音の 4 カテゴリーの単一音源の環境音をそれぞれ 6 つずつ用意し, その環境音に, ピッチ, 強度, 速度の調整を行った [4]. ピッチは周波数スペクトルを 4 半音上げた, また下げた音, 強度は  $\pm 5$  dB の強度の調整を行った音, 速度は 1.3 倍速, 0.7 倍速に調整した音を用意し, Russell の円環モデル [5] の快-不快・覚醒-眠気の 2 軸について印象評価を行った. その結果, より高いピッチ, より強い強度, より速い速度が, より高い快, より高い覚醒と関連していることを見出している. Russell の円環モデルは, 情動を快-不快・覚醒-眠気の二次元上で表せるとしたモデルで, 広く研究で用いられている.

また Fan らは, 6 秒のサウンドスケープ 600 個と, それらを混合した 613 個の合計 1,213 個のサウンドスケープ

を、快-不快・覚醒-眠気の2軸において順位付けし、“Emo-Soundscapes”データセットとして公開している [6]。そして、異なるサウンドスケープカテゴリに属する音声を混合した音声と、その元となる2つの音声により知覚される情動の間に、特定のパターンを見出している。また、所望する快-不快・覚醒-眠気の時間変化曲線とシナリオの文章を入力することで、それに応じたサウンドスケープを合成するシステムが公開されている。逆に環境音を解析して快-不快・覚醒-眠気の2軸を評価することも可能としている [7][8]。

そこで本研究は、より実際の日常生活で耳にしよう環境音についても同様に知覚される情動を変化させることができるのか、また環境音によって知覚される快-不快・覚醒-眠気の2軸を独立に制御できるのかに着目し、変換手法を検討する。

### 3. 環境音の変換手法

#### 3.1 環境音の分類法

環境音の分類法として、Schafer の分類法 [3] が広く使われている。Schafer は、サウンドスケープを、音響特徴量ではなく、音源の識別と聴取の文脈に基づいて6つのカテゴリに分類した。たとえば、Schafer の分類法では、静かな森の音は“Quiet and silence”の下に分類されるが、聴取の文脈がなければ、“Natural sounds”として分類することができる。これらのカテゴリは排他的ではない。

#### 3.2 音源

対象とする環境音として、Schafer の分類法のうち、日常生活の中で無意識のうちに耳にし、定常的である音として、自然、人、社会、機械カテゴリの音を対象とする。その中で「社会」が含まれる音として、人々の話し声を含む「喫茶店」の環境音、また「自然、人、機械」が含まれる音として、鳥の鳴き声、車のエンジン音、足音を含む「朝の雑踏」の環境音を選出した。これら2種類の環境音の音声データをインターネット上のフリー効果音配布サイト [9] から選出し、ダウンロードした。音声は Roland Recorder R-09 で録音され、チャンネル数は2、サンプリングレートは48,000Hz、ビットレートは160kbpsでMP3形式でフォーマットされていた。「喫茶店」の元音声データは41秒、「朝の雑踏」の元音声データは1分8秒の長さだった。

#### 3.3 環境音の変換

先行研究より、単一音源の環境音について、ピッチ、強度、速度を調整することで知覚される情動が変化することが見出されている [4]。また公共空間における放送設備において、騒音や残響があると聞き取りにくかったり不快に感じることがある。軽量床衝撃音に対する不快感が、残響が極端に長い音、短い音で増すことが分かっている [12]。さらに、精神作業時にうるささに関する心理的印象を低減するためには、無意味定常雑音であるピンクノイズが有効であることが分かっている [10][11]。

これらを踏まえ、多様な音源を含んだ実際の環境音を、空間の文脈を保持したまま違和感なく変換できるよう、PCに

て波形編集ソフト Audacity [13] を使い、以下の3種類の効果について、パラメータの異なる合計8条件の変換のうちから1条件ずつ施した音声データをそれぞれ作成し、原音と合わせて合計9条件の音声データを用意した。環境音の変換を行った条件は以下の通りである。各条件の右側に付した記号は、その条件のラベルとする。

