

博士論文

ヴァイオリンの演奏と歌声の声質に関する評価に利用可能な
音楽音響パラメータの設計

国立音楽大学大学院音楽研究科
音楽研究専攻 音楽学研究領域

桶本まどか

博士論文

ヴァイオリンの演奏と歌声の声質に関する評価に利用可能な
音楽音響パラメータの設計

国立音楽大学大学院音楽研究科
音楽研究専攻 音楽学研究領域

20GD0004 桶本まどか

目次

1 章 緒論.....	- 7 -
2 章 本論文で用いる基礎知識及び研究背景.....	- 9 -
2.1 緒言.....	- 9 -
2.2 感性の扱い.....	- 9 -
2.3 本論文の目的.....	- 10 -
2.4 本論文の方針及び概要.....	- 11 -
2.5 音楽メディアのデジタル化.....	- 12 -
2.5.1 音楽メディアのデジタル化による音楽聴取方法の変遷.....	- 12 -
2.5.2 音楽メディアの購入方法における時代変遷.....	- 13 -
2.5.3 音楽情報処理との関連.....	- 16 -
2.6 人工知能.....	- 18 -
2.6.1 人工知能とは.....	- 18 -
2.6.2 本論文における人工知能の定義.....	- 19 -
2.6.3 学習方法の分類.....	- 19 -
2.6.4 音楽における人工知能の利用.....	- 20 -
2.6.5 機械学習.....	- 21 -
2.6.5.1 機械学習の手法.....	- 22 -
2.6.5.1.1 線形回帰.....	- 22 -
2.6.5.1.2 ランダムフォレスト（波部，2016）.....	- 22 -
2.6.6 ディープラーニング.....	- 23 -
2.7 評価指標.....	- 23 -
2.7.1 F-measure（巢籠，2019 など）.....	- 23 -
2.8 音楽や音を対象とした感性に関する研究.....	- 24 -
2.10 結言.....	- 25 -
3 章 音響パラメータ.....	- 27 -
3.1 緒言.....	- 27 -
3.2 音響パラメータ.....	- 27 -

3.3	フーリエ変換（城戸，2007 など）	- 27 -
3.4	周波数解析を伴わない音響パラメータ	- 29 -
3.4.1	ゼロクロス	- 29 -
3.4.2	RMS	- 29 -
3.5	周波数解析を伴う音響パラメータ	- 30 -
3.5.1	スペクトルフラックス	- 30 -
3.5.2	スペクトルセントロイド	- 30 -
3.6	用途に応じた音響パラメータの設計	- 31 -
3.7	結言	- 31 -
4章	ヴァイオリン演奏の評価のための音響パラメータ	- 32 -
4.1	緒言	- 32 -
4.2	本章で扱うヴァイオリン演奏	- 32 -
4.2.1	演奏課題	- 32 -
4.2.2	演奏の記録	- 33 -
4.3	ヴァイオリン演奏に対する評価のための音響パラメータ	- 34 -
4.3.1	提案パラメータの概要	- 34 -
4.3.2	既存手法で用いられたパラメータ	- 35 -
4.3.3	提案パラメータの概要	- 35 -
4.3.4	提案パラメータの概要	- 36 -
4.3.4.1	基礎的な特徴量	- 37 -
4.3.4.2	2音間の音の滑らかさの特徴量	- 38 -
4.3.4.3	音の明るさを示す特徴量	- 39 -
4.3.4.4	音のばらつきを示す特徴量	- 40 -
4.3.4.5	F0 ずれに関する特徴量	- 40 -
4.3.4.6	ヴィブラートに関する特徴量	- 42 -
4.3.4.7	loct-vn 統計量	- 45 -
4.4	ヴァイオリン演奏の熟達度評価	- 46 -
4.4.1	熟達度評価実験	- 46 -
4.4.2	評価者間における評価スコアの共通性	- 50 -

4.4.3	パラメータの有効性	- 51 -
4.4.3.1	パラメータと平均評価スコアとの関係	- 51 -
4.4.3.2	機械学習による熟達度推定	- 56 -
4.4.4	先行研究との熟達度推定結果の比較	- 57 -
4.4.5	考察	- 58 -
4.4.6	ヴァイオリン演奏の熟達度評価に関するまとめ	- 59 -
4.5	ヴァイオリン演奏に対する熟達・技術・表現の評価と音響パラメータの関係 ..	- 59 -
4.5.1	目的	- 59 -
4.5.2	検証の方針	- 60 -
4.5.3	主観評価スコアの比較	- 60 -
4.5.3.1	主観評価の対象演奏と評価者	- 60 -
4.5.3.2	各評価者の評価結果	- 61 -
4.5.3.3	主観評価の分布	- 62 -
4.5.4	主観評価スコアの推定	- 63 -
4.5.4.1	音響パラメータ	- 63 -
4.5.4.2	各評価者の評価スコアの推定	- 64 -
4.5.4.3	パラメータの相対重み	- 65 -
4.5.4.4	ヴァイオリン演奏に対する熟達・技術・表現の評価と音響パラメータの関係の まとめ	- 72 -
4.6	種々の楽器に対する評価への利用可能性	- 72 -
4.7	結言	- 73 -
5章	J-POP 楽曲の歌声を対象とした嗜好性	- 74 -
5.1	緒言	- 74 -
5.2	日本語歌唱を伴う楽曲に対する嗜好性の評価要因（桶本ら，2021）	- 74 -
5.2.1	目的	- 74 -
5.2.2	調査する楽曲の4要素	- 74 -
5.2.3	調査の方針	- 75 -
5.2.3.1	AHP とは	- 75 -
5.2.3.2	代替案として用いた楽曲	- 75 -

5.2.4	実験の概要	- 76 -
5.2.4.1	楽曲の嗜好度に対する主観評価	- 76 -
5.2.4.2	重要度の対評価	- 76 -
5.2.4.3	4つの要素に基づいた各楽曲の評価	- 77 -
5.2.5	実験結果	- 77 -
5.2.5.1	嗜好度の主観評価の相関	- 77 -
5.2.5.2	重要度に関する結果	- 78 -
5.2.5.2.1	重要度の比に関する結果	- 78 -
5.2.5.2.2	重要度の比の比較	- 79 -
5.2.5.3	AHP分析によって得られた総合評価値	- 79 -
5.2.5.3.1	各聴取者における楽曲毎の総合評価値	- 80 -
5.2.5.4	嗜好度の主観評価値と総合評価値の相関	- 80 -
5.2.6	考察	- 81 -
5.2.7	嗜好度の評価要因に関するまとめ	- 82 -
5.3	歌声分離について	- 83 -
5.4	J-POP 楽曲の歌声を対象とした嗜好性評価のための音響パラメータ	- 84 -
5.4.1	倍音構造パラメータ (H)	- 85 -
5.4.1.1	F0を考慮した倍音構造パラメータ (H ₁)	- 85 -
5.4.1.2	メル尺度を考慮した倍音構造パラメータ (H ₂)	- 86 -
5.4.2	傾向パラメータ (T)	- 86 -
5.4.2.1	最大値 (T ₁)	- 86 -
5.4.2.2	セントロイド (T ₂)	- 87 -
5.4.2.3	周波数軸方向の差分の平均 (T ₃)	- 87 -
5.4.2.4	周波数軸方向の差分の標準偏差 (T ₄)	- 87 -
5.4.2.5	時間数軸方向の差分の平均 (T ₅)	- 87 -
5.4.2.6	時間軸方向の差分の標準偏差 (T ₆)	- 87 -
5.4.2.7	2次の曲線近似における x ² の係数 (T ₇)	- 87 -
5.4.2.8	2次の曲線近似における x の係数 (T ₈)	- 88 -
5.4.2.9	2次の近似曲線との差分の平均 (T ₉)	- 88 -

5.4.3	基礎分析パラメータ (B)	- 88 -
5.4.3.1	ゼロクロス (B ₁)	- 88 -
5.4.3.2	スペクトルスキューネス (B ₂)	- 88 -
5.5	嗜好性の評価実験	- 88 -
5.5.1	概要	- 89 -
5.5.2	聴取者	- 89 -
5.5.3	実験刺激	- 89 -
5.5.4	評価法	- 89 -
5.5.4.1	評価の妥当性	- 90 -
5.5.4.1.1	2段階評価の妥当性	- 90 -
5.5.4.1.2	評価人数の妥当性	- 90 -
5.5.4.1.3	評価の多様性	- 91 -
5.5.5	回答方法	- 91 -
5.5.6	回答の分布	- 91 -
5.6	機械学習による提案パラメータの有効性検証	- 93 -
5.6.1	検証の概要	- 93 -
5.6.2	検証結果	- 93 -
5.6.3	考察	- 95 -
5.6.4	歌声の声質における嗜好性の推定に関するまとめと今後の課題	- 99 -
5.7	楽曲推薦に関する技術	- 99 -
5.7.1	推薦技術のアルゴリズム	- 99 -
5.7.2	楽曲推薦に関する諸研究	- 100 -
5.7.3	歌声の特徴を考慮した楽曲推薦に関する研究	- 101 -
5.8	楽曲推薦への利用可能性	- 102 -
5.9	結言	- 102 -
6章	結論	- 103 -
	謝辞	- 106 -
	参考文献	- 108 -
	関連発表論文	- 114 -

1 章 緒論

音楽で奏でられる音をレコードや CD などにアナログまたはデジタルの形式で保存されたものを音楽メディアと呼ぶ。この音楽メディアのデジタル化やコンピュータ技術の発展に伴い、音楽をコンピュータで再生し分析することが容易となった。その結果、音楽メディアに含まれる波形をデジタル信号として表現した音楽音響信号から、音楽音響信号の特徴の一端を表す音響パラメータがコンピュータによって算出できるようになった。例えば、音の物理的な大きさに関するパラメータだけでなく、周波数分析により取得可能な音の高さに関するパラメータ、またそこから派生して、Cmajor, Cminorなどのコード名を表すパラメータなど様々な音響パラメータがある。このような音楽音響信号から、デジタル信号処理やフーリエ変換などの技術によって算出される音響パラメータを用いて、対象物に対しての是非を判断する基準として人間に内在する感性を具体的にデザイン要素に実現する工学的手法である感性工学分野、人間の知覚認知について調査や研究を行なう知覚認知分野などで、音や音楽における感性に関する研究がこれまでに多数報告されている。例えば、楽曲の印象評価に関する研究や、楽器音の印象評価に関する研究、また、報知音のデザインに関する研究など多岐に渡っている。これらの研究では、印象評価実験などで対象物に対する評価を取得し、次に音楽や音の音響波形から取得した音響パラメータを用いて、その評価についてコンピュータを用いて予測するという研究も多数存在する。現状では、ある音楽音響信号を入力し、その音楽音響信号に対する評価を推測するという技術については予測精度が低い、または妥当な予測が得られないなどの不十分な部分も多い。その理由の一つとして、音響パラメータの不足という点があると考えられ、この音響パラメータを補強、または拡充することで、音楽音響信号に対する評価をコンピュータによって予測する精度が向上させ、様々な工学的サービスや工学的応用を実現することとなり得ると考える。

そこで、本論文では音楽や音に関する感性工学や知覚認知分野の研究の一助となることを目的とした音響パラメータの設計を行なうことを目的とする。設計した音響パラメータの評価を行なう上で、どのような対象に対する音響パラメータを設計しているのかを明確にするために、本論文では、その対象を楽器演奏の熟達度評価と歌声の嗜好性評価という

点に着目し、音響パラメータの設計及びその評価を行なう。

まず、2章で本論文における基礎知識及び研究背景を述べ、本論文の目的、そして方針及び概要を述べる。次に、3章で音響パラメータの概要を述べた後、4章、5章で音響パラメータ設計とその評価について述べる。4章ではヴァイオリン演奏に対する評価のための音響パラメータを設計し、設計したパラメータがヴァイオリン演奏の評価における有効性を検証についてまとめる。5章では、J-POP 楽曲の嗜好性における評価要因に関してまとめ、その後 J-POP 楽曲の歌声の声質を対象とした嗜好性の評価のためのパラメータを設計し、その有効性評価を行なう。最後に、6章で本論文をまとめるとともに、今後の課題及び展望について述べる。

2 章 本論文で用いる基礎知識及び研究背景

2.1 緒言

本章では、まず本論文における感性の扱いについて述べ、続いて、本論文の目的と方針を述べる。その後、本論文に関する背景や基礎知識を述べる。

2.2 感性の扱い

本論文は、感性工学の一部として、人が音楽を聴取した際に生ずる音楽に対する評価を取得し、その評価をコンピュータによって自動推定を行なうという枠組みを設定する。本論文で述べる感性とは、哲学分野における感性的認識についての分野である感性学や美学とは直接的な関係を意図しているものではない。感性工学とは人間の感性やイメージを物理的なデザイン要素に翻訳して、感性に合った商品を生設計するテクノロジーとされている(長町, 1989)。また、長町はこの感性を感覚といった言葉などとは明確に区別している。図 2.1 に長町が示す感覚から感性までの関連(長町, 1989)を示す。

感性という言葉を用いた感性工学という研究分野は、1988年に設立された感性工学会がはじめであるとされている。また、感性工学会においては、“*Kansei Engineering International Journal*”という英文論文誌が出版されている(現在は *International Journal of Affective Engineering* に誌名が変更されている)。これは、感性という言葉に当てはまる適切な英語表現が存在しなかったためとされている。また、このような感性工学会の発足により、日本から感性工学という分野が世界に広がり発展した分野であるとされている。

一方、感性工学に類似した表現に「感性情報学」というものもある。感性情報学についてまとめた著書(坂本, 2018)では、「広辞苑で感性という言葉を引き出した際に、第一の意味の中に「感受性」という言葉があり、この感受性は大辞林で「外界からの刺激を深く感じ取り、心に受け止める能力」とあり、「感受性が鋭い」、「感受性が豊かだ」というように使うとされ、我々が「感性が鋭い」、「感性が豊かだ」という使う場合の意味と同様であり、つまり、「感性」とは「能力」であり、豊かであることや研ぎ澄まされていることがよいことであるものである」と述べている。

これらを踏まえて、本論文では感性を以下のように扱う。外界からの刺激を視覚、聴覚、触覚などの人間の感覚器を通じて脳に伝えられた刺激とし、この外界からの刺激を受け引き起こされた情動や感情、また、外界からの刺激を処理する能力であり、人間の知覚認知の一端もしくは知覚認知そのものを指すものとして扱う。感性については、「創造力の高さなどを表す際に高い感性の持ち主」などのように、能動的な能力を指す際に使用されることもある。このような能動的な能力については、過去に与えられた外界からの刺激を処理しているものとして扱う。

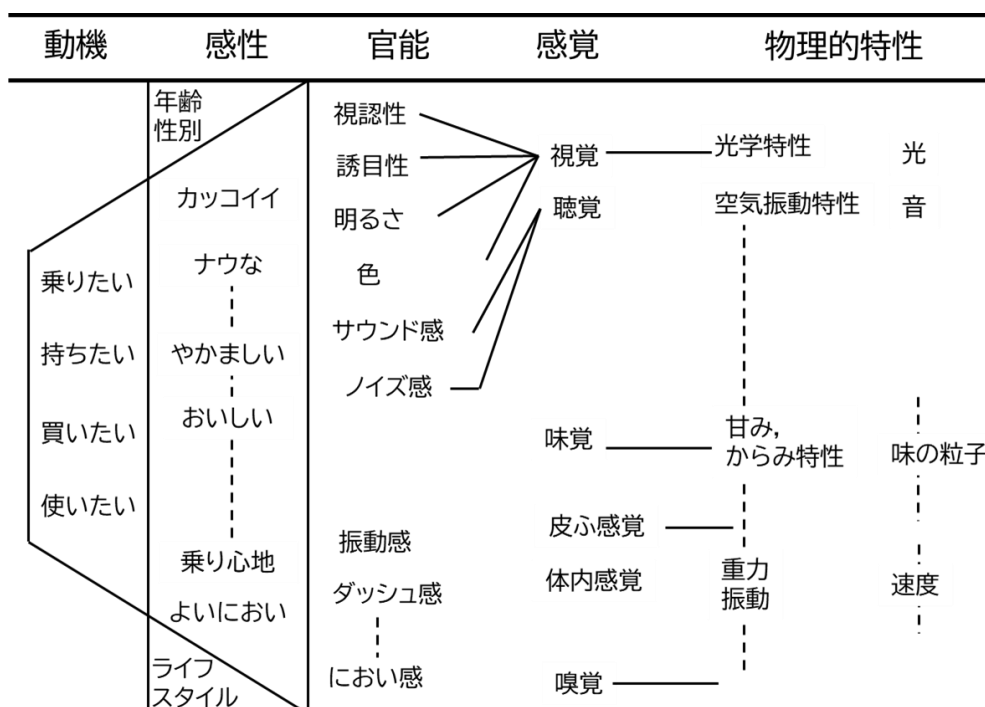


図 2.1 長町が示す感覚から感性までの関連を示した図（長町，1989 を参考に作成）

2.3 本論文の目的

音楽や音を人が聴取した際に、対象音または対象となる音楽から人間が何らかの心理現象が発生する上で、その享受した心的活動のことを感性と呼び、その感性を音楽音響信号から推定するという場合において、人間の感性に関する研究のアプローチとして、人間の感性は音楽音響信号に内在する無数の音響パラメータから構成されることを仮定した研究が存在する。この研究においては多数の音響パラメータを設計する必要があり、これまで

に多くの音響パラメータが設計されている。音楽や音に関するパラメータなどは多数存在するものの、まだまだ不足しているパラメータは数多く存在するため、音響パラメータを拡充することで、世の中のパラメータベースの研究・開発の一助となり得ると考えられる。

そこで、本論文では、音楽や音に関する感性工学や知覚認知の研究の一助となることを目的とした音響パラメータの設計を行なうことを目的とする。なお、前述のように、音響パラメータの設計においては、特定の研究に適した音響パラメータがあると考えられるため、すべての研究に適した音響パラメータとはならない可能性は考えられるが、対象の特性をパラメータの形式で表現するという点では共通であり、複雑な対象を具体的なパラメータの形式で表すことは、異なる対象に対する研究の一助となることが期待されるため、本論文ではそのような音響パラメータの設計を行なう。

2.4 本論文の方針及び概要

本論文の目的は、音楽や音に関する感性工学や知覚認知の研究の一助となることを目的とした音響パラメータの設計を行なうことであると述べた。設計した音響パラメータの評価を行なうという点でも、どのような対象に対する音響パラメータを設計したかが明確である方が望ましいと考え、本論文では、その対象を楽器演奏の熟達度評価と歌声の嗜好性評価という点に着目し、音響パラメータの設計、またその評価を行なう。

このとき、楽器の演奏については、特にヴァイオリン演奏に着目する。楽器演奏、特にヴァイオリン演奏に着目した理由として、これまでの演奏評価に関する研究の多くはピアノを対象としている（三浦ら，2010；Nonogaki et al., 2011 など）が、ピアノにおいては鍵盤を打鍵するとあらかじめ調律された音が鳴るという仕組みであるのに対し、楽器の中にはヴァイオリンのように演奏者が演奏時に微細なピッチの変化を行なうことが可能な楽器が多数存在する。本論文では、そのようなピッチの微細なコントロールが可能な楽器においても演奏評価を行なうための音響パラメータの設計可能性について検討するために、ポピュラな楽器のひとつであるヴァイオリン演奏を対象に検討を行なう。

また、歌声に着目した理由として、5章でも述べるが、J-POP 楽曲の嗜好性評価として、歌声の好みを重要視する場合もあり、歌声というのは J-POP 楽曲の嗜好性評価において重要な要素のひとつであるにも関わらず、歌声抽出の技術がこれまでに発展していなかった

ことなどから、歌声に着目したパラメータの有効性については十分な検討がなされていない。そこで、歌声の声質に着目した音響パラメータを設計することの有効性確認を行なうことを目的として歌声に着目した。

このような、パラメータを設計することで音楽配信サービスにおける楽曲推薦技術等への活用が期待される。例えば、ヴァイオリン演奏における演奏評価を基準に楽曲を推薦するサービスに発展する可能性、また、歌声の嗜好性に合わせた楽曲を推薦するサービスに発展する可能性もある。

2.5 音楽メディアのデジタル化

2.5.1 音楽メディアのデジタル化による音楽聴取方法の変遷

1877年に Thomas Alva Edison が初代蓄音機「フォノグラフ (phonograph)」を発表して以降、人間は音、ひいては音楽を記録する手段を有するようになった。それ以前の時代においては、音楽を聴取するにあたってはコンサートホールなどで生演奏を聴取するという方法しか採り得ることができなかった。それがフォノグラフをはじめとする蓄音機の登場で大きく変化した。フォノグラフの語源は、ギリシャ語で声もしくは音を意味する「フォネー」、そして同じくギリシャ語で書くことを意味する「グラフエー」を組み合わせで作られた造語であり、組み合わせ「音を書くことを」を意味する（秋吉, 2017）とされている。この言葉の意味する通り、フォノグラフの登場によって人々は音を記録することが可能となった。フォノグラフは、発表から 10 年ほど経った後に予め録音されたレコードを再生する機器として世の中に浸透したものの、当時のレコードは円筒管レコードであったため、複製が困難といった問題点もあった。この問題を解決したのは、1887年に Emil Berliner が発表した「グラムフォン (Gramophone)」である。グラムフォンは、円盤型のレコードを採用することで、円筒管レコードにおける複製が困難であるという問題を解消した。円盤型のレコードにおいては、金属の型を使用することで複製盤を作成することができる。レコードを用いて音楽を再生するという文化もこのあたりから徐々に栄え始めた。その後には、レコード再生機における回転モータの開発が行なわれたものの円盤型レコードという方式は登場後、多くのレコード再生機に引き継がれた。ここまで述べたレコードというのは、円盤（円筒管）レコードに記録された溝を掘ることで音を録音、またその溝を読み取るこ

とで音を再生するというものである。

一方、1958年に磁気テープを利用したカセットテープが登場した。1964年には、コンパクトカセットという規格が登場した。その後、1978年には、ポータブルカセットプレーヤーである「ウォークマン (Walkman)」が発表された。ウォークマンについては、今までのレコード再生機等と異なり、持ち運びが可能な音楽再生機であり、とうとう音楽が「いつでも」「どこでも」聴取できるという時代が始まった。そして、1982年にはCD (Compact Disc) の規格が発表され、音楽の記録媒体がアナログ信号からデジタル信号を扱う時代に突入した。その後MD (Mini Disc) の発表もあり、1990年代はデジタルで記録された記録媒体を用いて音楽を聴取することが主流となる時代となっていた。そして、2001年に iPod が発表され、音楽の聴取体験は更なる変化を遂げた。iPod のどこでも音楽が聴けるという利用スタイルはウォークマンと大差はないものの、iPod がウォークマンと大きく異なる点は、音楽を記録した媒体を持ち運ばなくとも、1000曲もの楽曲をいつでもどこでも聴取できるという点である。それまでのカセットやCDなどの記録媒体を使用したウォークマンなどのポータブルオーディオプレーヤーについては、いつでもどこでも音楽を聴取できるものの、聴取したい楽曲の収録された記録媒体を持ち運ぶ必要があった。例えば、CDについては1枚で記録できる時間は最大で80分程度である。つまり、1枚のCDで持ち運べる音楽は、仮に5分の曲であれば16曲程度である。それに対し、当時の iPod は iPod 本体に1000曲もの楽曲を記録し持ち運べることとなった。このような iPod の登場により、音楽は「いつでも」「どこでも」「大量の楽曲」を聴取できるようになった。

このように、音楽の聴取方法は、コンサートホールなどで生演奏を聴取するという方法しか採り得ることができなかった時代、言い換えると「場所」と「時間」の制約が強かったものの、蓄音機の発表から徐々に「場所」と「時間」の制約がなくなっていた。

2.5.2 音楽メディアの購入方法における時代変遷

「音楽メディアの購入」とは生演奏を聴取するために演奏会やコンサートなどのチケットの購入し視聴体験を得るということではなく、音楽が収録、記録された媒体の購入のことを指し、ここではレコード盤が販売されて以降の音楽メディアが記録された媒体の購入方法について触れる。一部、前述と重複する部分もあるが、ここでは購入ということに焦点を当て、再度整理を行なう。

音や音楽を記録,再生できるようになったのは,1877年のフォノグラフが始まりである。その10年ほど後に音楽を円筒管レコードという媒体に記録したものを販売するという動きが始まった。そして,レコードが円盤型レコードの開発により,音楽が記録された円盤型レコードという媒体が販売,購入するという動きが加速した。その後は,カセットテープやCD,MDなどの媒体,また,アナログからデジタルへと記録方法の変化などもあったものの,音楽が記録された何らかの媒体を販売し,大衆はそれを購入するという動きは,1990年代ごろまで続き,音楽メディアが記録された媒体の販売市場は右肩上がりになっていった。

また,1990年代には,デジタルデータのサイズとしては小さいものの,音楽を劣化の少ないフォーマットである「mp3」が登場した。このmp3は不可逆圧縮のオーディオデータフォーマットである。音楽がデジタルデータとして扱うことができる,かつデータサイズの小さなオーディオファイルとして扱えるようになったこと,さらには,インターネットの発展という様々な要素が絡み合い「Napster」に代表されるようなファイル共有サービスによる音楽データのファイル共有が行なれるようになった。音楽データの共有ができるということは,自らが購入していない音楽データを入手できるということであり,このようなサービスを利用した,所謂海賊版といわれるような違法コピーされた音楽データが浸透してしまうことで音楽業界に打撃を与えた。その後,音楽の著作権を保有する団体等からの訴訟により,Napsterのようなファイル共有サービスは事実上利用できなくなった。Napster等における問題点は,著作権等が考慮されておらず,海賊版が共有されていたという点であった。

しかし,この音楽メディアを媒体に記録せずに,音楽データ(オーディオファイル)として扱うという動きは,2003年に登場したiTunes Music Storeの登場をきっかけに大きく盛り上がり始めた。iTunes Music Storeは,iPodを開発したApple社によるサービスであり,そもそもはiPodで音楽を聴取するユーザ向けのサービスであった。このiTunes Music Storeでは,音楽データにプロテクトという処理を加えることで,音楽データを購入したユーザが音楽データを複製できない仕様となっており,著作権等を守り,著作権保有者にも対価が支払われる仕組みを持っていた。また,このiTunes Music Storeは,当初の計画よりも大きな売上を上げ,Napster等のファイル共有サービスの登場により落ち込んでいた音楽業界における新たな音楽メディアの販売,購入方法として注目を集めることとなった。所謂「音

「音楽配信サービス」のはじまりである。その後、世界では音楽配信サービスは、音楽ダウンロードサービス、そして、定額制音楽配信サービス（所謂、音楽サブスクリプション）へとつながる。音楽ダウンロードサービスは、ユーザが1曲200円などというように、楽曲毎、もしくはアルバム毎等に対価を支払い、音楽データを購入するサービスの形態である。一方、音楽サブスクリプションにおいては、ユーザが1か月2000円などのように、特定期間に対し一定金額を支払うことで各サービスが提供する楽曲を自由にストリーミング再生し、楽曲を聴取できるというサービスであり、ユーザは音楽データを保有しない音楽配信サービスの形態である。この音楽サブスクリプションについては、2008年に Spotify Technology によって開始された Spotify を皮切りに、2010年代半ばから様々なサービスが提供されるようになり、現在では、Amazon が提供する Amazon Prime Music、Apple が提供する Apple Music、Google が提供する Google Play Music、LINE MUSIC 株式会社が提供する LINE MUSIC など多岐にわたって展開されている。

ここまでは、どのような音楽メディアの媒体が登場したのかという点に焦点を当て述べたが、次は、音楽メディアの売上から音楽聴取メディアの変化について述べる。図 2.2 に国際レコード産業連盟 (International Federation of Phonogram and Videogram Producers ; 以下、IFPI) による 1999 年から 2021 年の世界におけるレコード音楽産業の売上高 (GLOBAL RECORDED MUSIC INDUSTRY REVENUES) (IFPI, 2022 参照) を示す。なお、図 2.2 において売上金額の単位は US ドルであり、横軸が西暦、縦軸が売上金額を示している。図 2.2 より、1999 年については、CD などの物理的媒体に音楽メディアが記録された Total Physical が主な売上であったものの、Total Physical の売上は徐々に減少し、2021 年では 1999 年に対し 5 分の 1 ほどの売上となっている。一方、音楽ダウンロードサービスを指す、Downloads & Other Digital については、2004 年頃の登場から 2012 年頃まで右肩上がりですべての売上が上昇しているが、2014 年頃から売上が減少している。これは、2005 年頃から徐々に売上が伸びた音楽サブスクリプションにあたる Total Streaming に押される形であったと考えられる。その音楽サブスクリプションにあたる Total Streaming については、2021 年では音楽レコード産業において最も売り上げのある音楽メディアの販売形式となっている。また、全体の売上については、1999 年から 2014 年頃までは減少傾向にあるものの、それ以降は増加傾向にある。音楽ダウンロードサービスの売上が増加しても、全体の売上は減少しているものの、音楽サブスクリプションの売上が増加するとともに音楽レコード産業全体の売上が増

加していることより、音楽配信サービスのうち音楽サブスクリプションが音楽レコード産業に大きな影響を与えたと考えられる。

これらをまとめると、音楽メディアの販売、購入にあたっては、1900年頃にレコードで始まった音楽メディアを記録した媒体の販売、購入という販売モデルが、記録メディアがCDなどに変わったという点はあるものの、1990年代頃まで主流であったが、それ以降は、音楽ダウンロードサービス、音楽サブスクリプションといったサービスが開始されることで、大きく様変わりし、現在では音楽サブスクリプションが音楽メディアの販売、購入方法として主流なものとなっている。このように音楽メディアの販売、購入方法は、21世紀以降大きな変化が起きた。

