

Title	音質評価指標を用いた音の快・不快の予測モデルの構築
Author(s)	松尾, 洋杜
Citation	
Issue Date	2026-03
Type	Thesis or Dissertation
Text version	author
URL	https://hdl.handle.net/10119/20528
Rights	
Description	Supervisor: 鷗木 祐史, 先端科学技術研究科, 修士(情報科学)

修士論文

音質評価指標を用いた音の快・不快の予測モデルの構築

松尾 洋杜

主指導教員 鵜木 祐史

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和8年3月

Abstract

Perceived impressions of sounds are determined not only by physical magnitude but also by sensory attributes. Even when sound levels are similar, impressions may differ due to temporal fluctuation, spectral balance, and tonal components. For practical applications including product sound quality design and assessment of acoustic environments, it is therefore important to predict pleasantness/unpleasantness from interpretable descriptors. Although data-driven approaches can yield accurate predictions, the relationship between acoustic factors and perceptual outcomes may be difficult to interpret, which limits their direct use for sound modification or design decisions. In contrast, psychoacoustically motivated sound quality metrics (SQMs) provide interpretable features that are closely related to auditory perception and can be linked to concrete modification strategies. The objective of this thesis is to construct an interpretable prediction model for sensory pleasantness and unpleasantness by integrating representative SQMs, and to evaluate whether the proposed framework is valid not only for controlled stimuli but also for unknown real environmental sounds.

This study was conducted as a continuation of prior work, and was organized around four research questions. First, it was examined whether sensory pleasantness and sensory unpleasantness can be represented on a single perceptual axis, rather than being treated as independent dimensions. Second, it was examined whether the stimulus conditions can be designed to cover a sufficiently wide perceptual range for discussing pleasantness/unpleasantness in a stable manner. Third, it was examined whether an SQM-based model can explain subjective evaluation results in an interpretable way. Fourth, it was examined whether the proposed model remains effective for unknown sounds, including real environmental sounds that are generally nonstationary and acoustically complex.

To address these questions, the study proceeded in three compact steps. First, previously obtained subjective data were analyzed to clarify the relationship between pleasantness and unpleasantness and to motivate unified modeling. Second, a listening experiment was conducted using systematically designed artificial stimuli. Sensory pleasantness was evaluated by a paired-comparison procedure based on a classical psychometric framework, and continuous perceptual scale values were derived from aggregated paired judgments. The stimulus set was constructed so that impressions would span from pleasant to unpleasant, including intermediate conditions, which is necessary for stable modeling and meaningful interpre-

tation. Third, an interpretable prediction model was formulated and evaluated. Five SQMs widely used in psychoacoustic modeling were integrated as explanatory variables: roughness, sharpness, tonality, loudness, and fluctuation strength. Following a conventional psychoacoustic modeling approach, an exponential-form representation was adopted so that the contribution of each metric is explicitly expressed through model parameters while allowing nonlinear sensitivity when required. Model parameters were estimated so that the model output reproduces the tendency of the subjective scale values across the designed stimulus conditions.

Model validity was examined from two viewpoints. For controlled stimuli, it was confirmed whether the proposed model can explain the subjective evaluation trend over the full range of designed conditions, and whether the fitted parameters provide a consistent interpretation of how each SQM contributes to pleasantness/unpleasantness. For generalization, the proposed model was applied to real environmental sounds that were not used during model construction, and model predictions were compared with subjective ratings obtained in an additional listening test. In addition, representative conventional models designed for separate targets, such as pleasantness-oriented or annoyance-oriented prediction, were used as references to examine the benefit of a unified SQM-based formulation on a common evaluation target. Furthermore, robustness was examined by computing SQMs under an alternative auditory filtering framework based on an ERB-related filterbank, because the numerical values of SQMs can depend on the auditory preprocessing model.

Overall, the investigations supported a unified treatment of sensory pleasantness and sensory unpleasantness and demonstrated that an interpretable SQM-based model can be constructed to explain subjective evaluation results and to provide meaningful predictions for unknown sounds. The proposed model is expected to be useful as a basis for future work toward practical sound modification and design, where perceptual improvements can be discussed in terms of specific psychoacoustic attributes. The results showed that,

- (1) Sensory pleasantness and sensory unpleasantness were supported to form a paired relationship on a common perceptual axis.
- (2) The stimulus conditions were designed to provide a sufficiently wide range for pleasantness/unpleasantness evaluation.
- (3) An interpretable prediction model was constructed by integrating five SQMs

and was shown to explain subjective results.

- (4) The proposed model was shown to remain effective for unknown sounds, including real environmental sounds, and to be robust to alternative SQM computation.

目次

第1章 序論	2
1.1 はじめに	2
1.2 本研究の背景	5
1.3 本研究の目的	6
1.4 本論文の構成	6
第2章 関連研究	9
2.1 機械学習を用いたアプローチ	9
2.2 物理量を用いた基礎的アプローチ	10
2.3 聴覚心理学に基づくアプローチ	11
第3章 本研究の方略	14
第4章 音の感覚的快さ・快くなさの関係	17
4.1 音の感覚的快さ	17
4.2 音の感覚的快くなさ	18
4.3 感覚的快さ・快くなさの関係	21
4.4 感覚的快さ・快くなさの一元評価	24
4.5 感覚的快さ・快くなさと快・不快の対応	28
第5章 音の快・不快を予測するための感覚的快さの主観評価実験	29
5.1 実験刺激	29
5.2 実験参加者	31
5.3 実験機器	31
5.4 実験方法	31
5.5 実験結果	35
5.6 考察	37
第6章 音の快・不快の予測モデル	38
6.1 提案法	38
6.2 モデル構築	38

6.3 各音質評価指標の寄与度	44
第7章 提案モデルの評価	47
7.1 音の快・不快の予測モデルの評価	47
7.2 感覚的快さの予測モデルとの比較評価	54
7.3 アノイアンスモデルとの比較評価	56
7.4 考察	59
第8章 結論	65
8.1 明らかにしたこと	65
8.2 残された課題	66
付録A 感覚的快くなさの推定	67
付録B 音の感覚的快さ・快くなさの予測モデルによる実環境音の予測	71
謝辞	73
参考文献	74
研究業績	80

図目次

1.1	身の回りの音の快・不快	4
1.2	本論文の構成	8
3.1	本研究の方略	16
4.1	感覚的快さの主観評価結果	20
4.2	感覚的快くなさの主観評価結果	20
4.3	感覚的快さと感覚的快くなさの関係	23
4.4	音質評価指標を利用した回帰モデルによる感覚的快さ・快くなさの 一元評価結果. 各プロットは, 図 5.4 のプロットおよび凡例の刺激 に対応する.	27
5.1	実験系の構成図	32
5.2	GUI	33
5.3	刺激配置	34
5.4	Table 5.1 の実験刺激に対する主観評価結果 (便宜上, 図 5.4 では, ラウドネス 21 sone の条件のみ塗りつぶし)	36
6.1	音の快・不快の予測モデルの構成図	42
6.2	音の快・不快の予測モデルによる快・不快の推定結果. 各プロット の色, 太さ, 記号は, 図 5.4 の凡例の刺激に対応する.	43
6.3	音の快・不快モデルによる音の快・不快の推定結果:(a) すべての SQMs を利用した場合, (b) ラフネスを利用しなかった場合, (c) シャープ ネスを利用しなかった場合, (d) トナリティを利用しなかった場合, (e) ラウドネスを利用しなかった場合, (f) 変動強度を利用しなかつ た場合の推定結果. 各プロットは, 図 5.4 のプロットおよび凡例の 刺激に対応する.	46
7.1	GUI	51
7.2	実環境音に対する提案モデルの予測値と主観評価値の関係	53
7.3	実環境音に対する感覚的快さの予測値と主観評価値の関係	55

7.4	実環境音に対するアノイアンスモデルの予測値と主観評価値の関係	58
7.5	実環境音に対する予測値（GTFB 利用）と主観評価値の関係	63
7.6	実環境音に対する予測値（GCFB 利用）と主観評価値の関係	64
A.1	音の感覚的快くなさのモデルによる感覚的快くなさの推定結果：(a) 全ての SQMs を利用した場合, (b) ラフネスを利用しなかった場合, (c) シャープネスを利用しなかった場合, (d) トナリティを利用しなかった場合, (e) ラウドネスを利用しなかった場合, (f) ラウドネスとラフネスのみ利用した場合. 各プロットは, 図 5.4 のプロットおよび凡例の刺激に対応する.	70
B.1	実環境音に対する感覚的快さ・快くなさの予測モデルとの比較評価	72

表 目 次

4.1	実験刺激の条件	19
4.2	感覚的快さ・快くなさの予測モデル（式 (4.4)）の各係数	26
5.1	実験刺激の条件	30
6.1	提案モデル（式 (6.1)）の各係数	41
7.1	音の快・不快の推定に利用した実環境音	49
7.2	提案モデル（式 (6.1)）における GTFB を利用した場合の各係数 . .	61
7.3	提案モデル（式 (6.1)）における GCFB を利用した場合の各係数 . .	62
A.1	Aures のモデルに基づく感覚的快くなさの予測モデル (式 (A.4)) の 各係数	69

略語	正式名称
CB	臨界帯域
Aures モデル	Aures の感覚的快さの予測モデル
アノイアンスモデル	Fastl & Zwicker のアノイアンスモデル
PT	純音
BN	帯域雑音
RMSE	二乗平均平方誤差
RMSE	二乗平均平方根誤差
Bark	臨界帯域幅尺度
ERB	等価矩形帯域幅
GTFB	ガンマトーン聴覚フィルタバンク
GCFB	ガンマチャープ聴覚フィルタバンク

変数名	定義
N	ラウドネス
S	シャープネス
R	ラフネス
K	トナリティ
FS	変動強度

第1章 序論

1.1 はじめに

音の快・不快は、誰しもが経験している感覚である。図 1.1 に示すように、我々は様々な音と共に生きている。また、その音に対する快や不快の知覚は、我々の生活に大きく影響を及ぼしている。例えば、図 1.1 の右側に示すように、鳥のさえずりや川のせせらぎ、優しい雨の音などの自然の心地よい音に癒されたることから、空間の居心地の良さや安心感を高め、製品やサービスの価値向上にも繋がりうる。図 1.1 の左側に示すように、工事現場の音や車両のアイドリング音、エンジン音などの不快な騒音により睡眠や集中が妨げられ、疲労やストレスを招き、生活の質を損なう要因となる。

このように、音の知覚は、我々の感情や行動と密接に関係している [1-3]。そのため、我々の日常生活において、快適な音の聴取のための音の設計や制御は非常に重要な課題である。具体例として、音のデザインでは好みで心地よい音をどのように生成するか、環境騒音の能動制御では不快で聞きたくない音をどのように抑制するか、などが挙げられる。一方で、我々はどのように音の快・不快を感じているのだろうか。この知覚の仕組みは、聴覚の構造や物理量を用いた研究をはじめとし、さらに聴覚心理学に基づく研究分野でも研究が進められている。

騒音評価では、等価騒音レベルと呼ばれる物理量が利用されている [4,5]。この物理量は、騒音に対する主観的なうるささに概ね対応し、騒音に対する心理的反応を見積もるのに有効であり、不快感の評価にも利用されている。一方で、等価騒音レベルが同じ値であったとしても、印象が異なることがある [6]。したがって、音の印象評価では、量的評価だけではなく質的評価を利用した音質評価指標が提案されており、ヒトの音色の感じ方を数値で計算できる [7-12]。音質評価指標を利用した評価として、例えば、家電製品では音の快適評価 [13] が、騒音では不快評価 [14-16] がある。以下に、両者で利用されるすべての音質評価指標について簡単にまとめる。

- ラウドネス： 主観的な音の大きさ
- ラフネス： 主観的な音の粗さ

- シャープネス：主観的な音の甲高さ
- トナリティ：主観的な音の純音感
- 変動強度：主観的な音の変動感の強さ

しかし、従来の研究では、音の快と不快は、独立した関係にあるのか、あるいは対極の関係にあるのかわかっていなかった。そのため、一般的に快と不快を一軸で扱う方法が確立されておらず、これらの快や不快を評価するモデルを音の不快感の見積りや不快音の快音化といった試みに利用できるかわからない。また、もしも、音の快・不快の両方を同じ軸上で表すことが可能であれば、単一のモデルでの快・不快評価を実現できる可能性がある。

単一のモデルで快・不快が対極にある評価軸を用いて評価すること（以後、一元評価と呼ぶ）ができれば、我々の生活を豊かにするための幅広い応用が期待される。例えば、物理量から見積もられる音質評価指標と快・不快を結びつけて評価することで、まるでリモコンを操作するかのように環境音に対して不快感を低減する能動的な制御や、嗜好に合わせた音のデザインなどが容易になると考えられる。具体的には、観測された音信号の音質評価指標を推定し、これらの分析により入力音の不快さの増大に関与する音質評価指標を特定する。この音質評価指標に関わる物理量を抑制する処理、あるいは快に関わる音響特徴量の増幅を行うことで快音化が実現される。これは、音を抑制する従来の騒音制御などとは異なり、不快音の快音化（または、音の不快音化）に繋がるという点で、新規性および、優位性を有する。

音に対する注意の向きやすさが不快さと関係している [17]。聞き漏らしを避ける必要がある Jアラートのような避難誘導音をはじめとする報知音に対して、音の不快音化を行うことで注意の向きやすさが向上するため、より効果的な報知音の設計への応用が期待される。

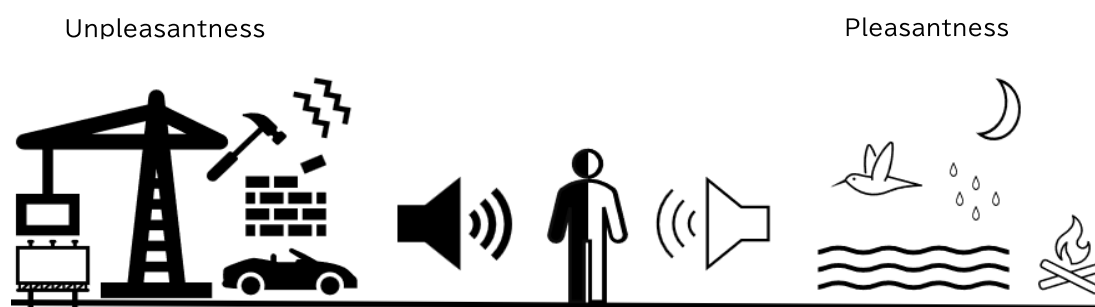


図 1.1: 身の周りの音の快・不快

1.2 本研究の背景

近年、快・不快を評価するために、機械学習を用いたアプローチがある。大量のデータセットを用いて訓練データとすることで、波形やスペクトログラムなどの特徴量から快や不快、または関連する特徴を直接推定することが試みられている。一方で、この方法は end-to-end の形であるため、評価の根拠がブラックボックスとなっている。音の制御やデザインを行う上でこのような情報は重要であるにも関わらず、評価の根拠がわからなくては、音響特徴量と快・不快を結び付けて議論することが難しい。本研究では、前節で述べたように、入出力だけでなく聴覚の仕組みに基づく中身で行われている音響特徴に関わる情報を操作することでの快音化を想定している。そのため、ヒトの聴知覚について、音響特徴と快・不快がどのように結びついているかを想定して検討を進める必要がある。また、評価の理由が分かることは、音デザインにおいて、考え方を整理するうえでも大切である。さらに、どの特徴がどの印象に効くのかが見えれば、その調整を行うときの手がかりにもなる。

聴覚心理に基づく音の知覚における印象評価では、音の快を評価する Aures の感覚的快さの予測モデル [13, 18]、音の不快を評価する Fastl & Zwicker のアノイアンスモデル（[14] の第 16 章 “Examples of Application”）がある。これらのモデルは、音の物理量から音質評価指標 [7] を見積もり、音の評価に利用する。また、両モデルにおいて、ラウドネス、ラフネス、シャープネスの三つの音質評価指標が共通して利用され、感覚的快さの予測モデルではトナリティが、アノイアンスモデルでは、変動強度が追加で利用されている。両モデルでは、快と不快に関わる音質評価指標が異なっていることがわかる。

先行研究 [19] では、音の快・不快を評価するための最初の試みとして、Aures の感覚的快さの予測モデルで利用される四つの音質評価指標（ラウドネス、ラフネス、シャープネス、トナリティ）を考慮して作成された様々な人口音の同一刺激に対して、感覚的快さと感覚的快くなさの主観評価実験を行った。また、その関係に対して、Aures の感覚的快さの予測モデルの係数を調整することで感覚的快さの結果を説明できること、定性的に相関があることを明らかにした。この結果は、本研究の基盤となる重要な知見である。しかし、定量的な議論までは実施できておらず、積み残し課題となっていた。

そこで、音の快・不快を一元的に評価するための予備的検討として、3.1 章では快・不快の関係を明らかにする。先行研究で取得した感覚的快さと感覚的快くなさの主観評価結果に対して、相関分析、および統計的検定を行った [20]。その結果として、両者の関係を同じ軸上で表せることを明らかにした。しかし、先の検

討 [19] [20] で利用した刺激条件の範囲は、快・不快の評価を行うためには十分ではなかった。

本研究ではこの問題点を考慮した上で、感覚的快さの予測モデルおよび、アノイアンスモデルで利用される五つの音質評価指標（ラウドネス、ラフネス、シャープネス、トナリティ、変動強度）をすべて考慮することで、音の快・不快の予測モデルの構築を試みる。

実際の環境音は時間的に変化する非定常音であり、先行研究で用いられた人工の定常音とは性質が異なる。そのため本研究では、提案したモデルに対して、汎化性能の観点から人工的に作成した音だけでなく、モデルにとって未知の実環境音にも適用し、モデルがどの程度有効であるかについても確かめる。未知の音に対しても同じように評価できれば、提案したモデルが実際の環境音に対しても課題解決に利用できる実現可能性が示される。

1.3 本研究の目的

本研究の目的は、音質評価指標を用いて音の快・不快を評価する予測モデルを構築することである。音質評価指標を考慮した快・不快に関する聴取実験を行い、得られた評価結果を精度よく説明できる予測モデルを提案する。特に、感覚的快さの予測モデル [13] とアノイアンスモデル [14] で利用される 5 つの音質評価指標（ラウドネス、シャープネス、ラフネス、変動強度、トナリティ）で構成される予測モデルに対して、モデル構築に用いた評価値に対する当てはまりを確認し、客観評価を行う。さらに、汎化性能も考慮し、提案モデルに対して未知の音を入力した場合に、その予測結果が同一音に対するヒトの評価結果をどの程度説明できるかを検証することで、主観評価を行う。本研究では、以上によって、音の快・不快の予測モデルの構築を行う。