- 原音 … O
- 原音+イコライザ (+6dB/oct) … H6
- 原音+イコライザ (+3dB/oct) … H3
- 原音+イコライザ (-3dB/oct) … L3
- 原音+イコライザ (-6dB/oct) … L6
- 原音+リバープ (プリディレイ 30ms) … R30
- 原音+リバープ (プリディレイ 60ms) … R60
- 原音+ピンクノイズ (S/N 比 12dB) … P12
- 原音+ピンクノイズ (S/N 比 18dB) … P18

なお、各効果における他の条件は、それぞれ以下のように統一した。

#### イコライザ

100Hz以下の周波数帯域の増幅率を0dBとし、100Hzから1kHzの周波数帯域において、1オクターブあたり+6dB、+3dB、-3dB、または-6dBの傾斜を持たせ、1kHz以上の周波数帯域において平坦とした増幅率の、合計4個のイコライザを形成した。

#### リバープの付加

残響を付加するため、プリディレイの値が30ms、または60msのリバープ効果をかけた。Audacityにおける他のパラメータは統一しデフォルトの通りとした。

#### ピンクノイズの付加

$1/f$  ゆらぎを含む無意味雑音として、ピンクノイズを選定した。S/N比が12dBまたは18dB、つまりピンクノイズの音圧が、元の音声データと比較して-12dB、または-18dBのピンクノイズを付加した。

また音声の長さについては、先行研究より各音声の性質を知覚するために必要な音声の長さは6秒とした。

全ての音声データについて、それぞれDCオフセットを削除した上で最大振幅が-1dBとなるようにノーマライズ処理を施した上で、音声データの開始12秒後から6秒間の区間のみを切り出し、PCM 16bit形式で保存したものを使用した。

### 4. 環境音の変換による雰囲気制御の検証実験

実際の日常生活で耳にしよう環境音についても、周波数スペクトル等の音響パラメータを操作することによって、印象評価を変えることができるのかを調査するため、3.3節で変換した環境音データを用いて、知覚される情動の変化との関係性を評価する検証実験を行った。

4.1 仮説

先行研究より、単一音源の環境音について、より高いピッチが、Russellの円環モデルにおいてより快、より覚醒した情動と関連していることが見出されていたため、本実験で用いた、より定常的な環境音についても、高音を増幅すると、より快、より覚醒していると知覚され、逆に高音を減衰させた場合、より不快、より眠気があると知覚されると考えた。また、ピンクノイズによりうるささに関する心理的印象が低減されるため、ピンクノイズを付加した場合は、「居心地が良い」「快適だ」と知覚されると考えた。リバーブを付加した場合は、先行研究で無意味雑音について残響が極端に長い場合は不快感が増しているため、より不快だと知覚されると考えた。

4.2 評価方法

変換された環境音に対する印象評価は、Schefféの対比較法を用いた質問紙によって行った。実験参加者は10名(20代前半9名, 30代前半1名, 男性9名, 女性1名)であった。

評価項目として、Russellの円環モデルの快-不快・覚醒-眠気の2軸から「快い」「覚醒した」を用いたほか、室内背景騒音や建築光環境の印象評価の先行研究[14][15]で用いられていた形容詞対のうちから、Russellの円環モデル上に対応する形容詞がないものとして、「居心地の良い」「違和感のない」「高級感のある」を追加し5項目とした。「違和感のない」は、変換された音声を聞いてもらうため、不自然さを問うため採用した。

1個の音源につき、原音と各効果を施した合計9個の音声データについて、対比較法によって評価した。1回の質問につき、音声AとBが表示され、各尺度につき、3(Aの方がとてもそうである)~0(同程度だ)~3(Bの方がとてもそうである)の7段階で評価した。

音声の再生と質問紙の実装は、オンラインでWebページに実装した。Webページを開いた参加者は、静かな部屋で、PCまたはスマートフォンに、イヤホンまたはヘッドホンを接続し、外部の音が聞こえない環境で答えることが求められた。性別と年齢、使用環境について回答し、「快い」と「覚醒した」の定義が示された後、印象評価を開始した。「快い」はポジティブで楽しい知覚を表し、その反対は、悲しい、不快な知覚を表すこと、「覚醒した」は、高活動でエネルギーであることを表し、その反対語の例には平穩、退屈、静か、眠気があることが伝えられた。質問紙は、図2のようなWebページで、HTML5 audio playerオブジェクトを通して音声を再生できるようになっている。参加者は音声を繰り返し再生することができる。2個の音声データを再生した後、5個の評価尺度についてそれぞれ7段階の選択肢から一つを選び、送信ボタンを押すと、次の質問が表示される。