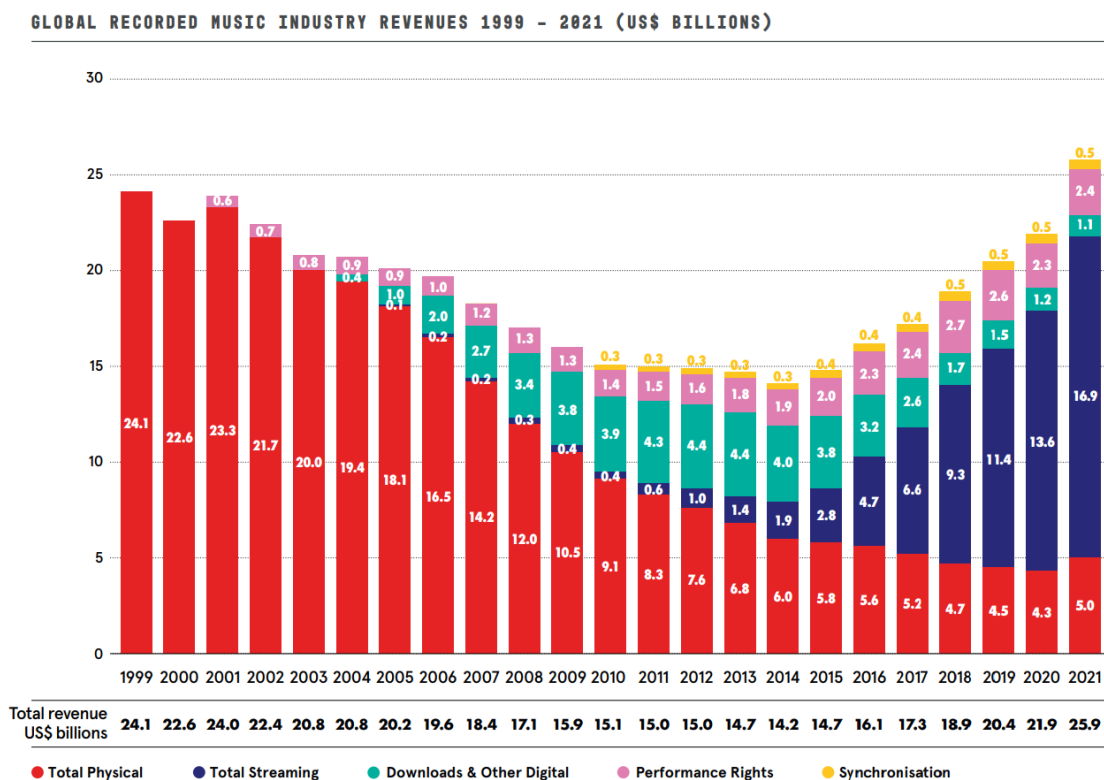


図 2.2 世界におけるレコード音楽産業の売上高 (IFPI, 2022 参照)

2.5.3 音楽情報処理との関連

音楽メディアのデジタル化によって、音楽をコンピュータによって扱うことが容易となった。

音楽メディアに関する技術において近年注目される技術のひとつは、楽曲推薦技術である。この技術が注目されるようになった一つの契機としては、音楽サブスクリプションの台頭であると考えられる。前述のように音楽サブスクリプションによって、ユーザは生涯の内にすべてを聴取することが困難ともいえる量の楽曲群から音楽を選出し聴取できるようになった。それにより、聴取する楽曲を選択するという作業が必須となり、音楽サブスクリプションというサービスを提供する側にとっても、ユーザが音楽を選択しやすいユーザインターフェースや枠組み等を設計することが必要となった。ユーザが音楽を選択しやすいユーザインターフェースを考慮しなければ、例えば、アプリなどの TOP 画面において「楽曲が発表された日時の順番に並んでいる」、「曲のタイトルが昇順で並んでいるだけ」という検索結果が表示されるだけでは、ユーザにとって満足度の高いとは言えないサービスになっているであろう。このような状況でユーザの満足度を高めるサービスのひとつとして提供されるのが楽曲推薦技術を利用した楽曲推薦と呼称されるサービスである。なお、ここでは楽曲推薦技術と楽曲推薦という言葉に分けて記述しているのは、ユーザからは、推薦された楽曲が楽曲推薦技術により自動で推移されたものか、人間が推薦した楽曲であるかの区別をつけることが明らかにすることはできないためである。

ここで、注目される技術のひとつとして楽曲推薦技術を挙げた理由は、楽曲推薦技術を利用している音楽サブスクリプションの市場が大きくなっているということだけではなく、コンピュータが音楽を理解するという技術が進歩し、その進歩した技術が楽曲推薦技術に利用される場合が多いためである。つまり、音楽情報処理の進歩の一側面が楽曲推薦技術には反映されている。例えば、Spotify という音楽サブスクリプションサービスでは、楽曲に対し様々な音響パラメータを算出し、それらを楽曲推薦技術に利用している。音響パラメータについての詳細は後述するものの、簡単に述べると、デジタル信号処理を用いて音楽メディアの音楽音響信号から信号処理技術によって音の特徴を算出した結果であり、例として音の大きさを表す RMS や高さを表す基本周波数などがある。音響パラメータは、音楽メディアのデジタル化によって音楽メディアをコンピュータが容易に扱うことができるようになったことで初めて実現した技術の一つである。近年では、ディープラーニングという技術が普及し、歌声や伴奏楽器などの混合音から歌声部分のみの音響信号の抽出などを行なうことができるようになるなど、コンピュータによるより高度な音楽の理解を行なうことができるようになった。このような技術を利用することで楽曲推薦技術の技

術がより向上することが期待される。このような技術の発展は広くは、人工知能という技術によって実現されたといえる。次のそのような人工知能について述べる。

2.6 人工知能

2.6.1 人工知能とは

人工知能という用語そのものは、1956年に開催された国際会議「ダートマス会議」にて John McCarthy により提案された造語であり、元来は人間と同様の知識を有した機械（コンピュータやロボットなど）を実現させることを目的とした研究分野であった(加納ら, 2017)。

昨今においては、「知識」という言葉の持つ意味の曖昧性により、「人工知能」の定義については研究者によって異なっている。表 2.1 に様々な研究者による人工知能の定義（総務省, 2022 参照）を示す。表 2.1 より、ドワンゴ社の山川氏は「計算機知能のうちで、人が直接・間接に設計する場合は人工知能と呼んでもよいのではないかと思う」と定義するのに対し、電気通信大学の栗原氏は「人工的に作られる知能であるが、その知能レベルは人を超えるものを想像している」と述べており、人工知能の能力の高さについても研究者によって定義が異なることが確認できる。また、公立ほこだて未来大学の中島氏と国立情報学研究所の武田氏は「人工的につくられた、知能を持つ実態あるいはそれをつくろうとすることによって知能自体を研究する分野である」と述べており人工知能の定義には研究分野も含まれるとしているのに対し、京都大学の長尾氏は「人間の脳活動を極限までシミュレートするシステムである」と述べており人工知能はシステムであるとしており、定義の範囲についても研究者によって異なる。また、大阪大学の浅田氏は「知能の定義が明確でないので、人工知能を明確に定義できない」と、そもそも定義が困難だとする研究者もいる。

表 2.1 様々な研究者による人工知能の定義（総務省，2022 を基に作成）

研究者	所属	定義
中島秀之	公立はこだて未来大学	人工的につくられた，知能を持つ実態，あるいはそれを作ろうとすることによって知識自体を研究する分野である．
武田英明	国立情報学研究所	
西田豊明	京都大学	「知能を持つメカ」ないしは「心を持つメカ」である
溝口理一郎	北陸先端科学技術大学	人工的につくった知的な振る舞いをするためのもの（システム）である
長尾真	京都大学	人間の脳活動を極限までシミュレートするシステムである
堀浩一	東京大学	人工的に作る新しい知能の世界である
浅田稔	大阪大学	知能の定義が明確でないので，人工知能を明確に定義できない
松原仁	公立はこだて未来大学	究極には人間と区別がつかない人工的な知能のこと
池上高志	東京大学	自然にわれわれがペットや人に接触するような，情動と冗談に満ちた相互作用を，物理法則関係なく，あるいは逆らって，人工的に作り出せるシステム
山口高平	慶応義塾大学	ヒトの知的な振る舞いを模倣・支援・超越するための構成的システム
栗原聡	電気通信大学	人工的につくられた知能であるが，その知能レベルは人を超えているもの
山川宏	ドワンゴ人工知能研究所	計算機知能の内で，人間が直接・間接に設計する場合を人工知能と呼んで良いのではないかと思う
松尾豊	東京大学	人工的につくられた人間のような知能，ないしはそれをつくる技術．人間のように知的であるとは，「気が付くことのできる」コンピュータ，つまり，データの中から特徴量を生成し現象をモデル化することのできるコンピュータという意味である．

2.6.2 本論文における人工知能の定義

研究者によって人工知能の定義が異なることを前述した．そのため，本論文における人工知能の定義についてここで整理を行なう．

2.5 で，音楽の聴取方法の変化，また聴取可能な音楽の曲数が膨大となっていることを述べた．このような状況において，楽曲推薦システムのような楽聴取体験を補助するシステムが台頭している．しかし，楽曲推薦等において，人間が満足する精度で推薦を行なうためには，人間の感覚，感性を実装したシステムが必要である．表 2.1 において，東京大学の池上氏は「自然にわれわれがペットや人に接触するような，情動と冗談に満ちた相互作用を，物理法則に関係なく，あるいは逆らって，人工的に作り出せるシステム」と定義している．このように，人工知能における知識の範囲に言及した定義も人工知能の定義として挙げられていることが確認できる．

これらを踏まえ，本論文では人工知能の定義を「人間の感性が実装されたコンピュータシステム」とする．

2.6.3 学習方法の分類

人工知能の分類法については，様々な切り口からの分類が存在するものの，ここでは，学習の際に学習器が算出する特徴量をあらかじめ定義するか否かという点に着目し分類を

行なう。

1 つ目の分類は、学習の際に学習器にパラメータを与える方法である。リンゴとミカンの識別を例にすると、学習器にリンゴとミカンの画像とそれらの正解ラベルを与えるだけでなく、全体の色が赤であればリンゴの可能性が高い、オレンジであればミカンの可能性が高い、また蒂の周りが凹んでいればリンゴの可能性が高い、凹んでいなければミカンの可能性が高いなどのように、リンゴとミカンの判断基準をパラメータとして学習器にあらかじめ与える方法である。

2 つ目の分類は、学習の際に学習器にパラメータを与えない方法である。学習器に与えるデータは、リンゴとミカンの識別を例にすると、学習器にリンゴとミカンの画像とそれらの正解ラベルを与えるだけとし、学習器自身にパラメータを設計させる方法である。

本論文では、前者をパラメータ有学習、後者をパラメータ無学習と呼称する。

2.6.4 音楽における人工知能の利用

音楽における人工知能の例として以下のものがある。

- 楽曲推薦システム
- カラオケの採点システム
- 歌声合成ソフト
- 自動作曲システム
- 演奏における表情付けシステム
- 歌声抽出

これらを大きく 2 つに分類することができる。1 つ目は、知覚認知に関する人工知能であり、楽曲推薦システム、カラオケの採点システム、歌声抽出などが分類される。その理由として、楽曲推薦においては楽曲から連想されるイメージとの関わり、カラオケの採点においてはどのような歌唱を行なうと上手いと判断されるのかという判断との関わり、歌声抽出においては伴奏音などが混合した音源から人間がどのように歌声を知覚しているのかということとの関わり、という側面において知覚認知に関する人工知能と位置付けた。2 つ目は創作に関する人工知能で、歌声合成システム、自動作曲システム、演奏における表情付けシステムなどが分類される。その理由として、歌声合成については人間がソフトウェアを調整することで歌をコンピュータに歌わせるという面、また自動作曲においては作曲

という創作活動をコンピュータが代替する面、演奏における表情付けにおいてはコンピュータに人間のような演奏を行なわせるという面から創作活動的な側面を持つと判断した。本論文では、前者の知覚認知に関する人工知能について触れる。

知覚認知に関する人工知能にも、パラメータ有学習とパラメータ無学習がある。与えない学習の例として、歌声抽出が挙げられる。歌声抽出とは、伴奏とともに歌唱した楽曲から歌唱パートのみを抽出する技術である。この歌声抽出において、パラメータ無学習を行なう理由のひとつとして、そもそも人間がどのようにして伴奏とともに歌唱した楽曲において歌唱パートの聴き分けができていたのかが明らかになっていないため、パラメータを設定するのが非常に困難なためだと考えられる。一方、人間が基準を定め得ることが可能なものが、楽曲推薦システムやカラオケの採点システムである（なお、この2つのシステムにおいては、詳細な仕様が明らかにされているわけではないので、パラメータ有学習であるか、パラメータ無学習であるのかは不明である）。例えば、楽曲の印象に基づいた楽曲推薦システムは、楽曲の音量、テンポ、歌詞などの基準、カラオケの採点システムにおいては、音高、声の大きさ、抑揚などの基準が考えられ、これらに関する音響パラメータを用いることで評価を行なっているということも考えられる。

2.6.5 機械学習

機械学習とは、コンピュータ等が与えられたデータからなんらかのルールやパターンを発見するというものであり、代表的なアルゴリズムには、ナイーブベイズ、決定木、回帰などがある。

次に機械学習と人工知能との関係について触れる。ディープラーニングは機械学習に含まれ、その機械学習は人工知能に含まれるという関係性であることが確認できる。この時、人工知能と機械学習が同一のものとなっていないのは、人工知能の中には、ルールベースと言われる技術が含まれるからである。ルールベースとは、人間が「if~then~else」のルールをコンピュータに明示的に実装する方式で、人工知能を実現するというものであり、ルールやパターンを自動で発見する機械学習とは異なるアプローチであるため、人工知能と機械学習は同一とはなっていない。

このような関係にある機械学習は、検索エンジンや交通渋滞予測、スパムメールの判定などに利用されており、非常に多くのサンプルデータを基に入力と出力を学習したもの（一

般社団法人日本ディープラーニング協会，2021) とされている。

また，この機械学習には，大きく3つの分類がある．教師あり学習，教師なし学習，強化学習である．これらの分類は，教師データと呼ばれる所謂正解にあたるデータの有無によって分けられる．教師データを使用するのが教師あり学習，教師データを使用しないのは教師なし学習，教師データは使用しないものの出力に応じて報酬（評価）が与えられるのが強化学習である．

2.6.5.1 機械学習の手法

機械学習には様々な手法がある．教師あり学習には回帰，決定木，ニューラルネットワークなど，教師なし学習には k-means，主成分分析など，強化学習にはマルコフ決定過程などが挙げられる．ここでは，本論文で使用する線形回帰とランダムフォレストについてその概要を述べる．

2.6.5.1.1 線形回帰

学習データにおける目的変数（教師データ）と説明変数の傾向を明らかにする手法を回帰といい，その傾向が線形であるものを線形回帰と呼ぶ．また回帰において，説明変数が1つの場合は単回帰，複数の場合は重回帰と呼ぶ．単回帰の式を式 (2.1) に示す．

$$Y_{linerreg} = \beta_0 + \beta_1 X_1 \quad (2.1)$$

この時， $Y_{linerreg}$ は単回帰の予測結果， β_0 は切片， β_1 は説明変数の係数， X_1 は説明変数である．

2.6.5.1.2 ランダムフォレスト（波部，2016）

ランダムフォレストとは，2001年に Leo Breiman によって提案された複数の弱識別器を利用し，それらを複数用いること高い予測性能を得るアンサンブル学習の一種であり，クラス分類などを行なうことができる．アンサンブル学習とは，複数の学習モデルの組み合わせによって構築されたモデルを使用する学習法のことである．ランダムフォレストにおいては決定木を弱識別器としている．なお，弱識別器とは，精度の低い識別器を指す．

ランダムフォレストでは，バイアスやオーバフィッティングといった問題を避け，精度を向上させるために決定木間の相関を低くするために決定木の構築法にランダム性を導入

している。このランダム性は、個々の決定木を学習する際の学習データのサンプリング、そして決定木の各ノードでの分割関数学習によって実現されている。

2.6.6 ディープラーニング

ディープラーニングとは、機械学習の方法の1種である多層パーセプトロンにおける隠れ層を増やすことで複雑な関数の表現を試みたものである。多層パーセプトロンとは、人間のニューロンを再現したパーセプトロンと呼ばれる人工ニューロンを使用したアルゴリズムである。1つ前のニューロンからの信号を基に現在のニューロンの動作を決めるというものである。現在では、画像からの物体検出などで利用されている。

2.7 評価指標

本節では、5章で使用する学習器の評価指標である F-measure について述べる。

2.7.1 F-measure (巢籠, 2019 など)

F-measure とは、適合率と再現率と呼ばれる指標の調和平均である。この適合率、再現率を説明するにあたっては、まず混合行列について説明する必要があるので、混合行列について述べる。

混合行列とは、学習器による分類結果を表 2.2 のようにまとめた行列である。TN が真陰性 (True Negative)、FN が偽陰性 (False Negative)、FP が偽陽性 (False Positive)、そして TP が真陽性 (True Positive) を表す。TN は真値が負例で予測値も負例、FN は真値が正例だが予測値が負例、FP は真値が負例だが予測値が正例、そして、TP は真値が正例で予測値も正例であることを示す。

この混合行列から算出される指標に、正解率 (Accuracy)、適合率 (Precision)、そして再現率 (Recall) の3つがある。この3つの指標は、混合行列で得られた TN, FN, FP, TP の値から算出することができる。まず正解率を求める式を式(2.2)に示す。

$$accuracy = \frac{TN + TP}{TN + FN + FP + TP} \quad (2.2)$$

正解率は、予測値と真値が一致した割合を示すである。次に、適合率を求める式を式(2.3)に示す。

$$precision = \frac{TP}{FP + TP} \quad (2.3)$$

適合率は、予測値が正例に対する真値の正例の割合を示す。最後に、再現率を求める式を式(2.4)に示す。

$$recall = \frac{TP}{FN + TP} \quad (2.4)$$

再現率は真値が正例に対する予測値の正例の割合を示す。

このように算出される、正解率、適合率、再現率のうち、トレードオフの関係にある適合率と再現率の調和平均から算出されるのが、F-measure である。この F-measure の算出式を式(2.5)に示す。

$$F - measure = \frac{2 \cdot precision \cdot recall}{precision + recall} \quad (2.5)$$

この F-measure では、学習器の分類・予測の「よさ」を測ることができる。

表 2.2 混合行列

		予測値	
		負例	正例
真値	負例	TN	FP
	正例	FN	TP

2.8 音楽や音を対象とした感性に関する研究

図 2.1 より、感性には、やかましい、おいしいなどの評価語が当てはまるとされている。この感性について音楽の場合でさらに考えると、演奏などの上手さ、楽曲の印象評価など多岐にわたる。

例えば、感性に基づくクラシック音楽の分類に関する研究（平江ら，2008）では、楽曲に対する印象の自動推定を行なう手法、またそのシステムを提案している。この研究では、クラシック楽曲に対する印象評価結果とクラシック楽曲から算出された音響パラメータ特

徴を用いて、予め印象評価を行なっていないクラシック楽曲についてもニューラルネットワークによって印象を推定するというものである。この研究では、印象の推定に用いる音響パラメータには、強度に関する特性、周波数に関する特性、そして、音響波形のエンヴェロープから算出されるリズムに関する特性に関する音響パラメータを用いている。この研究では音響パラメータとして、音響波形のエンヴェロープを考慮した時間軸上の振幅の時間変化に関するパラメータ、そしてドラム音から得られる周波数情報に関するパラメータを用いている。

また、楽曲だけではなく楽器音に関する研究もある。例えば、ドラム音の音色における感性情報と工学的パラメータとの対応付け（上野山ら，1993）では、ドラム音の音色に対する印象評価結果と、音響パラメータとの対応付けを行なっている。

また、音楽というよりも音という方が適切かもしれないが、感性の研究などをベースとして音のデザインを行なっているという事例も存在する。例えば、電気自動車の接近を知らせるための接近通報音のデザインに関する報告（山内，2011）がある。この報告では電気自動車の走行音が従来の自動車より静かなために、歩行者などが車の接近に気づきにくいという問題に対する動向が示されており、接近音の印象評価などの事例が紹介されている。その中では、実環境下で求められる接近報知音の音の大きさに関する事例や音響パラメータとして周波数変調に着目した事例など様々な事例が紹介されている。また、フランス国立音響音楽研究所（IRCAM）における音デザイン活動の紹介（川上，2011）では、電気自動車の音のデザインの事例や駅のホーム間移動のための誘導音の事例などが紹介されている。

このように、音楽や音を対象とした感性に関する研究は多岐に渡るが、音響パラメータを活用した研究では、それぞれの研究に適した音響パラメータを用いている。これらより、音楽や音に対する感性の研究において使用する音響パラメータについては、その対象に適した音響パラメータの設計が必要である。

2.10 結言

本章では、まず時代背景を述べ、続いて、本論文で用いる基礎知識を述べた。そして、本論文では、音楽に感性に関する感性の研究の一助となることを目的とした音響パラメー

タの設計を行なうことを目的とし、ヴァイオリンの演奏と歌声の声質に着目した音響パラメータを設計することで設計した音響パラメータの有効性検証を行なうことを述べた。

3 章 音響パラメータ

3.1 緒言

本章では、音響パラメータの概要を述べ、次にフーリエ変換について述べる。そして、音響パラメータの例とそのプログラムソースを示す。最後に用途に応じた音響パラメータを設計する必要性について述べる。

3.2 音響パラメータ

音響パラメータとは、音響波形から算出可能な音響的な特徴を数量化した値またはその算出法である。音響パラメータは、音楽音響分野の発展に伴い、音楽音響信号の分析や合成に関する技術が発展した中で、コンテンツベースの分析手法が広まり、従来は人間がハンドラベリングによって与えていたタグ情報を音響信号から人手を介することなく抽出し、分類や認識に用いるといった研究が盛んに行なわれるようになった研究を実現したものと述べられており、その特徴として、その音響信号を人間が直接聴取することなく、自動的に取得される点にあると述べられている（山田ら，2014）。

近年では、様々な音響パラメータが提案されている。例えば、MATLAB で計算可能な音響パラメータのライブラリである MIRtoolbox (Lartillot, O. et al., 2007) や C++ で計算可能な音響パラメータのライブラリである Essentia (Bogdanov, D. et al., 2013)、また、Python で計算可能な音響パラメータのライブラリである librosa (McFee, B. et al., 2015) など様々な言語を対象とした音響パラメータのライブラリがこれまでに提案されている。

3.3 フーリエ変換（城戸，2007 など）

フーリエ変換とは、物理的に観察される波形はすべてコサイン波とサイン波が組み合わせによるものという考えに基づいた、波形を構成するコサイン波とサイン波を分解することで波形に含有する各周波数の成分の強さを明らかにする手法である。物理的に観察される波形はすべてコサイン波とサイン波の組み合わせというのは、波形を $x(t)$ とした場合、

$$\begin{aligned}
x(t) = & A_0 + A_1 \cos\left(2\pi \frac{1}{T} t\right) + A_2 \cos\left(2\pi \frac{2}{T} t\right) + A_3 \cos\left(2\pi \frac{3}{T} t\right) + \dots \\
& + B_0 \sin\left(2\pi \frac{1}{T} t\right) + B_2 \sin\left(2\pi \frac{2}{T} t\right) + B_3 \sin\left(2\pi \frac{3}{T} t\right) + \dots
\end{aligned} \tag{3.1}$$

のようにコサイン関数とサイン関数の級数で表されるという原理を利用している。この時、 T は周期を表す。

ここで、複素指数関数でのフーリエ級数を考える。複素指数関数とは、複素数を指数とした指数関数である。なお、複素指数関数を導くために書き直した式(3.2)を用いる。

$$\begin{aligned}
x(t) = & \frac{1}{T} \left\{ a_0 + 2a_1 \cos\left(2\pi \frac{1}{T} t\right) + 2a_2 \cos\left(2\pi \frac{2}{T} t\right) + \dots + 2b_1 \sin\left(2\pi \frac{1}{T} t\right) + 2b_2 \sin\left(2\pi \frac{2}{T} t\right) + \dots \right\} \\
= & \frac{1}{T} a_0 + \frac{2}{T} \sum_{k=1}^{\infty} \left\{ a_k \cos\left(2\pi \frac{k}{T} t\right) + b_k \sin\left(2\pi \frac{k}{T} t\right) \right\}
\end{aligned} \tag{3.2}$$

また、オイラーの公式を利用した複素指数関数に対するコサイン波とサイン波は、

$$\exp(j2\pi ft) = \cos(2\pi ft) + j\sin(2\pi ft) \tag{3.3}$$

によって表される。式(3.2)、(3.3)を用いたフーリエ級数は、

$$x(t) = \frac{1}{T} \sum_{k=-\infty}^{\infty} X_k \exp\left(j2\pi \frac{k}{T} t\right) \tag{3.4}$$

で表され、その係数を求める式は、

$$X_k = \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \exp\left(-2j\pi \frac{k}{T} t\right) dt \tag{3.5}$$

である。ここで、積分区間を無限にした場合を考える。これによって、周波数の間隔が無限小となる。積分の範囲を $-\infty$ から $+\infty$ とした場合は、

$$X(f) = \int_{-\infty}^{\infty} x(t) \exp(-2j\pi ft) dt \tag{3.6}$$

となる。式(3.6)は複素指数関数を用いたフーリエ級数の式としてよく用いられ、この式によって連続スペクトルが求められる。

フーリエ変換は、対象波形とあらゆる周波数のコサイン波とサイン波形との積の積分なので、連続関数として実行することは、理論的に可能でも現実的ではない。そこで、一定間隔でサンプリングされたデータ列に対するフーリエ変換である離散フーリエ変換が提案されている。離散フーリエ変換 (discrete Fourier transform : DFT) の式を式(3.7)に示す。

$$X_k = \sum_{n=0}^{N-1} x_n \exp\left(-j2\pi \frac{nk}{N}\right) \tag{3.7}$$

式(3.6)は N 個の要素を持つ数列 x_n に対する離散フーリエ変換を表現する式である。

離散フーリエ変換を計算機で実行する場合には膨大な計算量が必要となる。公式通りに計算すると、積和の計算を N^2 回行なう必要があるためである。そこで、Cooley と Tukey によって高速フーリエ変換 (fast Fourier transform : FFT) (Cooley, J. W. et al., 1965) が発表された。この高速フーリエ変換による演算時間は、従来のフーリエ変換の演算時間に対し、数百分の一程度に短縮された。具体的には、 $O(N^2)$ の計算量が、 $O(N \log N)$ となる。なお、 $O()$ はアルゴリズムの計算量に関する指標である。このように計算量を減らすことができることは計算機科学では画期的なことであり、高速フーリエ変換によってフーリエ解析の工学応用が広がった。高速フーリエ変換は、フーリエ変換においては、バタフライ演算を行なうことで計算時間の短縮を行なっている。ただし、このバタフライ演算を行なうためには、高速フーリエ変換には、入力信号長を 2 のべき乗にしなければならないという制限がある。

3.4 周波数解析を伴わない音響パラメータ

本節では周波数解析を伴わない音響パラメータの例として、ゼロクロスと RMS (Root Mean Square) について述べる。

3.4.1 ゼロクロス

ゼロクロスは、音響信号の振幅値が 0 と交差する回数を表す音響パラメータである。従って、信号の平均位置を表すバイアス成分の影響を受ける。この *Zerocross* の算出式を式(3.8)に示す。

$$\text{Zerocross} = \sum_{i=\chi}^{\chi'} \theta \begin{cases} 1(x_i x_{i-1} < 0) \\ 0(x_i x_{i-1} > 0) \end{cases} \quad (3.8)$$

θ は交差していれば 1, 交差していなければ 0 となる。この時、 x_i は音楽音響信号の配列 x の i 番目の値、 χ は分析フレームの開始時刻、 χ' は分析フレームの終了時刻を表す。

3.4.2 RMS

RMS は音響信号の物理的な強度、振幅値の大きさに対応する音響パラメータである。RMS は実効値と呼ばれる場合もある。算出式を式(3.9)に示す。

$$RMS = \sqrt{\frac{1}{M} \sum_{i=1}^M x_i^2} \quad (3.9)$$

M は分析のサンプル数, x_i は音楽音響信号の配列 x の i 番目の値であることを表す.

3.5 周波数解析を伴う音響パラメータ

本節では周波数解析を伴う音響パラメータの例として, スペクトルフラックス (Spectral Flux) とスペクトルセントロイド (Spectral Centroid) について述べる.

3.5.1 スペクトルフラックス

スペクトルフラックスはスペクトルの変動量を表す音響パラメータである. 連続する 2 つの分析フレームにおける, パワースペクトルでの各周波数ビンの差についての和の絶対値である. 算出式を式(3.10)に示す.

$$SpectralFlux = \sum_{n=1}^N (L_h[n] - L_{h-1}[n])^2 \quad (3.10)$$

この時, N は分析フレーム内の周波数ビンの総数, $L_h[n]$ は第 h フレームにおける第 n 周波数ビンのパワーを表している. このスペクトルフラックスについては, 人間の聴覚特性 (等感度曲線) などが考慮されていないという問題点もある.

3.5.2 スペクトルセントロイド

スペクトルセントロイドはヒトが感じる「音の明るさ」によく対応することが知られている特徴量である. スペクトルセントロイドの算出式を式(3.11)に示す.

$$SpectralCentroid = \frac{\sum_{i=1}^N L_h[n] \times n}{\sum_{i=1}^N L_h[n]} \quad (3.11)$$

スペクトルセントロイドの値が高いほど, 明るい音であることを示す. この時, N は分析フレーム内の周波数ビンの総数, $L_h[n]$ は第 h フレームにおける第 n 周波数ビンのパワーを表している.