1.4 本論文の構成

本論文は、全 8 章である。図 1.2 に本論文の構成を示す。

第 1 章： 快・不快の評価の背景，その応用方法について述べる。また，本研究の目的について述べる。

第 2 章： 関連研究について述べる。

第 3 章： 本研究の方略として，目的に対して設けた四つのリサーチクエスションについて述べる。

第4章： 先行研究 [40] で得られた主観評価値を利用して，感覚的快さ・快くなさの関係について述べる．

第5章： 前章で明らかにした知見を踏まえ，感覚的快さの再評価を行う．

第6章： 音の快・不快の予測モデルを提案する．また，提案モデルが前節の評価結果を説明できるか検証することで，客観評価を行う．

第7章： 提案モデルが未知の音に対する快・不快の評価に有効であるか検証し，考察を行う．

第8章： 結論と今後の展望を述べる．

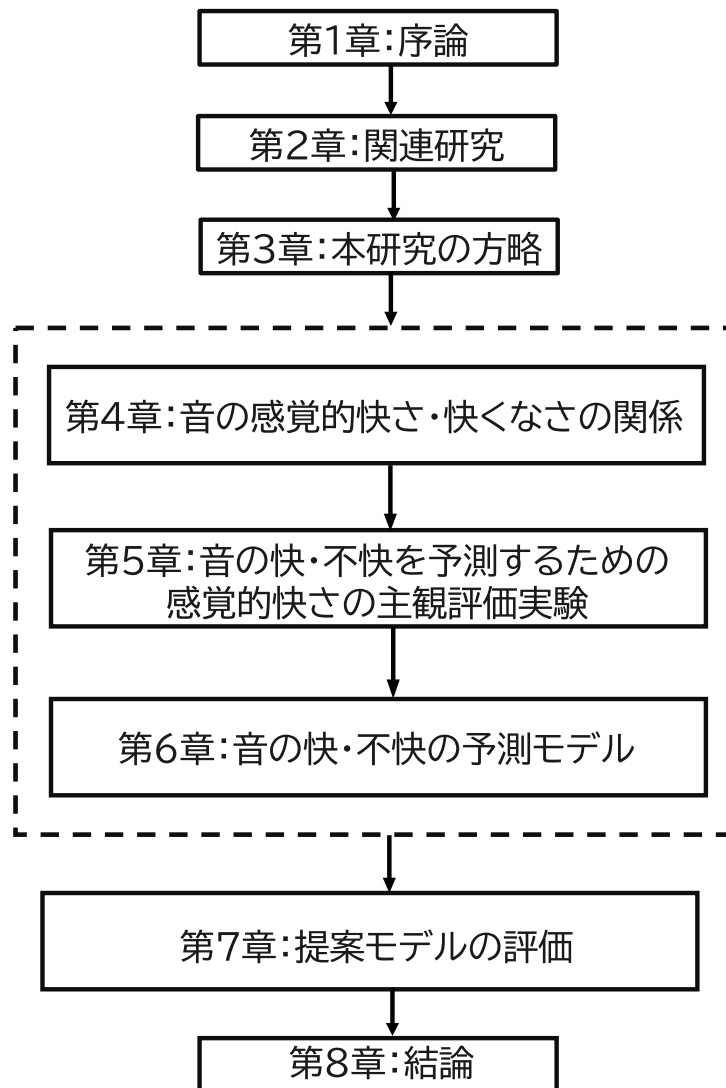


図 1.2: 本論文の構成

第2章 関連研究

2.1 機械学習を用いたアプローチ

音の快・不快を、波形やスペクトログラムなどの特徴量から直接推定する方法として、機械学習が広く用いられている。例えば、道路交通騒音の不快感に対して、主観評価に基づく推定を行う研究が報告されている [21]。Wang らは、主観評価を得る手法としての聴取実験と、客観的に予測する手法を整理した上で、深層学習に基づく客観的な不快感の評価方法を提案し、騒音と不快感の対応関係を主観評価結果に基づいて直接利用することを述べている [21]。また、同研究では、聴取実験のサンプル数が限られる場合に学習が不安定になり得る点について着目し、これを改善する学習方法について検討した [21]。

音環境の印象をデータから学習する試みは、サウンドスケープ研究の分野でも議論されている。Axelsson らは、サウンドスケープ知覚の基本次元を特定することを目的として、多数の属性尺度による評価結果に対して分析方法検討・適用し、pleasantness, eventfulness, familiarity の三成分を得たと報告している [22]。さらに、分析における各属性値と、サウンドスケープから受ける印象について、パス解析によってその関係が議論されている [22]。この検討では、活性度と感覚的快さに関係があることが示された。

加えて、Lionello らは、都市の環境音から受ける印象に対する予測モデルを対象とし、各サウンドスケープモデルにおける、印象の予測や音源の再現を行う機械学習モデルについて整理している [23]。ここでは、線形・非線形システムの性能差は大きくないとされる一方で、指標構造が複雑な場合においては、非線形性を考慮した機械学習モデルのほうがより適する可能性が示されている。

一方で、機械学習を利用した研究では、得られたモデルがどの情報を手がかりに判断しているかが分かりにくい場合がある。特に、快・不快を改善するために音を設計・制御したい場合において、end-to-end の構造で、中身がブラックボックス的なモデルでは、どの音響特徴量に対して操作を行うべきか特定しにくい。この点は、Rudin らによっても論じられており、ブラックボックス的な予測モデルは理解が難しい場合があることを指摘した上で、この分野に固有の制約や知識をモ

デルに組み込むことで、計算過程を追いやすくした解釈可能なモデルを用いる重要性を述べている [24].

2.2 物理量を用いた基礎的アプローチ

環境騒音の評価では、音圧の時間変動をエネルギー平均してまとめる等価騒音レベル $L_{Aeq,T}$ が用いられる [25, 26]. $T = t_2 - t_1$ とし、 $p_A(t)$ を A 特性で重み付けした音圧、 $p_0 = 20 \mu\text{Pa}$ を基準音圧とすると、

$$L_{Aeq,T} = 10 \log_{10} \left(\frac{1}{T} \int_{t_1}^{t_2} \frac{p_A^2(t)}{p_0^2} dt \right) \quad (2.1)$$

で定義される [25, 26]. この指標は騒音における曝露量や監視に利用される一方、音の印象を決める要因を細かく分けて扱うには限界がある.

そこで、製品音などでは、単なる音圧だけでなく音色や時間変化を含めた観点から音質を議論する枠組みが使われてきた. 自動車の分野では、車内音の音質改善研究が進み、デジタル信号処理と計算機処理を用いた評価システムが整備され、現実音に処理を加えたときその印象評価がどう変わるかを素早く調べられるようになってきたことが述べられている [27]. また、A 特性音圧レベルだけでは同一音に対する主観評価結果と対応しにくい例が多く、音の大きさについてはラウドネスレベルの方が良い対応が得られることが示されている [27]. さらに、自動車車内音の異音成分の例としてランブル音を取り上げ、時間波形が振幅変調音になっている場合に、時間包絡の波形に含まれている情報と主観評価結果の対応が良いことを述べている [28]. このように、単一のレベル指標だけでなく、時間変動の形そのものが印象に関係する. また、振幅変調により生じる不快な感覚を粗さとして説明し、滑らかな音は快く、滑らかでない音は粗く不快に感じられるため、粗さが自動車の音質の快適性を劣化させる要因となることを述べている [28]. SD 法と因子分析の例では、第 1 因子が快適性を表し、その尺度の中に粗い／滑らかといった粗さ感が含まれることを示している [28]. 以上は、快・不快の評価において、音の印象が重要であることを示唆する.

不快感を下げるための手段としては、遮音や吸音に加えて能動騒音制御 (Active Noise Control : ANC) がある. ただし、音圧を下げるだけでは不快感が十分に下らない場合があるため、聴覚マスキングを利用して聞こえ方を変える考え方もある. 西浦らは、工場騒音が作業員に強い不快感を与えること、工場騒音のスペクトルが低周波数帯域が支配的である一方で、高周波数帯域に鋭いピークを持つ場合があることを明らかにした [29]. また、耳栓などの受動的手法は音声伝達や本来聴きたい音を阻害する問題があるため、音声を阻害しない不快感低減を目的

として、能動騒音制御と聴覚マスキングの統合的方法を提案している [29].

提案方法として、参照信号 $u(t)$ に対し、聴覚マスキング由来の制御音と ANC の制御音を加算して最終的な制御音 $y(t)$ を生成し、

$$y(t) = \beta y_{\text{mask}}(t) + h(t) * u(t) \quad (2.2)$$

の形で表している (β は音圧調整係数, $h(t)$ は低域通過適応フィルタ, $*$ は畳み込み) [29]. さらに、聴覚マスキング側では入力信号から最も突出したピーク周波数を検出し、その周波数に対応する制御音源をデータベースから選ぶ流れを示している. また、低周波数帯域には音声も含まれるため ANC の制御帯域によっては音声も抑圧されうること、高周波数側のピーク周波数は作業条件で変化しうるためマスキングの制御帯域も変化し、それが不快感や音声了解性に影響することを述べている. このため、ANC と聴覚マスキングの制御帯域を変化させた条件で評価実験を行い、不快感低減量と音声了解性の両面からパラメータの検証を行った [29].

以上より、音の物理量は実用上の基盤とされており、制御は低減の手段として重要であるが、上述のように印象の改善を行うには、物理量だけでなく、さらに音色知覚と結びつけて扱う必要がある.

2.3 聴覚心理学に基づくアプローチ

ヒトの知覚は、刺激強度 I に対して感覚量 S が対数的に増えるという（ウェーバー・フェヒナーの法則）[30] が、以下の形で表される.

$$S = k \log \left(\frac{I}{I_0} \right) \quad (2.3)$$

音圧レベルが対数で定義されることも整合的であり、聴覚の再現では、非線形性を踏まえた検討が重要であると考えられる.

聴覚は周波数ごとに帯域分割して音进行处理するため、周波数軸を臨界帯域 [31] で表す考え方がある [32, 33]. Zwicker は、臨界帯域に基づき、可聴帯域を臨界帯域に分割し臨界帯域番号 (Bark 尺度) を [14] の第 6 章 “Critical Bands and Excitation” で提案している. 一方で、聴覚フィルタの帯域幅を等価矩形帯域幅で表す ERB 尺度は、聴覚フィルタ推定の心理音響実験に基づいて検討されており、聴覚フィルタ形状の非対称性を考慮できる点が特徴である [34–36]. 中心周波数 f (Hz) に対

する ERB-number は,

$$\text{ERBN}(f) = 21.4 \log_{10} \left(4.37 \frac{f}{1000} + 1 \right) \quad (2.4)$$

の形で表される．従来の研究では，Bark 尺度よりも ERB 尺度のほうが心理音響実験に基づくことに加えて，生理学の研究 [37, 38] ととも良い対応があることから，心理学的・生理学的の両面で有効であることが明らかになっている [36]．

ラウドネス，シャープネス，ラフネス，変動強度などの音色知覚の関係を物理量と結びつけて計算する音質評価指標は，聴覚フィルタバンクを介して計算される．磯山らは，聴覚フィルタバンクの計算に，従来モデルが用いてきた Bark 尺度よりも，ERB 尺度を利用したほうがヒトの知覚を正確に再現できるという考えのもと，入力信号を定常状態とした上で ERB 尺度に基づく時間領域あるいは，周波数領域の聴覚フィルタバンク（GTFB/GCFB）を利用したラウドネス，シャープネス，ラフネス，変動強度を計算するモデルを構築した [39]．

聴覚心理学に基づく予測モデルである，音の快を評価する Aures の感覚的快さの予測モデル [13, 18]，音の不快感を評価する Fastl & Zwicker のアノイアンスモデル [14] について以下に示す．

Aures は，ラフネス，シャープネス，トナリティ，ラウドネスが感覚的快さに与える影響を調査し，これらを説明変数とする式 (2.5) に示す感覚的快さの予測モデルを提案した [13, 18]．

$$W = e^{\alpha_1 R} e^{\alpha_2 S} (\alpha_3 - e^{\alpha_4 K}) e^{-(\alpha_5 N)^2} \quad (2.5)$$

ただし，ただし， R はラフネス（単位は asper）， S はシャープネス（単位は acum）， K はトナリティ（単位は tu）， N はラウドネス（単位は sone）である．重み係数は $\alpha_1 = -0.55$ ， $\alpha_2 = -0.113$ ， $\alpha_3 = 1.24$ ， $\alpha_4 = -2.2$ ， $\alpha_5 = 0.023$ である．

Fastl & Zwicker は，不快感を予測するラウドネス，シャープネス，変動強度，ラフネスに基づくアノイアンスモデルとして式 (2.6) に示す psychoacoustic annoyance (PA) を定義した [14]．

$$\text{PA} = N_5 \left(1 + \sqrt{w_S^2 + w_{FR}^2} \right) \quad (2.6)$$

$$w_S = \begin{cases} 0.25 (S - 1.75) \log_{10}(N_5 + 10) & S > 1.75 \\ 0 & S \leq 1.75 \end{cases} \quad (2.7)$$

$$w_{FR} = \frac{2.18}{N_5^{0.4}} (0.4FS + 0.6R) \quad (2.8)$$

ここで、 N_5 はラウドネスの 95 パーセンタイル値であり、式 (2.6) は式 (2.7)、式 (2.8) に基づいて計算される。ただし、 FS は変動強度（単位は vacil）であり、その他の記号は式 (2.5) と同義である。

さらに、音に対する注意の向きやすさである聴覚的顕著性の関係では、聴覚顕著性と音の大きさの関係間に良い相関が確認されたことが報告されており、これに対応して生理学的反応の変化も大きくなることが報告されている [17]。

第3章 本研究の方略

本研究は、先行研究 [19,40] を包含した上で、その継続課題として実施する。本研究の目的に対して、図 3.1 に示すように検討を行う。図 3.1 のピラミッド型は、下段の検討項目が基盤となって上位の検討が成立することを示しており、各項目を段階的に積み上げることで、最終的に提案モデルの有効性まで検証する。また、これに紐づき、(1)~(4) の大きく四つに分類したリサーチクエスチョンを順序立てて検討することで、目的の達成を試みる。以下に、その手順を示す。

(1) 音の感覚的快さと感覚的快くなさは同じ軸上で表すことができるのか

第4章では、音の感覚的快さと感覚的快くなさが同じ軸上にあるかを定量的に分析する。単一のモデルで音の快・不快を予測するためには、両者が独立しておらず同じ軸上にあることを示す必要があるためである。この、分析が本研究の土台となる。先行研究 [40] で得られた感覚的快さと感覚的快くなさに関する主観評価値を利用してこれを分析する。

ここで、感覚的快さ・快くなさを同じ軸上に表現できた場合、その関係を快・不快とみなす。しかし、利用した先行研究 [40] の刺激条件の範囲は、快・不快の評価を行うためには十分ではなかったことがわかった [20] ため、次の検討に進む。

(2) 音の快・不快の評価のために十分な刺激条件の範囲を設けられたのか

第5章では、上述の問題点を踏まえ先行研究 [40] の刺激条件の範囲は包含した上で、感覚的快さの再評価を行う。さらに、快・不快の評価のために十分な刺激条件の範囲であったかを検証する。ここでは前章で明らかにした同じ軸上に快・不快があることに基づき、感覚的快さに関する実験のみを行う。得られた主観評価値の折り返し部分が不快と仮定し、その結果を快・不快の主観評価値として扱う。

得られた主観評価値に対して、先行研究 [40] の範囲と比較を行い、刺激条件、およびこれに対応する主観評価値に関する問題点が解決したかを確認した上

で次の検討に進む.

(3) 提案モデルは、快・不快の評価結果を説明することができるのか

第6章では、得られた快・不快の主観評価値を利用して音の快・不快の予測モデルを提案し、客観評価を行う。提案モデルの構築は、感覚的快さの予測モデルとアノイアンスモデルで利用される五つの音質評価指標をすべて利用した音質評価指標と係数の指数表現の形式で行う。提案モデルの有効性を確かめるため、快・不快の主観評価値を用いて客観的に評価する。

(4) 提案モデルは未知の音に対しても有効なのか

第7章では、提案モデルの主観評価を行う。提案モデルに未知の音を入力し、同一音に対する主観評価値をどの程度の精度で予測できるかを検証する。また、快と不快を個別に評価する感覚的快さの予測モデル、アノイアンスモデルとの比較を行う。ここでは、快・不快を評価する提案モデルのみ主観評価値を説明できることが期待される。さらに、考察として、提案モデルに対して従来の音質評価指標の計算モデルに比べ、聴覚の再現により優れた磯山ら [39] が構築した音質評価指標モデルを導入しても予測の精度に問題がないかを検討する。

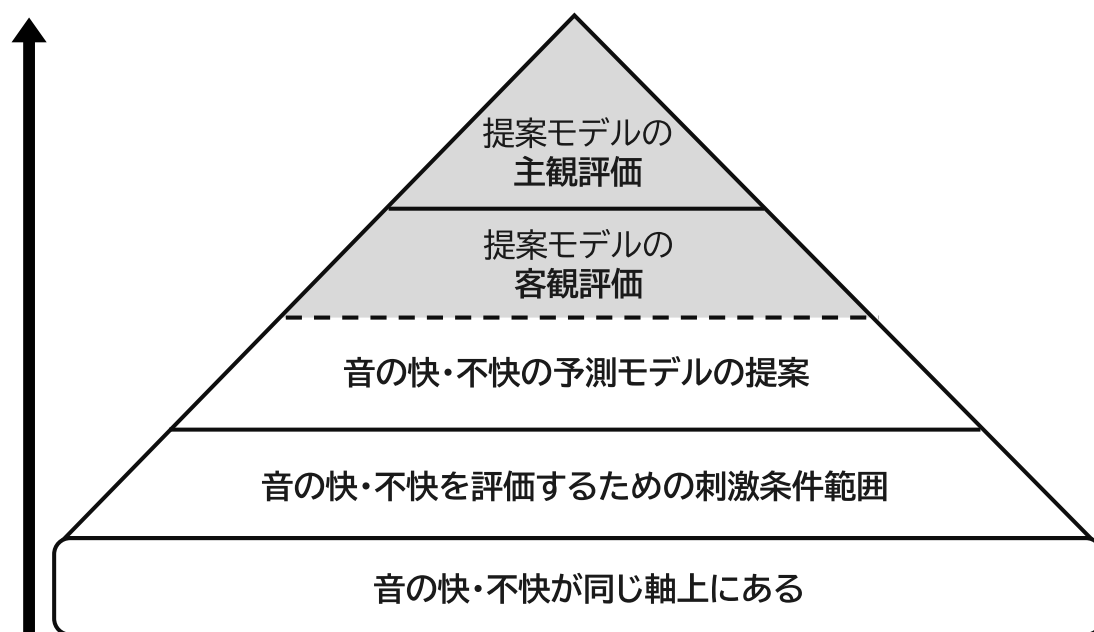


図 3.1: 本研究の方略

第4章 音の感覚的快さ・快くなさの 関係

本章では、谷口が行った感覚的快さと感覚的快くなさの主観評価実験 [40]，およびその結果について説明する [19]．また，そのデータを用いて感覚的快さと感覚的快さの関係を一元的に説明できるかを確認し，快・不快との対応を検討する．

4.1 音の感覚的快さ

谷口 [40] は，感覚的快さを調査するため，Aures の行った実験 [13] に倣い設定した様々な条件の人工音に対して主観評価実験を行った [19]．実験構成，およびその結果は以下のとおりであった．

実験には，正常聴力を有する日本語母語話者 12 名（大学院生，男性 8 名，女性 4 名）が参加した．実験は防音室で行われた．刺激の提示では，PC（Windows11, MATLAB R2023b, Mathworks），ヘッドホンアンプ（audio-technica Stereo Headphone Amplifier AT-HA21），ヘッドホン（SENNHEISER HDA200）を使用した．刺激の立上り・立下りに 100 ms のテーパーをかけた．実験参加者に負担がかからないように，十分に休憩を挟むように実験を行った．

表 4.1 に実験刺激の条件を示す．ここでは，ラウドネスが 2 sone と 14 sone，ラフネスに関わる振幅変調の変調度が 0.0, 0.5, 1.0，トナリティに関わる条件が単一周波数の純音と 1 kHz 帯域幅の帯域雑音，シャープネスに関わる周波数重心として 0.5 kHz, 1 kHz, 2 kHz, 4 kHz, 8 kHz，を想定した 60 種類の音刺激が準備された．音刺激の提示時間が感覚的快さや不快さに影響を与える可能性が指摘されているため，全ての音刺激の時間長は 1 秒に統一された．なお，順序効果を考慮して，快さの実験を行ったあとに，快くなさの実験を行ったグループと，快くなさの実験を行ったあとに快さの実験を行うグループの二つのグループに分けた．