2個の音源を用意したため、参加者は1個の音源につき36回、合計72回の比較評価が求められた。



図2: 質問紙のWebページ

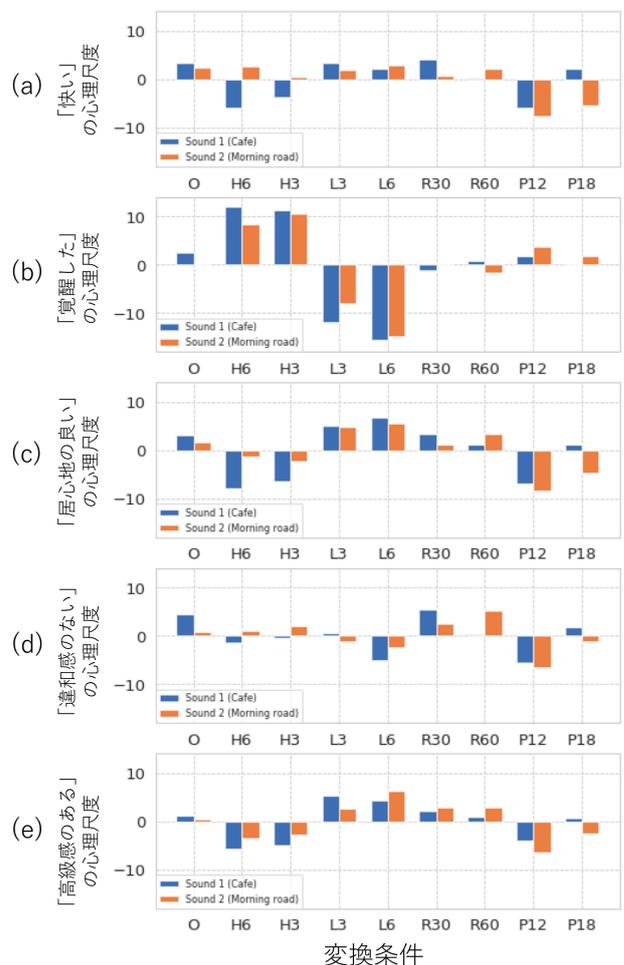


図3: 評価項目別の、各変換を施した環境音の、評価値から求めた心理尺度 (Sound 1: 「喫茶店」、Sound 2: 「朝の雑踏」、各変換条件は3.3節を参照)

4.3 実験結果

評価値を元に回帰分析を行った結果、評価項目別の、各変換を施した環境音の、評価値から求めた心理尺度は図3のようになった。

#### 4.4 考察

図 3(b) より, 高音を増幅するイコライザを施した変換条件 H6, H3 について, 先行研究 [4] と同様に, より覚醒していると知覚される傾向が示唆された. 逆に高音を弱くした L3, L6 について, より覚醒していないと知覚される傾向が示唆された.

しかし快-不快については先行研究 [4] と異なり, 音源依存性が見られた. 図 3(a), (c) で変換条件 H6 と H3 について Sound 1 と 2 の間で比較すると, 「喫茶店」の環境音は, 「朝の雑踏」と比較して, 高音増幅時, 不快, 居心地が悪いと知覚される傾向が示唆された. 高音を増幅して不快と知覚される結果は, 先行研究 [4] において環境音のピッチを高くした場合からの推論とは反対であるが, 「喫茶店」の音源では, 単一音源の環境音とは異なり, 人々の話し声が絶え間なく続いていたために, うるさいと知覚されたことに起因した, 音源依存の結果であった可能性がある.

また図 3(e) より, 低音が強い音のほうが高級感があると知覚される傾向が示唆された.

図 3(a), (c), (d), (e) の変換条件 P12 の項目に着目すると, S/N 比 12dB のピンクノイズを付加した音については, 不快, 居心地の悪い, 違和感のある, 高級感のないといった傾向が示唆された.