3.6 用途に応じた音響パラメータの設計

本節では、用途に応じた音響パラメータの設計の必要性について述べる。

スペクトルセントロイドは音の明るさを表す音響パラメータであることは 3.5.2 で述べた。この時、C4 と C5 のヴァイオリンの演奏音のスペクトルセントロイドを比較した場合は、C5 の演奏音の方がそもそもの音高が高いためスペクトルセントロイドの値も自ずと高くなる。演奏音の評価においては、従来のスペクトルセントロイドよりも、音の高さからの比率で表した方が演奏評価に適しているとつまり、従来の音響パラメータを用途に応じて適宜修正する必要がある。

本論文では、用途に応じた音響パラメータとして 4 章ではヴァイオリン演奏の評価、5 章では歌声の声質の評価に着目し、音響パラメータの設計、またその評価を行なう。

3.7 結言

本章では、音響パラメータの概要を述べ、次にフーリエ変換について述べた。そして、音響パラメータの例とそのプログラムソースを示した。最後に用途に応じた音響パラメータを設計する必要性について述べた。実際に設計したパラメータについては、4 章と 5 章でそれぞれ説明することを述べた。

4 章 ヴァイオリン演奏の評価のための音響パラメータ

4.1 緒言

本章では、まずヴァイオリン演奏の演奏特徴を捉えるパラメータを提案し、そのパラメータを熟達度推定によって評価した研究報告 (Okemoto et al., 2018 ; 桶本ら, 2020) について述べる。次に、提案したパラメータを用いて熟練・表現・技術の各指標についての評価事例 (Okemoto et al., 2019 ; 桶本ら, 2019) について述べる。最後に、提案パラメータの種々の楽器への適用可能性について議論する。

4.2 本章で扱うヴァイオリン演奏

本章で扱うヴァイオリン演奏の評価については、例えば音楽大学でヴァイオリン演奏を学ぶ学生や音楽大学の学生が出場するコンクールや音楽大学での演奏試験のようなある程度以上のレベルで熟練された演奏のように繊細な演奏表現の変化等を対象としたものではなく、音楽の演奏を専門としない高校生や大学生が趣味で演奏を楽しむような演奏から音楽大学で学ぶ学生や卒業生に演奏のようにスキル等の差が明確に表れやすい演奏を対象としている。また、それらの演奏者の演奏音に対して、音響パラメータによってヴァイオリン演奏の一側面をどの程度区別できるのか、またどの程度その熟練度を説明できるのかについて検討している。

4.2.1 演奏課題

本論文では、1 オクターブの上下行長音階にヴィブラートを付与した「ヴィブラートを付与した 1 オクターブ 上下長行音階 (以下, 1octvib 課題)」を演奏課題とする。図 4.1 に 1octvib 課題を示す。図 4.1 の 0 から 4 の 数字は運指番号を示す。なお、ヴァイオリン演奏時におけるボウイングについては、実験者が指示を行ない、全演奏者が同一のボウイングとなるようにした。本研究で、演奏課題として 1octvib を採用したのは、ノンフレット楽器の特徴的な演奏手法であるヴィブラートについても熟達度推定のパラメータとなると考え、音階教本 (Hrimaly, J 作曲, 岩船 編, 2007) を参考に、1octvib 課題を演奏課題とし

て採用した.

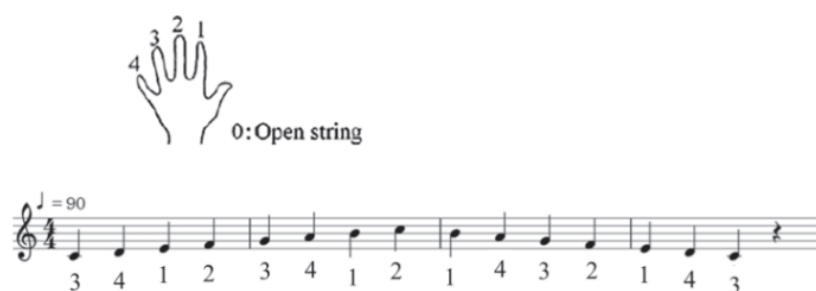


図 4.1 演奏課題

4.2.2 演奏の記録

本節では演奏の記録の概要を述べる。演奏の記録は、防音室内で、以下の機材を用いて行った。

- ヴァイオリン：Diolmaster（中国製の量産型ヴァイオリン）
- マイク：接近用 VO4099（DPA 社製），アンビエント用 SM58（Shure 社製）
- ウェブカメラ：HD WEBCAM C525
- 録音機：DR-680（TASCAM 社製）

演奏者は、演奏歴は 4～18 年（平均 14.0 年）、年齢は 19～22 歳（平均 21.0 歳）、男性 5 名、女性 5 名の計 10 名である。演奏者には、 $A4 = 440$ [Hz] でチューニングをさせ、1octvib を 10 回ずつ演奏させた。その演奏の様子をウェブカメラによって、運指やボウイングに誤りがないかを実験者が目視で確認し、誤りがあった際には再演奏を要求することで、演奏誤りのない演奏を各演奏者 10 回ずつ記録し、計 100 通り（10 人×10 トライアル）の演奏データを記録した。なお、演奏者にはヘルシンキ宣言に基づいた参加同意書を記入いただき、それに基づき演奏者の希望に応じて、演奏のやり直しや実験の中断を認めた。

なお、ここでの演奏歴は、演奏歴との比例関係を前提としたものではなく、あくまで演奏経験の一側面として記録している。その理由として、音楽大学の学生などのように毎日長時間にわたる練習を行ない、かつ、高度な教育を受けてきたような演奏者が経験した 10 年と、趣味で土曜日に 3 時間程度演奏をしているというような演奏者の 20 年では、演奏歴と熟達度について比例関係はないと考えるのが妥当であろう。しかし、実際にどんな教育を受けたなどの演奏経験を明確に数値化することも困難であるため、演奏経験の一側面

として演奏歴を記録した。

4.3 ヴァイオリン演奏に対する評価のための音響パラメータ

4.3.1 提案パラメータの概要

本論文で提案するヴァイオリン演奏を対象とした熟達度推定のためのパラメータの設計方針を説明する。ヴァイオリン演奏の熟達度評価においては、評価者に依存した様々な評価基準が存在する。例えば、指定テンポに対しテンポが揺らぐという芸術的逸脱 (Seashore, C. E., 1938) をより多く持つ演奏ほどより熟達しているという評価基準もあり得る。そのため、本論文における評価基準を明確にするために様々なヴァイオリン演奏について触れた知見を整理し、得られた結果を網羅したパラメータを設計する。二つの知見 (渋谷ら, 1996; 梅本, 1996) より、ヴァイオリンの演奏は「正しいピッチ」「長さ」「強さ」及び「美しい音色とヴィブラート」によって表現されると述べた。更に、ヴァイオリン演奏の研究 (渋谷ら, 1996) によると、「ピッチ」、「長さ」、「強さ」に関しては、それぞれが周波数、時間、基準となる音との強度比のように、物理量が存在し楽譜にもある程度の情報が記載されていると考えてよい。それに対し、音色は複数のパラメータが複雑に絡みあっていると考えられ、その情報は楽譜にもほとんど記載されていない。また、音色の表現語に対する因子 (日本音響学会監修, 1988) は、「迫力」、「金属製」、「美しさ」であるとされており、特に美しさについてはその定量化が本質的に困難である。更に、JIS 規格では音色は「聴覚に関する音の属性の一つで、物理的に異なる二つの音が、たとえ同じ音の大きさ及び高さであっても異なった感じに聞こえるとき、その相違に対応する属性」と定義されており、その物理的特徴がはっきりとは定まっていない。これらより、音色を表すパラメータの設計は本質的に困難であり、また、本論文のように生演奏を録音する実験においては、JIS の定義する音の大きさと高さが同じ演奏音源を収録し、音色の比較をすることも困難である。二つの知見 (渋谷ら, 1996; 梅本, 1996) の定義、音色を表すパラメータ設計の本質的な難しさより、本論文では、ヴァイオリンの音色の評価については重点を置かずに、特に梅本 (梅本, 1996) が述べている「正しいピッチ」と「ヴィブラート」に着目し、その特徴を取得するパラメータ群を提案する。

4.3.2 既存手法で用いられたパラメータ

ピアノの 1 オクターブ演奏に対する熟達度推定の手法（三浦ら，2010）（以下，1oct ピアノと呼称する）で用いられたパラメータについて述べる．1oct ピアノでは，以下の演奏情報（以下，1oct ピアノ特徴量とする）と，評価パラメータを組み合わせ独自のパラメータ群を提案している．用いた 1oct ピアノ特徴量は，打鍵タイミング，打鍵強度（ヴェロシティ），そして，押鍵時間長である．打鍵タイミングにおいてはメトロノームからの逸脱，打鍵強度においては演奏者の平均打鍵強度からの逸脱，押鍵時間長においては演奏課題における音価からの逸脱を評価するためにそれぞれのパラメータをスプライン補間し得られた傾向曲線からの逸脱を求めることで評価している．次に 1oct ピアノで用いられた各単一音における演奏情報に対する演奏全体の傾向を示すパラメータ（以下，1oct ピアノ統計量とする）により傾向曲線からの逸脱を評価している．本論文では，1oct ピアノ特徴量を基に，Onset, Duration, Velocity, Tempo に関するパラメータを提案する．また，1oct ピアノ統計量を応用した「1oct-vn 統計量」を提案する．

4.3.3 提案パラメータの概要

前述の特徴量に加え本論文で新たに提案する特徴量を説明する．本論文では，1oct ピアノ特徴量では算出できないヴァイオリンの演奏の 2 音間の滑らかさ，相対的な音の明るさ，音のばらつき，F0 ずれ，FM と AM のヴィブラートに注目した特徴量を算出する．それぞれの特徴量の概要について述べる．

- 2 音間の滑らかさを示す特徴量

ヴァイオリンは，弓を往復させ擦弦するボウイングという技術を用いて演奏を行なう．これにより，音の切り替わりの際，切れ目がなく演奏を行なうとは容易ではない．そのため，ヴァイオリン演奏において，音の滑らかさというのは演奏の特徴となり得ると考えられる．

- 相対的な音の明るさを示す特徴

音の明るさというのは，音色を表すパラメータの一種である．音色を単一のパラメータで表すことは困難であると前述した．そこで，音の明るさは既存のスペクトル重心と基準音の音の高さを組み合わせた特徴量を算出する．この特徴量は音色の特徴の一端を表す特徴量になると考えられる．

- 音のばらつきを示す特徴量

ヴァイオリン演奏において、ボウイング技術、弦の抑え方などにより単一音の安定性は揺らぐ。すなわち、音の安定性を確認することにより、ヴァイオリン演奏における、ボウイング技術、弦の抑え方などの技術を表す特徴量になり得ると考えられる。

- F0 に関する演奏特徴量

音高のコントロールは、フレットのないノンフレット楽器であるヴァイオリンにとって容易ではなく、重要な演奏の技術である。そのため、F0 パラメータはヴァイオリン演奏において重要な音高のコントロールを表す特徴量になり得ると考えられる。なお、本論文においては、基本周波数を F0 として表記する慣習にならない F0 特徴量と表現する。

- FM と AM のヴィブラートの特徴量

ヴァイオリンのように演奏音にヴィブラートを付与できる楽器において、演奏音にヴィブラートを付与することは重要な演奏表現のひとつである。そのため、これらの特徴量はヴィブラートの技術を表す特徴量になり得ると考えられる。また、歌声のヴィブラートに関する知見 (Sundberg, J. 著, 榊原健一 監訳, 2008) では、ヴィブラートの周波数の変調と、振幅の変調の関係について述べられている。そこで、本論文においてもヴィブラートを周波数と振幅の両側面に対し、パラメータを設計する。

4.3.4 提案パラメータの概要

本論文では、ここまで述べた音響特徴量と先行研究 (三浦ら, 2010) で用いられた 1oct ピアノ統計量を応用した 1oct-vn 統計量を組み合わせたパラメータを提案する。表 4.1 に提案する音響特徴量と 1oct-vn 統計量の組み合わせを示す。表 4.1 は、縦が音響特徴量、横が 1oct-vn 統計量を表し、それらの組み合わせによりパラメータを得る。また、表 4.1 における網掛けは先行研究で提案されたパラメータをもとにしたパラメータを示す。表 4.1 に示す音響特徴量のうち、基礎的な特徴量である A_0, A_1, A_2, A_3 は 4.3.4.1, 2 音間の音の滑らかさの特徴量である A_4, A_5, A_6 を 4.3.4.2, 音の明るさの特徴量である A_7 を 4.3.4.3, 1 である A_8 を 4.3.4.4, F0 ずれ特徴量である A_9, A_{10}, A_{11} を 4.3.4.5, FM と AM ヴィブラートの特徴量である $A_{12}, A_{13}, A_{14}, A_{15}, A_{16}, A_{17}$ を 4.3.4.6 にてそれぞれ述べる。また、表 4.1 に示す 1oct-vn 統計量 P_0 から P_5 を 4.3.4.7 にて述べる。

表 4.1 提案パラメータ

		平均	標準偏差	傾向曲線と 実演奏との 誤差の二 乗和平方 根	傾向曲線 の最大値と 最小値の 差	傾向曲線 の隣接 2 音間の階 差の二乗 和平方根	基準値と傾 向曲線との 誤差の和
		P_0	P_1	P_2	P_3	P_4	P_5
基礎的な特徴量	A_0	●	●	●	●	●	●
	A_1	●	●	●	●	●	●
	A_2	●	●	●	●	●	●
	A_3	●	●	●	●	●	●
2音間の音の滑らかさの 特徴量	A_4	●	●	●	●	●	●
	A_5	●	●	●	●	●	●
	A_6	●	●	●	●	●	●
音の明るさの特徴量	A_7	●	●	●	●	●	●
音間の F0 のばらつきを 示す特徴量	A_8	●	●	●	●	●	●
音間の F0 のばらつきを 示す特徴量	A_9	●	●	●	●	●	●
	A_{10}	●					
	A_{11}	●					
FMとAM ヴィブラートの特 徴量	A_{12}	●	●	●	●	●	●
	A_{13}	●	●	●	●	●	●
	A_{14}	●	●	●	●	●	●
	A_{15}	●	●	●	●	●	●
	A_{16}	●	●	●	●	●	●
	A_{17}	●	●	●	●	●	●

4.3.4.1 基礎的な特徴量

まず、1oct ピアノ（三浦ら，2010）の 1oct ピアノ特徴量を基に作成した基礎的な特徴量について述べる．これらの特徴量は単音ごとに算出される．このとき，式を用いて数量的な算出を行なう場合は，第 n 番目の音を表す添字を用いるが，特に数式を用いない説明においては添字を用いずに示す．

- Onset (A_0)

発音時刻を表す特徴量を算出する．本論文では，ハンドラベリングによって，音楽音響信号からオンセット時刻を取得した．このハンドラベリングデータは，3 名の被験者にそれぞれハンドラベリングさせ，得られた三つの時刻データを平均したものを Onset (A_0) として算出した．

- Duration (A_1)

発音時間長を表す特徴量を算出する。本論文では、ハンドラベリングによって、音楽音響信号からオンセット時刻を取得した。このハンドラベリングデータに基づき、オンセット時刻間の時間長を Duration (A_1)として算出した。

- Velocity (A_2)

発音強度を表す特徴量を算出する。この特徴量は各単一音の RMS (Root Means Square) 値から算出している。また、単位はデシベルとする。

- Tempo (A_3)

発音時刻を基に瞬時的な演奏のテンポを算出する。本論文では、Duration (A_1)に基づきテンポ値を算出した。単位は BPM(Beats Per Minute)とし、例えば発音時間長が 0.5 秒ならば、Tempo (A_3) の値は 120 となる。

4.3.4.2 2 音間の音の滑らかさの特徴量

次に、2 音間の音の滑らかさを示す特徴量について述べる。これらの特徴量は各単一音間（本論文の演奏課題の単一音は 15 のため、算出されるのは一つの演奏につき 14 個）で算出される。図 4.2 に音の滑らかさを示す特徴量の略図を示す。図 4.2 は、音の切り替わり時付近の波形のエンヴェロープ波形を表す。図 4.2 より、音の切り替わり時付近には 2 つのピークがあることが確認できる。ここでは、エンヴェロープの谷とその前後の二つのピークを算出し、その値を基に特徴量を算出するパラメータを提案する。このとき、二つのピークは計算機により自動で算出している。また、エンヴェロープは、ヒルベルト変換を用いた解析信号であり、50 [Hz]のローパスフィルタによって算出された。

- Peak to peak (A_4)

音の切り替わり時付近の二つのピークの時間長を算出する。この時間長が短いほどこの 2 音は滑らかに接続していると言える。

- Difference of two peaks (A_5)

音の切り替わり時付近の二つのピークの振幅差を算出する。この差が小さいほどこの 2 音は滑らかに接続していると言える。

- Difference from bottom (A_6)

連続する 2 つの音と音の間に存在する谷について着目し、エンヴェロープの谷とその

後に表れるピークとの差を算出する．小さいほどこの2音は滑らかに接続していると言える．

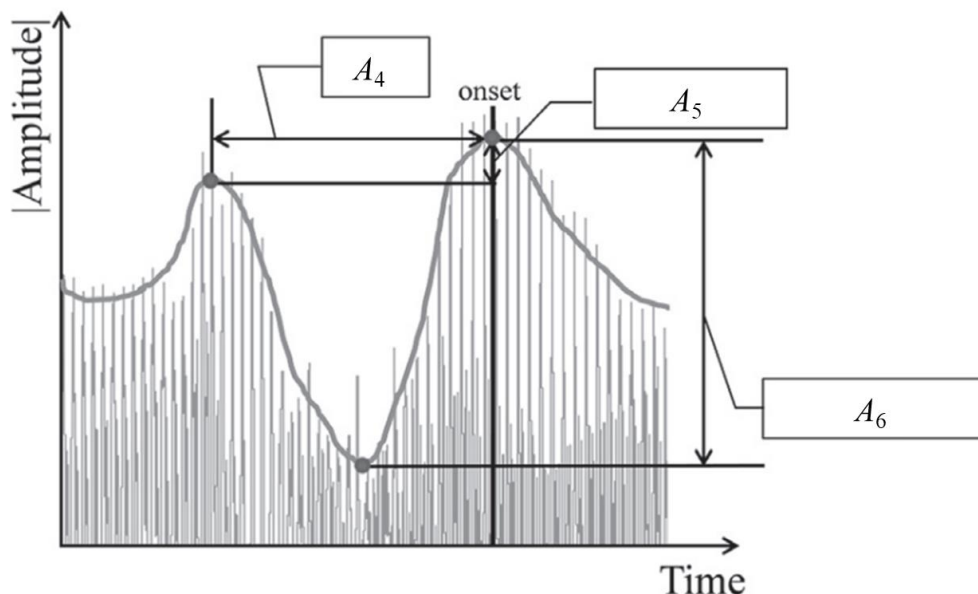


図 4.2 音の滑らかさを示す特徴量の略図

4.3.4.3 音の明るさを示す特徴量

次に、音の明るさを表す特徴量について述べる．この特徴量は単一音ごとに算出される．

- 相対スペクトル重心 (A_7)

音の明るさを示す特徴量として、スペクトル重心 (Spectral Centroid) がある．スペクトル重心 C の算出式を (4.1), (4.2) と (4.3) に示す．

$$C_n = 2^{m_n} \quad (4.1)$$

$$G_{m,i} = \frac{\sum_{b=1}^W \log_2 f[b] \cdot p_{n,i}[b]}{\sum_{b=1}^W p_{n,i}[b]} \quad (4.2)$$

$$m_n = \frac{\sum_{i \in n} G_{n,i}}{S_n} \quad (4.3)$$

ここで、 C_n は第 n 番目の音におけるスペクトル重心、 $G_{m,i}$ は第 n 番目の音における第 i 番目フレームの重心、 m_n は第 n 番目の音における全フレームの重心の平均であり、 f は周波数、 $p_{n,i}[b]$ は第 n 番目の音における第 i 番目フレームの周波数ビン b における振幅スペクトル、 S_n は第 n 番目の音に対する音響信号の総フレーム数、 W は周波数解

析に用いるサンプル点数である。このスペクトル重心を、 fn （第 n 番目の音に対する音高の周波数（ここでは譜面と 12 音平均律による））を対数軸上に変換したもので減算した相対スペクトル重心 (A_7) を算出する。相対スペクトル重心 A_7 の算出式を式 (4.4) に示す。

$$A_{7,n} = \log_2 C_n - \log_2 f_n \quad (4.4)$$

なお、2 を底とする対数軸を用いたのは、音高は周波数が 2 倍となると 1 オクターブ上昇するという性質を組み込むためである。値が大きいほど音は明るい傾向にあることを示す。

4.3.4.4 音のばらつきを示す特徴量

次に、単一音内でのスペクトルの変動を表す特徴量について述べる。この特徴量は単一音ごとに算出される。

- シグマフラックス (A_8)

音の変動量を表すパラメータとして、スペクトル変動量 (Spectral flux) がある。スペクトル変動量 U の算出式を式 (4.5) に示す。

$$U_{n,i} = \sum_{b=1}^{\frac{W}{2}} (p_{n,i}[b] - p_{n,i-1}[b])^2 \quad (4.5)$$

ここで、 i は音響信号に対して周波数解析を行なう際のフレーム ID である。各音に対してのスペクトル変動を算出し、単一音内のスペクトル変動の総和をフレーム数で除算し、更に対象音の RMS 値の平方根で除算することで、シグマフラックス（以後、 A_8 ）が得られる。シグマフラックス A_8 の算出式を式 (4.6) に示す。

$$A_{8,n} = \frac{\sum_{i \in n} U_{n,i}}{S_n R_n} \quad (4.6)$$

ここで、 n は各単一音に振られた ID であり（1 が開始音，終了音が 15 を表す）， R_n は第 n 番目の音の音響信号の RMS 値である。 A_8 の値が小さいほど、音のスペクトルの揺らぎは小さいことを表す。

4.3.4.5 F0 ずれに関する特徴量

次に、音階演奏全体における F0 のずれに関する特徴量について述べる。F0 ずれに関す

る特徴量は、単一音の F0 ずれ量(A_9)については単一音ごとに算出される。単一音の F0 ずれ量(A_9)以外についてはその特性上演奏音源から一つ算出される特徴量であり、1oct-vn 統計量とは組み合わせない。F0 ずれ特徴量は、単位時間ごとに基本周波数を算出できる YIN (De Cheveigné, A. et al., 2002) を用いて算出した第 n 番目の音に対する F0 を用いて特徴量を算出する。

- 単一音の F0 ずれ量 (A_9)

$A_4 = 440$ Hz を基準とし、C4 から始まる平均律また純正律の音高に従って算出した F0 (以後、 $I_{t,n}$, t は純正律又は平均律を表し、 n は各単一音に振られた ID であり 1 が開始音、終了音が 15 を表す) と、第 n 番目の演奏音から計測された F0 との差分 (以後、 $M_{t,i}$, i はフレーム ID) をフレーム内 F0 ずれ量と呼称する。フレーム内 F0 ずれ量はフレームごとの特徴量であるため、単一音当たりの F0 ずれ量 (A_9) を算出するために単一音内におけるフレーム内 F0 ずれ量を平均する。単一音に対する F0 ずれ量 A_9 の算出式を式 (4.7) と (4.8) に示す。

$$A_{9,t,n} = \frac{\sum_{i_n}^{i_{n+1}} q}{i_{n+1} - i_n} \text{ [cent]} \quad (4.7)$$

$$q = 1200 \times \log_2 \frac{M_i}{I_{t,i}} \text{ [cent]} \quad (4.8)$$

ここで、 i_n は第 n 音目の発音時刻におけるフレーム ID、 q はフレーム内 F0 ずれ量である。 $A_{9,t,n}$ は値の絶対値が小さいほど、正確なピッチで演奏できているという特徴を持つ。

- F0gap 傾向 (A_{10})

図 4.3 に A_{10} の概略図を示す。図 4.3 は、横軸は音 ID、縦軸は A_9 を表し、この傾向直線の傾きを A_{10} とする。 A_{10} を算出することで、演奏音のオクターブ伸縮を表すことができる。オクターブ伸縮とは、譜面上の 1 オクターブ間隔 j が、周波数上のオクターブの関係である 2 倍よりも高くなったり、あるいは低くなったりする現象を示すものである。すなわち、 A_{10} の傾きにより演奏音のオクターブ伸縮を得ることができる。

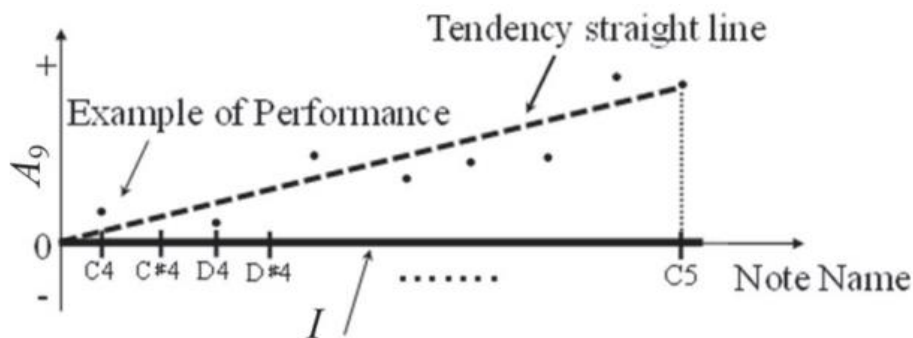


図 4.3 F0gap 傾向

● F0gap 傾向からのばらつき (A_{11})

A_{10} からの差分を算出することで、演奏音の F0gap 傾向(A_{10})からの周波数のばらつき (A_{11}) を算出できる。 A_{11} の算出式を式 (4.9) に示す。

$$A_{11,n} = \sqrt{\frac{1}{N} \sum_{n=1}^N (A_{9,t,n} - l_n)^2} \quad (4.9)$$

ここで、 N は音数、 l_n は第 n 番目の音に対する音響信号における傾向直線の値とする。
 A_{11} の値が小さいほど、F0gap 傾向からの周波数のばらつきが小さいことを表す。

4.3.4.6 ヴィブラートに関する特徴量

次に、FM と AM ヴィブラートに関する特徴量について述べる。これらの特徴量は単一音ごとに算出される。まず、FM ヴィブラートに関する特徴量について述べる。FM ヴィブラート特徴量は、F0 ずれ特徴量と同様に YIN (De Cheveigné, A. et al., 2002) により得られた F0 を用いて特徴量の算出を行なう。また、AM ヴィブラートについては、波形のエンヴェロープを用いる。このとき、F0 とエンヴェロープの変動に対し、それぞれ三つの特徴量を算出する。図 4.4 にこれらの特徴量のイメージを示す。図 4.4 より、波線が F0 エンヴェロープの揺らぎもしくは波形のエンヴェロープ、一点鎖線が揺らぎの傾向、丸が揺らぎの起伏の回数、点線矢印が揺らぎの幅を模擬している。まず、F0 の揺らぎを用いた FM ヴィブラートに関する特徴量について述べる。FM ヴィブラート特徴量は、F0 ずれ特徴量と同様に YIN (De Cheveigné, A. et al., 2002) により得られた F0 を用いて特徴量の算出を行なう、

- F0 カーブの傾向直線の傾き (A_{12})

A_{12} の算出式を式 (4.10) に示す.

$$\begin{pmatrix} A_{12,n} \\ Y \end{pmatrix} = (F^T F)^{-1} F^T y_n \quad (4.10)$$

ここで、 Y は傾向直線の切片であり、 F は第 1 列に YIN の解析結果のフレーム番号、第 2 列に「1」を持つフレーム数×2 の行列、 y_n は第 n 番目の音の音響信号に対する YIN の出力結果を要素として持つフレーム数×1 の列ベクトルである。 A_{12} により、単一音の FM ヴィブラートによる F0 の上昇あるいは下降の傾向を表すことができ、 A_{12} の値が正であれば F0 が上昇傾向であることを、負であれば下降傾向であることを表す。

- F0 カーブの起伏の回数 (A_{13})

A_{13} を算出するために、傾向変化を持たない F0 の揺らぎ γ を得る。 γ の算出式を式 (4.11) に示す.

$$\gamma_n = y_n - (A_{12,n} \times F_\alpha + Y \times e_n) \quad (4.11)$$

ここで、 F_α はベクトル F の第 1 (α) 列、 e_n は第 n 番目の音の音響信号のフレーム数×1 のすべて「1」が格納された列ベクトルである。式(4.11)で得られた γ_n に対し、FFT を行ない、パワースペクトルの最大ピークより FM ヴィブラートによる F0 の起伏の回数 A_{13} を算出できる。 A_{13} の値が大きいほど FM ヴィブラートの起伏の回数が多いことを表す。

- FM ヴィブラートの起伏の幅 (A_{14})

式 (4.12) に A_{14} の算出式を示す.

$$A_{14,n} = \max(\gamma_n) - \min(\gamma_n) \quad (4.12)$$

ここで、 $\max(\gamma_n)$ と $\min(\gamma_n)$ はそれぞれ第 n 番目の演奏音における γ の最大値と最小値である。 A_{14} の値が大きいほど起伏の幅が大きい FM ヴィブラートであることを表す。

次に、AM ヴィブラートに関する特徴量について述べる。AM ヴィブラート特徴量は、音響信号のエンヴェロープに対し FM ヴィブラート特徴量と同様の計算によって特徴量を算出する。このとき、エンヴェロープは、バターワースフィルタによる 20 [Hz] のローパスフィルタによって算出された。

- エンヴェロープの傾向直線の傾き (A_{15})

A_{15} の算出式を式 (4.13) に示す.

$$\begin{pmatrix} A_{15,n} \\ E \end{pmatrix} = (V^T V)^{-1} V^T e_n \quad (4.13)$$