表 4.1 の刺激条件に対して感覚的さの実験が行われた．実験では，刺激の感覚的快さを評価するために，サーストンの一対比較法を利用した．実験参加者には，一つの刺激対として第 1 音，第 2 音の順に提示して，どちらの音が感覚的に快くなく感じたかを選択させた．実験刺激は，第 1 音が 1 秒間再生されたのち，0.8 秒

の休止時間があり、その後第2音が1秒間再生された。第1音および第2音はランダムに決定された。刺激対は合計1770対であった。1日当たり600回の比較を行わせた。1セット150回として、1日4セット実施した。セットごとに10分の休憩を挟んだ。実験は休憩を含めて概ね90分程度を要した。これを3日間実施した。

表4.1の刺激条件に対する感覚的快さの主観評価結果を図4.1に示す。横軸は刺激間の相対距離を示し、正の値に大きくなるほど快いということを表している。縦軸は、実験刺激（純音PTか帯域雑音BN）と条件（3種類の変調度）に合わせて並べている。図中の記号の太さは、ラウドネスが2 sone（細い）と14 sone（太い）に対応する。記号の形は中心周波数に対応し、0.5 kHz：◇、1 kHz：○、2 kHz：□、4 kHz：△、8 kHz：▽とした。記号の色は信号の種類と変調度を表す。信号の種類は、純音：暖色、帯域雑音：寒色で二種類を区別し、変調度は、色の濃さに対応し濃くなるにつれて、0、0.5、1とした。

図4.1から、ラウドネスが大きいときよりも小さいときのほうが、感覚的に快いと判断していることがわかる。また、純音では基本的に変調度が小さいときよりも大きいときのほうが、快いと判断していることがわかるが、純音のラウドネスが2 sone、中心周波数8 kHzの場合はこの関係が逆転していることがわかる。また、主観評価値の値の範囲は、尺度全体の範囲 $[-1.5, 1.5]$ （全体3.0）に対して、最も快いと評価された音源の最大値が1付近、その反対である最小値が -1.4 付近となっていることが確認された。

4.2 音の感覚的快くなさ

谷口は、感覚的快くなさを調査するため前節と同一の音刺激に対して主観評価実験を行った[40]。実験構成は前節と同様であり、結果は以下のとおりであった。

表4.1の刺激条件に対する感覚的快くなさの主観評価結果を図4.2に示す。横軸は刺激間の相対距離を示し、正の値に大きくなるほど快くないということを表している。図のプロットの記号や形、太さは前節と同様である。なお、感覚的快くなさは塗りつぶして表示している。

図4.2から、ラウドネスが小さいときよりも大きいときのほうが、感覚的に快くないと判断していることがわかる。また、純音では基本的に変調度が小さいときよりも大きいときのほうが、快くないと判断していることがわかる。また、主観評価値の値の範囲は、尺度全体の範囲 $[-1.5, 1.5]$ （全体3.0）に対して、最も快くないと評価された音源の最大値が1.2付近、その反対である最小値が -0.8 付近となっていることが確認された。

以上の結果の傾向より、定性的に、感覚的快さと感覚的快くなさは対の関係にあることが示唆された。

表 4.1: 実験刺激の条件

信号の種類	中心周波数 (kHz)	変調周波数 (Hz)	変調度 (MI)	ラウドネス (sone)
純音 (PT)	0.5, 1, 2, 4, 8	70	0, 0.5, 1.0	2, 14
帯域雑音 (BN)	0.5, 1, 2, 4, 8	70	0, 0.5, 1.0	2, 14

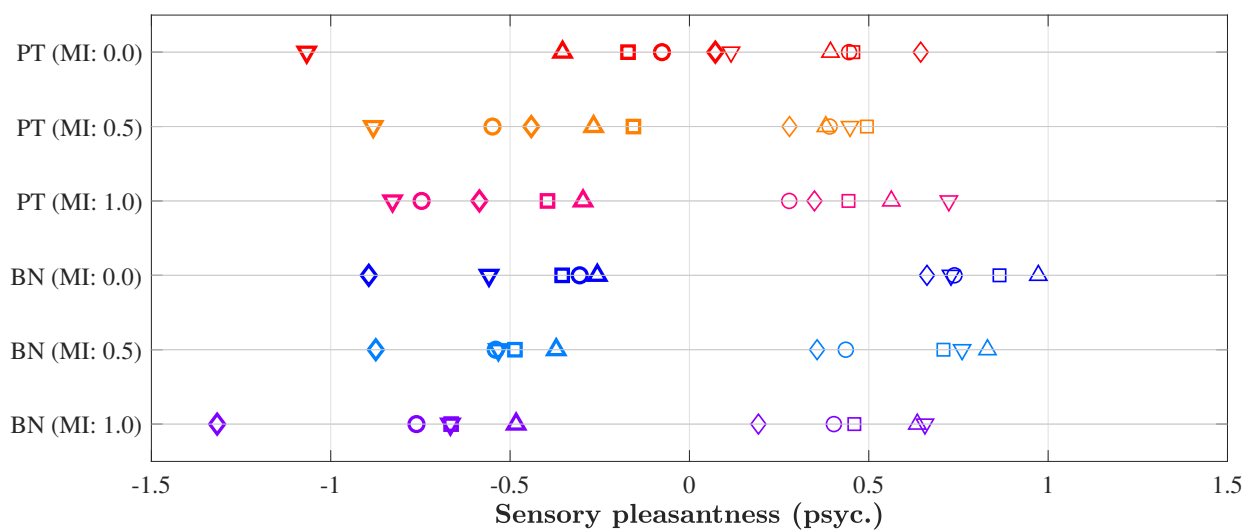


図 4.1: 感覚的快さの主観評価結果

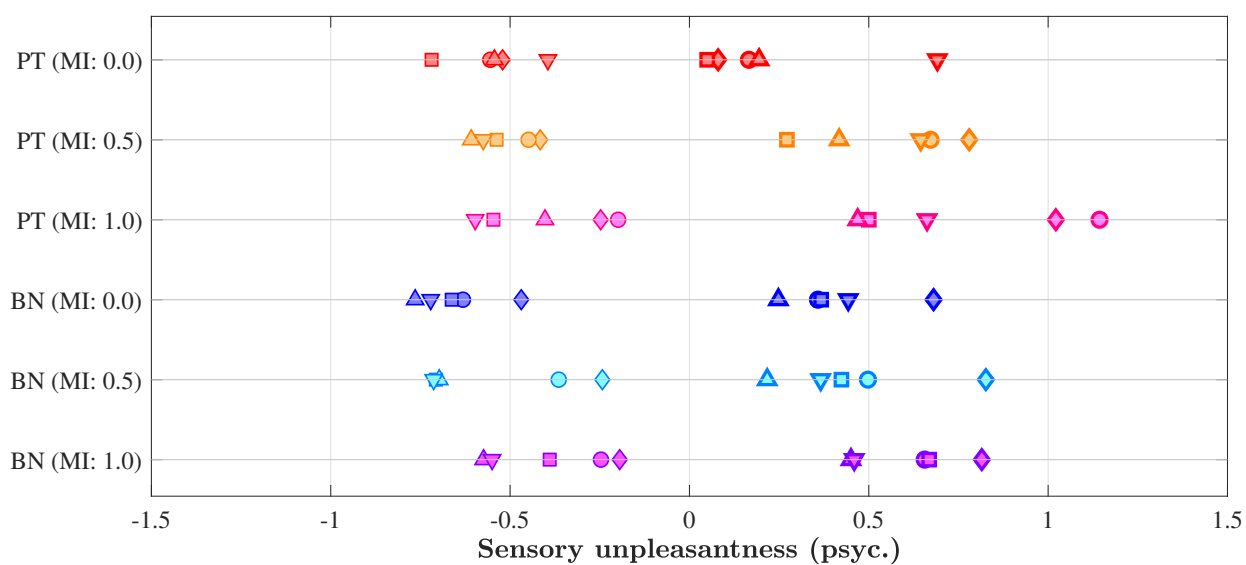


図 4.2: 感覚的快くなさの主観評価結果

4.3 感覚的快さ・快くなさの関係

本節では、上述で説明した感覚的快さと感覚的快くなさの主観評価値を用いてその関係について検討する。特に、両者が同じ軸上で表せるかについて、統計的方法によって定量的に分析する。

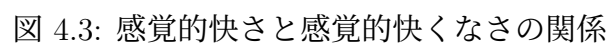
図 4.3 に評価結果を示す。図の横軸は、感覚的快さ、縦軸は、感覚的快くなさであり、60 刺激に対してそれぞれの評価結果が対応している。各プロットは、[20] の図 1 の凡例に対応する。また、図中の破線はこの回帰直線に対する 95%信頼区間を表す。両評価値の関係について、独立しているのか、それとも同じ評価軸上に現れるのかを調べた。ここでは、両評価値が同じ評価軸上で表されることをもって、その関係を快・不快とみなす。

感覚的快さと感覚的快くなさの主観評価値に対して、その関係が独立した評価尺度ではなく、互いに対応する関係にあるかを調べるために、相関分析を行った。相関分析の結果、相関係数は -0.95 であった。また、Pearson の相関係数に対する無相関の帰無仮説 ($r = 0$) を検定した結果、 $p = 4.32 \times 10^{-32}$ であり、この相関は統計的に有意であった ($p < 0.001$)。このことから、ある音に対する感覚的快さの評価が高いほど、その音に対する感覚的快くなさの評価が低くなる傾向にあることがわかった。また、これらは、95%信頼区間内におおよそ収まっており、大きな外れ値は見られない。このことから、音の感覚的快さと感覚的快くなさは、独立しておらず、対極の関係になっていることが示唆された。

感覚的快さと感覚的快くなさは強い負の相関関係があることがわかったが、相関分析の結果だけでは、同一の評価軸上に存在することを示すには十分とはいえない。そこで、両者に対してノンパラメトリックの統計的検定を行い、同一の評価軸上で対の関係であることを検証する。ここでは、対応のある 2 群の差の中央値が 0 であるという帰無仮説のもとで、ウィルコクソンの符号付順位検定を行った。ここで用いたウィルコクソンの符号付順位検定は、対応のある 2 群の差に着目し、その差の中央値が 0 であるという帰無仮説を検定するノンパラメトリック検定である。MATLAB では、`signrank` 関数として実装されており、2 標本の場合は両者の差の中央値が、2 標本のデータの分布から得られたという仮説を検定する。その結果、統計的に有意な差は認められなかった ($p = 0.98$)。統計検定量 z 値は 0.022 と非常に 0 に近いことから、感覚的快さと快くなさの分布は正規分布とのずれが非常に小さく、一貫した感覚であり、対極に分布していることがわかった。

以上より、感覚的快さ・快くなさ（以降、快・不快）は、互いに独立した評価尺度ではなく、同一の評価軸上の対に現れるものであることがわかった。しかし、感覚的快さと感覚的快くなさの二つの実験から得られた結果を、両者の軸の向き

を揃え（感覚的快くなさのデータを反転させた）、快・不快のデータとして検討した際、快・不快の接合付近（中央）のデータが、原点に向けて十分に寄っていないことがわかった。快・不快を連続した軸で扱う上で、接合付近のデータ不足しているということは、両者の刺激条件の範囲が十分ではなかったことを示している。



4.4 感覚的快さ・快くなさの一元評価

前節の結果から、感覚的快さと快くなさは独立した関係ではなく、同一の評価軸上の対の関係にあることがわかった。そこで、左側に感覚的快くなさ、右側に感覚的快さとなるように軸を調節し、一元的な快さ－快くなさの評価軸を構成する。本節では、この評価軸を用いて、Aures のモデルに基づく回帰モデルが前節の感覚的快さと感覚的快くなさの関係を説明可能かを検討する。

次式のように、感覚的快さと感覚的快くなさの要素を統合した感覚値 W_i を定義する。

$$W_i = [-W_u W_p] \quad (4.1)$$

ここで、対数表現を用いたモデルを検討するにあたり W_i には負の値が含まれるため、そのままでは対数スケールでの計算が行えない。そこで、次式のように、 W_i 全体をシフトさせ、すべての値が正となるように処理を行う。具体的には、最小値の絶対値に 1 を加えた量を全体に加算し、対数スケール上で 0 から始まるように変換する。

$$W_i + |\min(W_i)| + 1 \quad (4.2)$$

Aures のモデルに基づく四つの音質評価指標を利用した回帰モデル（以後、感覚的快さ・快くなさの予測モデルと呼ぶ）は、次式のように構成される [19]。

$$\log(\hat{W}_p) = \beta_{p,0} + \beta_{p,1}R + \beta_{p,2}S + \beta_{p,3}K + \beta_{p,4}N^2 \quad (4.3)$$

一元的な快・不快の推定では、と式 (4.3) と同じ形式で係数のみ調節した次式の形式で構成される。

$$\log(\hat{W}_i) = \beta_{i,0} + \beta_{i,1}R + \beta_{i,2}S + \beta_{i,3}K + \beta_{i,4}N^2 \quad (4.4)$$

ここでは、次式の RMSE が最小になるように、係数 $\beta_{i,k}$ ($k = 0, 1, \dots, 4$) を推定する。

$$\text{RMSE} = \frac{1}{I} \sum_{i=1}^I |\log(W_i) - \log(\hat{W}_i)|^2 \quad (4.5)$$

得られた係数はそれぞれ、表 4.2 に示すとおりであった。バイアス項である $\beta_{i,0}$ は除き、ラウドネス、ラフネス、シャープネスの重み係数は負の値、トナリティの係数のみ正の値となった。このことから、感覚的快さの予測モデルと同様に、トナリティ以外の値が大きくなると不快、トナリティが大きいと快と計算されるモデルであることがわかる。ただし、バイアス項の働きがあるため、出力される値

自体はそうにならない可能性もある。

主観評価値のスケールと一致させるため、式 (4.4) の両辺に対して指数変換を行い、元のスケールで得られた推定結果を図 4.4 に示す。図の横軸は式 (4.1) に示す、感覚的快さと、符号反転した感覚的快くなさの主観評価値を表し、縦軸は式 (4.4) を用いた推定値を表す。実験結果と推定結果に対して推定誤差を計算した結果、RMSE は 0.203 であった。0.203 は、感覚的快くなさのデータの範囲 $[-1.5, 1.5]$ (全体 3.0) に対して、10% 程度の推定誤差である。

この結果から、音質評価指標を利用した感覚的快さ・快くなさの予測モデルが快・不快の主観評価結果を十分に説明できていることがわかる。以上の結果から、感覚的快さ・快くなさを一元的な評価軸上で扱うことが統計的にも適切であり、音質評価指標に基づく回帰モデルによって、その関係を説明できることがわかった。

表 4.2: 感覚的快さ・快くなさの予測モデル (式 (4.4)) の各係数

係数	値
$\beta_{i,0}$	1.08
$\beta_{i,1}$	-0.00170
$\beta_{i,2}$	-0.154
$\beta_{i,3}$	0.0349
$\beta_{i,4}$	-0.00210

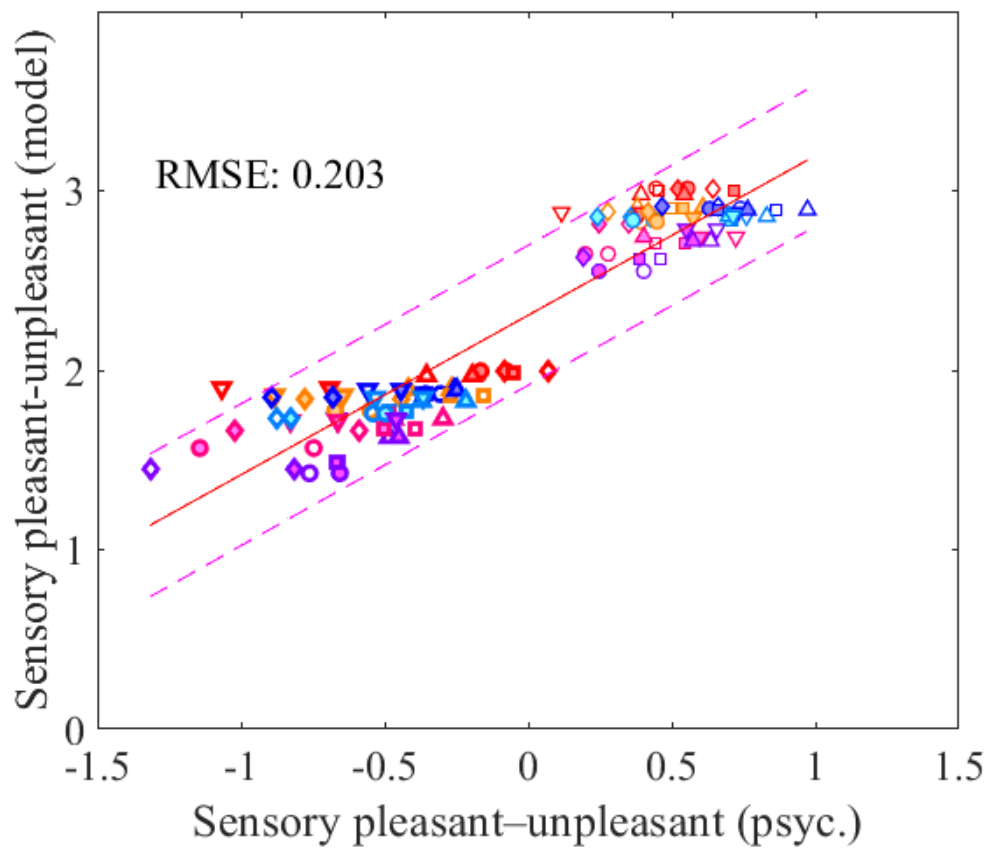


図 4.4: 音質評価指標を利用した回帰モデルによる感覚的快さ・快くなさの一元評価結果. 各プロットは, 図 5.4 のプロットおよび凡例の刺激に対応する.