図 3(a)~(e) を通し, 変換条件 R30, R60 の項目より, 残響を付加した場合の心理尺度は比較的, 原音との差が小さく, 効果の量が不十分であった可能性がある.

#### 5. まとめと今後の展望

本研究では, 実空間で収録された環境音に対し, 3 種類の変換手法により音響パラメータを操作し, 知覚される情動の差について評価した. 音源として「喫茶店」と「朝の雑踏」の 2 個の環境音の音声データを用意し, それぞれにイコライザをかける, リバーブを付加する, またはピンクノイズを付加するという 3 種類の変換を施した. 2 個の音源についてそれぞれ, 同一音源の変換されたデータ内で比較し, 知覚される情動について 5 個の評価語を用い, 主観評価を行った.

その結果, 高音を増幅するイコライザを施した場合, より覚醒していると知覚される傾向が示唆された一方, 快さは, 音源が「喫茶店」の場合に, より不快であると知覚される傾向が見られ, 快さの知覚には音源依存性があることが示唆された. また低音増幅により高級感が高いと知覚される傾向が示唆された.

環境音の変換手法と知覚される情動の変化の関係が明らかになれば, 実空間において自在に環境音に対する印象を制御することが可能となり, 個人の嗜好や心理状態と空間の文脈の両方を考慮した雰囲気への制御への応用が期待できる.

今後はより精密な検証と, 異なる音源間での比較を行い, 2 個の音源から知覚される情動を逆転させることができるかを含め検証する. そして実空間において環境音を遮断せず, 空間の文脈を保持したまま環境音を補正することで, 違和

感なく環境音に対する印象を制御する手法の構築を目指す.

#### 参考文献

- [1] Västfjäll, Daniel: Emotion induction through music: A review of the musical mood induction procedure, *Musicae Scientiae* 5.1\_suppl, pp. 173–211, 2001.
- [2] 川井敬二 他: 環境音の印象評価構造に関する研究: 被験者自身の言葉に基づいた評価構造の抽出, *日本音響学会誌* Vol. 60.5, pp.249–257, 2004.
- [3] Schafer, R. Murray: *The soundscape: Our sonic environment and the tuning of the world*, Simon and Schuster, 1993.
- [4] Weiyi Ma, William Forde Thompson: Human emotions track changes in the acoustic environment, *Proceedings of the National Academy of Sciences*, Vol. 112.47, pp.14563–14568, 2015.
- [5] Russell, J. A.: A circumplex model of affect., *Journal of personality and social psychology*, Vol. 39, No. 6, pp. 11–61, 1980.
- [6] Jianyu Fan, Miles Thorogood, and Philippe Pasquier: Emo-soundscapes: A dataset for soundscape emotion recognition, 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE, 2017.
- [7] Jianyu Fan, Miles Thorogood, and Philippe Pasquier: Automatic soundscape affect recognition using a dimensional approach, *Journal of the Audio Engineering Society* Vol. 64.9, pp. 646–653, 2016.
- [8] <http://www.audiometaphor.ca/> (最終閲覧日: 2019年7月18日)
- [9] 小森平: 街の環境音 1 (日常系) ~ フリー効果音・無料効果音素材, <https://taira-komori.jp/environment01.html> (最終閲覧日: 2019年7月18日)
- [10] 泉山中三: アメニティの視点からみた音響の環境利用について, *研究紀要* Vol. 13, pp. 1–17, 1989.
- [11] 為末隆弘 他: 定常および変動音を用いたマスキング効果によるうるささの低減, *電子情報通信学会技術研究報告*. EA, 応用音響 Vol. 103.471, pp. 13–18, 2003.
- [12] 豊川勝生 他: ファジィ AHP 法による残響特性の快適感の評価, *ランドスケープ研究* Vol. 59.5, pp. 193–196, 1995.
- [13] <https://www.audacityteam.org/> (最終閲覧日: 2019年7月18日)
- [14] 中村 寛: 室内背景騒音の聴感印象に関する実験的検討, *東京大学工学部 建築学科卒業論文梗概集*, 2003.
- [15] 小椋和彦: 建築光環境の空間的評価指標に関する研究.