ここで、 E は傾向直線の切片であり、 V は第 1 列 にエンヴェロープ波形のフレーム番号、第 2 列に「1」を持つフレーム数×2 の行列、 e_n は第 n 番目の音のエンヴェロープ波形を要素として持つフレーム数×1 の列ベクトルである。 A_{15} により、単一音の AM ヴィブラートによるエンヴェロープの平均的な増減の傾向、すなわち AM ヴィブラートによる振幅の変化を表すことができ、 A_{15} の値が正であれば振幅が大きくなる傾向であることを表す。

- エンヴェロープの起伏の回数 (A_{16})

A_{16} を算出するために、傾きを持たないエンヴェロープを得る。 ε の算出式を式 (4.14) に示す。

$$\varepsilon_n = e_n - (A_{15,n} \times A_\alpha + E \times e_n) \quad (4.14)$$

ここで、 A_α はベクトル A の第 1 列、 e_n は第 n 番目の音響信号のフレーム数×1 のすべて「1」が格納された列ベクトルである。式 (4.14) で得られた ε に対し、FFT を行ない、パワースペクトルの最大ピークより AM ヴィブラートによるエンヴェロープの起伏の回数 A_{16} が算出できる。 A_{16} の値が大きいほど AM ヴィブラートの起伏の回数が多いことを表す。

- エンヴェロープの起伏の幅 (A_{17})

式 (4.15) に A_{17} の算出式を示す。

$$A_{17,n} = \max(\varepsilon_n) - \min(\varepsilon_n) \quad (4.15)$$

ここで、 $\max(\varepsilon_n)$ と $\min(\varepsilon_n)$ はそれぞれ第 n 番目における ε_n の最大値と最小値である。 A_{17} の値が大きいほど起伏の幅が大きいエンヴェロープであることを表す。

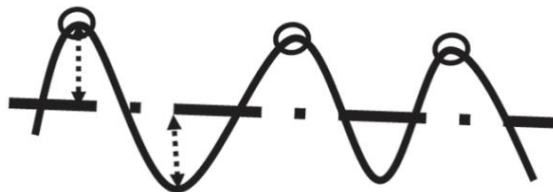


図 4.4 ヴィブラートパラメータの概略

4.3.4.7 loct-vn 統計量

続いて、loct-vn 統計量について述べる。loct ピアノ（三浦ら，2010）で用いられた loct ピアノ統計量を応用する。loct ピアノ統計量は、平均、標準偏差（東京大学教養学部統計学教室，1991；小島，2006；など）の他に、単一音ごとに算出された特徴量から傾向曲線を算出し、特徴量と傾向曲線の間を統計量としている。loct ピアノ統計量をもとに、本論文においては、傾向曲線を 5 次のスプラインカーブとすることとした。

- 平均 (P_0)

P_0 の算出式を式 (4.16) に示す。

$$P_0 = \frac{\sum_{n=1}^N x_n}{N} \quad (4.16)$$

- 標準偏差 (P_1)

P_1 の算出式を式 (4.17) に示す。

$$P_1 = \sqrt{\frac{\sum (x_n - \bar{x})^2}{N - 1}} \quad (4.17)$$

- 傾向曲線と実演奏との誤差の二乗和平方根 (P_2)

P_2 の算出式を式 (4.18) に示す。

$$P_2 = \sqrt{\sum_{n=1}^N (x'')^2} \quad (4.18)$$

- 傾向曲線の最大値と最小値の差 (P_3)

P_3 の算出式を式 (4.19) に示す。

$$P_3 = \max(\hat{x}') - \min(\hat{x}') \quad (4.19)$$

- 傾向曲線の隣接 2 音間の階差の二乗和平方根 (P_4)

P_4 の算出式を式 (4.20) に示す。

$$P_4 = \sqrt{\sum_{n=2}^N (\hat{x}'_n - \hat{x}'_{n-1})^2} \quad (4.20)$$

- 基準値と傾向曲線との誤差の和 (P_5)

P_5 の算出式を式 (4.21) に示す。

$$P_5 = \sum_{n=1}^N \hat{x}'_n \quad (4.21)$$

ここで、 x_n は、 n における音響特徴量値、 \hat{x}' は傾向曲線の値、 x'' は実演奏と傾向曲線の誤差、 \bar{x} は実演奏の平均値、 N は入力された音の総数（本論文においては15）を表す。

4.4 ヴァイオリン演奏の熟達度評価

4.4.1 熟達度評価実験

なお、ここでの熟達度評価とは、例えば音楽大学でヴァイオリン演奏を学ぶ学生や音楽大学の学生が出場するコンクールや音楽大学での演奏試験のようなある程度以上のレベルで熟練された演奏のように繊細な演奏表現の変化等を対象としたものではなく、音楽の演奏を専門としない高校生や大学生が趣味で演奏を楽しむような演奏から音楽大学で学ぶ学生や卒業生に演奏のようにスキル等の差が明確に表れやすい演奏を対象とした評価である。本論文において熟達度推定は、説明変数をパラメータ、目的変数を評価スコアとする線形回帰によって実施する。このため、熟達度推定における目的変数を取得するため記録実験により記録されたヴァイオリン演奏に対する評価実験を行なった。実験刺激には、接近用マイクとアンビエント用マイクにより得られた音を1:1の比率で加算平均した音を用いる（パラメータ算出においては、接近用マイクの音源のみを用いた）。評価者は、演奏歴は21～50年（平均39.8年）、年齢は28～54歳（平均46.8歳）、女性5名（E1～E5）である。いずれもヴァイオリン専攻で音楽大学を卒業し、現在はヴァイオリンの演奏、指導を主な活動としている、プロのヴァイオリン演奏の専門家である。評価者には、各演奏音に対し、1から10の10段階（10が最もよい）で評価させ、「録音環境や楽器に起因する音色については評価しない」、「発音のタイミングが完全に均一でなくとも熟練した演奏だと感じられれば高い評価を、逆にタイミングが均一であっても熟練していない演奏だと感じられれば低い評価をする」という2点の評価基準のみを設けた。この教示を行った理由は、発音のタイミングや欠落音や挿入音な等の単純な評価ではなく、抑揚やヴィブラートといったもう少し複雑な現象による評価を行なうことのできる熟達度推定を目指すためである。以上の2点の基準に従い、各評価が主観的に評価を行なった。なお、すべての評価者には、ヘルシンキ宣言に基づく参加同意書に記入いただいた。

なお、評価においては、評価者がどのような時代の楽曲を主に専門として演奏するかなどによって基準が分かれる部分はあるかと思うが、細かな評価基準について評価者に判断を任せているので、詳細はここでは言及できない。しかし、用いた課題が1オクターブ上下行長音階ということもあり、専門にかかわらない基礎演奏としての評価を行なっているとみなしている。

図 4.5 に、得られた評価スコアを示す。図 4.5 の縦軸は評価スコアの値、横軸が演奏 ID を表し、各系列が評価者を表している。この評価スコアを確認すると、全体で2から9の範囲でばらつきあることが確認できる。人間の心理として、10や1などの極端な数値を避けたということもあると考えられるため、評価の素点については、ある程度のばらつきがあったということが示唆される。また、図 4.6 (a) から (e) に各評価者の評価スコアのヒストグラムを示す。図 4.6 の縦軸は各評価者の各評価スコアの値の回答数、横軸が評価スコアの値を表している。

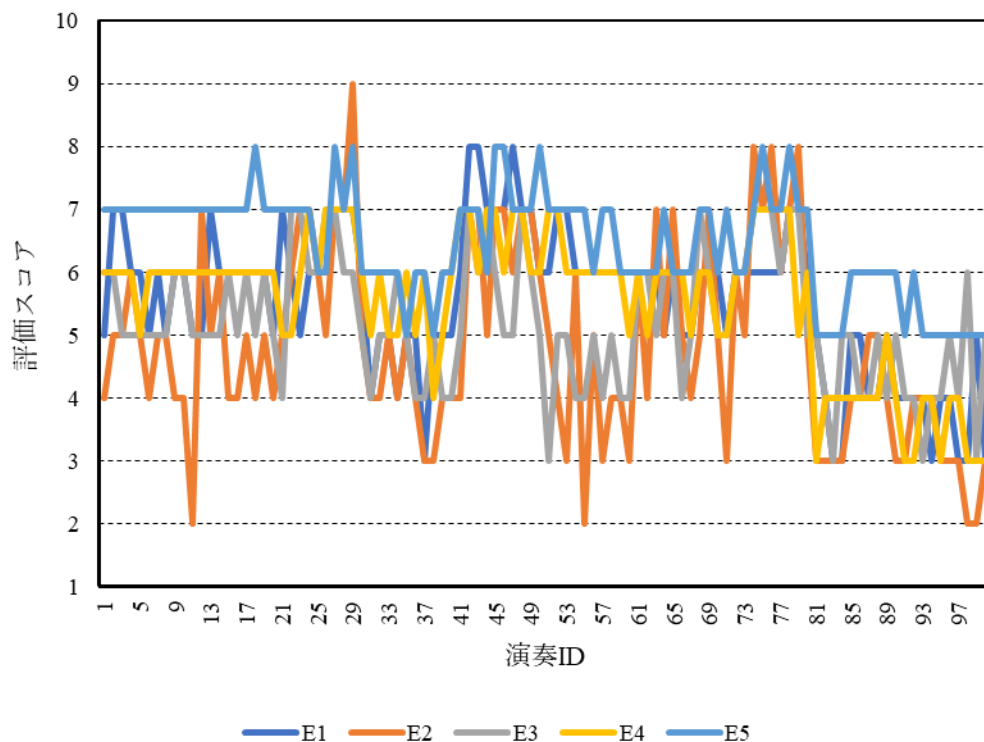
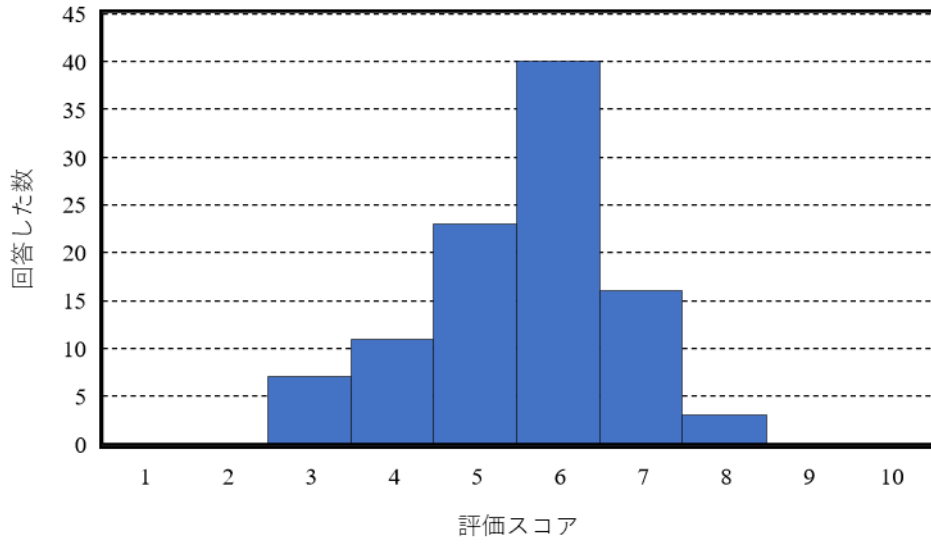
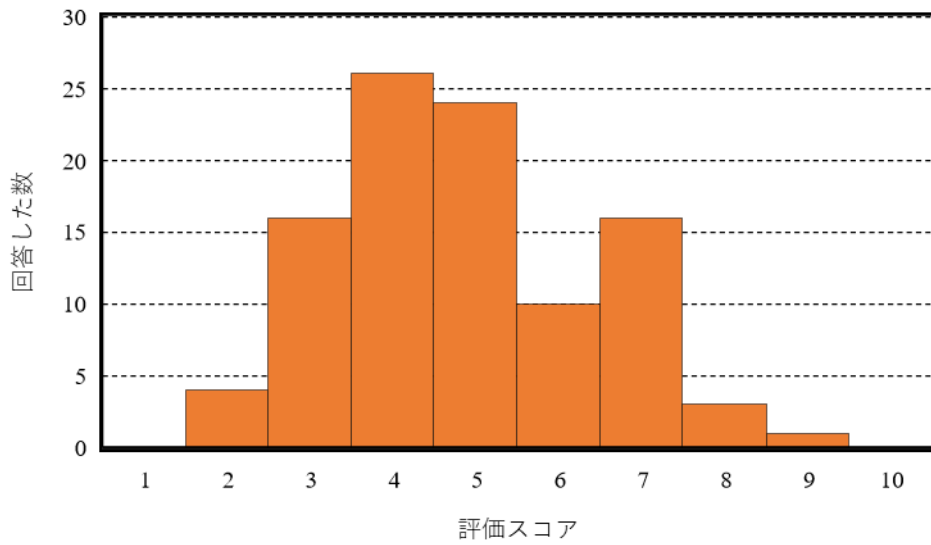


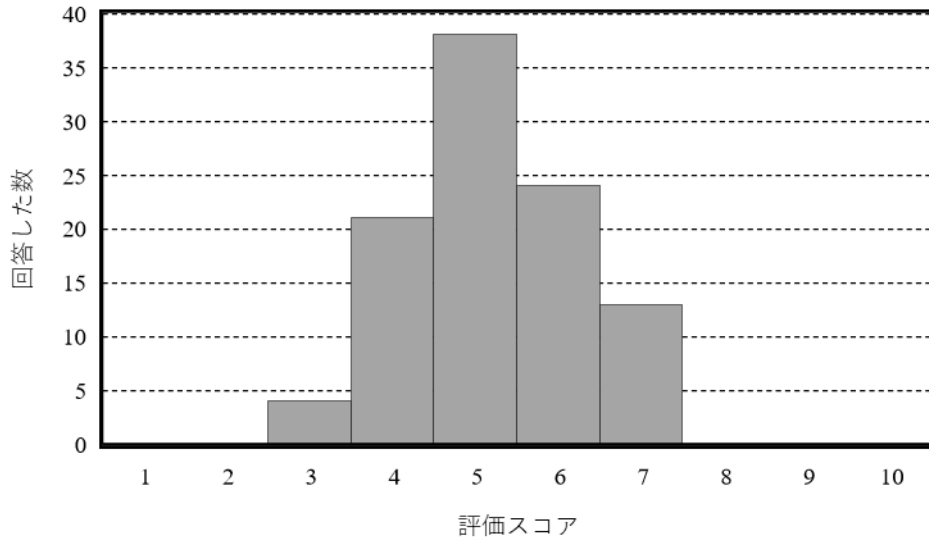
図 4.5 各評価者の評価スコア



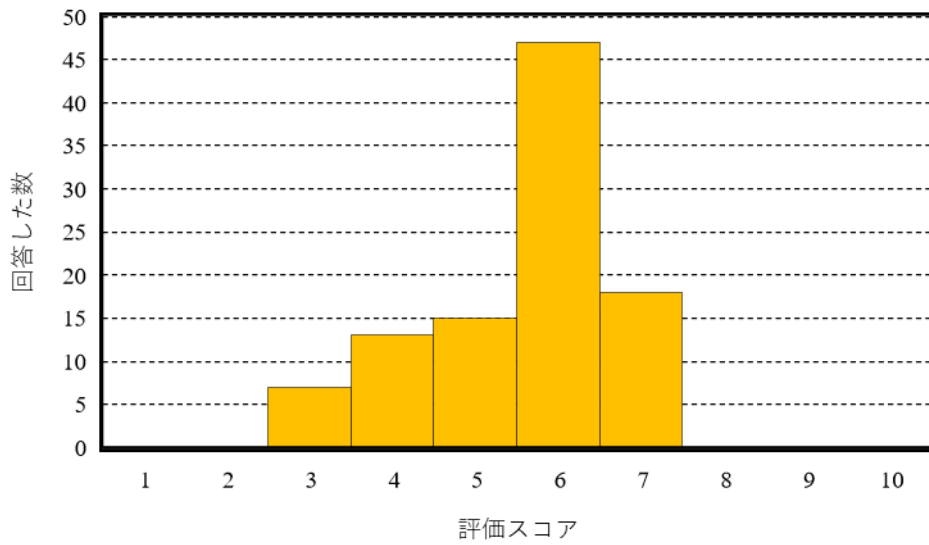
(a) 評価者 E1



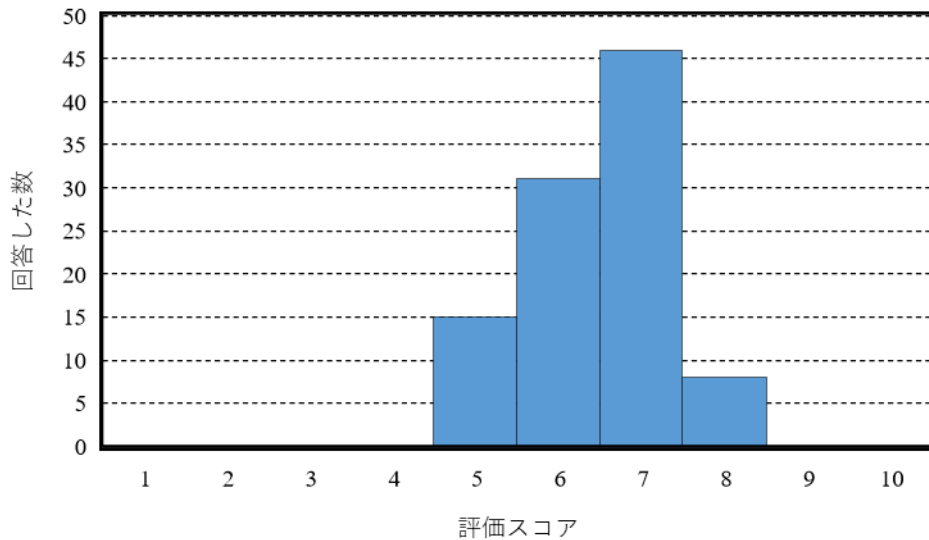
(b) 評価者 E2



(c) 評価者 E3



(d) 評価者 E4



(e) 評価者 E5

図 4.6 各評価者の評価スコアのヒストグラム

4.4.2 評価者間における評価スコアの共通性

評価実験により得られた 5 人の評価スコアの共通性について検討を行なう。本論文においては、熟達度推定に有効なパラメータを提案することが目的であるため、評価者ごとの評価スコアをそれぞれ推定するのではなく、評価者全員から得られた評価スコアの平均によって検証することが、提案パラメータの有効性の議論に相応しいと言える。よって、各個人の評価スコアを推定するのではなく、評価者全員の評価スコアを用いて推定する。しかし、ここでの評価スコアは序数尺度であり、等間隔である保証がないため、間隔尺度に変換する必要がある（難波ら，1998）。そのため、取得された 5 人の評価スコアに大きな偏りがあることは望ましくなく、5 人の評価スコアが平均をとることのできる程度に相関関係があるかということを相関係数によって検証する。表 4.2 に、評価実験により得られた 5 人の評価スコアを z スコアに変換した評価者 z スコアの評価者間相関を示す。表 4.2 より、評価者間の相関係数は $0.45 \leq r \leq 0.72$ ($r = 0.60$, $h = 100$) であることが確認された。また、 $h = 100$ のときには、相関係数の絶対値が 5%有意水準の棄却限界となる相関係数 $r = 0.20$ より大きい場合は無相関とは言えない（市原，1995）ことより、これらの評価スコアには、相関関係が認められると言える。また、図 4.7 に先行研究における評価者間の評価スコアの相関係数を示す。図 4.7 より、先行研究と比較しても十分な有効性があると考

えられる。これより、本論文では、5 人の評価者による各評価スコアに対し、平均 0, 分散 1 の正規分布に従った z スコア（難波ら, 1998）に変換し、それらを平均した平均 z スコアを平均評価スコアとして用いる。

表 4.2 評価者 z スコアの相関係数

	E1_z	E2_z	E3_z	E4_z	E5_z
E1_z	1.00				
E2_z	0.60	1.00			
E3_z	0.48	0.60	1.00		
E4_z	0.72	0.61	0.51	1.00	
E5_z	0.70	0.57	0.46	0.74	1.00

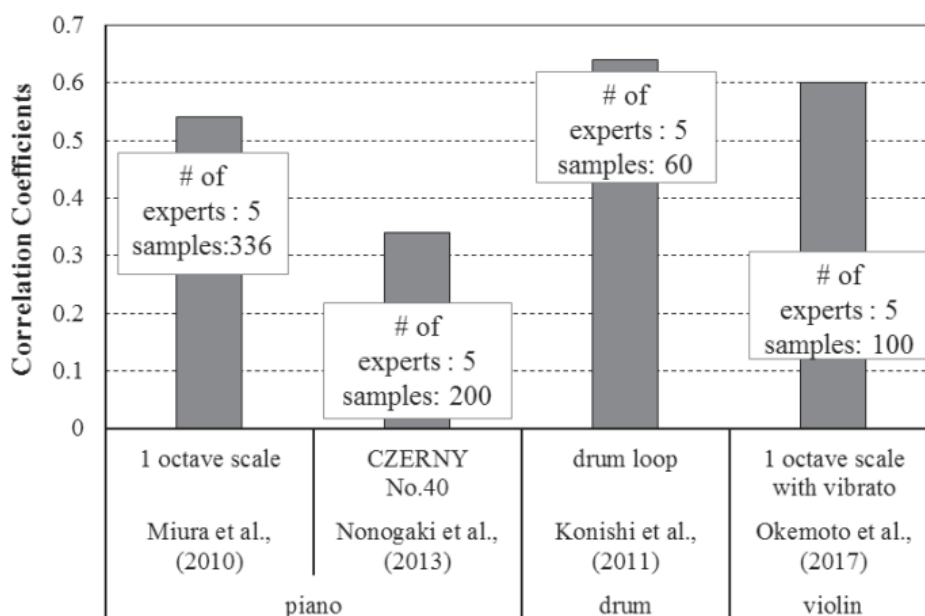


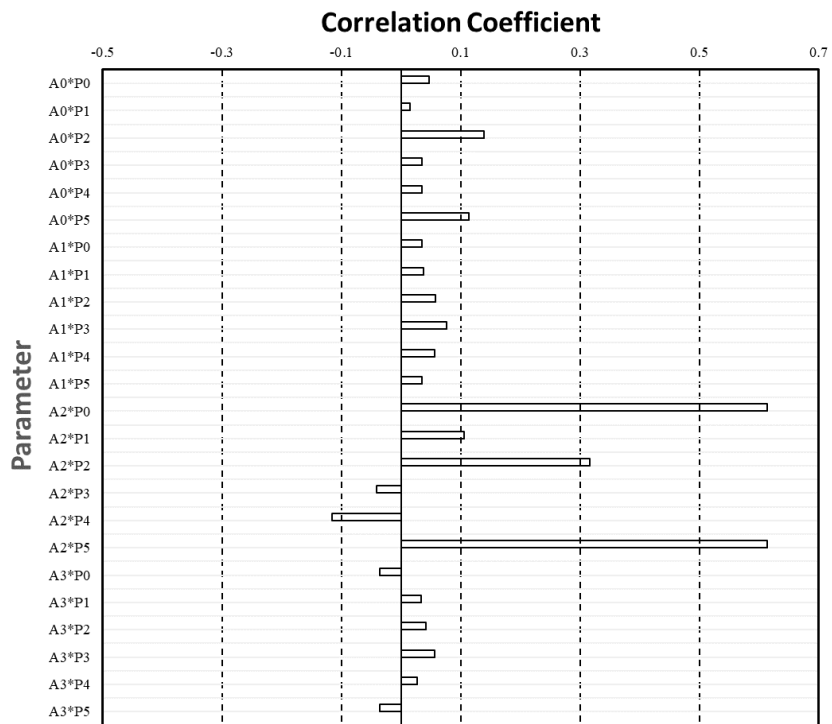
図 4.7 熟達度推定における評価者間の評価スコアの一貫性

4.4.3 パラメータの有効性

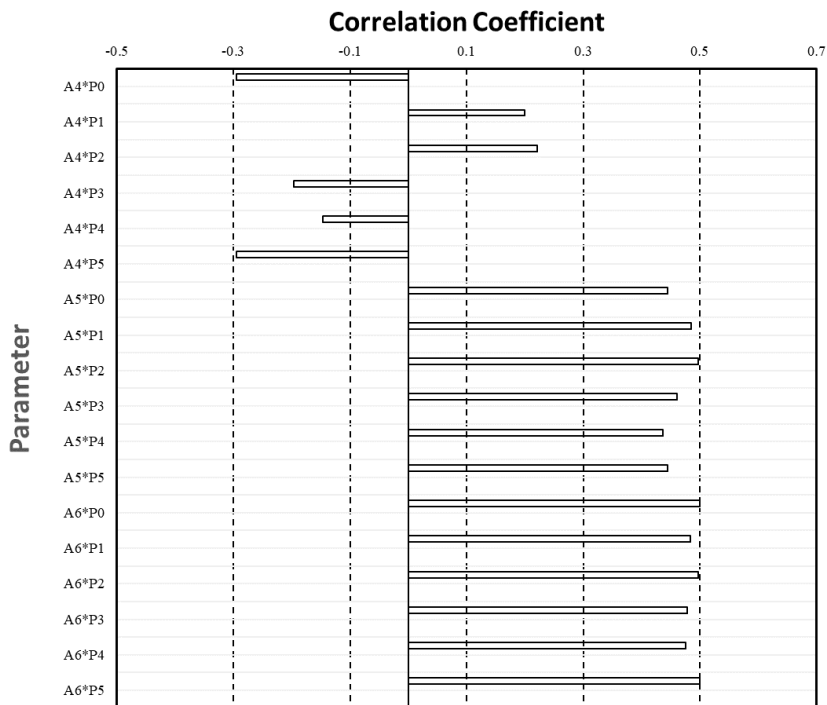
4.4.3.1 パラメータと平均評価スコアとの関係

本節では、提案パラメータと平均評価スコアとの相関係数を示し、提案パラメータによるヴァイオリン演奏の熟達度推定に対する有効性を検討する。収録した 100 通りのヴァイ

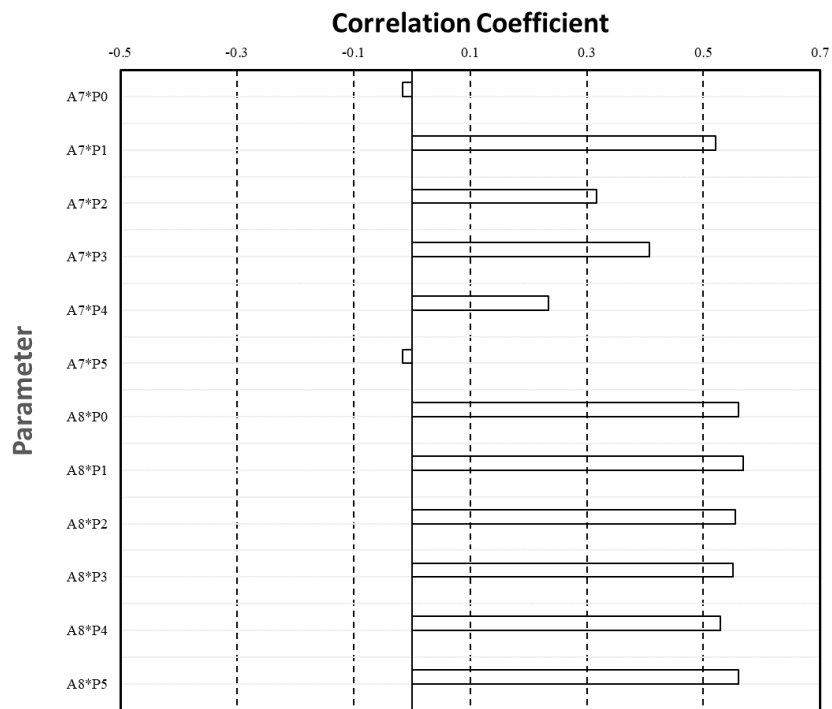
オリン演奏に対し特徴量を算出し、その算出した特徴量から得られたパラメータと平均評価スコアの相関係数を確認することで、各パラメータの熟達度推定への有効性を確認する。図 4.8 に、提案パラメータによる特徴量と平均評価スコアの相関係数を示す。図 4.8 において、(a)は基礎的な特徴量に関するパラメータ、(b)は 2 音間の音の滑らかさの特徴量に関するパラメータ、(c)は音のばらつきを表す特徴量に関するパラメータ、(d)は F0 ずれの特徴量に関するパラメータ、(e)は FM ヴィブラートの特徴量に関するパラメータ、(f)は AM ヴィブラートの特徴量に関するパラメータについて示している。図 4.8(a)から(f)より、 $-0.44 \leq r \leq 0.61$ 程度の相関であることが確認できる。図 4.8(b)より、特に音の滑らかさを表す特徴量 から得られるパラメータの相関が非常に高く、音の滑らかさがヴァイオリン演奏の熟達度を評価するために重要なパラメータであることが示唆された。一方、図 4.8(a)より A_0 などのピアノの先行研究（三浦ら，2010）を基にした特徴量から得られるパラメータの相関が低い傾向にあると確認できる。これは、評価者に設けた「録音環境や楽器に起因する音色については評価しない」、「発音のタイミングが完全に均一でなくとも熟練した演奏だと感じられれば高い評価を、逆にタイミングが均一であっても熟練していない演奏だと感じられれば低い評価をする」という 2 点のうち、後者の評価基準が影響した可能性も考えられる。しかし、ピアノの先行研究（三浦ら，2010）においては、発音時刻や発音強度など音の基本的な情報から得られるパラメータのみであり、本論文では、より複雑な演奏表現を持つ熟達度推定に有効なパラメータを提案できているとも言える。また、 A_5 や A_6 などの 2 音間の音の滑らかさのパラメータの相関係数が高い傾向にあることが確認できる。これより 2 音間の音の滑らかさのパラメータ群は、本論文における評価者の評価スコアに対する熟達度推定については有効である可能性が示唆される。更に、 A_8 は音のばらつきを推定する指標のため、本論文の課題はヴィブラートが付与されていたため、相関が正になったものと考えられる。F0 ずれ特徴量に対し、平均を算出したパラメータについては、平均律を基準としたものは強い負の相関、純正律を基準としたものは弱い正の相関であることが確認でき、平均律を基準としたパラメータ群の方が有効である可能性が示唆される。更に、ヴィブラートの特徴量については、FM と AM ヴィブラートのいずれにおいても、ヴィブラートの幅に関するパラメータが比較的強い相関であることが確認でき、相関の高いパラメータを聴取者が重要視している演奏の特徴であるならば、本論文における評価者はヴィブラートについてはヴィブラートの幅を重視していた可能性が示唆される。