4.5 感覚的快さ・快くなさと快・不快の対応

本章におけるこれまでの検討の結果、感覚的快さ・快くなさは同じ軸の対の関係にあることがわかった。特に、検定統計量の z 値が0.0221とおおよそ0であったことから、感覚的快さと快くなさの分布は正規分布とのずれが非常に小さく、一貫した感覚であり、対称的に分布していることがわかった。これらのことから、本研究では、先行研究 [40] で得られた感覚的快さと感覚的快くなさの関係について、快・不快とみなせる対応関係があると仮定する。また、この関係を利用することで、符号反転などの適切な変換処理は必要であるが、快さを予測するモデルを用いた不快さの予測、ならびに快・不快の一元的な評価が可能であると考えられる。

ところで、4.1節、4.2節の結果で述べたように、図4.1、図4.2を確認すると、両主観評価値ともに尺度全体の範囲に対して、両端のデータが不足していることが確認できる。また、図4.3より、感覚的快さと感覚的快くなさの二つの実験から得られた結果を、快・不快のデータとして検討した際、快・不快の接合付近（中央付近）のデータが、原点に向けて十分に寄っていないことがわかる。快・不快を連続した評価軸として扱うためには、快の値と不快の値がそれぞれ両端側まで十分に広がり、評価値の範囲が両端まで十分に確保されている必要がある。

これらのことから、接合付近のデータが十分ではなかったということは、快と不快の両端の範囲に該当する主観評価値に対応する刺激条件の範囲を十分に設定できていないということを示している。そのため、先行研究で設定した刺激条件の範囲は、快・不快を評価するためにはその刺激条件の範囲が十分であったとはいえない。

第5章 音の快・不快を予測するための 感覚的快さの主観評価実験

本章では、前章の問題点を踏まえ感覚的快さの再評価を行う。ただし、ここまでの検討に基づき得られた再評価結果の折り返し部以降が不快と仮定し、その結果を快・不快の主観評価値として扱う。この結果に対して、前章で用いた刺激条件が快・不快の評価のために十分な範囲であったかを検証する。

5.1 実験刺激

Table 5.1 に、実験刺激の条件を示す。ここでは、不快側に評価値が広がることをねらいとし、先の検討で用いた刺激条件 [20] は包含した上で、規定に基づく十分に安全が保障される限界までのラウドネスの条件の拡張、アノイアンスモデルで利用されている変動強度に関わる条件を追加した。具体的な刺激条件は、ラフネスに関わる振幅変調の変調周波数が 70 Hz、変調度が 0%, 50%, 100%。シャープネスに関わる周波数重心として 0.5 kHz, 1 kHz, 2 kHz, 4 kHz, 8 kHz。トナリティに関わる条件が単一周波数の純音と 1 kHz 帯域幅の帯域雑音。ラウドネスが 2 sone, 14 sone, 21 sone。変動強度に関わる振幅変調の変調周波数が 4 Hz、変調度が 0%, 50%, 100%、を想定した 150 種類の音刺激を準備した。また、表 5.1 において、先の検討から追加した条件は、太文字で示した。ここでは、音刺激の提示時間が感覚的快さや不快さに影響を与える可能性が指摘されているため、すべての音刺激の時間長を 1 秒に統一した。

表 5.1: 実験刺激の条件

信号の種類	中心周波数 (kHz)	変調周波数 (Hz)	変調度 (%)	ラウドネス (sone)
純音 (PT)	0.5, 1, 2, 4, 8	4, 70	0, 50, 100	2, 14, 21
帯域雑音 (BN)	0.5, 1, 2, 4, 8	4, 70	0, 50, 100	2, 14, 21

5.2 実験参加者

実験には、正常聴力を有する日本語母語話者 12 名（大学院生 21~26 歳，男性 6 名，女性 6 名）が参加した。

本実験ではサーストンの一対比較法を採用し，各比較を二者強制選択 (A が選ばれる確率を p とする) として扱う。各刺激対について実験参加者 1 名につき 1 回の判断が得られるため，1 対比較あたりの試行回数は $n = 12$ となる。ここで，差がない場合の選好確率を $p = 0.5$ とする帰無仮説 $H_0 : p = 0.5$ に対し，有意水準 $\alpha = 0.05$ ，検出力 0.8 を目標としてパワーアナリシスを行った結果， $n = 12$ では真の確率が概ね $p \geq 0.829$ の差を十分な検出力で捉えられることが確認された。以上より，本実験の刺激条件で想定される差を検出する上で，実験参加者数の 12 名は信頼のある主観評価結果が得られる人数である。

5.3 実験機器

実験系の構成を図 5.1 に示す。実験は防音室で行われた。刺激の提示では，PC (Windows11, MATLAB R2024b, Mathworks)，オーディオインターフェース (RME Fireface UCX)，ヘッドホン (SENNHEISER HDA200) を使用した。刺激の立上り・立下りに 100 ms のテーパーをかけた。実験参加者に負担がかからないように，十分に休憩を挟んで実験を行った。

5.4 実験方法

実験では，刺激の感覚的快さを評価するために，サーストンの一対比較法を利用した。実験参加者には，一つの刺激対として第 1 音，第 2 音の順に提示して，どちらの音が感覚的に快く感じたかを選択させた。実験刺激は，図 5.3 に示すとおり，第 1 音が 1 秒間再生されたのち，0.5 秒の休止時間があり，その後に第 2 音が 1 秒間再生された。第 1 音および第 2 音はランダムに決定された。刺激対は合計 11175 対であった。1 セット 300 回として，1 日あたり 4 セット，計 1200 回の比較を行わせた。セットごとに 10 分の休憩を挟んだ。実験は休憩を含めて概ね 120 分程度を要した。これを 10 日に分けて実施した。なお，実験最終日のみ 1 セットのみの実施で刺激対の総比較は完了した。

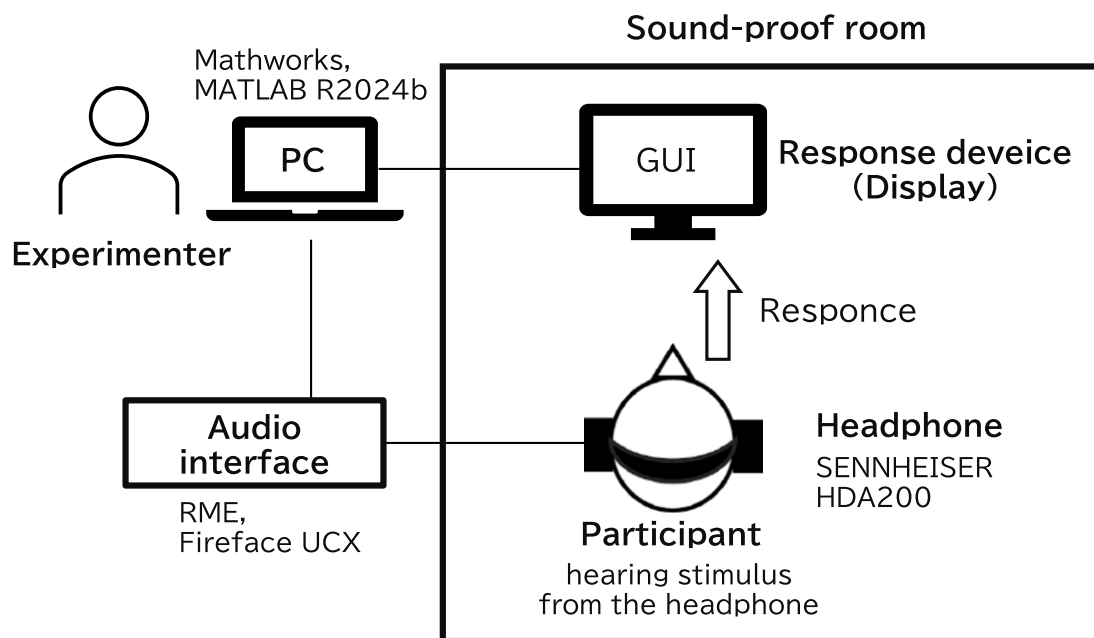


図 5.1: 実験系の構成図

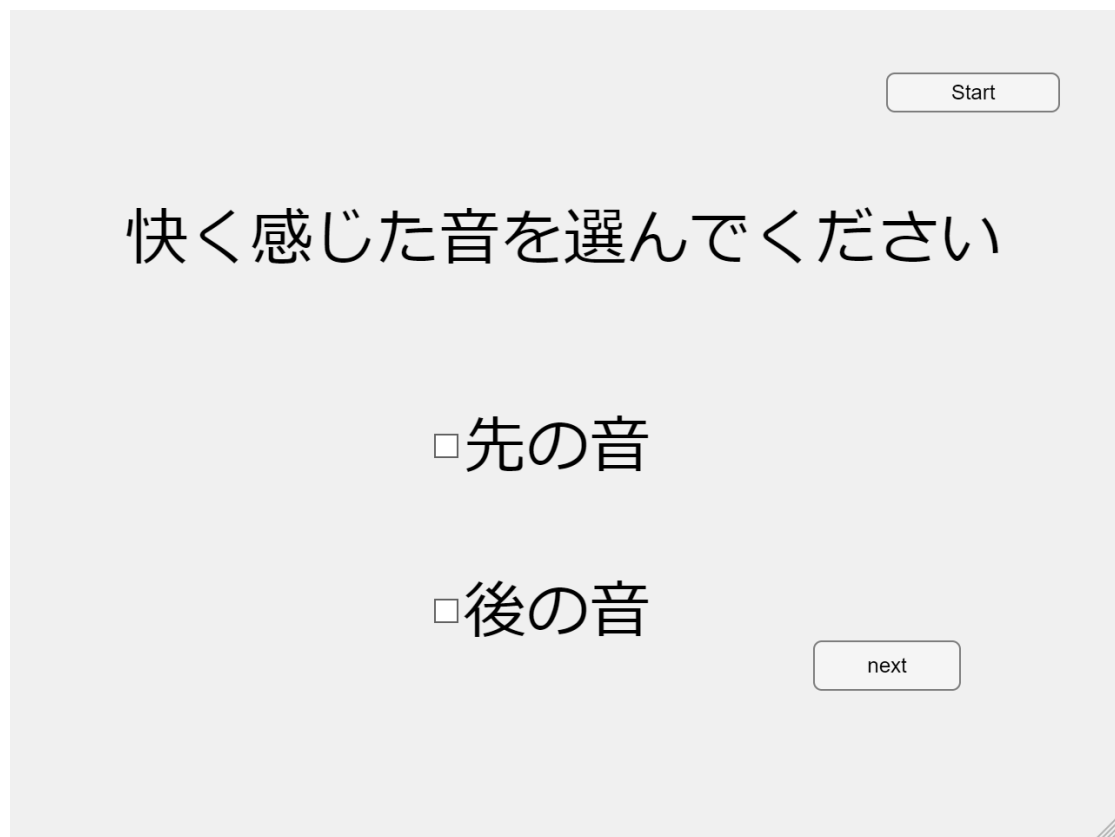


図 5.2: GUI

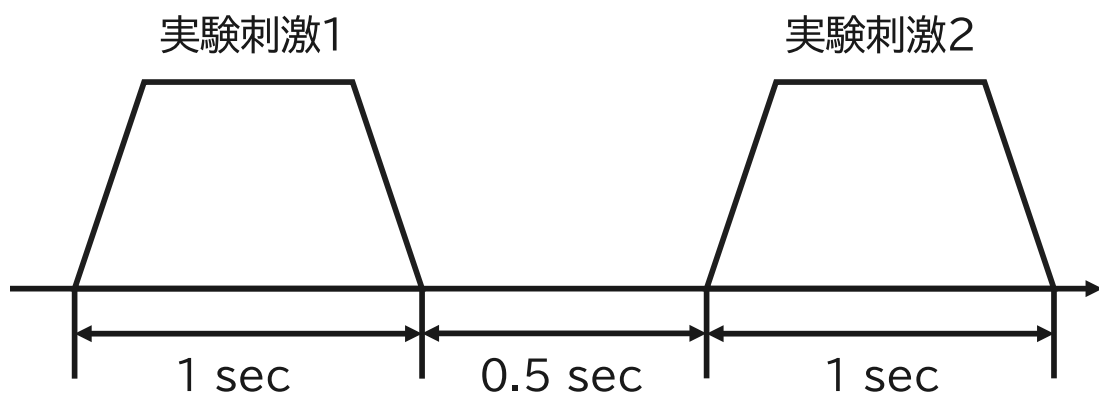


図 5.3: 刺激配置

5.5 実験結果

Table5.1 の実験刺激に対する主観評価結果を図 5.4 に示す．ここでは，この結果を快・不快の評価値として扱う．横軸は刺激間の相対距離を示し，正の値に大きくなるほど快，負の値に大きくなるほど不快であることを表しており，横軸上で値が離れるほど，主観的な印象の差が大きいことを意味する．縦軸は，実験刺激の種類（純音：PT か帯域雑音：BN）と，変調条件（2 種類の変調周波数，3 種類の変調度）に対応するように並べている．図中の記号の太さはラウドネスの大きさに対応し，記号の形と色は中心周波数および変調度に対応する．なお，説明を簡単にするため，ラウドネス 21 sone の条件のみ塗りつぶして表示した．この表示により，先行研究と，本実験間の結果を同一図内で比較できる．

図 5.4 より，ラウドネスが大きいときよりも小さいときのほうが，感覚的に快と判断されやすいことがわかる．また，純音では，変調周波数 70 Hz の際，変調度が大きいときよりも小さいときのほうが快に評価されやすい傾向が見られる．一方で，刺激条件によっては，これらの関係が逆転し，不快に評価が寄る場合があることも確認できる．このことは，快・不快の判断が単一の要因のみで決まるのではなく，各音質評価間の複数の要因の組合せによって変化する可能性を示唆している．

さらに，評価値は快・不快の接合付近（中央）にも十分に分布しており，先行研究で確認された偏りが改善されたことがわかる．これは，主観評価結果が快あるいは不快のどちらか一方に評価が偏る状況が解決され，連続的な快・不快の軸として扱う上で必要な中央付近のデータが得られたことを示す．また，本実験では先行研究で用いた刺激条件をすべて包含する刺激条件を設定したため，先の検討との評価結果を同一の枠組みで比較することが可能である．範囲を拡張した刺激条件とそれ以外の条件（塗りつぶしのある点とない点）を比較すると，評価値の分布範囲が広がっていることが確認できる．特に，塗りつぶした条件に着目すると，期待されたとおり，不快側に評価値が拡充されており，刺激条件の拡張によって不快側のデータが補われたことがわかる．以上より，本実験では，快・不快を評価するための刺激条件の範囲を設定できたことがわかる．

信号の種類	変調度	色
純音 (PT)	0	赤
純音 (PT)	0.5	橙
純音 (PT)	1	桃
帯域雑音 (BN)	0	青
帯域雑音 (BN)	0.5	水
帯域雑音 (BN)	1	紫
中心周波数 (kHz) 記号		
0.5		◇
1		○
2		□
4		△
8		▽
ラウドネス (sone) 太さ		
2		細
14		太
21		極太

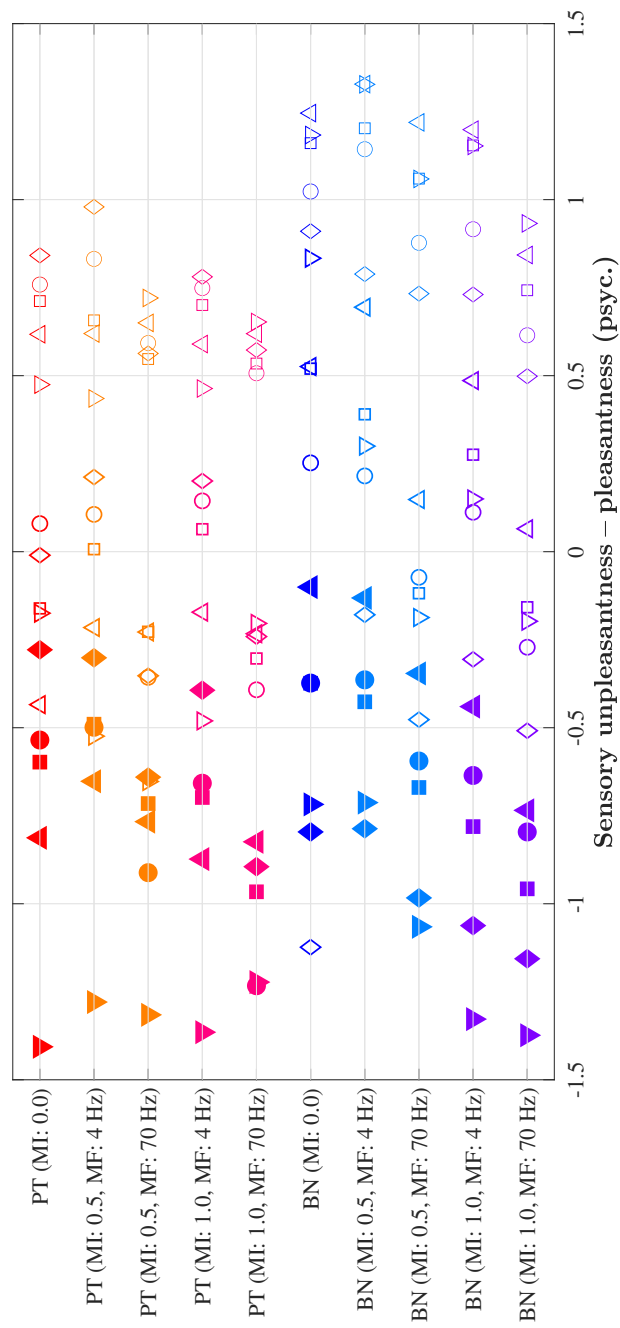


図 5.4: Table 5.1 の実験刺激に対する主観評価結果（便宜上，図 5.4 では，ラウドネス 21 sone の条件のみ塗りつぶし）

5.6 考察

図 5.4 に示した主観評価結果より、快・不快の評価には特にラウドネスとラフネスに関する影響が大きいことがわかった。この結果は、第 2 章で述べた関連研究における知見と整合する。例えば、音の大きさが不快感に強く関係するという報告は騒音分野や、聴覚心理学の分野などで従来より報告されている [17]。アノイアンスモデル [14] と感覚的快さの予測モデル [13] では、ラウドネスが大きくなるほど不快に、あるいは感覚的な快さが小さくなるように設計されており、本実験で得られた主観評価結果と整合がある。

聴覚的顕著性 (auditory salience) の研究では、音の大きさと不快感に相関があることが報告されており [17]、主観的に目立つ音ほど生理反応（瞳孔や脈拍の変化）が大きくなることが示されている。これは、本実験で得られた、ラウドネスが大きいほど不快と評価される傾向と一致する、ラウドネスの大きさと聴覚顕著性に密接な関係があることが示唆された。これらの知見は、緊急性の高い聴覚顕著性が求められる報知音などの設計に紐づくことが期待される。

自動車騒音 [28] の研究では、粗っぽい音は聴覚の快適性と密接に関係し、ほとんどの場合高い相関が確認されたことが報告されている。これは、本実験で得られた、ラフネスの値が大きくなるにつれ不快と評価される傾向と整合がある。さらに、同研究 [28] では、音刺激の各位相間の相関についても言及されており、位相の相関が高いほど快と評価されることが報告されている。

ラフネスと変動強度の計算において、その重みには各位相間の相関の良さに関わる係数が利用されている [10]。各位相間の相関が良いほど、両者の重み係数は大きくなり値が大きく出力される。例えば、ラフネスの計算において、ホワイトノイズのような音は、ラフネスが最大値となる変調周波数 70 Hz 相当の包絡構造を与えた場合、全体域に渡って位相の情報を持つにも関わらず各位相間の相関が極めて小さいため、その粗さはそれほど大きく算出されない。本実験で用いた刺激条件では、位相構造そのものを独立に操作したわけではないが、ラフネスや変動強度に関連した物理量の操作によって、位相間相関の違いが間接的に主観評価結果に反映されている可能性がある。

第6章 音の快・不快の予測モデル

本章では、前章で用いた刺激条件の範囲に問題がないことがわかったため、五つの音質評価指標に基づく音の快・不快の予測モデルを提案する。特に、提案モデルに対して、構築に用いた評価値をどの程度説明できるかを確認し、各指標の重要性についても検討する。

6.1 提案法

本節では、音の快・不快の予測モデルの構築に際し、利用する関数モデルの形式について述べる。

提案モデルの構築では、関連研究で紹介したフェヒナーの法則に基づくこと、および、指数関数を対数変換することで、各音質評価指標を分離して扱うことが可能であるという点から、感覚的快さの予測モデル [13] に倣う。各音質評価指標と係数の指数表現を用いて提案モデルを構築する。構築には、感覚的快さの予測モデル [13] と、アノイアンスモデル [14] で利用される五つの音質評価指標（ラフネス、シャープネス、トナリティ、ラウドネス、変動強度）を統合的に考慮する。

6.2 モデル構築

本節では、五つの音質評価指標（ラフネス、シャープネス、トナリティ、ラウドネス、変動強度）と係数の指数表現を用いて提案モデルを構築する。

構築した提案モデルの構成図を図 6.2 に示す。また、式 (6.1) に提案モデルの式を示す。

$$\hat{V}_p = e^{\alpha_{p,0}} e^{\alpha_{p,1}R} e^{\alpha_{p,2}S} e^{\alpha_{p,3}K} e^{\alpha_{p,4}N^2} e^{\alpha_{p,5}FS} \quad (6.1)$$

ただし、 R はラフネス（単位は asper）、 S はシャープネス（単位は acum）、 K はトナリティ（単位は tu）、 N はラウドネス（単位は sone）、 FS は変動強度（単位は vacil）である。ここでは、式 (6.2) の平均二乗誤差（MSE）が最小になるよう

に、係数 $\alpha_{p,k}$ ($k = 0, 1, \dots, 5$) を推定する.

$$\text{MSE} = \frac{1}{I} \sum_{i=1}^I |V_{p,i} - \hat{V}_{p,i}|^2 \quad (6.2)$$

ここで、 $V_{p,i}$ は快・不快の主観評価値、 $\hat{V}_{p,i}$ は刺激 i に対するモデル推定値、 I は刺激数である.