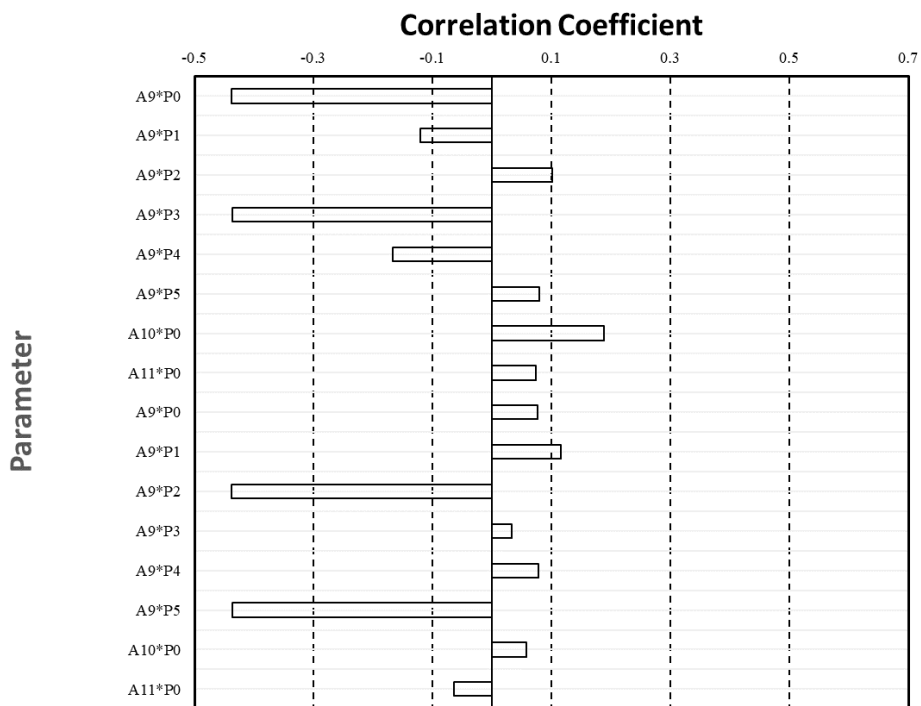
(a) 基礎的な特徴量に関するパラメータ



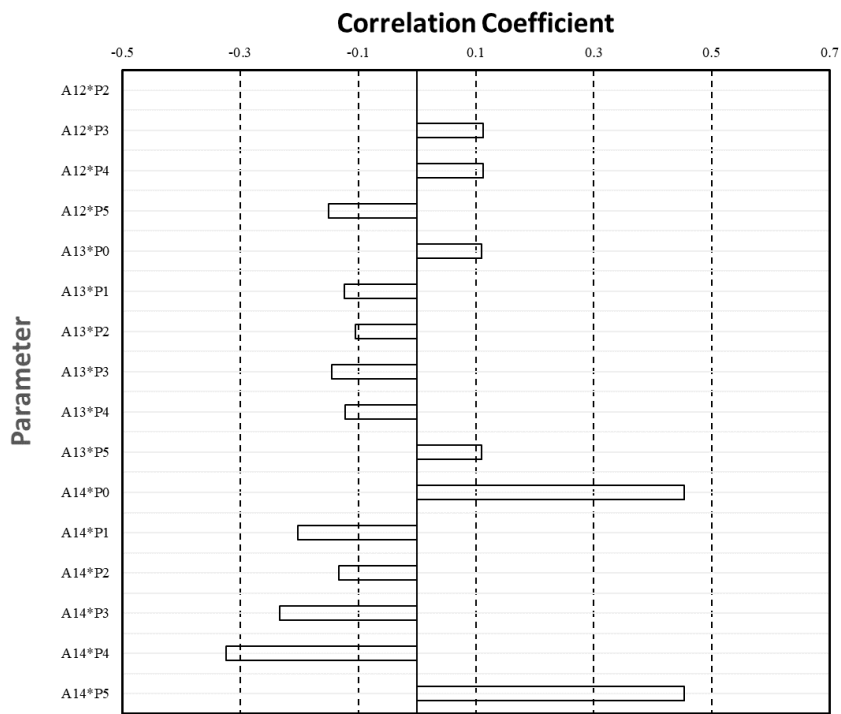
(b) 2音間の音の滑らかさの特徴量に関するパラメータ



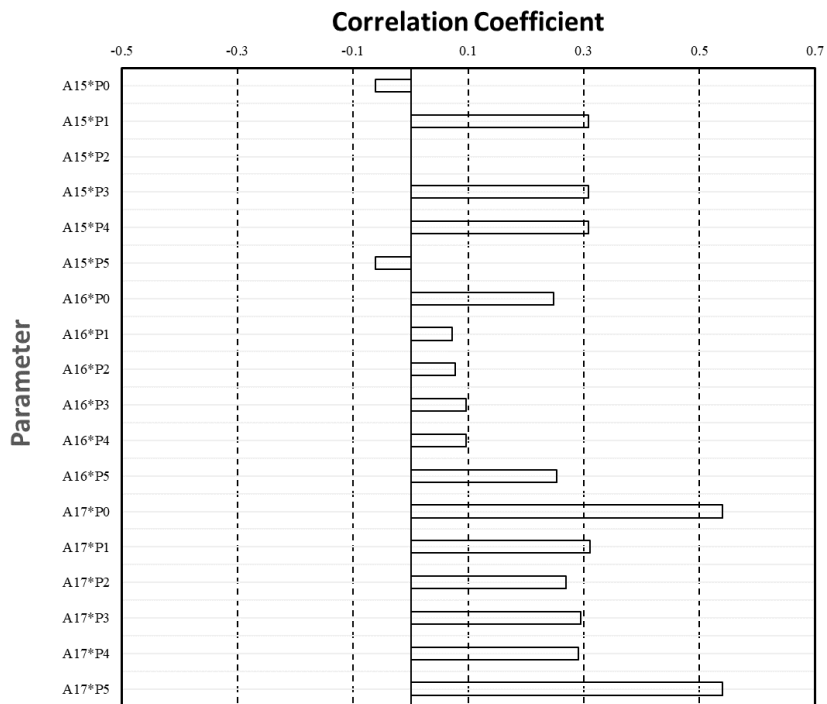
(c) 音のばらつきを表す特徴量に関するパラメータ



(d) F0 ずれの特徴量に関するパラメータ



(e) FM ヴィブラートの特徴量に関するパラメータ



(f) AM ヴィブラートの特徴量に関するパラメータ

図 4.8 パラメータと平均評価スコアの相関

4.4.3.2 機械学習による熟達度推定

ここでは、提案パラメータによる熟達度推定を行ないパラメータの熟達度推定に対する有効性を検証する。本論文は、ヴァイオリン演奏の特徴を表すパラメータを提案することを目的としてきた。そして、提案パラメータと平均評価スコアを用いて熟達度推定を行なうことにより、提案パラメータの有効性が間接的にはあるが検証できると考えられる。本論文では、提案パラメータの有効性を検証するために、記録演奏から算出された提案パラメータ群を説明変数、評価スコアを目的変数とした 10fold-CV (荒木, 2018) で線形回帰による推定を行なった。なお、この推定によって得られたスコアを熟達度スコアと呼称する。このとき、提案パラメータ群は、すべての提案パラメータを用いた「loctvib パラメータセット」、すべての提案パラメータの内 F0 ずれパラメータを平均律で算出したパラメータセット（「loctvib eq パラメータセット」と呼称）、F0 ずれパラメータを純正律で算出したパラメータセット（「loctvib pure パラメータセット」と呼称）の 3 通りをパラメータ群として用いた。図 4.9 に、熟達度スコアの推定精度として線形回帰による熟達度スコアと平均評価スコアの相関係数と決定係数 R^2 を示す。なお、線形回帰では、パラメータの多重共線性の問題を解消するために、パラメータに対し主成分分析 (佐藤, 1971; 堤ら, 1998; 吉村ら, 2011) を行なった。図 4.9 より、三つの提案パラメータ群を用いた相関係数は、 $0.76 \leq r \leq 0.78$ であり、最も相関係数が高いのは相関係数が 0.78 の loctvib パラメータセットを用いた線形回帰による推定手法であることが確認できる。

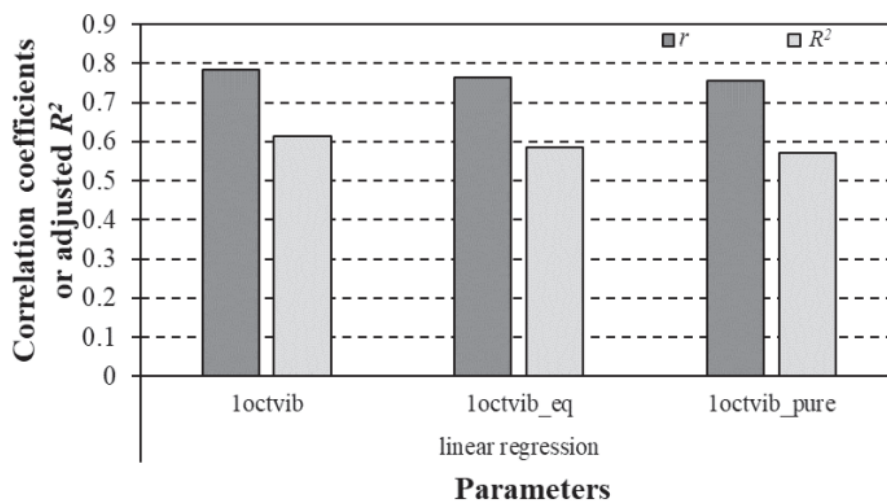


図 4.9 熟達度の推定精度

4.4.4 先行研究との熟達度推定結果の比較

4.4.3.2の結果を、楽器演奏を対象とした熟達度推定の結果と比較し、提案パラメータの有効性について検討する。4.4.3.2において最も推定率が高かった 1octvib パラメータセットを用いた線形回帰による熟達度推定の結果との比較を行なう。図 4.10 に、先行研究と本論文での熟達度の推定精度を示す。図 4.10 中の、チェロの演奏に対する熟達度の推定の研究 (Yamasaki et al., 2014) は提案パラメータ群から F0, FM と AM ヴィブラートパラメータを除いたパラメータ群によってチェロの演奏に対し熟達度の推定を行なっている。このチェロの研究 (Yamasaki et al., 2014) と比較すると熟達度の推定率が上昇しており、F0 ずれ, FM と AM ヴィブラートパラメータの有効性が示唆される。図 4.10 より、本論文での推定精度は先行研究と同等の推定精度を示していることが確認できる。また、図 4.7, 図 4.10 より、本論文での推定精度は、評価スコアの相関と比較し同等以上の値であることが確認できることより、人間にとっても複雑である評価を計算機が行なうことができているということが示唆される。

さらに、 $h=100$ の場合は、相関係数の絶対値が 0.20 を超えていれば無相関であると言えない (市原, 1995) ことより、熟達度スコアと評価スコアの相関は無相関ではないということが確認できる。これより、ヴァイオリン演奏の熟達度推定、言い換えるとヴァイオリン演奏の評価スコアの推定のために提案したパラメータ群は熟達度推定に有効であることが示唆される。

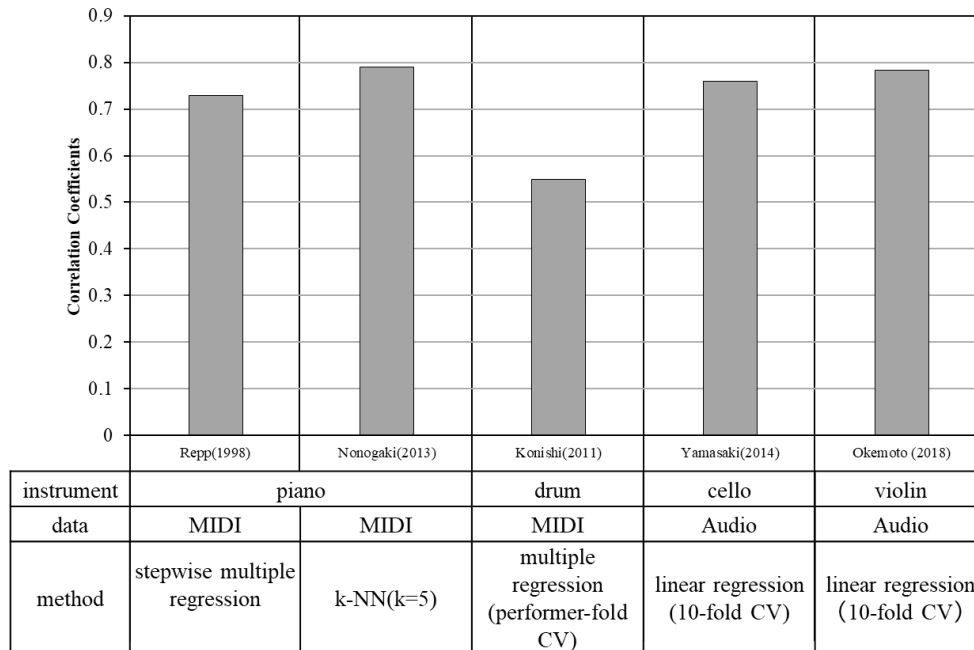


図 4.10 先行研究との比較

4.4.5 考察

これまでの結果より、目的であるヴァイオリン演奏の音の長さ、音の高さ、音の強さに焦点を当てたパラメータの設計、そしてその有効性が説明できたと考えられる。これは、提案パラメータを用いた熟達度推定の結果が先行研究と比較し、高い水準であるためである。今後の展望として、特徴量ごとの評価、例えばヴィブラートの評価、音高に対する評価などを行うことが望まれる。また、本論文における提案パラメータを応用することで、ヴァイオリン以外の楽器による熟達度の推定が行なえとされる。ピアノやドラムを対象とした研究（三浦ら，2010；Nonogaki et al., 2011；小西ら，2011）においては、楽器演奏の情報が MIDI 信号から得られているが、本論文では音響信号を対象としているため、MIDI 情報を用いることのできない楽器、例えば吹奏楽器などに対して、提案パラメータを用いることで熟達度の推定が高い水準で行うことができると期待される。しかしながら、本論文において、ヴァイオリンの音色については考慮したパラメータは設計できていない。これは、前述の要因のためであるが、楽器演奏において音色は考慮すべき問題であると考えられる。楽器の音色に関する研究（安藤，2003）では、様々な楽器の音色について述べられており、その中のバロック音楽のリコーダの音色評価に関する知見では、リコーダ演奏の専門家が音楽の時代を特定すれば評価が可能だと述べたということが記載されており、

楽器演奏において、時代によって求められる音色は変化するということが示唆される。ヴァイオリン演奏においても同様のことが示唆されると考えられるため、今後は演奏課題を楽曲とし、その演奏課題の音楽の時代において求められる音色を考慮した演奏の熟達度推定手法が望まれる。しかし、本論文における評価実験ではヴァイオリン演奏の音色に対する評価を行なっておらず、音色に特化した評価実験が必要である。

4.4.6 ヴァイオリン演奏の熟達度評価に関するまとめ

本論文では、音楽の演奏を専門としない高校生や大学生が趣味で演奏を楽しむような演奏から音楽大学で学ぶ学生や卒業生に演奏のようにスキル等の差が明確に表れやすい演奏を対象としてヴァイオリン演奏の熟達度推定に用いるパラメータ群の提案を行なった。提案パラメータは、ヴァイオリン演奏の特徴を考慮した音の長さ、音の高さ、音の強さ、ヴィブラートに焦点を当てたパラメータの設計を行なった。また、ヴァイオリン演奏の録音とその演奏に対する熟達度の評価実験を行なった。そして、ヴァイオリン演奏の音響信号と熟達度の評価実験より得られた平均評価スコアを用いて、提案パラメータを用いた熟達度推定を行なった。その結果、先行研究と比較しても高い水準で熟達度推定ができており、提案パラメータの有効性が示唆された。しかしながら、提案パラメータは音色に対して十分な考慮ができていないパラメータであるとは言えないため、今後は音色を考慮したパラメータ設計が望まれる。

4.5 ヴァイオリン演奏に対する熟達・技術・表現の評価と音響パラメータの関係

4.5.1 目的

4.4 までで、ヴァイオリン演奏を対象とした楽器演奏の熟達度を計算機によって客観的に評価するための手法について述べた。これらの研究では、演奏の音響信号からその音響的な特徴を表すパラメータを算出し、あらかじめ取得した主観評価スコアを用いて、演奏の熟達度を推定する手法が提案されている。また、これらの研究では、楽器演奏の熟達度という点についてのみ言及しており、楽器演奏の技術力、表現力といった側面には言及されていない。しかし、楽器演奏において、例えば「演奏技術に対して批判があるものの、

表現に対しては称賛されている」という演奏は存在する。つまり、演奏においては、技術と表現という2つの側面からについても評価を行なうことがあり得る。そこで本報告では、演奏に対する熟達、技術、表現の評価の関係について調査する。特に、演奏における熟達、技術、表現の評価のスコアに対する比較、音響パラメータを用いて各評価スコアと物理特徴との関係に対し線形回帰による調査を行なう。

4.5.2 検証の方針

本報告では、4.4 述べたヴァイオリン演奏の熟達度評価 (Okemoto et al., 2018) で用いられたプロのヴァイオリン演奏家による演奏の熟達、技術、表現に対する主観評価を用いて、評価者内と評価者間の各評価の共通性を相関係数によって評価する。続いて、各評価の主観評価を、音響信号から取得した音響パラメータによって推定し、その結果から、主観評価の各評価スコアと物理特徴を表す音響パラメータとの関係を明らかにする。

4.5.3 主観評価スコアの比較

まず、プロのヴァイオリン演奏家による、演奏の熟達、技術、表現に対する主観評価 (Okemoto et al., 2018) の評価スコアを分析する。ここでは、評価者内と評価者間の評価を相関係数によって分析する。これによって、各評価者内の各評価スコアに対する共通性と、各評価に対する各評価者間の共通性を確認する。

4.5.3.1 主観評価の対象演奏と評価者

主観評価の対象となる演奏課題として4.2と同様の1octvibを用いた。1octvibを10人のアマチュア演奏者に10回ずつ演奏させ、計100通りの演奏を録音した。録音の際にはマイクは接近とアンビエントの2種類を用い、評価の際にはそれらを合成した音源を用いた。評価者は5名で、いずれも音楽大学を卒業し、現在は演奏活動やヴァイオリン演奏の指導を行なっている。評価の際には、熟達、表現、技術の評価に対しては、それぞれが独立性を保った評価になるように指示した。例えば、技術を評価する際は熟達、表現については考慮しないようにするよう指示をした。この指示のもと、熟達、表現、技術に対し、それぞれ1-10の10段階(10が最も良い)で評価させた。なお、この評価者は4.4で述べた熟達度評価の評価者と同一である。

この評価において、各評価者がどのような基準で、熟達、表現、技術の3項目について評価を行なったのかは言及することができず、独立性がある指標であるのかという評価も困難な側面はある。また、演奏課題が1オクターブ上下行長音階という基礎的なこともあり、熟達、表現、技術という基礎評価とは言いづらい指標をどこまで評価できるのかという側面もある。しかし、3つの評価については、実際の楽曲においても評価の基準として有する指標であると考えられ、これらの評価が実際の演奏についても評価できる可能性を検討するためにも、本章ではまず1オクターブ上下行長音階という課題でこの評価の推定可能性や重要視するパラメータの調査を行なう。

4.5.3.2 各評価者の評価結果

各評価者の熟達評価、表現評価、技術評価に対する共通性を調査する。同一の演奏に対する評価者内における各評価の共通性を確認するために、各評価者内の各評価スコアに対し、相関係数を求めた。表4.3に、5人の評価者内での各評価スコアの相関係数を示す。表4.3より、いずれの評価者内においても熟達、表現、技術に対する評価スコアは、相関係数が0.65を上回っており、一定の共通性を持っていることが確認できる。特に、熟達と技術に対する評価スコアの相関は平均0.88であり、全体平均の0.80よりも高いことが確認される。

表 4.3 各評価者の各評価スコア間の相関

(a) 評価者 E1

	熟達	表現	技術
熟達	1.00		
表現	0.85	1.00	
技術	0.86	0.85	1.00

(b) 評価者 E2

	熟達	表現	技術
熟達	1.00		
表現	0.76	1.00	
技術	0.88	0.83	1.00

(c) 評価者 E3

	熟達	表現	技術
熟達	1.00		
表現	0.65	1.00	
技術	0.79	0.71	1.00

(d) 評価者 E4

	熟達	表現	技術
熟達	1.00		
表現	0.82	1.00	
技術	0.86	0.86	1.00

(e) 評価者 E5

	熟達	表現	技術
熟達	1.00		
表現	0.65	1.00	
技術	0.99	0.65	1.00

4.5.3.3 主観評価の分布

各評価スコアに対する評価者間における評価スコアの共通性を調査する。その確認のため、各評価スコアにおける各評価者スコアに対する相関係数を求める。表 4.4 に、各評価スコアに対する各評価者間の相関係数を示す。表 4.4 より、相関係数の平均は、熟達度評価については 0.60、表現評価については 0.48、技術評価については 0.63 であった。これより、熟達と技術の評価スコアと比べ、表現の評価スコアの相関係数の平均は低い傾向であることが確認された。

表 4.4 各評価に対する評価者間の相関

(a) 熟達

	E1	E2	E3	E4	E5
E1	1.00				
E2	0.60	1.00			
E3	0.48	0.60	1.00		
E4	0.72	0.61	0.51	1.00	
E5	0.70	0.57	0.46	0.74	1.00

(b) 表現

	E1	E2	E3	E4	E5
E1	1.00				
E2	0.45	1.00			
E3	0.55	0.48	1.00		
E4	0.68	0.18	0.45	1.00	
E5	0.59	0.51	0.50	0.44	1.00

(c) 技術

	E1	E2	E3	E4	E5
E1	1.00				
E2	0.62	1.00			
E3	0.58	0.67	1.00		
E4	0.83	0.55	0.50	1.00	
E5	0.75	0.62	0.44	0.76	1.00

4.5.4 主観評価スコアの推定

演奏に対する各評価スコアと音響パラメータの関係を調査する。本論文では、この調査のために、評価スコアを目的変数、音響パラメータを説明変数とし、線形回帰による各主観評価スコアの推定を行ない、線形回帰により得られた相対重みにより、各評価スコアと音響パラメータとの関係について確認する。各評価スコアは、正規された評価スコアを目的変数として用いた。なお、線形回帰では M5 アルゴリズム (Quinlan, J. R. et al., 1992) によるパラメータ削除を行なっている。

4.5.4.1 音響パラメータ

本論文では、4.3, 4.4 のヴァイオリン演奏の熟達度評価 (Okemoto et al., 2018 ; 桶本ら, 2020) で用いられた演奏特徴パラメータと統計パラメータを組み合わせた 106 の音響パラメータによるパラメータセットを用いた。このパラメータセットのうち、楽譜に示された音高と実際の音高の差を表す音響パラメータには、平均律と純正律に着目したパラメータが含まれており、いずれかが採用される。そのうち、本論文では、平均律に着目した音響パラメータを採用し、純正律に対するパラメータは除外した 98 の音響パラメータを用いる。これは 4.3, 4.4 のヴァイオリン演奏の熟達度評価 (Okemoto et al., 2018 ; 桶本ら, 2020)

において、平均律の音響パラメータを用いた場合の熟達度推定の推定精度が高かったためである。用いたパラメータセットは熟達度推定と同様の 4.1 示す。

4.5.4.2 各評価者の評価スコアの推定

各評価スコアの推定を行なう。ここでは、各評価者の各評価スコアを目的変数、各音響パラメータを説明変数とした線形回帰によって推定を行ない、得られた推定結果と目的変数である各評価スコアの相関係数で推定精度を評価する。

図 4.11 に各評価者の各スコアに対する推定結果を示す。図 4.11 は、横軸が評価者 ID、縦軸が推定精度を表し、各ロッドが各評価を示している。図 4.11 より、E1 から E5 の各評価スコアを平均すると、熟達は 0.44、表現は 0.47、技術は 0.55 である。また、評価者 E1, E2, E3, E4 においては、熟達度評価、表現評価、技術評価の 3 者比較ではいずれも技術評価が最も高い。これより E1 から E4 の 4 名においては、熟達、表現の評価スコアと比べ、技術の評価スコアに対する推定精度が最も高いことが確認できる。これより、5 人の各評価スコアの平均は技術評価に対する推定精度が最も高く、E1 から E4 においてもその傾向は保たれている。

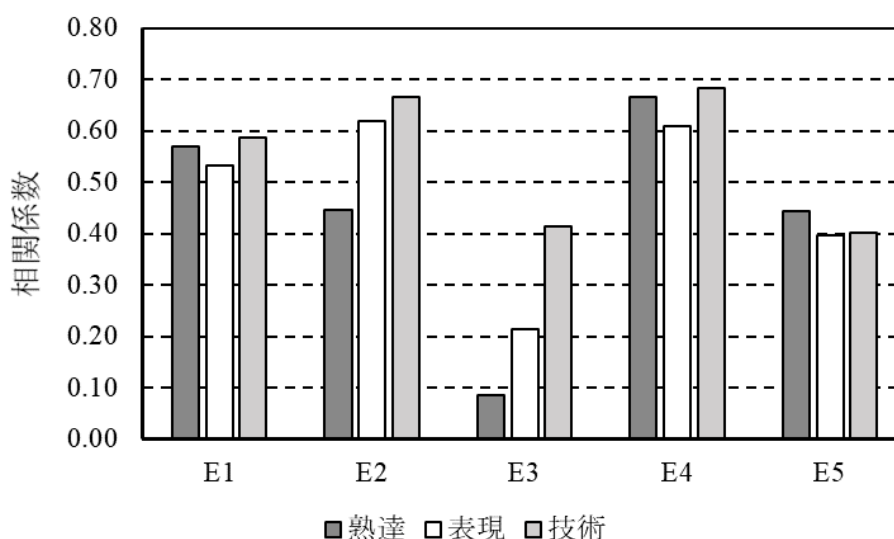
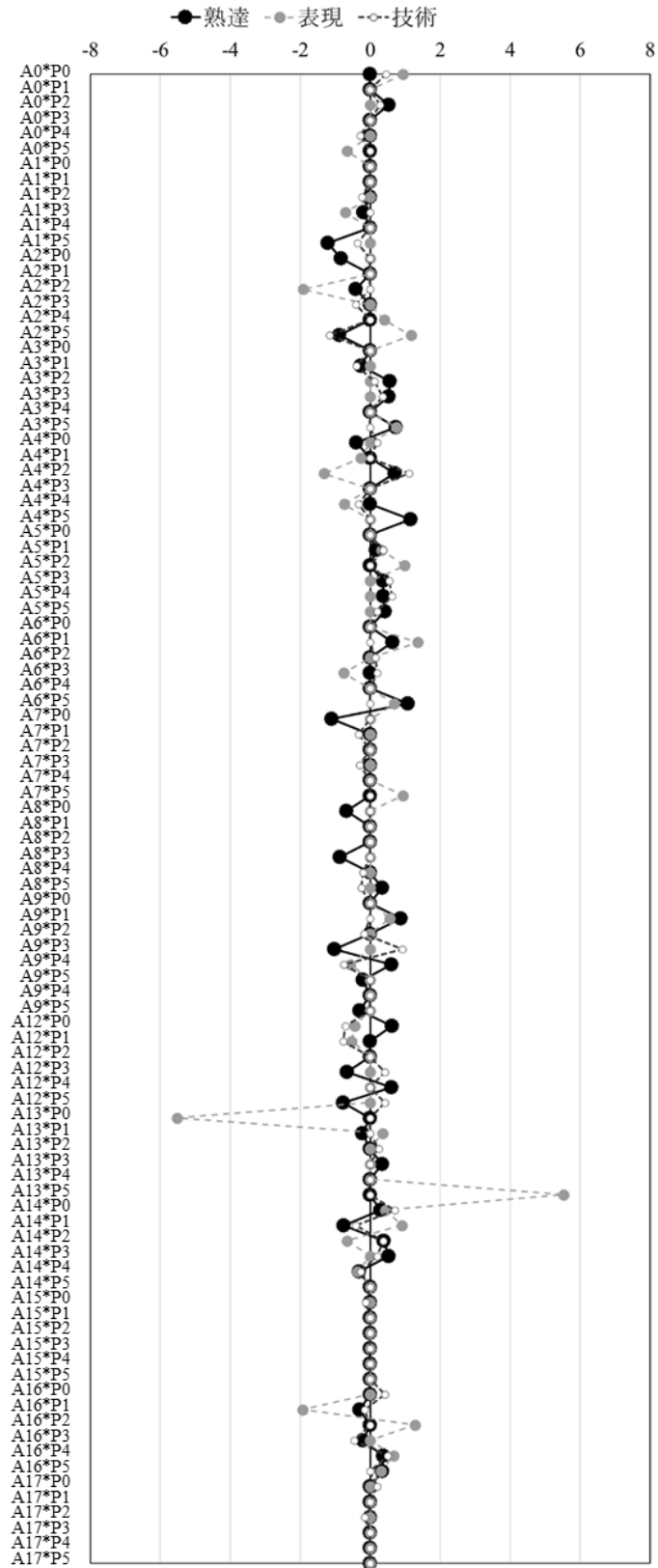


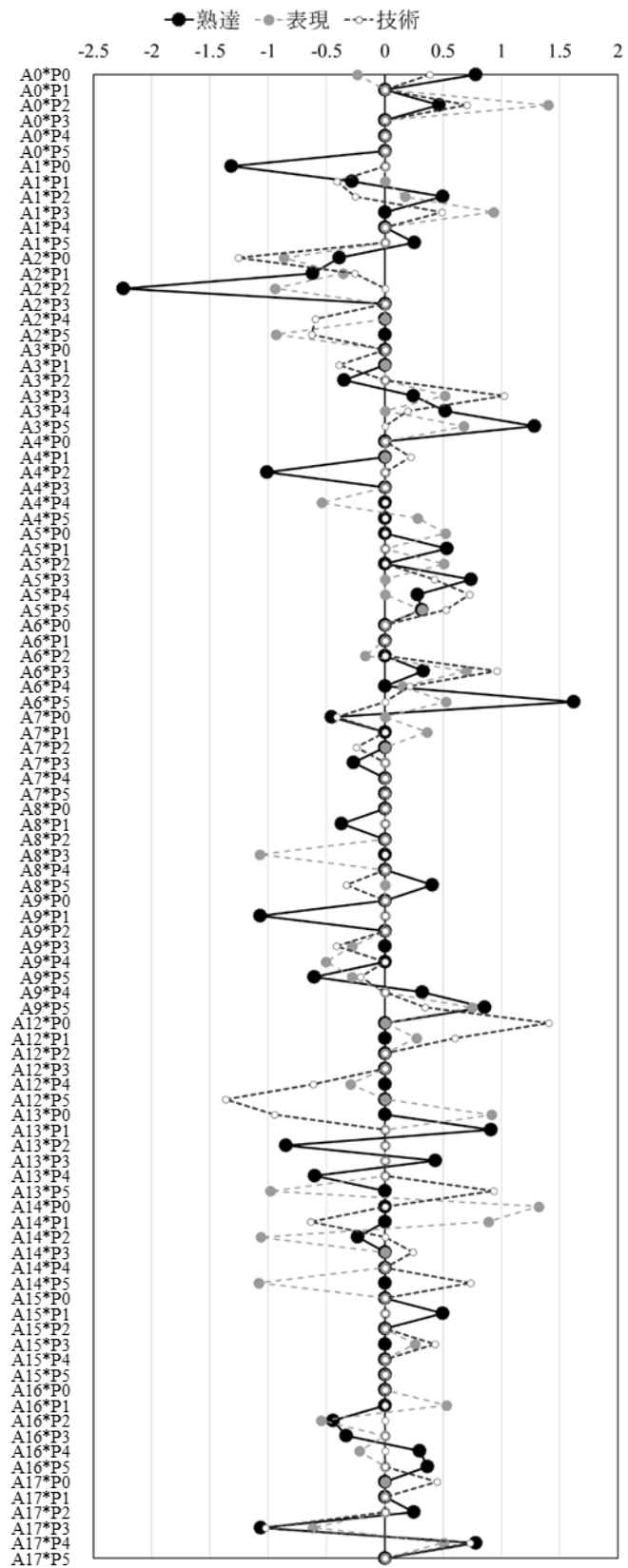
図 4.11 各評価者の各評価スコアと対する推定精度

4.5.4.3 パラメータの相対重み

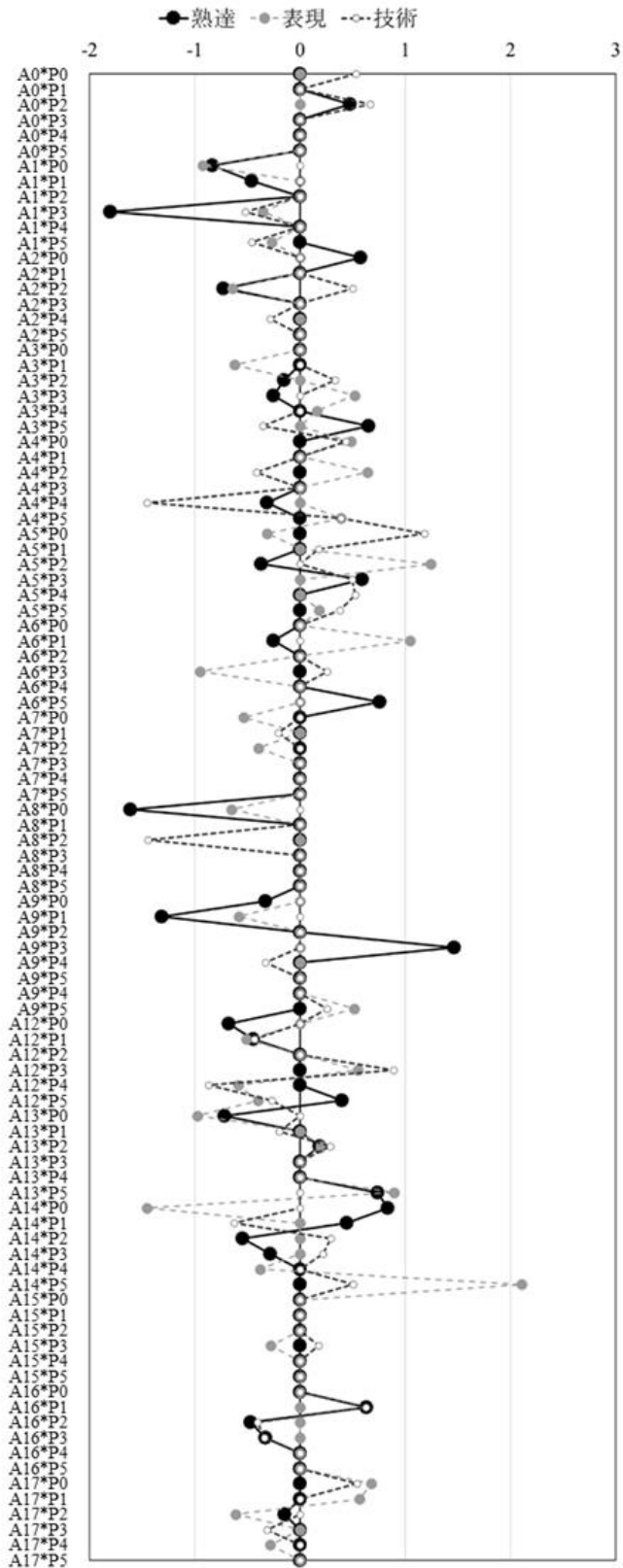
各評価者が各評価スコアに対し評価を行なう際に、演奏のどのような特徴に着目しているのかということ、4.5.4.2で行なった線形回帰により得られた相対重みから検証を行なう。線形回帰の相対重みを調べることにより、各評価者の各評価に内在する評価の基準を明らかにすることができるため、ここでは線形回帰の相対重みを調査する。図 4.12 に各評価者の各評価に対するパラメータの相対重みを示す。図 4.12 において縦軸はパラメータ名、横軸が相対重みを表す。なお、4.5.4.2 の線形回帰では M5 アルゴリズム (Quinlan, J. R. et al., 1992) によるパラメータ削除を行なっているため、相対重みが 0 のパラメータも複数存在する。また、図 4.12 における各評価者内の各評価スコアを推定した際の相対重みの共通性を確認するため、各評価スコアの相関係数を算出した。表 4.5 に各評価者の各評価スコアを推定した際の相対重みの相関係数を示す。図 4.12, 表 4.5 より、各評価者内で熟達度評価、表現評価、技術評価に対し、相対重みの値は必ずしも一致しないことが確認できる。例えば、E1 は、表現評価では FM ヴィブラートの振幅回数の平均に関するパラメータの相対重みの値が他とのパラメータの相対重みの絶対値と比較し大きいものの、他の評価者ではそのような傾向は確認できない。つまり、この相対重みが評価者の重要視する程度を表すとすれば、重要視するパラメータは評価者に依存すると言える。図 4.12(b)より、E2 における熟達度の評価では、 $A_2 * P_3$ や $A_3 * P_5$ のように 4.3.4.1 で述べた基礎的なパラメータの相対重みの絶対値が、 A_{16} のような 4.3.4.6 で述べたヴィブラートに関するパラメータなどの相対重みの絶対値と比較し大きい傾向にあり、熟達度評価の際には音の大きさやテンポなどの基礎的な演奏スキルを重要視していたと考えられる。図 4.12(c)より、E3 の熟達度評価においては、E2 と異なり、 A_{16} のような 4.3.4.6 で述べたヴィブラートに関するパラメータを含む幅広いパラメータで同程度の相対重みの絶対値となっており、E2 と E3 においては、熟達度の評価基準が異なる傾向にあることが確認できる。表 4.4 より、E2 と E3 の熟達度の主観評価スコアの相関は 0.60 とある程度の相関関係にあることが確認できる。これより、演奏全体への評価は近い傾向であっても、各評価者に内在する評価基準は異なる傾向を持つ可能性が示唆される。



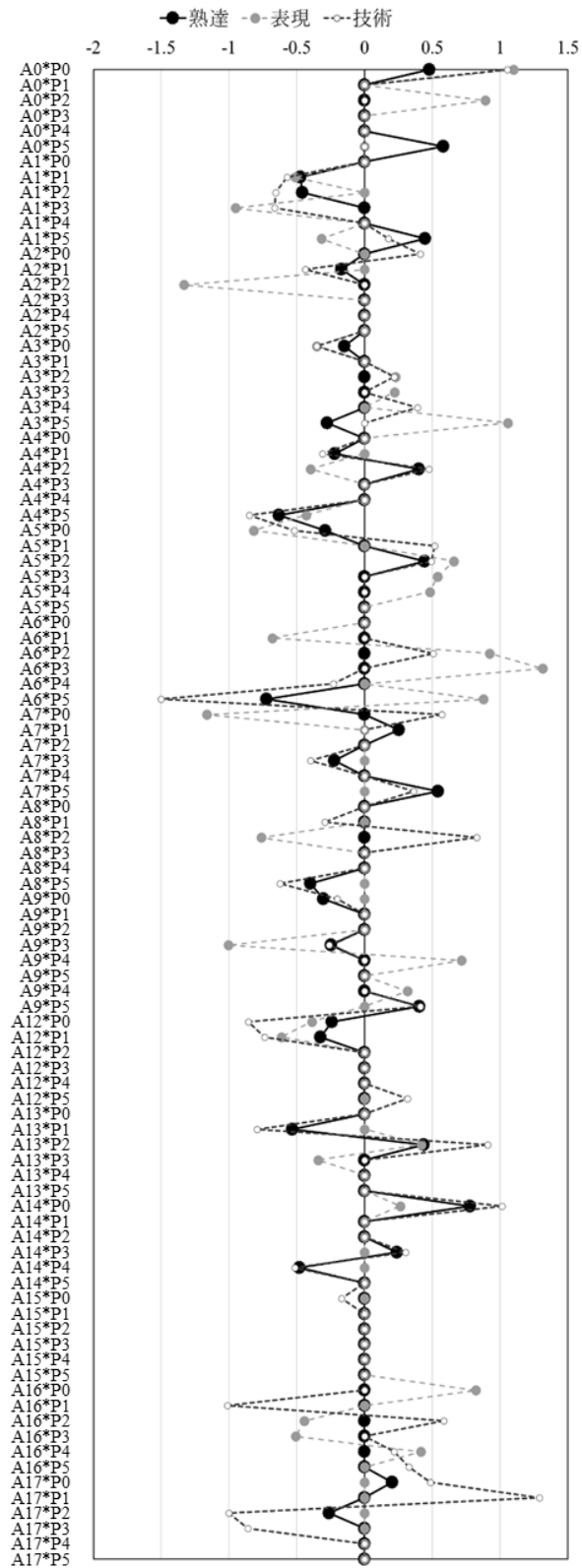
(a) 評価者 E1



(b) 評価者 E2



(d) 評価者 E4



(e) 評価者 E5

図 4.12 各評価者の各評価に対するパラメータの相対重み

表 4.5 各評価者の各評価スコアを推定した際の相対重みの相関

(a) 評価者 E1

	熟達	表現	技術
熟達	1.00		
表現	0.03	1.00	
技術	0.15	-0.02	1.00

(b) 評価者 E2

	熟達	表現	技術
熟達	1.00		
表現	0.37	1.00	
技術	0.24	0.18	1.00

(c) 評価者 E3

	熟達	表現	技術
熟達	1.00		
表現	-0.06	1.00	
技術	0.61	-0.06	1.00

(d) 評価者 E4

	熟達	表現	技術
熟達	1.00		
表現	0.17	1.00	
技術	0.12	0.16	1.00

(e) 評価者 E5

	熟達	表現	技術
熟達	1.00		
表現	0.14	1.00	
技術	0.71	0.14	1.00

4.5.4.4 ヴァイオリン演奏に対する熟達・技術・表現の評価と音響パラメータの関係のまとめ

ヴァイオリン演奏の熟達度推定で用いたパラメータを説明変数、ヴァイオリン演奏の熟達者による熟達・技術・表現という指標に関する評価を目的変数とした各指標の推定を行った。そして推定時の各パラメータの相対重みから各推定における有意なパラメータを調査した結果、各指標や各評価者間で有意なパラメータが異なることが確認された。また、演奏全体への評価は近い傾向であっても、各評価者に内在する評価基準は異なる傾向を持つ可能性が示唆された。

4.6 種々の楽器に対する評価への利用可能性

熟達度推定に関するまとめで述べたように、本論文で提案するパラメータ群、またそれぞれの単一パラメータについては、その他の楽器の演奏の特徴抽出に有効であると考えられる。

F0 を伴うパラメータに関しては、ピアノやオルガンなどの鍵盤楽器と異なり、クラリネットなどの吹奏楽器などのように F0 が演奏者自身によって調整が可能な楽器の演奏については、F0 の取得精度も大きくパラメータの計算に影響をする。しかし、今回対象としたヴァイオリン演奏については、YIN という F0 推定技術を用いて、F0 を十分な精度で取得できたものの、ヴァイオリン以外については YIN による十分な精度の F0 の取得可能性については検討できていないため、F0 の取得法から検討が必要である。

また、FM ヴィブラートを行なった際の知覚上のピッチをある一点に定める場合については、FM ヴィブラートの高い頂点を基準とする場合、低い頂点、また、F0 揺らぎの中心など様々なパターンが考えられる。本論文では物理的な計測によって FM ヴィブラートを扱ったが、心理的な現象としての FM ヴィブラートに対する知覚に着目していないので、ヴィブラートを伴った場合の各音のピッチを考慮したパラメータについても今後検討が必要であると考えられる。FM ヴィブラートの知覚上のピッチについては、演奏する/される楽器の違い、音楽の教育経歴の違いなどによっても異なる可能性も考えられるため、今後更なる検討が必要であると考えられる。

ヴァイオリン演奏の熟達度推定に関する部分で述べたように、純正律と平均律を基準と

した場合の推定精度の差から、評価者が評価する際の基準が説明できること、また熟達、表現、技術の推定の部分では指標（熟達・表現・技術）や評価者によって有意なパラメータが異なるということを述べたが、これらのように、音響パラメータから人間の知覚認知メカニズムが説明できる可能性も示唆される。

また、本論文においては、ヴァイオリン演奏の1オクターブ上下行長音階を対象としており、和音のように同時に複数の音高の音を鳴らす演奏を対象としていない。クラリネットやフルートなどのような楽器の性質上同時に複数の音高の音を鳴らすことを想定していない楽器の演奏音について現状のF0ずれに関する特徴量などの利用可能性はあると考えられる。しかし、同時に異なる音高を鳴らすピアノやヴァイオリンなどについては、和音の演奏を想定した場合において、各音を鳴らすタイミングについてはすべての音を同時に鳴らす、またいずれかの音を早くもしくは遅く鳴らすといった技術、またそれぞれの音の大きさについてはすべての音を同じ大きさを演奏するいずれかの音だけ大きくもしくは小さくするなどといった点についても演奏において重要なスキルであると考えられる。ピアノについては、これはMIDIピアノなどを用いることで、MIDIで取得可能なパラメータとして打鍵タイミングや打鍵速度についての調査は可能であると考えられるが、実際に聞こえる音に関してもパラメータの取得法の検討が必要である。さらに、ヴァイオリンのような音を同時に鳴らすだけではなく音高を演奏者が調整する必要がある楽器の演奏においても、同時に演奏した際のF0についてもパラメータを取得する必要があると考えられるが、そのためには複数音から複数のF0を取得する手法について検討を行なう必要がある。

4.7 結言

本章では、ヴァイオリン演奏に着目し、その演奏の特徴を捉えるパラメータの提案し、熟達度推定の結果からパラメータの有効性が示唆されることを述べた。次に、このパラメータを用いて熟達・技術・表現の各指標を対象とした推定結果から、各指標や各評価者によって推定に重要なパラメータが異なり、重要視する演奏の特徴が異なる可能性についても示唆された。最後に、提案パラメータが種々の楽器に対する演奏評価への利用可能について述べた。

5 章 J-POP 楽曲の歌声を対象とした嗜好性

5.1 緒言

本章では、まず JPOP 楽曲の嗜好性における評価要因について調査結果（桶本ら，2021）について述べる。続いて、歌声の声質に着目した JPOP の嗜好性に関する研究（桶本ら，2022）に関して調査結果を述べる。最後に、調査結果の楽曲推薦への利用可能性について述べる。

5.2 日本語歌唱を伴う楽曲に対する嗜好性の評価要因（桶本ら，2021）

5.2.1 目的

人間が音楽を聴取した際に、「この曲は好みだ」、「この曲は好みではない」といった主観評価を行なうことが多い。またその評価は「歌詞が好きだから」、「このメロディが好きだから」などのように様々な判断基準の下で行なっている（高校生新聞 ONLINE, 2021 参照）。このように音楽を評価する際には、音楽を構成する要素に分割し、要素に対する評価に基づいて総合的に行なっている側面があると考えられる。楽曲の構成要素の例としては歌詞があり、例えば楽曲配信サービスでは、「さくらに関する歌詞の楽曲」のように検索キーワードに基づいた推薦システムはあるものの、歌手と歌詞に対するユーザの好み、例えば「特定の歌手によるさくらに関する歌詞の楽曲」のような複数の要素を総合的に判断した楽曲推薦システムは存在しない。

そこで、楽曲の嗜好度における判断基準となり得る複数の要素を用いた楽曲推薦のための基礎的検討として、楽曲の嗜好度を構成する要素を 4 つ選出し、その 4 つの要素に対する評価と楽曲の嗜好度の評価の相関関係について調査する。

5.2.2 調査する楽曲の 4 要素

楽曲を構成する要素として、

- ・歌手の好み（歌のうまさ，声質の好み，等を含む）
- ・メロディの好み（テンポ，調，コード進行，等を含む）

- ・伴奏音（メロディ以外のパート）の好み（ハモリ，ベースなどのパート，等を含む）
- ・歌詞の好み

を選定した。これらは，好きな楽曲に関する調査（高校生新聞 ONLINE，2021 参照）等を参考にして，著者が選出したものである。

5.2.3 調査の方針

4 つの基準による楽曲の嗜好度の評価メカニズムを解明するために，複数の基準による意思決定メカニズムの解明においてマーケティングなどでよく用いられる階層化意思決定法（Analytic Hierarchy Process，以下 AHP）（Saaty, T. L., 2008）を用いる。

5.2.3.1 AHP とは

AHP とは，ある目的の評価を行なう際に，代替案に対する評価を，その評価に内在する各要素がどの程度重要視するのかを明らかにする手法である。例えば，マイホームの購入における業者の選択という目的に対し，その目的に内在する要素は，値段，部屋の数，アクセス，また，代替案が A 社，B 社，C 社があると想定する。要素（値段，部屋の数，アクセス）に関する重要度の一対比較，また，各代替案（A 社，B 社，C 社）について各要素の評価を行なわせ，AHP 分析を行なうことで，マイホームの購入における評価の階層化構造，総合評価を算出することを可能とする手法である。

本報告では，この AHP を用いて，目的を音楽の嗜好度，要素を 2 で説明した 4 要素，代替案として後述の 2.3.2 で示す 6 つの日本語歌唱を伴うポピュラ楽曲として調査を行ない，AHP 分析によって得られた代替案に対する総合評価値と主観評価による楽曲の嗜好度を比較し，列挙した 4 要素が嗜好性を構成する要因としての有効性について検証する。ここで，列挙した 4 要素による AHP 分析によって得られた総合評価値と主観評価値の相関が高ければ，4 要素の楽曲の嗜好性判定の判断基準としての有効性が示唆される。

5.2.3.2 代替案として用いた楽曲

AHP 分析において代替案として使用する楽曲は，Billboard Japan 2021 年上半期チャート（Billboard Japan，2021 参照）から計 6 曲を選択した。選択に際しては，単一楽曲における歌唱者が複数ではない，歌唱者の重複をしないなどの点を考慮した。表 5.1 に選択した楽

曲を示す.

表 5.1 代替案として用いた楽曲

楽曲 ID	楽曲名	アーティスト
M1	ドライフラワー	優里
M2	炎	LiSA
M3	夜に駆ける	YOASOBI
M4	うっせえわ	Ado
M5	廻廻奇譚	Eve
M6	虹	菅田将暉

5.2.4 実験の概要

AHP 分析のためのアンケート形式の評価実験を行なった. 実験は回答用紙を聴取者に電子的に配布し, 回答後に実験実施者に返送していただく形式で実施した.

聴取者は著者を含む本学学生 6 名 (L_{5a1} から L_{5a6}) であり, そのうち L_{5a2} については日本語を母語としないが日本語歌詞が理解できる方である. 各楽曲について評価を行なう際は, 既知の楽曲については聴取を求めなかったが, 未知の楽曲については, 回答用紙に記載された YouTube リンクを参考に, ワンフレーズ以上聴取するように指示を行なった. また, 分析のために普段の音楽聴取に関する設問も回答していただいた.

5.2.4.1 楽曲の嗜好度に対する主観評価

楽曲嗜好度については, 「絶対的に好みではない」, 「好みではない」, 「どちらかといえば好みではない」, 「普通」, 「どちらかといえば好み」, 「好みである」, 「絶対的に好みである」の 7 段階で評価を行なっていただいた. この嗜好度を各楽曲について評価していただいた.

5.2.4.2 重要度の一対評価

歌手の好み, メロディの好み, 伴奏音の好み, 歌詞の好みの 4 要素を両軸にとり, 「左がかなり重要」, 「左が重要」, 「どちらかといえば左が重要」, 「左右が同じくらい」, 「どちらかといえば右が重要」, 「右が重要」, 「右がかなり重要」という 7 段階で評価を行なっていただいた.

5.2.4.3 4つの要素に基づいた各楽曲の評価

表 5.1 の 6 つの楽曲を両軸にとり、「左がかなり良い」、「左が良い」、「どちらかといえば左が良い」、「左右が同じくらい」、「どちらかといえば右が良い」、「右が良い」、「右がかなり良い」という 7 段階で、5.2.4.2 の 4 項目について評価を行なっていただいた。

5.2.5 実験結果

5.2.5.1 嗜好度の主観評価の相関

聴取者間相関を算出するために、5.2.4.1 で示した 7 段階の名義尺度を、1 から 7 の数値に置き換えて相関係数を算出した。表 5.2 に、主観評価に対する聴取者間の相関を示す。また、表 5.2 内の $SAve$ を構成する各楽曲に対する全体の主観評価値の平均 $save_m$ については、式(5.1)で算出される。

$$save_m = \sum_{l=1}^6 S_{m,l} \quad (5.1)$$

この時、 m は楽曲 ID、 l は聴取者 ID、 $S_{m,l}$ は楽曲 m に対する聴取者 l の主観評価値である。また $SAve = \{save_m\}(m=\{1,2,\dots,6\})$ を 6 名の平均的な評価を表す量として用いる。この $SAve$ と L_{5a1} から L_{5a6} の相関を求めることで、聴取者全体に対する各評価者の主観評価における特殊性を確認できる。ここでは楽曲毎の主観評価値に対する相関係数を求めているため、サンプル数は 6 であり、0.81 以上の場合に 5% 有意水準で有意となる。表 5.2 より、0.81 以上の相関を確認することはできなかった。L_{5a1} と L_{5a5} の主観評価値の相関係数 0.78 が最も高いことが確認できる。

表 5.2 主観評価に対する聴取者毎の相関

	L1	L2	L3	L4	L5	L6	Save
L1	1.00						
L2	0.16	1.00					
L3	-0.44	-0.55	1.00				
L4	0.27	-0.54	0.70	1.00			
L5	0.78	-0.36	0.09	0.62	1.00		
L6	-0.45	0.33	0.05	-0.21	-0.55	1.00	
Save	0.50	-0.12	0.40	0.80	0.68	0.11	1.00

5.2.5.2 重要度に関する結果

AHP では、評価基準毎に一対比較表を作成し、その一対比較表の行方向の幾何平均を求め、幾何平均の合計を、各評価基準の幾何平均で割ることで、重要度を算出できる。例えば、歌手の好みとメロディの好みの比較の場合、歌手の好みが非常に重要と判断された場合は、歌手の好みのメロディの好みに対するスコアが 4 点、メロディの好みの歌手の好みに対するスコアが 1/4 点となるような一対比較表を作成する。また一対比較表に対し、行方向に幾何平均を求め、4 要素に対する幾何平均の合計で除算することで、各要素の重要度を算出する。

5.2.5.2.1 重要度の比に関する結果

図 5.1 に、AHP 分析によって得られた聴取者毎の 4 つの要素に対する重要度の比を示す。図 5.1 は、横軸が聴取者 ID、縦軸が重要度の比を表しており、各棒については下から歌手、メロディ、伴奏、歌詞を表している。例えば、図 5.1 より、聴取者 L_{5a1} の歌手の重要度は、0.30 であることを表している。

図 5.1 より、L_{5a1}, L_{5a2}, L_{5a3}, L_{5a5} についてはメロディの好み、L_{5a4}, L_{5a6} については伴奏の好みの重要度が最も高いことが確認できる。

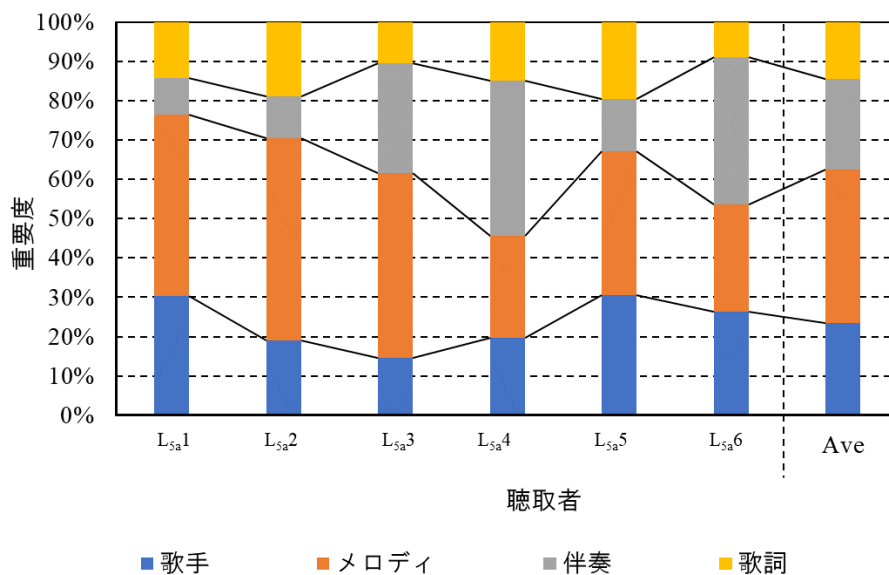


図 5.1 聴取者毎の 4 つの要素に対する重要度の比

5.2.5.2.2 重要度の比の比較

表 5.3 に AHP 分析によって得られた聴取者毎の 4 つの要素に対する重要度の比に対する被験者間の相関を示す。また、表 5.3 の $FAve$ を構成する、各楽曲に対する全体の主観評価値の平均 $fave_f$ については、式(5.2)で算出される。

$$fave_f = \sum_{l=1}^6 I_{l,f} \quad (5.2)$$

この時、 f は 4 つの要素 ($f=\{1,2,3,4\}$) を表しており、 $I_{l,f}$ は聴取者 l における要素 f の重要度を表している。また $FAve = \{fave_f\}(f=\{1,2,3,4\})$ を 6 名の平均的な評価を表す量として用いる。この $FAve$ と L1-6 の相関を求めることで、聴取者全体に対する各評価者による重要度の比における特殊性がわかる。ここでは、4 つの要素に対する相関係数を求めており、サンプル数は 4 なので、0.95 以上で 5%有意水準での有意となり、有意と認められた場合は*で示す。

表 5.3 より、 $L_{5a} 1$ と $L_{5a} 5$ については、5%有意水準で有意であることが確認できる。また、次点で重要度の相関が高いのは、 $L_{5a} 1$ と $L_{5a} 2$ の 0.90 であることが確認できる。

表 5.3 重要度の被験者間における相関

	$L_{5a} 1$	$L_{5a} 2$	L_{35a}	$L_{5a} 4$	$L_{5a} 5$	$L_{5a} 6$	$FAve$
$L_{5a} 1$	1.00						
$L_{5a} 2$	0.90	1.00					
$L_{5a} 3$	0.63	0.77	1.00				
$L_{5a} 4$	-0.24	-0.16	0.50	1.00			
$L_{5a} 5$	0.98*	0.83	0.46	-0.40	1.00		
$L_{5a} 6$	0.05	-0.05	0.52	0.89	-0.06	1.00	
$FAve$	0.85	0.85	0.93	0.30	0.73	0.47	1.00