得られた各係数は表 6.1 に示すとおりであった. バイアス項である $\beta_{i,0}$ は除き、ラウドネス、ラフネス、シャープネス、トナリティの係数全てが負の値となった. これは、トナリティの係数のみ正の値である、感覚的快さの予測モデルと相違する. これは、刺激の条件を増やしたことにより、トナリティが高い音が返って、耳障りに評価されたことを示唆する. 実験後の簡単なインタビューでは、純音間の強い音より、雑音間の強い音の方が落ち着くという意見もあった. さらに、トナリティは、音刺激が純音であるか帯域雑音であるかといったスペクトル形状だけでなく、振幅変調における変調周波数および変調度によっても変化する. 一般に、振幅変調を行うと、周波数スペクトル上では中心周波数の両側にサイドバンドが生じる. このとき、サイドバンドの中心周波数からの周波数間隔は変調周波数に、サイドバンドの振幅（高さ）は変調度に対応する. トナリティの算出においては、サイドバンドが中心周波数からどれだけ離れているか、あるいはサイドバンドの振幅が小さい場合に、中心周波数成分が相対的に際立つため、純音感（トナリティ）が高く計算される. すなわち、ラフネスや変動強度と統合的に評価が行われる. また、全ての音質評価指標の各係数が負の値になったことに対し、バイアス項は感覚的快さ・快くなさの一元表示モデルと比べて 0.2 ほど大きくなった.

各音質評価指標の分析には、Sound Quality Analysis Toolbox (SQAT) を使用した [41]. 具体的には、Daniel & Weber のラフネスの計算モデル [42], Aures のシャープネスの時変動計算モデル (time varying method) [13], Aures のトナリティの計算モデル [13], Zwicker のラウドネスの時変動計算モデル (time varying method) [43], Osses の変動強度計算モデル [44] を用いた.

構築した提案モデルの説明精度を確認するため、主観評価値とモデル予測値の平均二乗平方誤差を算出する. ただし、式 (6.1) は指数関数の積であるため、推定値 $\hat{V}_p$ は常に 0 より大きい値のみを出力する. 一方で、本研究の快・不快評価は負の値を含む尺度であり、そのままでは尺度の扱いが一致しない. そこで、評価値全体を正の領域へ平行移動し、すべての値が正となるように前処理を行う. 具体的には、式 (6.3) に示すように、評価値ベクトル $\mathbf{V}_p$ の最小値の絶対値を全体に加

算し、最小値が0となるように変換する.

$$V_p + |\min(V_p)| \quad (6.3)$$

図6.2に、快・不快の主観評価値と提案モデルによる推定値の関係を示す. 図の横軸は、快・不快の評価値、縦軸は、提案モデルによる推定値であり、150刺激に対してそれぞれの評価と推定値が対応している. また、図中の破線はこの回帰直線に対する95%信頼区間を表す. 変換後の評価値と推定結果に対して推定誤差を計算した結果、RMSEは0.37であった. 0.37は、快・不快の尺度全体の範囲3.00に対して約10%程度である. この結果から、提案モデルが主観評価結果を高い精度で説明できることがわかった. 以上より、構築した提案モデルを客観的に評価した結果その妥当性が示された.

表 6.1: 提案モデル（式 (6.1)）の各係数

係数	値
$\alpha_{p,0}$	1.31
$\alpha_{p,1}$	-0.162
$\alpha_{p,2}$	-0.0245
$\alpha_{p,3}$	-0.108
$\alpha_{p,4}$	-0.00140
$\alpha_{p,5}$	-0.0300

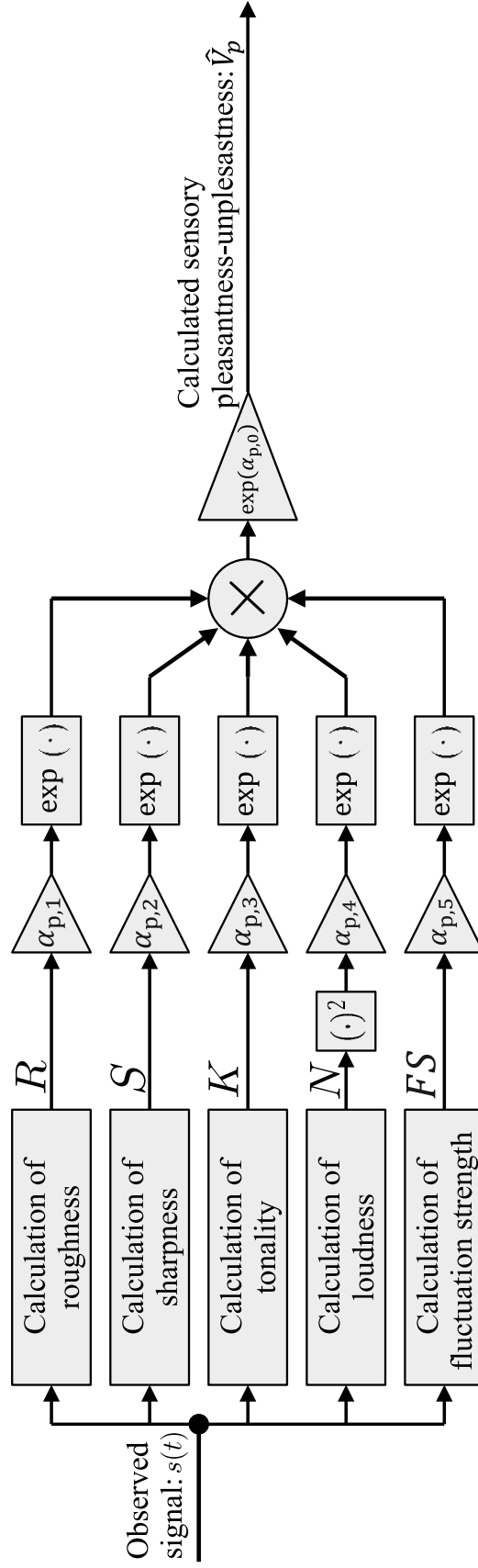


図 6.1: 音の快・不快の予測モデルの構成図

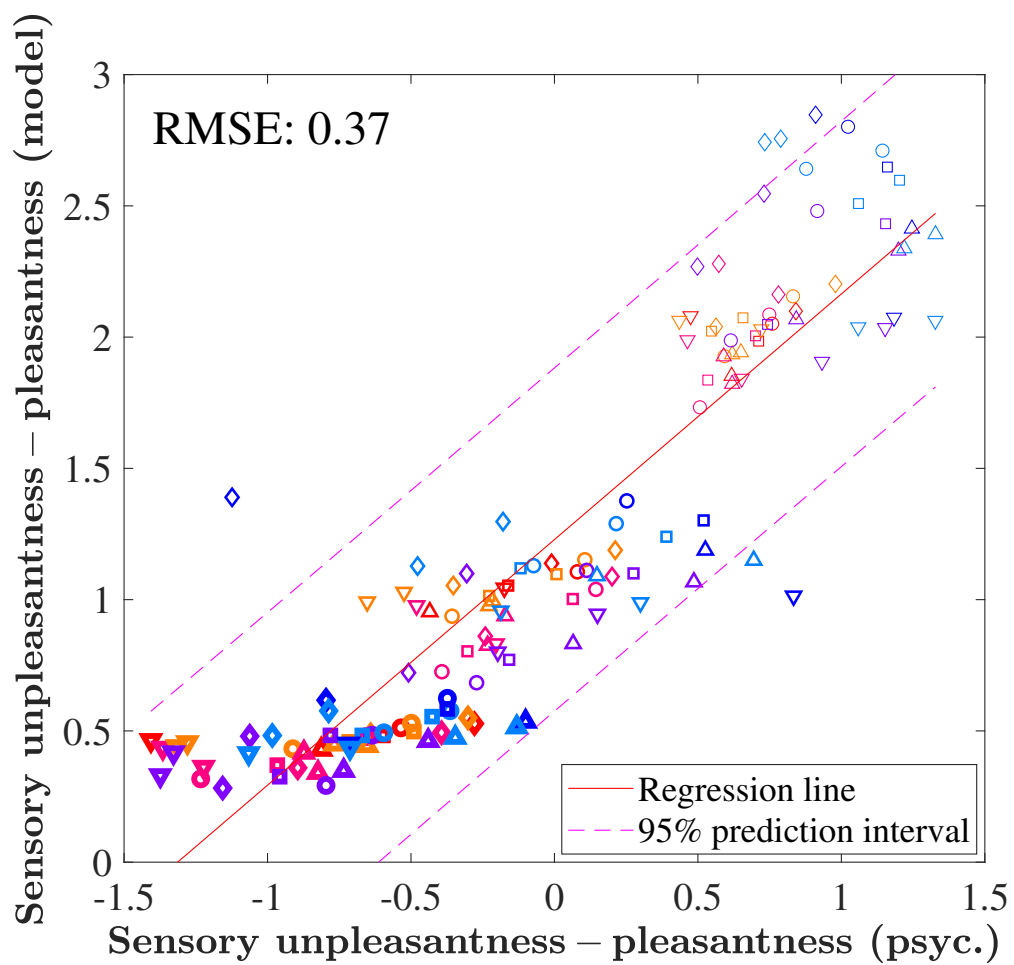


図 6.2: 音の快・不快の予測モデルによる快・不快の推定結果. 各プロットの色, 太さ, 記号は, 図 5.4 の凡例の刺激に対応する.

6.3 各音質評価指標の寄与度

本節では、提案モデルにおいて、各音質評価指標が快・不快の推定にどれだけ寄与しているかを調べる。寄与度の評価方法としては、(1) ラフネスを利用しない場合 ($\alpha_{p,1} = 0$)、(2) シャープネスを利用しない場合 ($\alpha_{p,2} = 0$)、(3) トナリティを利用しない場合 ($\alpha_{p,3} = 0$)、(4) ラウドネスを利用しない場合 ($\alpha_{p,4} = 0$)、(5) 変動強度を利用しない場合 ($\alpha_{p,5} = 0$) の5条件を設定し、それぞれの推定精度の変化を比較する。ここで、利用しないとは、各音質評価指標に対応する係数を0にしたうえで、他の係数を平均二乗誤差が最小となるように推定し直すことを指す。

式(6.1)に示す形式の指数表示の状態では各項を分離して扱うことが出来ないため、各項を分離して扱うためにこの検討は対数スケールで行うものとする。また、寄与度の比較を元のスケール（線形スケール）で行い、評価値 V_p と推定値 $\hat{V}_p$ の二乗平方誤差を最小化すると、絶対誤差を最小化することになるため、寄与度の大きい指標に関わる誤差が目的関数を支配しやすい。その結果、評価値が大きい刺激に対する当てはまりを主に改善する指標（例えばラウドネスが予想される）の影響が過大に反映され、その他の指標の寄与が相対的に現れにくくなる可能性がある。また、

$$\log(V_{p,i}) - \log(\hat{V}_{p,i}) = \log\left(\frac{V_{p,i}}{\hat{V}_{p,i}}\right) \approx \frac{V_{p,i} - \hat{V}_{p,i}}{V_{p,i}} \quad (6.4)$$

式(6.4)より、テイラー展開によって対数上での差は両者の比率に近似されるため、対数空間での誤差は絶対差ではなく何倍ずれているかを示す。 $\hat{V}_{p,i}$ が $V_{p,i}$ に近いときは、元のスケールでの相対誤差に近似できる。したがって、対数上で係数の推定を行うことで、評価値の大小にかかわらず、広い範囲の刺激に対する当てはまりを均等に反映しやすく、寄与度の比較に適していると考えられる。

以上を踏まえ、本節のMSEは対数スケール上、すなわち $\log(V_p)$ と $\log(\hat{V}_p)$ の差に基づいて算出した。そこで本節では、提案モデルを対数変換した表現に基づき、対数スケール上で寄与度の調査を行った。ここでは、上述のとおり、対数スケール上において(6.5)の形で係数の適合を行った。評価値についても同様に対数を取り $\log(V_p)$ と $\log(\hat{V}_p)$ の差の二乗和

$$\text{MSE} = \frac{1}{I} \sum_{i=1}^I |\log(V_{p,i}) - \log(\hat{V}_{p,i})|^2 \quad (6.5)$$

を最小化することに対応する。

以上の処理によって求めた $\log(\hat{V}_p)$ に対して指数変換を行い、元のスケールで得られた快・不快の推定結果を図6.3に示す。図の横軸は第4節で行った聴取実験によ

り得られた快・不快の主観評価値を表し、縦軸は推定した快・不快の推定値を表す。150 刺激に対してそれぞれの評価結果が対応している。また、図中の破線は回帰直線に対する 95%信頼区間を表す。図 6.3(a) は指標すべてを利用した場合、その他はそれぞれ上述の (1)～(5) に対応しており、図 6.3(b), 図 6.3(c), 図 6.3(d), 図 6.3(e), 図 6.3(f) に示す。結果として、ラウドネスのみを利用しない場合、RMSE は 0.29 から 0.57 まで増加した。次いでラフネスのみを利用しない場合、RMSE は 0.29 から 0.32 まで増加した。次いでトナリティのみを利用しない場合、RMSE は 0.29 から 0.31 に増加した。次いでシャープネスのみを利用しない場合、RMSE は 0.29 から 0.30 に増加した。変動強度のみ利用しない場合、RMSE は増加しなかった。以上より、今回の主観評価結果については、ラウドネス、ラフネス、トナリティ、シャープネスの順に寄与度が高いことがわかった。また、RMSE の増加量が特に大きかったラウドネスは、本実験において快・不快の判断に支配的に作用しており、この結果は図 5.4 の定性的な傾向の分析とも整合性が認められる。

一方で、増加量が小さい指標については、寄与が相対的に小さいことを示唆する。特に変動強度は、本研究で用いた感覚的快さの評価では変化の影響が表れにくく、主観評価に対する寄与が小さくなった。変動強度は聴覚心理額に基づく関連研究において、音の不快を評価するアノイアンスモデルで用いられている一方、快さを評価する感覚的快さの予測モデルでは用いられていない。今回実施した実験は、先行研究のデータを利用して明らかにした快・不快が同じ軸上にあるという仮定に基づき、感覚的快さの評価しか行っていないため、不快に関する変動強度の影響が表れにくかった可能性が考えられる。

図 6.2 より、感覚的快さ・快くなさ一元評価モデルの推定結果（図 4.4）で観察された、上限と下限が一定になる傾向が緩和されていることが分かる。これは、ラウドネスが推定において支配的になっていることから、ラウドネスの依存による現象と考えられ、本推定結果では他の音質評価指標の寄与度が増加したため、緩和されたのではないかと考えられる。実際に、前章では、ラウドネスとラフネスのみを利用したモデルで、全ての音質評価指標を利用した場合と同等の精度で感覚的快さ・感覚的快くなさの関係を説明可能であったが、本推定結果を確認すると、上述の各音質評価指標を利用しなかった場合の RMSE より、ラウドネスとラフネスのみで、全ての音質評価指標を利用した場合と同等の精度での推定はできないことがわかる。

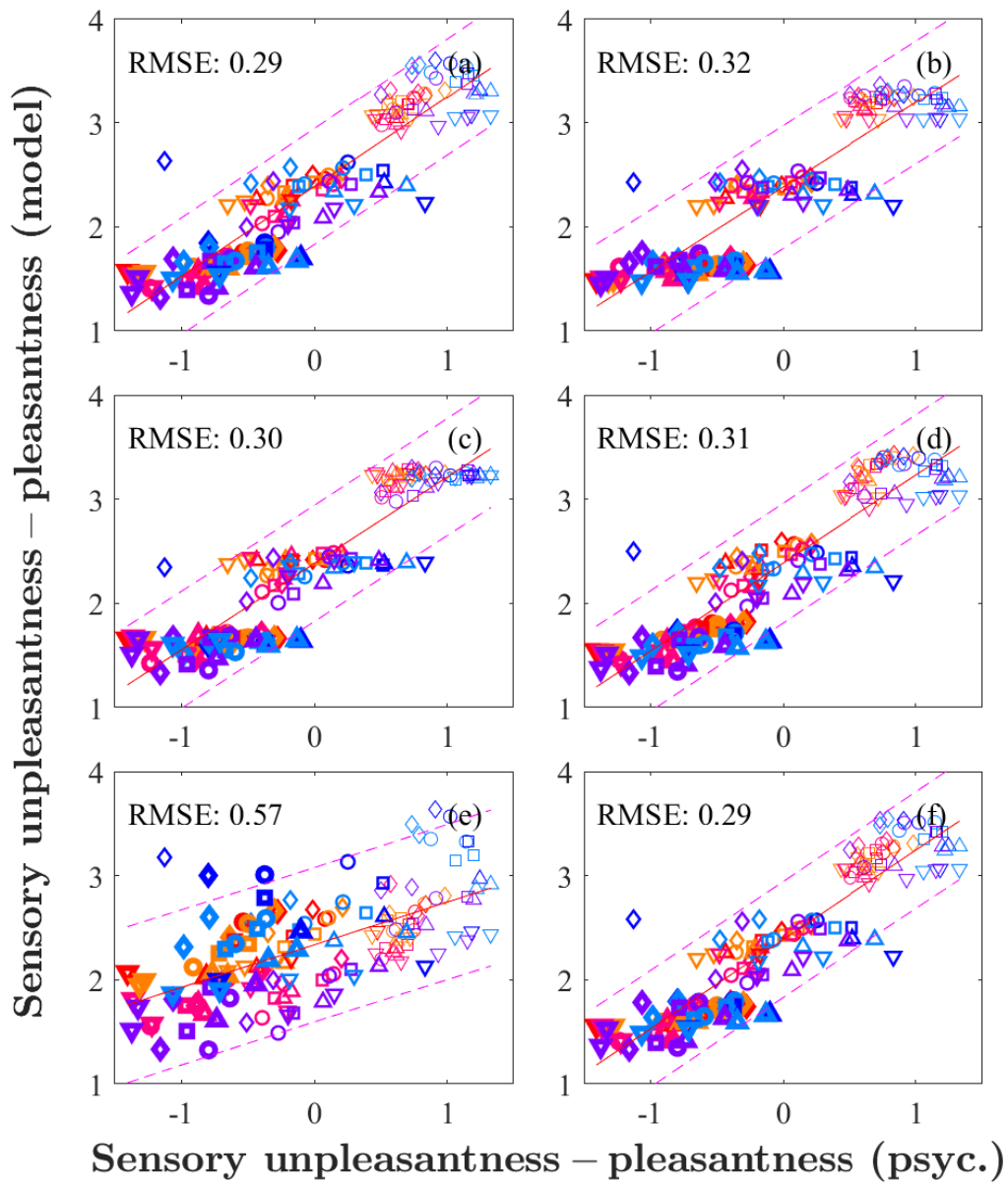


図 6.3: 音の快・不快モデルによる音の快・不快の推定結果：(a) すべての SQMs を利用した場合，(b) ラフネスを利用しなかった場合，(c) シャープネスを利用しなかった場合，(d) トナリティを利用しなかった場合，(e) ラウドネスを利用しなかった場合，(f) 変動強度を利用しなかった場合の推定結果．各プロットは，図 5.4 のプロットおよび凡例の刺激に対応する．