5.2.5.3 AHP 分析によって得られた総合評価値

AHP 分析における総合評価値は、5.2.5.2 の重要度と同様に、要素毎の各楽曲に対する嗜好性評価より算出できる。まず、5.2.5.2 と同様に一対比較表を各代替案の各評価項目に対する重要度に対して作成し、同様の手順で幾何平均と幾何平均に対する割合を算出する。これが、要素毎の各楽曲の重要度である。そして各楽曲に対する総合評価値 $AHPscore$ は、式(5.3)より算出される。

$$AHPscore_{l,m} = \sum_{f=1}^4 Iex_{l,m,f} * I_{l,f} \quad (5.3)$$

この時、 $Iex_{l,m,f}$ は聴取者 l が楽曲 m の要素 f に関して評価した重要度、 $I_{l,f}$ は聴取者 l における要素 f の重要度を表している。

5.2.5.3.1 各聴取者における楽曲毎の総合評価値

図 5.2 に各楽曲に対する各聴取者の総合評価値を示す。図 5.2 は、横軸が聴取者 ID、縦軸が AHP 分析により得られた総合評価値、各棒が楽曲 ID を示している。また、Ave については、全聴取者の総合評価値を楽曲毎に平均した値が示されている。例えば、図 5.2 より、楽曲 M1 の聴取者 L1 における総合評価値は、0.10 であることを表している。

図 5.2 より、聴取者によって、最も評価の高い楽曲が異なることが確認できる。楽曲毎の平均より、M3 の評価が高い傾向にあることが確認できる。

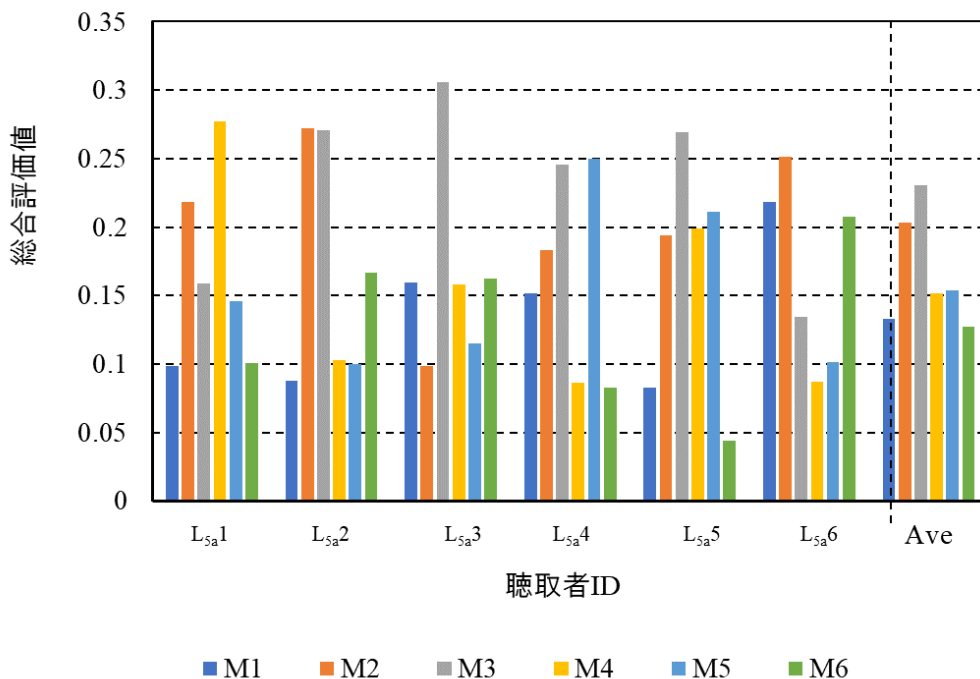


図 5.2 各聴取者における楽曲毎の総合評価値

5.2.5.4 嗜好度の主観評価値と総合評価値の相関

図 5.3 に聴取者毎の主観評価値と総合評価値の相関係数を示す。これより、聴取者が評

評価基準を意識せずに評価した各楽曲の主観評価と4つの評価基準を対象とした各楽曲の評価を比較することができ、今回選定した4つの評価基準の妥当性を確認する。図5.3は、横軸が聴取者ID、縦軸が主観評価値と総合評価値の相関係数を示す。また、Aveについては、全聴取者の主観評価値と総合評価値を楽曲毎に平均した値の相関係数が示されている。ここでは、楽曲毎の主観評価値と総合評価値の相関係数を求めているため、サンプル数は6であり、0.81以上で5%有意水準での有意となり、有意な棒には斜線を付与した。

図5.3より、L_{5a1}、L_{5a4}、L_{5a5}とL_{5a6}については、5%有意水準で有意であることが確認できる。また、L_{5a3}が最も相関が低いことが確認できる。

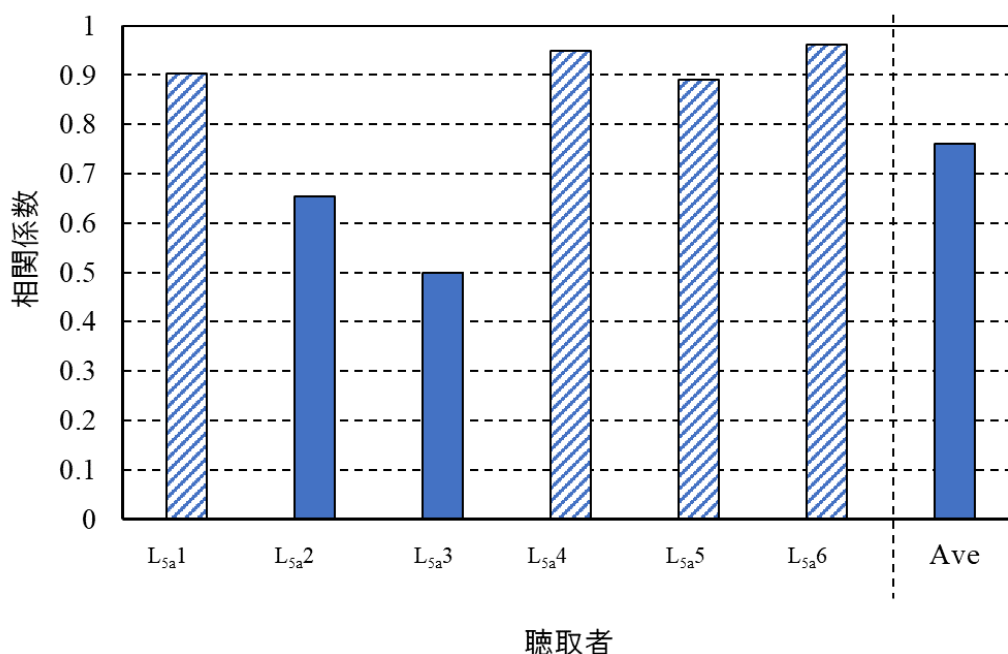


図 5.3 聴取者毎の主観評価値と総合評価値の相関

5.2.6 考察

図5.3より、楽曲毎の主観評価値と総合評価値の相関係数が有意なL_{5a1}、L_{5a4}、L_{5a5}、L_{5a6}について検討する。L_{5a1}、L_{5a4}、L_{5a5}とL_{5a6}の主観評価値と重要度がそもそも似通っていた群であった可能性について検討を行なう。表5.2のL_{5a1}、L_{5a4}、L_{5a5}、L_{5a6}とSAveが交差するセルより、L_{5a6}の主観評価値と全体の主観評価値の相関(0.11)が特に低い傾向に

あることが確認できる。さらに、表 5.3 の L_{5a1} , L_{5a4} , L_{5a5} , L_{5a6} と $FAve$ が交差するセルより、重要度についても全体の平均に対して相関 L_{5a4} と L_{5a6} (0.30 と 0.47) が低い傾向にあることが確認できる。一方、前述の通り、図 5.3 より L_{5a1} , L_{5a4} , L_{5a5} , L_{5a6} における主観評価値と総合評価値の相関係数は 5%有意水準で有意である。これより、評価基準や主観評価値は全体と異なるものの、個人内で明確な基準がある聴取者の群が存在することを確認できる。

また、図 5.3 より、相関が低い L_{5a2} と L_{5a3} (0.65 と 0.50) については、各重要度が AHP 分析の総合評価値で計算される加算結合でない可能性があり、例えば積算結合によって総合評価値が得られる可能性も考えられる。言い換えると「歌詞は好きだけど歌手が好みではない」といった場合に、加算結合では歌詞の好みに対するスコアと歌手のスコアを単純に加算した場合もあれば、歌手が好みでない場合には歌詞が好きであっても全体評価が下がるという評価もあり得るということである。このような算出基準の違いを調査するには、例えばファジィ AHP (杉山ら, 1993) などを用いて積算結合での総合評価値の算出可能性についての検討を行なう必要性が示唆される。それだけでなく、主観評価では、「その楽曲に対する思い入れ」のような個人依存性が高く、本報告では未検討の要素を考慮していた可能性も示唆される。

さらに、図 5.1 において、歌手の重要度が比較的高い L_{5a1} , L_{5a5} , L_{5a6} については、実験内の歌声合成ソフトウェアを使用した楽曲、またアニメやゲームの楽曲をよく聴取すると回答している。これらの楽曲は、同じ楽曲を異なる歌手が歌うという状況が、他のジャンルと比較し多い傾向にある。これより、普段聴取する音楽ジャンルの文化が、音楽を聴取する際の各要素の重要度と関係する可能性が示唆される。これに関連して、音楽大学で学ぶ学生に聴取者として参加いただいたものの、いずれも演奏を主専攻とする方ではなかったため、音楽に演奏者として関わる方と、聴取者として関わる方によっても、重要視する項目、その傾向が異なる可能性も考えられるため、今後は、普段の音楽との関わり方を考慮した実験を行なう必要性があると考えられる。

5.2.7 嗜好度の評価要因に関するまとめ

日本語歌唱を伴う楽曲に対する嗜好性の評価要因に関する検討を行なった。これより、被験者毎に評価基準が異なる可能性が示唆された。なお、4つの要素を考慮した AHP によ

る楽曲の嗜好性に対する評価基準の有意な有効性が聴取者全員に対しては確認できなかった。その理由として、各重要度が AHP 分析の総合評価値で計算される加算結合でない可能性、さらには要素が十分でない可能性が考えられ、今後はファジィ AHP などを用いて積算結合での総合評価値の算出について検討を行なう。また、今回の 4 要素については、メロディを重要視する群と伴奏の要素を重要視する群の存在が確認でき、メロディを重要視する傾向の方が強いことが確認された。さらには、普段聴取する音楽ジャンルと重要視する要素との関係性があり得ることが示唆された。

5.3 歌声分離について

歌声の声質を考慮した音響パラメータを設計するにあたり、音響パラメータの算出対象とする音響信号は歌声のみの音響信号が望ましい。そこで、まず、伴奏音を伴う楽曲から歌声を抽出する必要がある。

歌声抽出に関する技術を含む音源分離技術については、2007 年から開催されてきた SiSEC (Signal Separation Evaluation Campaign) などで盛んに報告されている。近年では、音響信号そのものを用いた機械学習が可能となり、その結果精度の高い歌声抽出が行なわれるようになった。例えば、代表的なものとしては Tensorflow を利用した Spleeter (Prétet, et al., 2019) がある。

本論文では、ソースが公開されており利用可能な歌声抽出として、GitHub にて公開されている vocal-remover v4.0.0 (Turumeso, 2020) (以下, vocal-remover) を用いる。vocal-remover を採用した理由としては、歌声抽出技術の精度が他のものと比較し優位だと、著者が判断したためである。vocal-remover によって抽出された歌声に対し、音響パラメータの算出を行なう。

本章において歌声とは、この歌声抽出によって抜き出された音源を歌声として扱う。歌声を扱うにあたっては、歌声と話し声の定義をそれぞれ述べ、その差などについての調査も必要ではあるものの、この差には音楽と音の差を明らかにするのと同じような性質の複雑な知覚認知メカニズムを明らかにする必要がある。音楽の定義において、定性的側面については起源から述べようとする場合など様々な議論があるものの統一的な見解にまとまるまでには至っていない。定量的な側面においては、定性的にも定まっていないこともあ

り、音響パラメータに落とし込めるまでには定まっていない。そこで、本章では、vocal-removerによって抽出された音源を歌声として扱う。

5.4 J-POP 楽曲の歌声を対象とした嗜好性評価のための音響パラメータ

本論文では、歌声の声質のみに着目した音響パラメータ、つまり、歌い方といった歌声における声質以外の特徴が影響しない音響パラメータの設計を行なう。歌声に着目した理由としては、5.2の結果のうち、図5.2を確認すると、歌手の重要度がメロディの次に高い聴取者もいるため、歌手の特徴のひとつである歌声の声質に特化したパラメータの設計意義があると感じたためである。また、歌声の声質においては、倍音構造が重要であると考え、歌声の音源を対象として歌声の声質に着目した音響パラメータとして、倍音構造などに関する18種類の音響パラメータを設計した。この18種類の音響パラメータは、周波数とメル周波数の側面から算出された2つの倍音構造パラメータを算出し、それらの傾向を表す9つの傾向パラメータを組み合わせることによって算出を行なう。また、提案パラメータの他に2つの基礎的な信号分析に用いられるパラメータ（以下、基礎分析パラメータと呼称）を組み合わせ、計20個の音響パラメータ群を歌声の声質の特徴として提案する。表5.4に提案パラメータ群の一覧を示す。

表 5.4 提案パラメータ

		F0を考慮した 倍音構造パラ メータ	メル尺度を考慮 した倍音構造パ ラメータ	ゼロクロス	スペクトルス キューネス
		H ₁	H ₂	B ₁	B ₂
最大値	T ₁	○	○	○	○
セントロイド	T ₂	○	○		
周波数軸方向の 差分の平均	T ₃	○	○		
周波数軸方向の 差分の標準偏差	T ₄	○	○		
時間数軸方向の 差分の平均	T ₅	○	○		
時間軸方向の 差分の標準偏差	T ₆	○	○		
2次の曲線近似 におけるx ² の係数	T ₇	○	○		
2次の曲線近似 におけるxの係数	T ₈	○	○		
2次の近似曲線 との差分の平均	T ₉	○	○		

5.4.1 倍音構造パラメータ (H)

本論文では、倍音構造パラメータとして、 F_0 を考慮したパラメータとメル尺度（古井，1992；日本音響学会，2011；など）を考慮したパラメータの2種類を提案する．ここで倍音構造を考慮したパラメータを提案する理由について述べる．歌声における声の個人性は、イントネーションなどで評価する側面もあるとは思うが、主にスペクトル構造が重要であると考えたためである．例えば、同じ音高の歌唱であってもよりキンキンとした歌声が好きといった嗜好性もあると考えられる．これらには歌声のスペクトル構造が関わっているため、声のスペクトル構造に関する倍音構造を考慮したパラメータを提案する．また、この倍音構造の時間的変化にも着目するために統計パラメータと組み合わせて用いる．図 5.4 に倍音構造パラメータ算出の概要を示す．

5.4.1.1 F_0 を考慮した倍音構造パラメータ (H_1)

このパラメータは、 F_0 から 3 オクターブ分の倍音構造を、分析フレーム毎に算出するパラメータであり、音高に影響されない歌声の倍音を取得することができる．算出方法は、歌声抽出音源の無音区間を除き、続いて、YIN (Cheveigné, et al., 2002) に基づいて作成された F_0 算出プログラムを用いて、 F_0 カーブを取得する．次に、取得した F_0 から 3 オクターブのパワースペクトルを取得するバンドパスフィルタを通過させることで F_0 を考慮した倍音構造パラメータを算出している．この時、サンプリング周波数 44100Hz で記録された信号に対して、4096 点で FFT を行ない、無音区間を除いた歌声音源の長さに対応した F_0 から 3 オクターブ分の倍音を格納したベクトルが出力される．

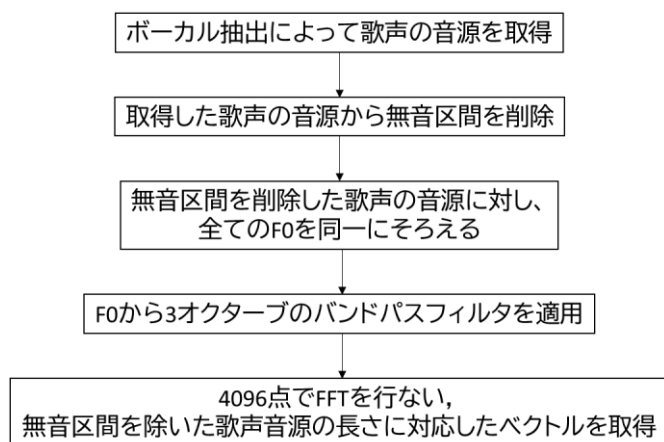


図 5.4 F_0 を考慮した倍音構造パラメータ算出の概要

5.4.1.2 メル尺度を考慮した倍音構造パラメータ (H_2)

5.4.1.1 で述べたパラメータと異なり、こちらにはバンドパスフィルタを通過させず、また F_0 を考慮せず、バンドパスフィルタ外のスペクトル成分が取得可能であり、具体的には、 F_0 以下の周波数帯域、また F_0 から 3 オクターブより上の周波数帯域の成分も考慮される。この時、サンプリング周波数の 3% の窓長でメルスペクトログラムを算出し、無音区間を除いた歌声音源の長さに対応したベクトルが出力される。また、人間の聴覚構造を考慮したメル尺度を適用したパラメータとすることで、高周波数の成分の抑制などが行なわれている。

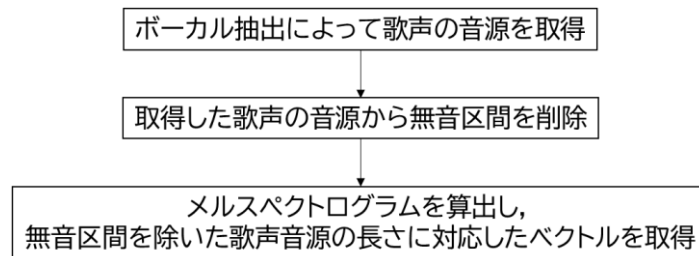


図 5.5 メル尺度を考慮した倍音構造パラメータ算出の概要

5.4.2 傾向パラメータ (T)

5.4.1 で取得した倍音構造パラメータに対し、4 章のヴァイオリン演奏の評価に使用したパラメータなどを参考にして、9 つの傾向パラメータを算出し、歌声の声質に着目した音響パラメータを算出する。これらの傾向パラメータを使用することで歌声の多能の特徴ではなく、特徴の傾向を取得できる。

5.4.2.1 最大値 (T_1)

倍音構造パラメータの最大値を算出する。各倍音構造パラメータが有するベクトルの最大値を取得することで、倍音構造を表すベクトルの局所的な強度を確認することができる。 F_0 にあたる部分のパワーが最も大きくなるが多いため、 F_0 のパワー強度を確認していることにあたる場合が多い。

5.4.2.2 セントロイド (T_2)

倍音構造パラメータの重心を算出する。各倍音構造パラメータが有するベクトルのセントロイドを取得することで、倍音構造を表すベクトルの周波数上でのパワーバランスの一端を確認することができる。

5.4.2.3 周波数軸方向の差分の平均 (T_3)

倍音構造パラメータの周波数軸方向の差分を取得し、その平均値を算出する。周波数分析区間内で隣り合う周波数ビンのパワーの差分を取ることで、周波数軸上でのパワーの変化量の大きさを確認することができる。

5.4.2.4 周波数軸方向の差分の標準偏差 (T_4)

倍音構造パラメータの周波数軸方向の差分を取得し、その標準偏差（東京大学教養学部統計学教室，1991；小島，2006；など）を算出する。周波数分析区間内で隣り合う周波数ビンのパワーの差分を取ることで、周波数軸上でのパワーの変化量のばらつきを確認することができる。

5.4.2.5 時間数軸方向の差分の平均 (T_5)

倍音構造パラメータの時間軸方向の差分を取得し、その平均値を算出する。隣り合う周波数分析区間の差分を算出することで、各周波数ビンの時間的なパワーの変動の大きさを確認することができる。

5.4.2.6 時間軸方向の差分の標準偏差 (T_6)

倍音構造パラメータの周波数軸方向の差分を取得し、その標準偏差を算出する。隣り合う周波数分析区間の差分を算出することで、各周波数ビンの時間的なパワーの変動のばらつきを確認することができる。

5.4.2.7 2 次の曲線近似における x^2 の係数 (T_7)

倍音構造パラメータに対し 2 次の曲線近似を行ない、 x^2 の係数を取得する。各分析区間での周波数分布について 2 次の曲線近似を行なうことで、周波数分析のパワー分布の一

端を確認することができる。

5.4.2.8 2 次の曲線近似における x の係数 (T_8)

倍音構造パラメータに対し 2 次の曲線近似を行ない、 x の係数を取得する。各分析区間での周波数分布について 2 次の曲線近似を行なうことで、周波数分析のパワー分布の一端を確認することができる。

5.4.2.9 2 次の近似曲線との差分の平均 (T_9)

倍音構造パラメータに対し 2 次の曲線近似を行ない、得られた近似式より、近似曲線を算出する。算出された近似曲線と倍音構造パラメータとの差分の平均値を算出する。近似曲線と各倍音構造パラメータのベクトルとの差分を算出することで、対象となった歌声の倍音における 2 次曲線に近い倍音構造であるかどうかを確認することができる。

5.4.3 基礎分析パラメータ (B)

Zerocrossings (以下、ゼロクロス) (山田ら, 2014) と Spectral Skewness (以下、スペクトルスキューネス) (Peeters, 2004) を算出する。音楽音響信号を対象とした楽曲の年代を推定した研究では、RMS (Root Means Square) などの音響パラメータが用いられている (岡田ら, 2018) もの、声質のみに着目したパラメータにおいては不適切と判断した。

5.4.3.1 ゼロクロス (B_1)

音響信号の振幅値が 0 と交差する回数を算出する。

5.4.3.2 スペクトルスキューネス (B_2)

音響信号のスペクトル歪度を算出する。

5.5 嗜好性の評価実験

5.4 で述べた音響パラメータの妥当性を評価するためには、歌声の好みに対する印象評価結果が必要である。この印象評価結果を取得するために、印象評価実験を行なった。

5.5.1 概要

印象評価実験では、5 秒の伴奏音を伴う楽曲を連続聴取させ、その歌声の声質が「好みである/好みでない」を 2 件法で回答いただいた。なお、本実験は、国立音楽大学研究倫理委員会の（主に）人を対象とする研究に関する研究計画等審査で承認されている（受付番号：2108）。

5.5.2 聴取者

聴取者は著者を含む国立音楽大学、もしくは、国立音楽大学大学院に所属する 10 名であった。以後、L_{5b1} から L_{5b10} と表記する。聴取者には、ヘルシンキ宣言に基づく実験参加同意をいただいた。

5.5.3 実験刺激

前述の通り、伴奏音を伴う楽曲を聴取させ、その歌声の声質について主観評価をさせた。この時、聴取させた楽曲のアーティストのイメージが評価に影響を与える可能性があると考え、VOCALOID 等の歌声合成ソフトによるオリジナル楽曲を人間がカバー歌唱している楽曲、計 70 通りを用いた。その 70 通りの楽曲からサビ部分を 5 秒抜粋し、正規化を行ない、冒頭と終端に 0.5 秒のテーパーを付与した実験刺激を作成した。そして、刺激番号を英語で読み上げた音声、歌唱音、読み上げ音声、実験刺激が交互になるような実験刺激を合成した。実験開始前に、テスト音源を聴取し音量調整を行なっていただき、以後は音量調整を行わないように指示した。

5.5.4 評価法

評価は、歌唱の声質が「好みである/好みでない」の 2 件法とした。注意事項として、歌い方などではなくあくまで声質にのみに着目し回答するように、また、回答に迷った際には、直感で判断を行なうように指示をした。さらに、今回の好みに対する判断において何らかの基準があった場合はその基準を訪ねる自由記述欄を回答用紙の最下部に設定した。

5.5.4.1 評価の妥当性

5.5.4.1.1 2段階評価の妥当性

関連研究における評価尺度について述べる．絶対音感に関する調査を行なった研究 (Hedger, S. C., et. al., 2013) では、ピッチに対する評価を3段階で行なっている．また、歌唱力の評価に関する研究 (中野, 2007) では、従来の英語発音の評価などでは n 段階評価 ($n=5,7$ など) が採用されているが聴取者の音楽経験などの違いによって評価基準が異なる可能性や基準が同一でもその距離感が聴取者によって異なる可能性を述べている．

上記も考慮し、本論文で歌声の嗜好性を2段階で評価した理由を述べる．1つ目の理由として、YouTubeなどの世間で利用されるシステムなどにおいて嗜好に関する評価については、good、もしくは評価しないのように2段階で評価されることが多いというものである．2段階評価は聴取者にとっても普段から使用するような慣れ親しんだスケールの評価であると考えられる．2つ目は楽曲への応用を想定しているという点である．実際の楽曲推薦を考えるとユーザは、ある特定の音楽に対し、聴きたい/聴きたくないといういずれかの判断を下すと想定し、その中間となる、聴きたいと聴きたくないとの間の選択肢は設定することはないのではないかと、また楽曲推薦を考慮した場合における嗜好というのは、聴きたいもしくは聴きたくないという評価の言い換えにあたると思った．これより、歌声の嗜好度を2段階で評価することとした．3つ目は、尺度の度数を増やすと各度数の評価基準が聴取者によって異なるという問題が存在するためである．例えば、10段階にした場合において、1が好みでない、10が好みであるという評価尺度として設定したと仮定すると、1と2の程度差は被験者によって異なる．ヴァイオリン演奏の評価においてはその問題を解消するために、各評価者の10段階評価によるスコアを z スコアに置き換えた．ヴァイオリン演奏においては、線形回帰による評価を想定していたため、 z スコアにする問題というものは存在しない．しかし、今回は、嗜好度をクラスタリングすることを想定しているため、 z スコア化するとクラスタリングにおける正解ラベルの設定が困難となる．これら3点を考慮し、本論文で歌声の嗜好性の評価は2段階とした．

5.5.4.1.2 評価人数の妥当性

本章では、聴取者の人数を10人としている．まず、本実験では、各個人の嗜好性をまとめた回答を推定するのではなく、各個人の嗜好性そのものを推定する．また、歌声におけ

る声質の嗜好性評価においては、声質に着目したパラメータを使用した場合と使用していない場合の結果を比較することで、パラメータの有効性を検証するものである。以上より、評価者の人数が 10 人であることは妥当であると考えられる。

5.5.4.1.3 評価の多様性

嗜好性といっても、その評価基準などの個人差を有する。例えば、生まれ育った地域、これまでの音楽教育、生まれた時代など様々な要因でその個人差が生ずる。そのため、評価された嗜好性には、全聴取者では統一的な基準を定めることは非常に困難である。またその基準は聴取者自身も説明できない場合もあると考えている。例えば、フォントを例にすると明朝体フォントが好みであるかどうかを好みである/好みでないで回答させた場合、いずれかを選択することはできるかと思うが、その理由を説明することは困難であろう。嗜好性の評価基準は、個人差を有すること、何かしら定めることは困難であるということ踏まえ、本論文では歌声の嗜好性については、全体の評価をまとめた評価ではなく、各被験者の評価を推定の対象とする。聴取者が 3 人いたとして、ある楽曲に 1 人が好みである、2 人が好みでないとした場合に、過半数を超えているのでこの曲の歌声は好みでないというラベルを付与するというものではないということである。

5.5.5 回答方法

本実験は、ファイル共有サービスにて、実験刺激、回答用紙を配布して、回答を行なっていた。回答終了後、回答用紙は再度ファイル共有サービスにアップロードしていただいた。

5.5.6 回答の分布

図 5.5 に 10 人の聴取者の回答の分布を積み上げ折れ線グラフで示す。また、表 5.5 に聴取者の回答の相関係数を示す。なお、相関係数の算出には、回答をそれぞれ、「好みである」を 1、「好みでない」を 0 に置き換えることで算出を行なった。図 5.5 の横軸は刺激 ID を示し、縦軸は各聴取者が好みであると回答された数を示している。例えば、刺激 ID S01 については、L_{5b1}, L_{5b3}, L_{5b4}, L_{5b7}, L_{5b10} の計 5 名が好みであると回答していることを表し、折れ線は被験者番号の若い (L_{5b1}, L_{5b2}, L_{5b3}, …) 回答から積み上げられるので、い

ずれの刺激においても最上段は評価者番号の若い評価者（例えば L_{5b}10）が配置される。

表 5.5 より、聴取者の回答における相関の絶対値の平均は 0.20 であった、図 5.5 と表 5.5 より、相関係数は全体として低い傾向にあるものの、4 つの実験刺激に対してはすべての聴取者が好みであると回答している。これより、歌声の声質の好みについては、一貫性は見いだせないものの、その一部においては何らかの共通認識を有する可能性も示唆される。一方、5.5.4.1.3 で述べたように評価基準をそれぞれの聴取者がどのように定めたのかというの、聴取者自身も説明できない可能性のあるという性質を持つため、相関が高くとも基準は全く異なる可能性も捨てきれない。