第7章 提案モデルの評価

本章では、提案モデルが未知の音の予測に対して有効であるか、主観評価実験を利用して検証する。また、快と不快を個別に評価する既存のモデルとの比較評価および、ヒトの聴覚の再現により優れたモデルを利用し問題がないかを検討する。

7.1 音の快・不快の予測モデルの評価

本節では、未知の音として用いる音源データセットおよび、提案モデルの評価に利用する主観評価値の取得方法を説明した後、提案モデルの評価について述べる。

まず、提案モデルの評価に利用した音源は、UrbanSound8K [45], ESC50 [46], Freesound.org [47–54], J-Alert 音源 [55,56] から引用した。UrbanSound8K は、都市音に分類される 10 種の音（クラクション、工事音など）を対象とした、4 秒以下のラベル付き音声ファイルを収録している。本データセットは、前景音、または、背景音など各方位や各時間帯から収録されている。ESC50 は、UrbanSound8K と同様の形式であり、収録されている音源の種類が異なる様々な種類の音源が収録されている。Freesound.org は、黒板、皿を掻く音源など身の回りにある様々な種類の音源が収録されている。J-Alert 音源では、緊急地震速報の音源を引用した。上述のデータセットから、快音（波音、鳥の声など）と不快音（黒板を掻く音、警報音など）を想定し、それぞれ 12 音源、計 24 音源準備した。また、音刺激の時間長は第 3 節と同様に 1 秒に統一した。これらの音圧レベルは、ラウドネスが 2, 7, 14, 21 sone となる 4 条件に振り分けた後、調節した。各ラウドネス条件には快音 3 音源、不快音 3 音源（計 6 音源）を割り当て、振り分けはランダムに行った。

ここで、上述の音源データセットはいずれも音の快・不快の評価のために収録されたものではなく、主に機械学習のための大量のデータセットである。特に、UrbanSound8K と ESC50 は訓練データとテストデータとして利用できるような配布フォーマットがある。本研究では、これらの音源に対して音質評価指標を見積もり、その音質評価指標を利用して提案モデルによる予測値を計算する。音質評価指標は、衝撃騒音のようにインパルスに似た特性を持つ音源に対して、正確な計算結果が保証されていない。帯域上のラウドネス密度分布や時間方向の平滑化を前提に設計されているため、衝撃騒音のように立ち上がり之急で極めて短時間

の非定常音では、不安定な計算がその結果が強く依存し、知覚と安定して対応しない。特にシャープネスでは、高域側の寄与を重み係数で反映し、総ラウドネスで正規化して算出される実装が多いため、ラウドネス密度が小さい区間では値が過大あるいは、発散してしまう。さらに、インパルスのような音は時間的に鋭い変化を持つため、スペクトル上は広帯域に成分が分布する。これにより、スペクトルを用いた計算も正しくされない。以上より、引用した音源に含まれる衝撃音に対して判定を行い、該当する音源があれば除外する必要がある。このことは、変動強度を計算するモデルを構築した論文の中でも触れられており、実環境音に対して計算を行う際、アヒルの鳴き声などに対しては計算結果が正しく計算できておらず、異常値のため考慮しないことが述べられている [44]。

判定方法は、50 ms のフレームを 25 ms ごとにシフトさせその範囲での実効値 (RMS) を計算する。計算した各範囲における最大値と最小値の差が 10 dB または、変動係数が 10% 以上の音源は衝撃音とした。変動係数は、全体に対して音源がどれだけ突出しているかを表し、以下式 (7.1) で計算される。以上の方法は、Towsey らの方法を参考にした [57]。

$$CV = \frac{\sigma}{\mu} \quad (7.1)$$

σ は実効値の標準偏差、 μ は実効値の平均値である。確認の結果、選定した 24 の音源は衝撃騒音とは判定されなかったため、選定した音に対して、音の快・不快の予測を行う。なお、引用した合計 4000 ほどの音源に対し、同様の判定を行ったところ犬の吠え声や、銃声のような音、子供の数瞬の声などを含んだ音源は衝撃騒音と判定された。選定した 24 の音源と設定したラウドネス、またこれに対応する音圧レベルについて表 7.1 に示す。

表 7.1: 音の快・不快の推定に利用した実環境音

音源	ラウドネス (sone)	音圧レベル (dB)
Forest wind	2	44.9
Rain	2	36.8
Windowchime	2	45.2
Cleaner	2	35.2
Laundry	2	38.6
Stryform	2	41.5
Bell	7	68.4
Bird 1	7	73.1
Little rain	7	50.8
Blackbord	7	51.8
Construction	7	51.4
Drill	7	50.9
Bird 2	14	70.2
Leaves	14	73.1
Riverbabbling	14	62.4
Chainsow	14	65.2
Folkscratch	14	82.8
Silen	14	63.7
Rin	21	78.9
Sea	21	68.9
Water	21	68.7
Jackhammer	21	71.0
Kinkyu jisin alert	21	74.3
Machine	21	70.2

次に、提案モデル評価のために表 7.1 で示す音源に対して主観評価値を取得する。評価値は各音源ごとに平均し、MOS (Mean Opinion Score) として集計した。MOS 実験には、正常聴力を有する日本語母語話者 20 名（大学院生 21~26 歳，男性 10 名，女性 10 名）が参加した。[58,59] より，7 段階リッカート尺度でのやかましさの印象調査の例において，180 人と 20 人での実験結果の間に良い相関があったことから，著者は評価尺度法において 20 人程度の被験者を推奨している。以上の理由から実験参加者の人数を設定した。実験系は，図 5.1 と同様である。

実験では，音刺激の快・不快を評価するために，五段階リッカート尺度を利用した。実験参加者には，刺激音の快・不快を 1：不快，3：どちらでもない，5：快の五段階で選択させた。なお，2 と 4 は，やや不快などの副詞による，実験参加者のバイアスが指摘されているため，教示文字は空欄とした [59]。各実験刺激は，1 秒間再生され，実験参加者が評価を入力した後，次の刺激が再生された。刺激音の提示順番はランダムに提示された。刺激は合計 24 であった。実験は概ね 60 分程度を要した。なお，実験参加者には，提示順による影響を抑えるため，直前・直後の音刺激の印象からできるだけ切り離して各音源を独立に評価するよう教示し，ヘッドフォンから再生される当該音そのものに対する快・不快の印象を回答するよう言及した。実験前に数回練習試行を行わせた。以上の MOS とリッカート尺度を利用する方法は，国際規格電気国際通信連合にて音声や製品の品質評価でも実際に利用されている [60–62]。

提案モデルによる予測では，以下のことが予想される。評価に利用する音源として快い音と不快な音を想定したそれぞれ 12 音源ずつを選定した。一方で，主観評価結果はその通りに二つの群に分かれるのではなく，ラウドネスに依り両者が混在すると予想される。このことから，ラウドネスを説明変数とし，その寄与度が大きい提案モデルとの相関は良い対応が現れることが期待される。

Start

1 / 24

再生された音の快・不快を五段階で評価してください

不快

どちらでもない

快い

1

2

3

4

5

○

○

○

○

○

次へ

被験者ID

19

図 7.1: GUI

最後に、表 7.1 に示した音源に対する、提案モデルを用いた推定結果と主観評価結果の関係を図 7.2 に示す。図の横軸は、快・不快の主観評価値、縦軸は、提案モデルによる推定値であり、24 刺激に対してそれぞれの評価と推定値が対応している。プロットの色はラウドネス（2, 7, 14, 21 sone）の大きさに対応しており、ラウドネスが大きいほど濃くなるよう 4 段階で示した。また、図中の破線はこの回帰直線に対する 95%信頼区間を表す。評価値と推定値に対して相関係数を計算した結果、相関係数の値は 0.75 であった。期待されたとおり、ラウドネスによる影響が強い傾向があり、高い相関が確認された。

さらに、モデル推定値と主観評価値間の RMSE を算出した。ここでは、両者の尺度が異なるため、モデル推定値を MOS 尺度に線形変換して計算した。なお、変換係数の推定と RMSE の算出を同じデータで行うことによる誤差の過小評価が指摘されているため、各サンプルを 1 つずつ除外して変換係数を推定し、除外したサンプルを予測する操作を全サンプルで繰り返した。得られた各サンプルの予測誤差から主観評価値と提案モデルによる予測値間の RMSE を算出した。主観評価値とモデル推定値間の RMSE は 0.66 であった。0.66 は 1 未満であることから、本尺度上の快・不快の値をまたがない。また、外れ値がないことから、良好な予測が出来ていることがわかる。以上の結果より、提案モデルは、相関係数、RMSE 共に良好な説明精度を示し、未知の音に対する快・不快を正確に予測可能であること、および、単一のモデルで音の快・不快を評価可能な新たな予測モデルとして優れていることが示された。

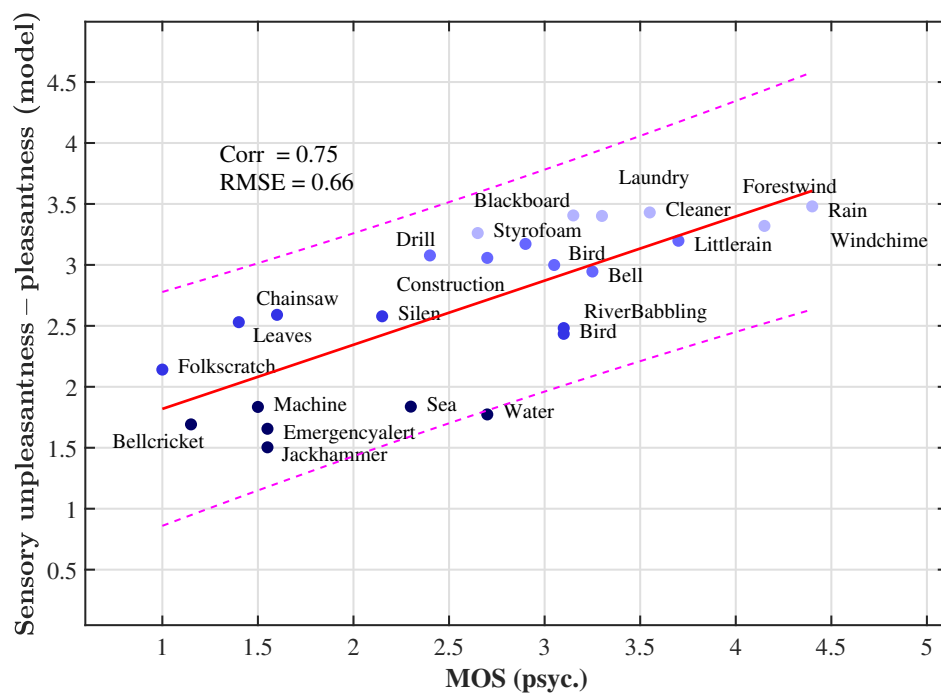


図 7.2: 実環境音に対する提案モデルの予測値と主観評価値の関係

7.2 感覚的快さの予測モデルとの比較評価

本節では、表 7.1 の音源に対する、提案モデルと第 2 節で示した感覚的快さの予測モデル [13, 18] の快・不快の説明精度を比較する。なお、感覚的快さの予測モデルは快・不快両方の評価を目的としたものではなく、音の快を評価するために構築された。このため、提案モデルのみが優れた説明精度を示すことが期待される。

表 7.1 で示した音源に対して、感覚的快さの予測モデル（式 (2.5)）を利用して予測した結果を図 7.3 に示す。図の横軸は、快・不快の主観評価値、縦軸は、各予測モデルによる推定値であり 24 の刺激に対するそれぞれの評価と推定値が対応している。図中の破線はこの回帰直線に対する 95%信頼区間を表し、各プロットの色は図 7.2 に対応する。各音質評価指標の計算は、第 4 節と同様の方法で行った。

図 7.3 では、推定値と評価値の間に明確な相関は観察されず、信頼区間幅も大きい。また、ばらつきも規則性がなく偏差が大きいことがわかる。ことから予測の不確かさが大きく説明精度は低いと考えられる。この傾向は、図 7.2 では確認されず、定性的には提案モデルのみ良好な説明精度を示したといえる。

主観評価値と推定結果に対して相関係数を計算した結果、感覚的快さの予測モデルは 0.46 であった。提案モデルよりも、0.20 小さい値となった。さらに、主観評価値とモデル推定値間の RMSE を前節と同様の方法で算出した。感覚的快さ予測モデルの RMSE は、0.88 であり、提案モデルよりも 0.2 以上大きくなった。比較結果として、提案モデルの相関係数が高く、RMSE が小さくなった。

一方で、式 (2.5) ではラウドネスに関わる項が二乗されており、その影響が予測において大きいことがわかる。実際に図 7.3 では、縦軸に着目すると、ラウドネスの傾向は凡そ捉えられているように見える。

以上の結果より、期待されたとおり、感覚的快さの予測モデルは、主観評価値と予測結果間に高い相関があるとはいえない。また、RMSE は、大幅に提案モデルよりも大きくなったことから提案モデルと比較して未知の音に対する快・不快を良い精度で予測できていないことがわかった。さらに、予測におけるラウドネスの寄与は提案モデル、感覚的快さの予測モデルにおいて共に大きいことが、両者の予測精度に大きな差が見られたことから、その他の音質評価指標の重要性が示唆された。

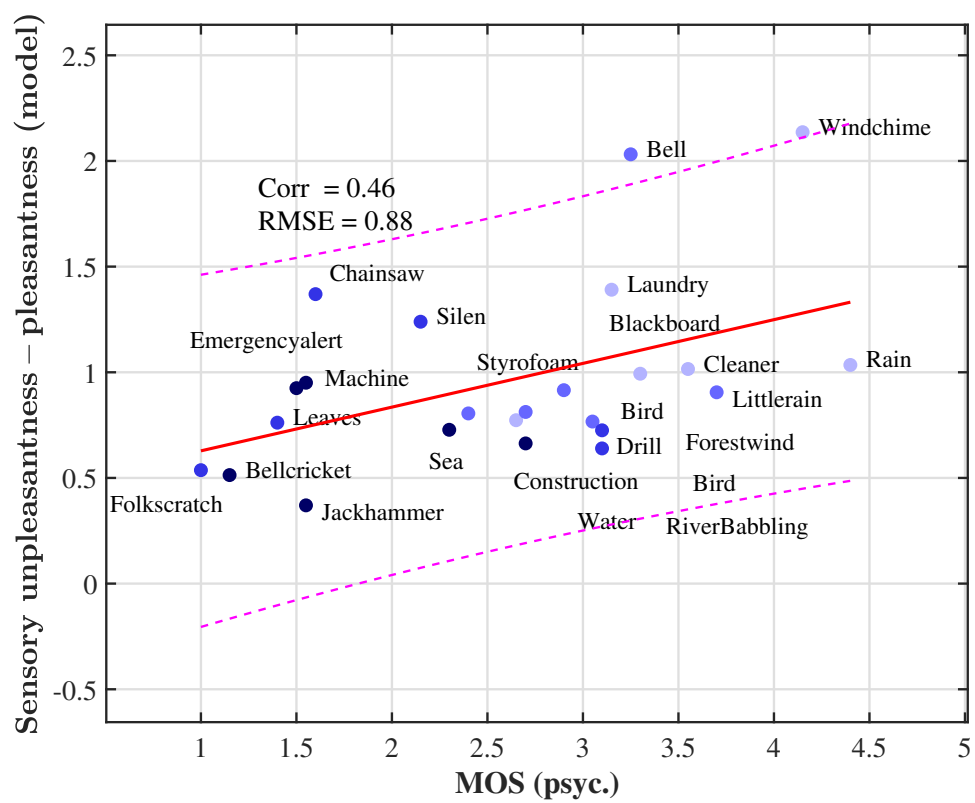


図 7.3: 実環境音に対する感覚的快さの予測値と主観評価値の関係

7.3 アノイアンスモデルとの比較評価

本節では、表 7.1 の音源に対する、提案モデルと第 2 節で示した、アノイアンスモデル [14] の快・不快の説明精度を比較する。なお、アノイアンスモデルは快・不快両方の評価を目的としたものではなく、音の不快を評価するために構築されたモデルである。このため、前節と同様に、快・不快の予測において提案モデルのみが優れた説明精度を示すことが期待される。

表 7.1 で示した音源に対して、アノイアンスモデル（式 (2.6)）を利用して予測した結果を図 7.4 に示す。図の横軸は、快・不快の主観評価値、縦軸は、各予測モデルによる推定値であり 24 の刺激に対するそれぞれの評価と推定値が対応している。図中の破線はこの回帰直線に対する 95%信頼区間を表し、各プロットの色の濃さは図 7.2 に対応する。各音質評価指標の計算は、第 4 節と同様の方法で行った。

図 7.4 では、モデル推定値が下限 (0 付近) で飽和している傾向があること、ならびに左上に大きな外れ値が存在することが確認できる。以上の傾向は、図 7.2 では確認されず、定性的には提案モデルのみ良好な説明精度を示したといえる。

主観評価値と推定結果に対して相関係数を計算した結果、アノイアンスモデルは -0.72 であり、提案モデルよりも 0.03 小さくなった。なお、アノイアンスモデルの相関係数が負となるのは、出力が不快の大きさを表す一方、主観評価値は値が大きいほど快を表すためであり、ここでは符号は考慮しないものとする。さらに、主観評価値とモデル推定値間の RMSE を前節と同様の方法で算出した。アノイアンスモデルの RMSE は、 0.86 であり、提案モデルよりも 0.20 以上大きくなった。比較結果として、提案モデルの相関係数が高く、RMSE が小さくなった。

ここでは、感覚的快さの予測モデルとは逆に、式 (2.6) ではラウドネスが大きい時、シャープネスに関わる重み係数も大きくなるような計算方法がとられるため、ヒトの評価よりも過度に大きな値が出力され外れ値、および飽和部分が発生したことが考えられる。

以上の結果より、期待されたとおり、アノイアンスモデルは、主観評価値と予測結果間に高い相関があるものの大きな外れ値やモデル出力の下限部の飽和現象が見られた。また、RMSE は、大幅に提案モデルよりも大きくなったことから、提案モデルと比較して未知の音に対する快・不快を良い精度で予測できていないことがわかった。

7.2 節、7.3 節の比較結果より、期待されたとおり提案モデルは、音の快・不快の評価において唯一高い予測精度を示すことがわかった。本稿では、快・不快の評価のため感覚的快さの予測モデルおよびアノイアンスモデルで利用される五つの音質評価指標をすべて利用して予測モデルを構築した。未知の音に対する快・不

快の推定において、本モデルのみの予測精度が優れていたことから、提案モデルの整合性が確認され、快・不快の評価において、五つの音質評価指標（ラフネス、シャープネス、トナリティ、ラウドネス、変動強度）を統合的に考慮することが重要であることが示唆された。

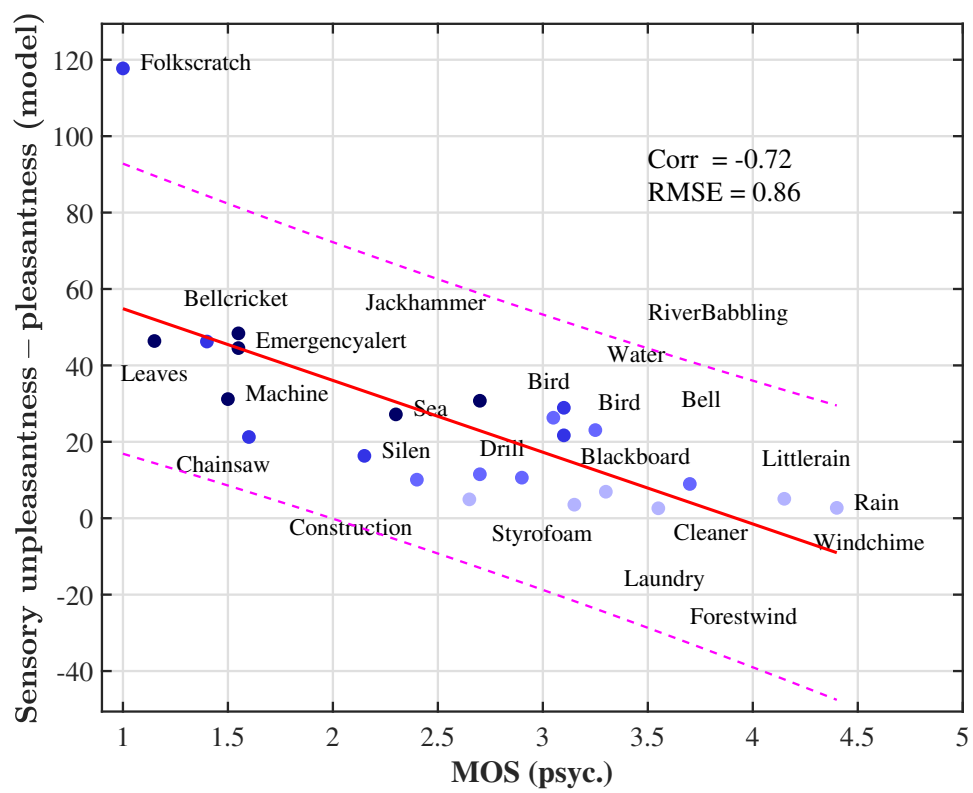


図 7.4: 実環境音に対するアノイアンスモデルの予測値と主観評価値の関係

7.4 考察

本節では、ERB 尺度を利用して計算された音質評価指標を、提案モデルへ組み込み、その予測精度に問題がないかを確かめる。

以上では、先行研究との比較を可能にするため、Bark 尺度に基づく音質評価指標の計算モデルを用いてきた。一方で、本研究では、第2章で述べたとおり、Bark 尺度よりも聴覚特性の再現性が高いとされる ERB 尺度に基づく指標設計が有効であるという立場をとる。そこで、磯山ら [39] が構築した GTFB および GCFB に基づく音質評価指標を本研究で構築した提案モデルに組み込む。ただし、これまでの検討で用いてきた音質評価指標の計算モデルは時間変動 (time vary method) であり、磯山ら [39] が構築した音質評価指標の計算モデルでは、入力された信号を定常音として計算する。そのため、両者を直接比較することは難しく、本節では、比較は行わず、予測精度に問題ないか確認のみ行う方針とする。

提案モデルにおける係数推定の流れは第3章と同様である。以下に、それぞれの推定した係数を示す。GTFB を利用した音質評価指標を用いた場合、各係数は表 7.2 に示すとおり値となった。

GTFB を利用した音質評価指標を用いた場合、各係数は表 7.3 に示すとおり値となった。両者とも、ラウドネスに関わる係数が小数点第4位以下と小さい値となった。また、全ての音質評価指標に関わる係数が負の値となっていることがわかる。

これらを用いて、表 7.1 に対する予測を行った結果を、それぞれ GTFB を利用した音質評価指標を用いた場合を図 7.5 に、GCFB を利用した音質評価指標を用いた場合を図 7.6 に示す。それぞれ、図の横軸は、快・不快の主観評価値、縦軸は、GTFB および、GCFB を利用して計算した音質評価指標を組み込んだ提案モデルによる推定値であり、24 刺激に対してそれぞれの評価と推定値が対応している。プロットの色はラウドネス (2, 7, 14, 21 sone) の大きさに対応しており、ラウドネスが大きいほど濃くなるよう4段階で示した。また、図中の破線はこの回帰直線に対する 95%信頼区間を表す。定性的な分析では、両者ともに外れ値はなく、全ての予測値が 95%信頼区間に収まっていることが分かる。また、GTFB を利用した場合よりも、GCFB を用いた場合のほうがより信頼区間の幅が狭まっていることが分かる。これは、GCFB の音圧依存性を考慮している点がより主観評価値と対応がとれていることを示唆していると考えられる。

評価値と推定値に対して相関係数を計算した結果、GTFB を利用した音質評価指標を用いた場合は 0.66、GCFB を利用した音質評価指標を用いた場合は 0.77 であった。両者ともに高い相関が確認され主観評価値と良い対応が取れていること

がわかった。

さらに、モデル推定値と主観評価値間の RMSE を算出した。ここでは、前章と同様の方法で両者の尺度が異なるため、モデル推定値を MOS 尺度に線形変換して計算した。主観評価値とモデル推定値間の RMSE は、GTFB を利用した音質評価指標を用いた場合は 0.75、GCFB を利用した音質評価指標を用いた場合は 0.64 であった。両者ともに、RMSE が 1 未満であることから、本尺度上の快・不快の値をまたがない。RMSE について、より GCFB を利用した場合の方が小さい理由は、上述と同様、音圧依存性による影響であると考えられる。以上の結果より、ERB 尺度を利用して計算された音質評価指標を提案モデルへ組み込んだ際に、予測精度に問題ないことがわかった。特に、音圧依存性を考慮した GCFB を利用して計算した音質評価指標を組み込んだ場合に、より予測精度が向上することがわかった。

表 7.2: 提案モデル（式 (6.1)）における GTFB を利用した場合の各係数

係数	値
$\alpha_{p,0}$	1.39
$\alpha_{p,1}$	-0.222
$\alpha_{p,2}$	-0.0595
$\alpha_{p,3}$	-0.152
$\alpha_{p,4}$	-0.000100
$\alpha_{p,5}$	-0.0363

表 7.3: 提案モデル（式 (6.1)）における GCFB を利用した場合の各係数

係数	値
$\alpha_{p,0}$	1.40
$\alpha_{p,1}$	-0.288
$\alpha_{p,2}$	-0.0731
$\alpha_{p,3}$	-0.202
$\alpha_{p,4}$	-0.000700
$\alpha_{p,5}$	-0.0520

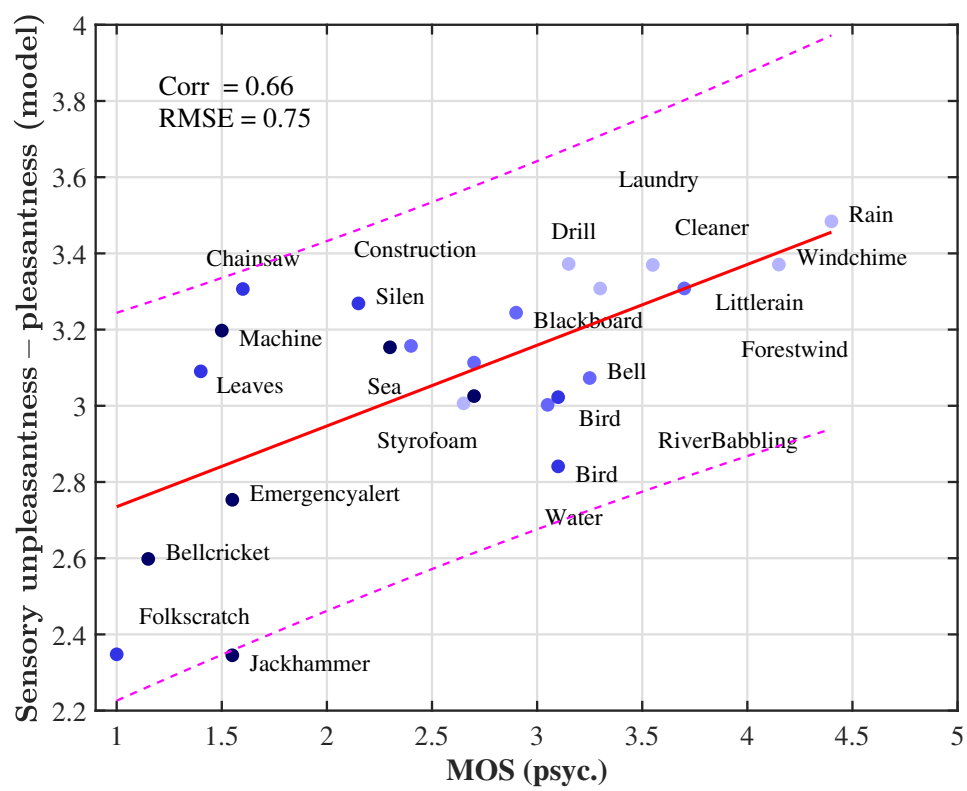


図 7.5: 実環境音に対する予測値（GTFB 利用）と主観評価値の関係

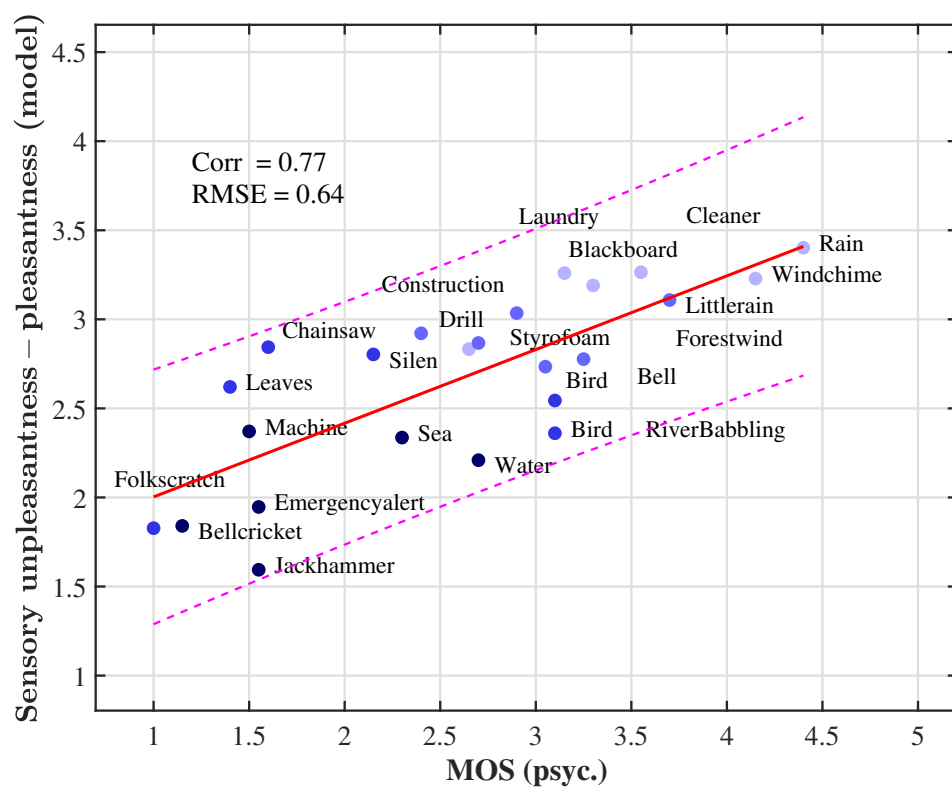


図 7.6: 実環境音に対する予測値（GCFB 利用）と主観評価値の関係

第8章 結論

8.1 明らかにしたこと

本研究では、音質評価指標を用いた音の快・不快の予測モデルの構築を行った。特に、図 3.1 に示す以下の四つのリサーチクエスチョンを設け、順序立てて検討を行った。具体的には、まず、先行研究で得られた感覚的快さと感覚的快くなさの主観評価値を用い、両者が同じ軸上にあるかを定量的に分析した。同じ軸上で表せる場合、その関係を快・不快とみなした。ただし、刺激条件の範囲が十分ではないことが分かったため、次の検討に進んだ。次に、感覚的快さの再評価を行い、快・不快の評価に十分な刺激条件の範囲であったかを検証した。同じ軸上に快・不快があることに基づき感覚的快さのみを評価し、主観評価値の折り返し部を不快と仮定して、再評価結果を快・不快の主観評価値として扱った。さらに、得られた快・不快の主観評価値を用いて五つの音質評価指標を利用した音の快・不快の予測モデルを提案した。また、快・不快の主観評価値を用いて提案モデルを客観的に評価した。最後に、提案モデルに未知の音を入力し、主観評価値をどの程度の精度で予測できるかを検証した。その結果、次のことを明らかにした。

Q.1 音の感覚的快さ・快くなさは、同じ軸上で表すことができるのか。

A.1 音の感覚的快さ・快くなさは、同じ軸上の対極の関係にあること。

Q.2 快・不快の評価を行うために十分な刺激条件の範囲を設けられたのか。

A.2 快・不快の評価を行うために十分な刺激条件の範囲を設定できたこと。

Q.3 提案モデルは、快・不快の評価結果を説明することができるのか。

A.3 提案モデルが、快・不快の評価結果を高い精度で説明できること。

Q.4 提案モデルは、未知の音に対する快・不快の評価に有効なのか。

A.4 提案モデルが、未知の音の快・不快を正確に予測可能であること。

8.2 残された課題

残された課題として以下の三点が挙げられる．

- 不快に関する聴取実験

本研究では，先行研究のデータの定量的分析の結果を踏まえ，感覚的快さの評価のみ行った．一方，6.3 節の調査では変動強度の影響が確認できなかったため，同一刺激条件に対して不快に関する聴取実験も実施し，快・不快の関係をさらに検討する余地がある．また，本刺激条件の設計では先行研究に倣い Bark 尺度を用いた計算モデルを利用したが，本研究では ERB 尺度を用いた計算モデルでも問題がないことを確認できた．そのため，次の実験では刺激設計に用いる計算モデルを慎重に検討する必要がある．

- 関数モデルの形式についての検討

提案モデルに用いた関数モデルの形式について，今回は感覚的快さの予測モデルに倣った形式でのみ検討を行った．一方で，一次関数に基づく形式や，アノイアンスモデルのように比例・反比例を取り入れた形式など，複数の関数形が考えられる．そのため，今後は他の関数モデルも含めて比較検討を進めることで，予測精度の向上に繋がる可能性がある．

- 快音化についての検討

本研究の目的は音の快・不快の予測モデルの構築であり，提案モデルを用いた快音化については検討できていない．今後，実用化に向けた次のステップとして，提案モデルを利用した快音化についての検討が期待される．

付 録 A 感覚的快くなさの推定

ここでは、谷口ら [19] の報告で検討した Aures に基づく回帰モデルが、感覚的快くなさを説明できるかを検討する。

先の報告 [19] で検討した、四つの音質評価指標に基づく感覚的快さ W_p を推定する回帰モデルは、次式のように構成される。

$$\log(\hat{W}_p) = \beta_{p,0} + \beta_{p,1}R + \beta_{p,2}S + \beta_{p,3}K + \beta_{p,4}N^2 \quad (\text{A.1})$$

感覚的快くなさ W_u の推定では、式 (A.1) と同じ形式で係数のみ調節した次式の形式で構成される。

$$\log(\hat{W}_u) = \beta_{u,0} + \beta_{u,1}R + \beta_{u,2}S + \beta_{u,3}K + \beta_{u,4}N^2 \quad (\text{A.2})$$

ただし、主観評価値 W_u には負の値が含まれるため、そのままでは対数スケールでの計算が行えない。そこで、次式のように、 W_u 全体をシフトさせ、すべての値が正となるように処理を行う。具体的には、最小値の絶対値に 1 を加えた量を全体に加算し、対数スケール上で 0 から始まるように変換する。各指標の計算は 4.3 節と同様である。

$$W_u + |\min(W_u)| + 1 \quad (\text{A.3})$$

次に、式 (A.2) の係数 $\beta_{u,k}$ ($k = 0, 1, \dots, 4$) を、次式の二乗平均平方根誤差 (RMSE) が最小になるように推定する。ここで、得られた係数を表 A.1 に示す。

$$\text{RMSE} = \frac{1}{I} \sum_{i=1}^I |\log(W_{u,i}) - \log(\hat{W}_{u,i})|^2 \quad (\text{A.4})$$

式 (A.2) の両辺に対して指数変換を行い、元のスケールで得られた感覚的快くなさの推定結果を図 A.1(a) に示す。図の横軸は聴取実験によって得られた感覚的快くなさの主観評価値を表し、縦軸は式 (A.2) を用いて推定した感覚的快くなさの推定値を表す。60 刺激に対してそれぞれの評価結果が対応している。また、図中の破線はこの回帰直線に対する 95% 信頼区間を表す。実験結果と推定結果の推定誤差を確認した結果、RMSE は 0.177 であった。0.177 は、感覚的快くなさのデータ

の範囲 $[-0.765, 1.143]$ (全体 1.908) の 10%以下である。この結果から、音質評価指標を利用した回帰モデルが主観評価結果をよく説明できていることがわかる。

式 (A.2) のモデルにおいて、各音質評価指標がどれだけ感覚的快くなさの推定に寄与しているかを調べるために、(1) ラフネスを利用しない場合 ($\beta_{u,1} = 0$)、(2) シャープネスを利用しない場合 ($\beta_{u,2} = 0$)、(3) トナリティを利用しない場合 ($\beta_{u,3} = 0$)、(4) ラウドネスを利用しない場合 ($\beta_{u,4} = 0$) での推定精度を検討した。ここでは、すべての音質評価指標を利用したモデルの感覚的快くなさの推定結果に対して、四つの条件での RMSE を調べることで、それぞれの音質評価指標が感覚的快くなさにどれだけ寄与しているかを評価した。

これら四つの条件における推定結果を図 A.1(b)、図 A.1(c)、A.1(d)、図 A.1(e) に示す。結果として、ラウドネスのみを利用しない場合、RMSE は、0.177 から 0.451 まで増加した。次いでラフネスのみ利用しない場合、RMSE は、0.177 から 0.203 まで増加した。残るシャープネスとトナリティに関しては、変化がなかった。これらのことから、今回の主観評価結果については、ラウドネスとラフネスのみ利用したモデルでも十分に説明できることがわかる。

そこで、四つの条件に加えて寄与度の高い上位二つの指標のみ利用することで、どの程度感覚的快くなさを正しく説明出来るか評価した。結果を図 A.1(f) に示す。RMSE は全ての音質評価指標を利用した推定精度と同等の値となった。予想されたとおり、ラウドネスとラフネスを利用した回帰モデルで十分に結果を説明できることがわかった。

表 A.1: Aures のモデルに基づく感覚的快くなさの予測モデル (式 (A.4)) の各係数

係数	値
$\alpha_{p,0}$	0.2051
$\alpha_{p,1}$	0.142
$\alpha_{p,2}$	-0.0245
$\alpha_{p,3}$	-0.0290
$\alpha_{p,4}$	0.00290

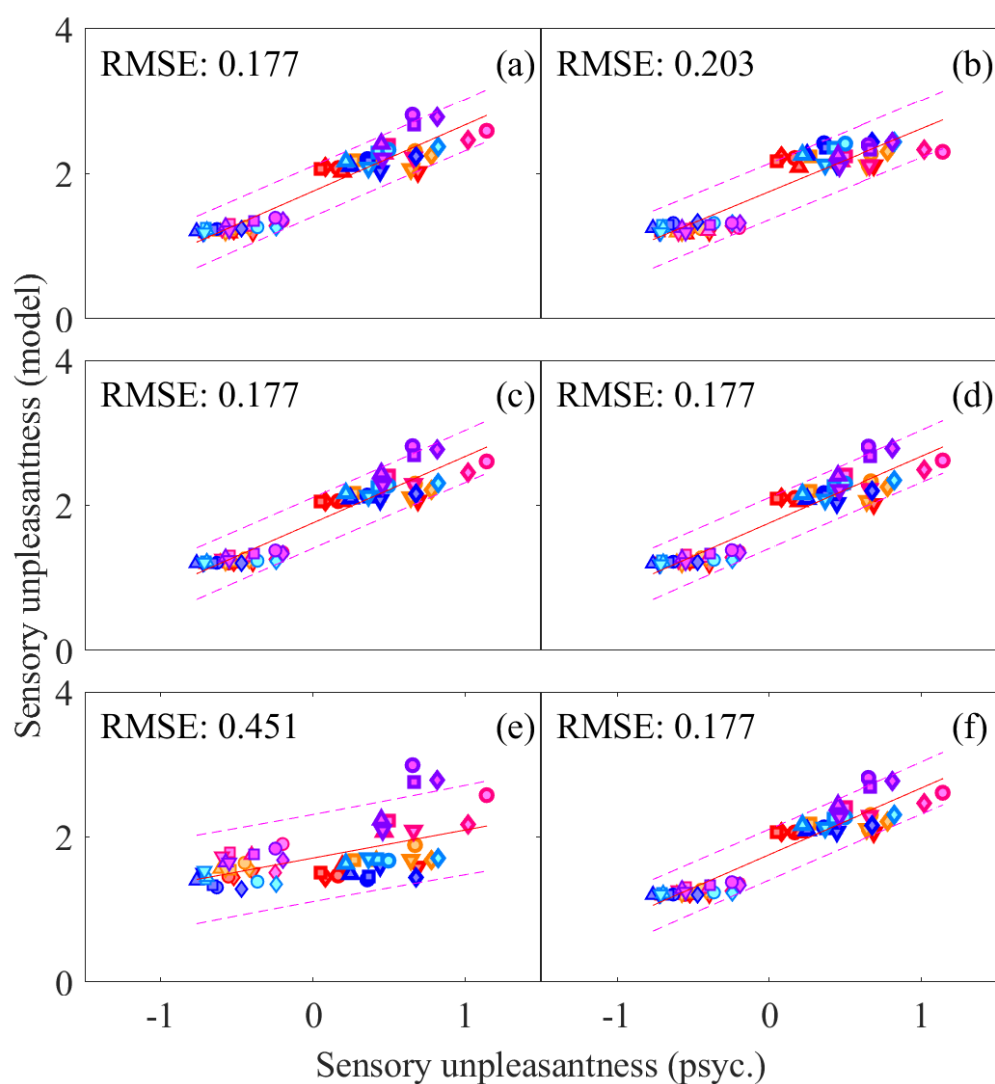


図 A.1: 音の感覚的快くなさのモデルによる感覚的快くなさの推定結果: (a) すべての SQMs を利用した場合, (b) ラフネスを利用しなかった場合, (c) シャープネスを利用しなかった場合, (d) トナリティを利用しなかった場合, (e) ラウドネスを利用しなかった場合, (f) ラウドネスとラフネスのみ利用した場合. 各プロットは, 図 5.4 のプロットおよび凡例の刺激に対応する.

付 録 B 音の感覚的快さ・快くなさ の予測モデルによる実環境 音の予測

本章では、先行研究の主観評価値を 4.2 で構築した音の感覚的快さ・快くなさの予測モデル（式 (4.4)）を用いて、表 7.1 に示す音源に対して予測を行った結果を示し、提案モデルと比較する。予測結果を図 B.1 に示す。図の横軸は、快・不快の主観評価値、縦軸は、感覚的快さ・快くなさの予測モデルによる推定値であり、24 の刺激に対するそれぞれの評価と推定値が対応している。図中の破線はこの回帰直線に対する 95%信頼区間を表し、各プロットの色は図 7.2 に対応する。定性的な分析では、外れ値はなく、全ての予測値が 95%信頼区間に収まっており、提案モデルとの大きな差異はみられない。

評価値と予測値に対して相関係数を計算した結果、相関係数の値は 0.73 であった。提案モデルの予測における相関値は 0.75 であり、僅かに提案モデルのほうが優れている。一方、第 7.2 節より、感覚的快さの予測モデルの予測における相関値は 0.46、アノイアンスモデルの予測における相関値は -0.72 であり、これらと比較すると、感覚的快さ・快くなさの予測モデルの相関係数は提案モデルに次いで高く、感覚的快さの予測モデルより大きい値となった。また、アノイアンスモデルと感覚的快さ・快くなさの予測モデルと同程度であることがわかった。

モデル推定値と主観評価値間の RMSE を算出した。ここでは、両者の尺度が異なるため、7.1 章と同様の方法でモデル推定値を MOS 尺度に線形変換して計算した。主観評価値とモデル予測値間の RMSE は 0.68 であった。提案モデルの予測における RMSE は 0.66 であるため、僅かに提案モデルのほうが優れている。さらに、第 7.2 節より、感覚的快さ予測モデルの予測における RMSE は 0.88、アノイアンスモデルの予測における RMSE は 0.86 であり、それぞれ感覚的快さ・快くなさの予測モデルよりも 0.20 以上大きくなった。以上の比較より、相関係数は提案モデルが最も高く、RMSE は提案モデルが最も小さいことから提案モデルが最も優れた説明精度を示したことがわかった。

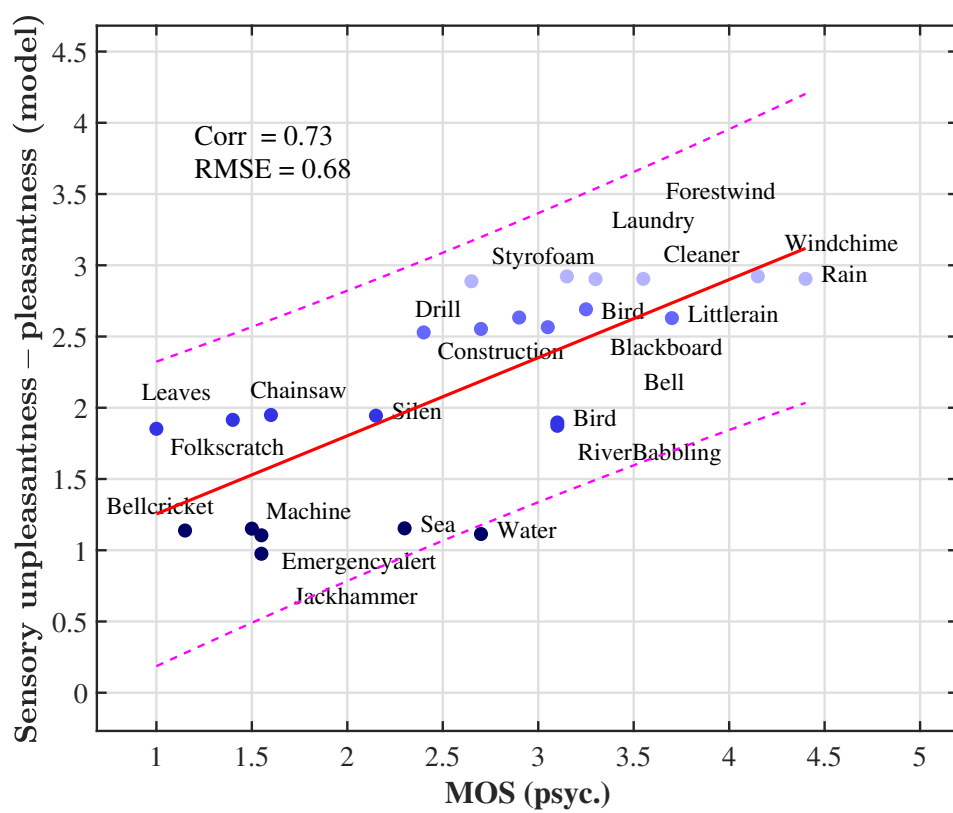


図 B.1: 実環境音に対する感覚的快さ・快くなさの予測モデルとの比較評価

謝辞

大学院生としての研究を行うにあたり，多大なるご指導，及びご鞭撻をいただいた鵜木祐史教授に深く感謝いたします。また，木谷俊介先生，上江洲安史先生には，研究を行うにあたり重要なアドバイスを，行き詰まった時には気分転換の機会もいただきました。心より感謝いたします。加えて，ラボメンバーはいつでも快く，様々な面で支えていただきました。大変感謝しております。

さらに，長期間にわたる実験に不快な顔をせず参加いただいた実験参加者のみなさまにも，誠に感謝しております。

最後に，家族の皆がこれまでの生活を暖かく支えて続けてくれたことを幸せに感じます。この経験を活かし，恩を返せるよう努めてまいります。

参考文献

- [1] 難波精一郎, “音の嗜好: 個人差と文化による差,” 音響誌, Vol. 46(9), 776–780, 1990.
- [2] 島井哲志, 田中正敏, “環境音の快–不快評価と音圧の関係,” 音響誌, Vol. 49(4), 243–252, 1993.
- [3] 古川茂人, 聴覚, コロナ社, 2021.
- [4] 橘秀樹, “環境騒音の評価指標,” 騒音制御, Vol. 31(6), 421–425, 2007.
- [5] 矢野隆, “騒音のアノイアンス,” 音響誌, Vol. 77(8), 529–536, 2021.
- [6] 三宅 晋司, 田中 豪一, 斎藤 和雄, “不快音の脳波に及ぼす影響,” 衛生誌, Vol. 39(2), 523–534, 1984.
- [7] 高田正幸, “音質評価指標の計算法と適用事例,” 音響誌, Vol. 75(10), 582–589, 2019.
- [8] 岩宮眞一郎, “音質評価指標-入門とその応用-,” 音響誌, Vol. 66(12), 603–609, 2010.
- [9] 鈴木陽一, 赤木正人, 伊藤彰則, 佐藤 洋, 荳木 禎史, 中村 健太郎, 音響学入門, コロナ社, 2011.
- [10] 岩宮眞一郎, 音色の感性学, コロナ社, 2010.
- [11] 平松幸三, 小林聡, 松井利仁, 高木興一, 山本剛夫, “生活環境音を用いた大きさ・うるささ・不快感の評定実験,” 音響誌, Vol. 44(5), 1988.
- [12] 難波精一郎, “精神物理学と音色,” 日本学士院紀要, Vol. 71(2), 53–83, 2018.
- [13] von Aures, W, “A model for calculating the sensory euphony of various sound,” Acustica, Vol. 54, 130–141, 1985.

- [14] Zwicker, E., Fastl, H., *Psychoacoustics: Facts and models*, 3rd ed., Springer, Berlin, Germany, 1999.
- [15] Jae, K. L., Jaewoong, J., Seo, I. C., Soo, I. L., “Annoyance modeling using personal and situational variables for construction site noise in urban areas,” *Applied Acoustics*, Vol. 182, 108256, 2021.
- [16] Guo-Qing, A., Xing-Wang C., Bing, Z. b., Chun-Ming, P. b, “Improvement of Zwicker ’ s psychoacoustic annoyance model aiming at tonal noises,” *Applied Acoustics*, Vol. 105, 164–170, 2016.
- [17] 古川茂人, 高橋宏和, 柏野牧夫, “聴覚的顕著性の操作に基づく、音響情報の選択的強調技術の創出,” *NTT テクニカルジャーナル*, Vol. 13(11), 2015.
- [18] von Aures, W, “Der sensorische Wohlklang als Funktion psychoakustischer EmpfindungsgroBen,” *Acustica*, Vol. 58, 282–290, 1985.
- [19] 谷口亮太郎, 磯山拓都, 上江洲安史, 木谷俊介, 鶴木祐史, “音質評価指標を用いた感覚的快さの評価に関する検討,” *音響学聴覚研資*, Vol. 54(4), 423–428, 2024.
- [20] 松尾洋杜, 磯山拓都, 谷口亮太郎, 上江洲安史, 木谷俊介, 鶴木祐史, “音質評価指標を用いた感覚的快さの評価に関する検討,” *音響学聴覚研資*, Vol. 55(4), 191–196, 2025.
- [21] Wang, J., Wang, X., Yuan, M., Hu, W., Hu, X., Lu, K., “Deep Learning-Based Road Traffic Noise Annoyance Assessment,” *Int. J. Environ. Res. Public Health*, Vol. 20(6), 5199, 2023, doi: 10.3390/ijerph20065199.
- [22] Axelsson, Ö., Nilsson, M. E., Berglun, B., “A principal components model of soundscape perception,” *J.Acoust.Soc.Am*, Vol. 128(5), 2836–2846, 2010, doi: 10.1121/1.3493436.
- [23] Lionello, M., Aletta, F., Kang, J., “A systematic review of prediction models for the experience of urban soundscapes,” *Applied Acoustics*, Vol. 170, 107479, 2020, doi: 10.1016/j.apacoust.2020.107479.

- [24] Rudin, C., “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead,” *Nature Machine Intelligence*, Vol. 1(5), 206–215, 2019, doi: 10.1038/s42256-019-0048-x.
- [25] World Health Organization, “Environmental noise,”
in *Compendium of WHO and other UN guidance on health and environment, 2022 update*, Chapter 11, Geneva, 2022 (WHO : HEP22.01).
- [26] International Organization for Standardization, “Acoustics — Description, measurement and assessment of environmental noise — Part 2: Determination of environmental noise levels,” Edition 2, Geneva, 2007.
- [27] 橋本竹夫, “機械騒音の音質改善の手法,” *音響誌*, Vol. 53(6), 450–455, 1997.
- [28] 橋本竹夫, “自動車における音質評価,” *騒音制御*, Vol. 16(1), 4–7, 1992.
- [29] 伊藤瑠美, 若林佑幸, 福森隆寛, 中山雅人, 西浦敬信, “能動騒音制御と聴覚マスキングに基づく工場騒音の不快感低減,” *音響学誌*, Vol. 73(4), 226–229, 2017.
- [30] F.-G. Zeng, A unified law for sensation and perception: Weber–Fechner and Stevens’ laws, *PNAS*, 2020.
- [31] Fletcher, H, “Auditory patterns,” *Rev. Mod. Phys.*, Vol. 12, 47–65, 1940.
- [32] Moore, B., Glasberg, B., “Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns,” *Hearing Research*, Vol. 28(2), 209–225, 1987.
- [33] Moore, B., Glasberg, B., “Derivation of auditory filter shapes from notched-noise data,” *Hearing Research*, Vol. 47(1), 103–138, 1990.
- [34] Irino, T., Unoki, M., “A time-varying, analysis/synthesis auditory filterbank using the gammachirp,” In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP ’98* (Cat. No.98CH36181), Vol. 63, 3653–3656, 1998.
- [35] Irino, T., Unoki, M., “An analysis/synthesis auditory filterbank based on an iir implementation of the gammachirp,” *Journal of the Acoustical Society of Japan (E)*, Vol. 20(6), 1999.

- [36] 赤木正人, “聴覚フィルタとそのモデル,” IEICE, Vol. 77(9), 948–956, 1994.
- [37] Greenwood, D, “Critical Bandwidth and the Frequency Coordinates of the Basilar Membrane,” *Acustica*, Vol. 33(10), 1344–1356, 1961.
- [38] Greenwood, D, “A cochlear frequency - position function for several species – 29 years later,” *Acustica*, Vol. 87(6), 2592–2605, 1990.
- [39] 磯山拓都, 木谷俊介, 鵜木祐史, “聴覚フィルタバンクを用いた音質評価指標の計算モデルの応用,” *音響学論文集*, Vol. 152, 801–804, 2024.
- [40] 谷口亮太郎, “音質評価指標を用いた感覚的快さと感覚的快くなさの評価に関する調査,” 北陸先端科学技術大学院大学 課題論文, 2024.
- [41] Greco, F. G., Merino, M. R., Osses, A., “ggrecow/SQAT,” <https://github.com/ggrechow/SQAT> (2025 年 11 月アクセス).
- [42] Daniel, P., Weber, R., “Psychoacoustical Roughness: Implementation of an Optimized model,” *Acustica*, Vol. 83(1), 113–123, 1997.
- [43] ISO 532-1, “Acoustics – Methods for calculating loudness level,” ISO 532-1:1975(E), 1975.
- [44] Osses, A., García, R., Kohlrausch, A., “Modelling the sensation of fluctuation strength,” *Proc. Mtgs. Acoust.*, 28(1), 050005, 2016.
- [45] Salamon, J., Jacoby, C., Bello, J. P., “A dataset and taxonomy for urban sound research,” *Proc. ACM Int. Conf. Multimedia*, Vol. 22, 1041–1044, 2014.
- [46] Piczak, K. J., “ESC: Dataset for Environmental Sound Classification,” *Proc. ACM Int. Conf. on Multimedia*, Vol. 23, 1015–1018, 2015.
- [47] Ooi, K., Ong, Z. T., Watcharasupat, K. N., Lam, B., Hong, J. Y., Gan, W. S., “Large-Scale Dataset and Baseline models of Affective Responses to Augmented Urban Soundscapes,” *IEEE Transactions on Affective Computing*, Vol. 15(1), 105–120, 2024.
- [48] Ooi, K., Ong, Z. T., Watcharasupat, K. N., Lam, B., Hong, J. Y., Gan, W. S., “Large-Scale Dataset and Baseline models of Affective Responses to Aug-

- mented Urban Soundscapes,” <https://doi.org/10.21979/N9/9OTEVX> (2025 年 10 月アクセス).
- [49] Autumn, T, “Nails on Chalkboard,” Freesound, <https://freesound.org>. (2025 年 10 月アクセス).
- [50] humanoide9000, “Fork scraping metal 1,” Freesound, <https://freesound.org>. (2025 年 11 月アクセス).
- [51] SDLx, “Bowed-Styrofoam-5.aiff,” Freesound, <https://freesound.org>. (2025 年 11 月アクセス).
- [52] Robinhood76, “09285sea calm waves loop,” Freesound, <https://freesound.org>. (2025 年 11 月アクセス).
- [53] SterckxS, “Relaxing wind chime,” Freesound, <https://freesound.org>. (2025 年 11 月アクセス).
- [54] electronixtech, “babbling-brook,” Freesound, <https://freesound.org>. (2025 年 11 月アクセス).
- [55] 大分県臼杵市役所臼杵庁舎, 防災行政無線.
- [56] 富山県市公式 web サイト, 防災行政無線.
- [57] Towsey, M., Parsons, S., Sueur, J., “Ecological acoustic signal processing tools (Ecosounds),” *Ecological Informatics*, Vol. 21, 110–119, 2014.
- [58] 大串清宏, 音響聴覚心理学, 誠信書房, 2019.
- [59] 難波精一郎, 桑野園子, 音の評価のための心理学的測定法, コロナ社, 1998.
- [60] ITU-T, “Methods for subjective determination of transmission quality,” Recommendation P.800, International Telecommunication Union, Geneva, 1996.
- [61] ITU-R, “General methods for the subjective assessment of sound quality,” Recommendation ITU-R BS.1284-2, International Telecommunication Union, Geneva, 2019.

- [62] Çakır A. D, Yılmaz S, “ Assessment of sound environment pleasantness by sound quality metrics in urban spaces, ” ITU A—Z, Vol. 13(2), 87–99, 2016.

研究業績

口頭・査読無し

1. 松尾洋杜, 磯山拓都, 谷口亮太郎, 上江洲安史, 木谷俊介, 鵜木祐史, “音質評価指標を用いた感覚的快さの評価に関する検討,” 日本音響学聴覚研究会資料, Vol. 55(4), 191–196, 2025.
2. 松尾洋杜, 磯山拓都, 谷口亮太郎, 上江洲安史, 木谷俊介, 鵜木祐史, “音質評価指標を用いた感覚的快さ・快くなさの一元評価,” 電気・情報関係学会北陸支部連合大会, H1-1, 2025.
3. 松尾洋杜, 上江洲安史, 木谷俊介, 鵜木祐史, “音の快・不快を予測するための評価実験とその結果に基づく予測モデルの検討,” 日本音響学会春季研究発表会, 2026(発表予定).
4. 松尾洋杜, 磯山拓都, 谷口亮太郎, 上江洲安史, 木谷俊介, 鵜木祐史, “音の快・不快の予測モデルの構築とその評価に関する検討,” 日本音響学聴覚研究会, 2026(発表予定).

受賞

1. 電気・情報関係学会北陸支部連合大会, 音響部門優秀発表賞, 2025.