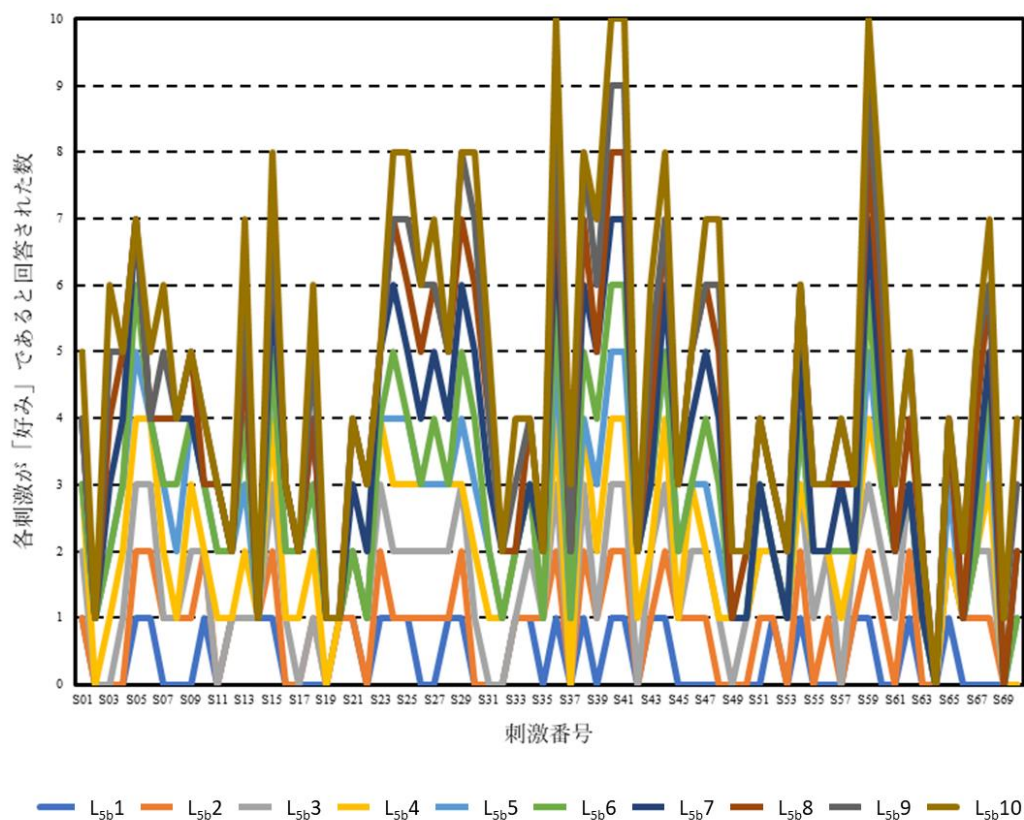


図 5.5 聴取者の回答の分布

表 5.5 回答結果の相関係数

	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
L1	1.00									
L2	0.10	1.00								
L3	0.27	0.22	1.00							
L4	0.07	0.09	0.13	1.00						
L5	-0.01	0.13	-0.20	0.11	1.00					
L6	0.20	0.35	0.19	0.19	0.37	1.00				
L7	0.00	0.25	0.14	0.33	0.30	0.34	1.00			
L8	0.13	0.15	0.77	0.08	-0.07	0.19	0.16	1.00		
L9	0.05	0.16	-0.07	0.20	0.12	0.24	0.19	0.02	1.00	
L10	0.13	0.13	0.08	0.41	0.39	0.39	0.41	0.17	0.36	1.00

5.6 機械学習による提案パラメータの有効性検証

有効性検証として、提案パラメータ群を説明変数、歌声の声質の好みに対する主観評価結果を目的変数とした機械学習を行なう。

5.6.1 検証の概要

本論文では、5.4 で述べたパラメータ群をすべて使った ALL 条件、H と T の組み合わせからなるパラメータ群を用いた HT 条件、B のみを用いた B 条件でそれぞれ機械学習を行ない提案パラメータの有効性を検討する。この時、Closed テスト（クローズドテスト）と、Open テストとして 10 fold-CV（荒木，2018）による検証を行なう。機械学習のアルゴリズムとしてランダムフォレスト（random forest）（Breiman, 2001）にて分類を行ない、F-measure（荒木，2018）によって推定精度の評価を行なう。B 条件と ALL 条件や HT 条件を比較することで、提案パラメータの有効性を検証する。

5.6.2 検証結果

図 5.6 に 3 条件に対するクローズドテストの推定結果、図 5.7 に 3 条件に対する 10 fold-CV の推定結果を示す。図 5.6, 5.7 はそれぞれ横軸が聴取者、縦軸が F-measure であり、各聴取者の回答を 3 条件のパラメータでランダムフォレストを行なった際の F-measure が表されている。また、グラフ中の CL はチャンスレベルを指しており、歌声の声質に対し、「好みである／好みでない」の 2 件法で回答させたので、チャンスレベルは 50%であることを表している。そして、図 5.6, 5.7 のグラフ右側の Avg. については、条件ごとに F-

measure の値を平均したものである。

図 5.6 より、クローズドテストの場合は、いずれの場合においても、F-measure の値は、1.00 であった。図 5.7 より、L_{5b}8 は最も F-measure の値が高い条件 HT の場合においても 0.47 であるものの、その他の聴取者においては、提案パラメータを含む条件で F-measure の値はいずれもチャンスレベルを超えていることが確認できる。また、聴取者ごとに最も F-measure の値が高い条件は、ALL 条件が 3 名、HT 条件が 6 名、ALL と HT 条件が同等であるのが 1 名であった。これより、提案パラメータを用いた方が、F-measure の値が向上する傾向が確認できた。F-measure の値が高い音響パラメータの組み合わせが異なるのは、歌声の声質に対する評価基準の差が影響していると考えられる。さらに、Avg.については、条件 B の場合は F-measure の値は 0.49 であり、チャンスレベルを下回っていたのに対し、最も F-measure の値が高い HT 条件では F-measure の値は 0.58 であり、F-measure の値が大幅に向上していることから提案パラメータの有効性が示唆される。

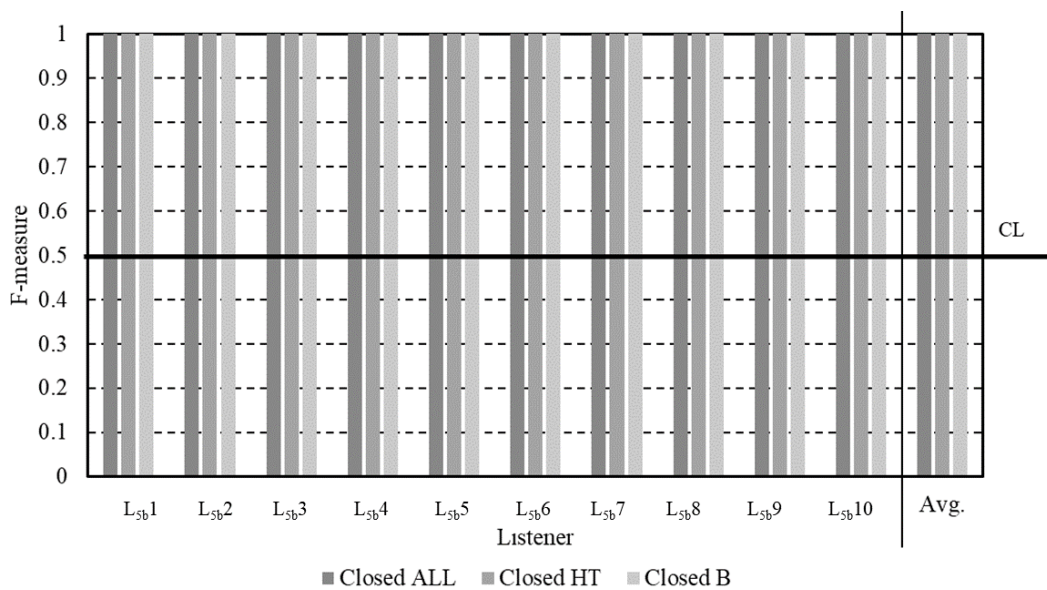


図 5.6 クローズドテストによる推定結果

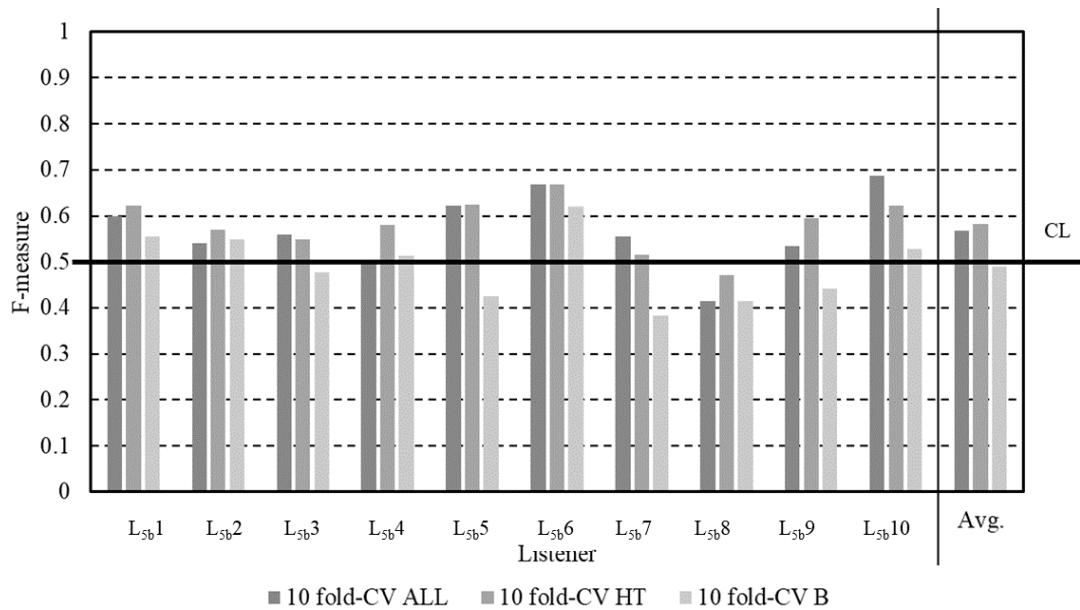


図 5.7 10 fold-CV による推定結果

5.6.3 考察

提案パラメータの有効性は示唆されたものの、聴取者ごとに、最も F-measure の値が高い条件も異なっていたことについて更なる検証を行なうため、各聴取者の回答と 20 個の音響パラメータの相関係数を確認する。このとき、表 5.5 の相関係数を算出した際と同様に、各被験者の回答については「好みである」を 1、「好みでない」を 0 に置き換えた。図 5.7 に各音響パラメータに対する聴取者ごとの回答の相関係数を示す。図 5.7 は、横軸が相関係数、縦軸が音響パラメータ、棒が各音響パラメータとの比較対象となった聴取者 ID を表している。例えば、 H_1T_1 の音響パラメータに対し、聴取者 L5b1 の回答の相関係数は 0.29 であることが表されている。図 5.7 より、 H_1T_1 の音響パラメータを確認しても、L5b3 の相関係数は 0.35 であるのに対し、L5b2 については -0.03 であり、その相関係数の傾向が異なることが確認できる。同様に、 H_1T_3 などの音響パラメータについても、相関係数の傾向が異なる音響パラメータが確認できる。これより、歌声の声質の好みについては、聴取者ごとにパラメータ群を調整することが望ましいことが示唆される。このことは歌声の好みに関する自由記述欄の回答でも示唆された。回答欄に記述のあった 10 名中 4 名の回答は次に示す。L5b1 については「中性的な声が好み」、L5b2 については「無理なく出ている声、作っていない声が好み」、L5b5 については「聴いていて高音が痛く感じないか、ストレートな声か」、L5b6 については「喉ががらがらした音を含んでいる、また高くきんきんしている、

中性的な声は苦手」というような記述があった。これをみても、「高音が痛い、きんきんしている」といった声の倍音成分に関係するような特徴を着目している場合、また「喉ががらがらした音を含んだ声」、「ストレートな声」といった歌声のラフネス (Terhardt, 1974) や Fluctuation Strength (Fastl, 1990) に関係がありそうな特徴に着目するなど、様々な基準で好みの評価を行なっていることが確認できる。更に、自由回答に記述があった L_{5b} 1, L_{5b} 2, L_{5b} 5 と L_{5b} 6 の 4 名の回答と図 5.7 の結果を比較すると、L_{5b} 1, L_{5b} 5, L_{5b} 6 の声の中性感、高音など声の倍音成分に関係するような特徴を着目している聴取者については推定精度が高い傾向にあり、L_{5b} 2 の「無理なく出ている声、作っていない声が好き」といった提案パラメータとの関係が弱いと考えられる要素を重要視している聴取者については推定精度が低い傾向にあることが確認できる。また、L_{5b} 1 と L_{5b} 6 では、中世的な声に関しては異なる嗜好であり、このことから歌声の声質の好みについては、聴取者ごとにパラメータ群を調整することが望ましいことが示唆される。これより、更に音響パラメータを拡充させ、聴取者ごとにパラメータ群を調整することにより、F-measure の値がより高い推定を行なうことができる可能性が示唆される。

本論文では、歌声の声質の嗜好性を対象とした音響パラメータの設計を試みた。表 5.5 に示した通り、好みについて被験者間の回答に対する相関係数が低い傾向にあり、その要因としてその判断基準が多様であるからだと述べた。このような好みと被験者間の回答の傾向に対しては、顔の好みに関する報告 (中山ら, 2012) でも、被験者間の相関係数は 0.27 であったと報告されており、本論文の好みに対する評価の被験者間相関が低いということと同様の傾向を示している。また、布の好みの個人差について分析を行なった研究 (市原, 1996) では、布に対し 20 の印象評定項目で 7 段階評価を行なわせ、その結果から評定の個人差について言及している。その結果、若々しいやさわやかといった印象評定については比較的個人差が少なかったものの、好きの印象評定については強い個人差がみられたと報告されている。更に、この研究では印象評定値を用いて因子分析を行なうことで布の印象を構成する因子を取得し、布の印象評価モデルを構築、検討を行なっている。これにより、人間が布に対し印象評価を行なう際の認知過程の一端が明らかになっている。歌声の声質の好みについても、その認知過程が明らかになれば、音響パラメータの設計に対し有用性が期待されると考えられ、歌声の声質における好みについても適用可能であると考えられる。

以上、個人差のさらなる検討は必要であるものの、歌声の好みに対して第一次近似的な

観点から言って「F0からの倍音構造」や「メル尺度の倍音構造」に関する音響パラメータの有効性が示された。これは、歌声の好みという複雑な現象において、そのスペクトルに含まれる倍音という成分が少なくとも好みに寄与している可能性が示され、今後の研究へと適用される可能性が見いだされた。

図 5.8 より、各音響パラメータの相関は聴取者によって異なる、言い換えると聴取者によって重要な音響的特徴は異なることが確認された。ヴァイオリン演奏の評価の場合は音響パラメータと評価基準の関係から、熟達度などの評価基準についてその一端を確認できる可能性について述べた。同様に、歌声の嗜好についても、図 5.8 よりその一端が確認できる。

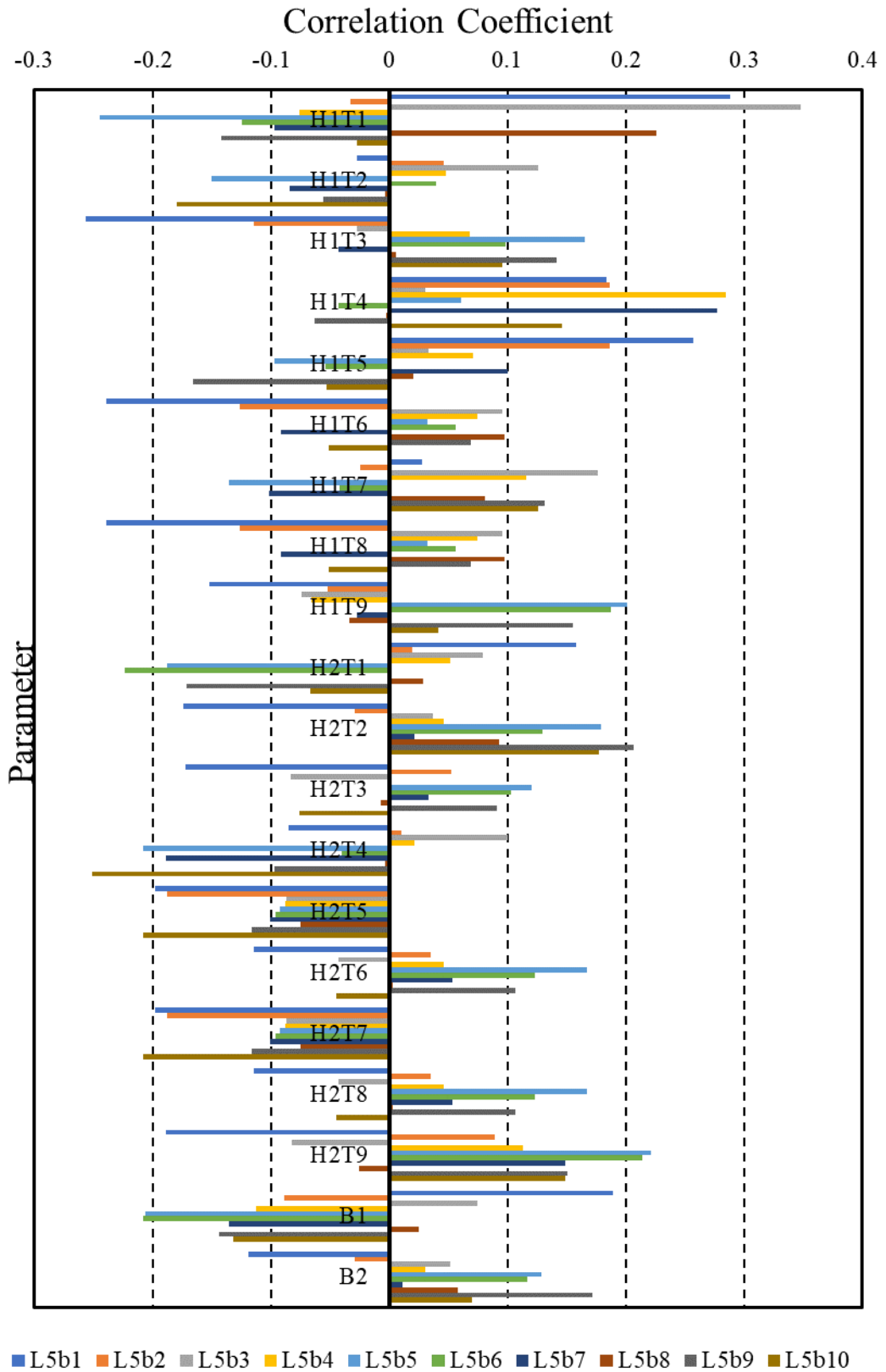


図 5.8 各音響パラメータに対する聴取者ごとの回答の相関係数

5.6.4 歌声の声質における嗜好性の推定に関するまとめと今後の課題

本論文では、歌声の声質の嗜好性を推定するための音響パラメータの設計を試みた。10 人間の聴取者による歌声の声質に好みの評価に対して、各聴取者の回答を目的変数、提案する音響パラメータを説明変数とし、ランダムフォレストによる分類を行なった結果、提案する倍音構造パラメータによる分類において、F-measure の値の平均はチャンスレベルである 50%より高い 0.58 であることが確認され、提案する音響パラメータの有効性が示唆された。つまり、歌声に含まれる倍音という特徴が好みを決定づける要因であることが第一次近似的に示された。一方、聴取者の判断基準のいくつかが音響パラメータに反映できていない可能性が高く、それらの音響パラメータを設計することでより F-measure の値が高い分類を行なうことが可能であることが示唆された。さらに、歌声の声質の好みについて、その判断基準を明文化できない聴取者も存在するので、SD 法などの印象評価の側面から好みの認知過程を明らかにすることによって音響パラメータの設計に活かすことができる可能性も示唆された。

5.7 楽曲推薦に関する技術

2.5 では、音楽デジタル化という側面から音楽サブスクリプション等の普及により楽曲推薦技術が普及したと述べたがここでは、楽曲推薦技術に関する技術的側面を中心に先行研究などを踏まえて、整理を行なう。

5.7.1 推薦技術のアルゴリズム

楽曲推薦技術については、音響パラメータを用いる推薦技術、視聴履歴などを用いる推薦技術の大きく 2 通りの手法、またその組み合わせによって行なっているとされている。そのうち、後者のあるユーザの好みの楽曲を推薦するために、似た好みを持つ別のユーザの視聴履歴を参考に推薦する技術は協調フィルタリングと呼ばれている。協調フィルタリングについては、楽曲推薦技術以外にも様々なサービスで利用されており、例えば、Amazon の商品推薦においてもアルゴリズムが採用されていると考えられている。この手法は、本、映画などコンテンツに問わず、行動履歴があれば採用できるため、多くの研究が行なわれた技術ではあるものの、こと楽曲推薦においてはユーザの行動履歴の蓄積が十分でない楽

曲が推薦されにくいという、コールドスタート問題があるとされている（吉井ら，2009）．この問題を克服するために音響パラメータといった楽曲の内容に関する情報に基づいたフィルタリングによってユーザの好みの楽曲と類似する楽曲を推薦する技術が盛んに研究されているものの，この手法は多くのユーザが好むなどの楽曲の人気度合いを考慮することができないという問題があるとされている（吉井ら，2009）．このように，これらの2つの手法の問題点を解消するために，近年ではこれらを組み合わせたハイブリッド型の推薦技術に注目が集まっている．

本論文では，音響パラメータを用いた楽曲推薦技術に利用可能な提案パラメータの有効性検討を行なっている．前述の通り，ハイブリッド型の推薦技術に近年注目が集まっているものの，協調フィルタリングのための大量の行動履歴，いわゆるビッグデータを取得するのが困難であるため，行動履歴を伴わない場合での有効性検討を行なった．

5.7.2 楽曲推薦に関する諸研究

楽曲推薦に関する研究は，MIRの研究分野の発展と共に盛んに研究が行なわれている．例えば，ユーザの気分に適した楽曲推薦技術，歌詞情報を考慮した楽曲推薦技術など，様々な目的に対し，様々な手段によって楽曲推薦技術の提案が行なわれている．なお，ここでは，協調フィルタリングに関する技術を含まない楽曲推薦技術について触れる．また，ここでは楽曲の印象などの推定に関する研究も楽曲推薦技術の関連技術として取り扱う．

楽曲推薦技術において，盛んに取り上げられるテーマの1つは，楽曲の印象に基づいた推薦技術がある．この技術は楽曲に印象に関するメタ情報が付与されていなければ実現が難しい．また，その際に，すべての印象を人間の手によって付与することも困難である．そこで，音響分析によって音響パラメータを取得し，印象評価実験等で得られた楽曲に対する印象とのマッチングを取り，そのマッチングの結果を基に未知の楽曲に対し，印象のメタ情報を付与する楽曲推薦の基礎技術となり得る研究が行なわれている（伊藤ら，2011；西川ら，2011；など）．これらの研究では，印象評価実験に用いた印象評価の基準に沿った形式の印象が推定される．例えば，音楽ゆらぎ特徴を用いた楽曲印象値の推定手法を提案している研究では，SD法（Osgood, et al., 1957）に基づいた印象評価，楽曲の印象軌跡に関する研究では，RussellのValence-Arousalの感情平面（Russell, 1980）に基づいた印象評価を行っており，同じ印象を推定する研究といっても，その推定の対象となる印象の形式

は異なる。印象を求める際には、楽曲の印象軌跡に関する研究では、音響パラメータだけでなく歌詞情報も用いており、既に利用可能となっているメタ情報を利用する場合もある。

その他にも、楽曲の時代ごとの特徴に関する研究 (Serrà, et al., 2012 ; 岡田ら, 2018 ; など)、動画像に適した楽曲に関する研究 (桐本ら, 2008 ; 追木ら, 2018 ; など) など、一重に楽曲推薦技術といっても、その目的、手法、対象は多種多様となっており、歌声のみに注目した楽曲推薦に関する研究も存在している。

5.7.3 歌声の特徴を考慮した楽曲推薦に関する研究

歌声の特徴を考慮した楽曲推薦に関する研究では、声質の類似度に基づく楽曲推薦に関する研究 (藤原ら, 2007) がある。この研究では、伴奏音つきの音源から伴奏音抑制技術を用いて取得した音響パラメータを取得する手法、また相互情報量を用いて 2 つの特徴ベクトルの類似度を算出する手法が提案され、これらを組み合わせることで歌声の類似度に基づく楽曲推薦を行なっている。また、得られた推薦結果を被験者に評価させた結果、比較手法に対する、提案手法の有効性が示されている。

声質の類似度に基づく楽曲推薦では、用いる音響信号は伴奏音つきの音源であるため、その音響信号に対し、伴奏音抑制を行なっている。まず、歌声における基本周波数を PreFEst (Goto, 2005) によって推定し、次に推定された F0 に基づいて歌声の調波構造を抽出、最後に正弦波重畳モデル (Moorer, 1977) に基づいて音響信号の再合成を行なう、伴奏音抑制技術が用いられている。この伴奏音抑制技術によって再合成された音響信号を用いて、LPC メルケプストラム係数 (徳田ら, 1988) と $\Delta F0$ (Ohishi, et al., 2005) の 2 つの音響パラメータを算出している。なお、この時、PreFEst の特性上、楽曲の間奏部分、また伴奏音のダイナミクスが大きい場合については、再合成された音響信号が歌声の特徴を表す音響パラメータに相応しくない音響信号となっている場合があるため、分析に適さない区間に対し算出された音響パラメータを類似度計算から除外する処理を加え、その後、類似度計算を行なっている。この研究では、音響パラメータを算出する際に、楽曲を歌唱する際の旋律情報として変化する F0 が考慮されていないという問題があり、歌声の類似度において、その声質の類似度と、歌い方の類似度のいずれがどの程度の割合で考慮されているのかが明確にされていない。

5.8 楽曲推薦への利用可能性

本論文の歌声に関するパラメータ設計については、課題は多数残るものの、この課題の解消することによって、楽曲推薦システムへの利用可能性は十分に有すると考えられる。

楽曲推薦システムにおいては、朝聴きたい曲、作業中に聴きたい曲、眠りたいときに聴きたい曲など、音楽を聴取するシチュエーションに合わせた推薦を行なう楽曲推薦システムがある。歌声についてもシチュエーションの影響を受ける可能性が考えられるため、今後はシチュエーションの影響についても調査を行なう必要があると考えられる。

また、本論文では歌声のみの嗜好度について調査を行なったものの、楽曲全体に対する嗜好度を考慮した際には、歌声と楽曲の組み合わせが楽曲全体の嗜好度に影響を与える可能性も考えられる。ある歌唱者の声が好みであったとしても、ある曲調の楽曲についてその歌唱者が歌唱した場合は好みでないという場合もあり得るが、歌声だけに着目した場合についてはこの組み合わせ効果を反映できないので、今後はこの歌声と楽曲の組み合わせについても検討する必要がある。

シチュエーションの影響、組み合わせの影響など、検討を行なうべき項目は多数存在するものの、歌声の特徴を捉えるために設計されたパラメータはこれらの検討を行なう上でも有意なパラメータであると考えている。

5.9 結言

本章では、まず JPOP 楽曲の嗜好性における評価要因について調査結果（桶本ら，2021）について述べた。続いて、得られた結果から歌声に着目した JPOP の嗜好性に関する研究（桶本ら，2022）に関して調査結果を述べた。最後に、調査結果の楽曲推薦への利用可能性について述べた。

6章 結論

本論文では、音楽や音を対象とした感性に関する研究の一助となることを目的とした音響パラメータの設計を行なうことを目的とした。設計した音響パラメータの評価を行なうという点でも、どのような対象に対する音響パラメータを設計したかが明確である方が望ましいと考え、本論文では、その対象を楽器演奏の熟達度評価と歌声の嗜好性評価という点に着目し、音響パラメータの設計、またその評価を行なった。

ヴァイオリン演奏の評価について述べた4章においては、ヴァイオリン演奏の評価に実用可能なパラメータを設計し、その有効性を検証した。なお、本論文でのヴァイオリン演奏とは、基礎的な演奏課題であるヴィブラートを付与した1オクターブ上下行長音階とした。まず、ヴァイオリン演奏の熟達度推定について述べた。本論文において熟達度推定は、熟達者による1octvibに対する熟達度の評価スコアを目的変数、パラメータを説明変数とし、10-fold-CVで線形回帰による推定を行なった。熟達度推定において、設計したパラメータは、音響パラメータと統計パラメータの組み合わせから成る106のパラメータであった。なお、推定精度を評価スコアと推定スコアの相関によって行なったところ、相関係数は0.78であることが確認された。なお、この時、提案パラメータ、特に2音間の遷移に関するパラメータなどが有効である可能性が示唆された。続いて、この音響パラメータを用いて、ヴァイオリン演奏における熟達・技術・表現における評価要因の違いについて調査した。なお、演奏課題は熟達度推定と同様の1octvibである。評価者内と評価者間の各評価の共通性を相関係数によって評価した。続いて、主観評価を目的変数、音響信号から取得した音響パラメータを目的変数として10-fold-CVで線形回帰による推定を行ない、その結果から、主観評価の各評価スコアと物理特徴を表す音響パラメータとの関係を明らかにするという方針で検証を行なった。その結果、いずれの評価者内においても、熟達・技術・表現に対する評価スコアの相関は0.65を上回り、特に、熟達と技術の評価スコアの相関が高い傾向にあることが確認できた。続いて、線形回帰による推定における各パラメータの相対重みに着目した検証を行なった。その結果、各パラメータの相対重みがそのパラメータの表す音響的特における熟達者の重要視する程度を表すならば、重要視するパラメータは評価者、またその評価内容に依存する可能性が示唆された。

続いて、J-POP 楽曲の歌声を対象とした嗜好性に関して述べた 5 章では、日本語歌唱を伴う楽曲に対する嗜好性の評価要因について述べたのちに、歌声における声質の嗜好性評価のための音響パラメータ設計を行なった。日本語歌唱を伴う楽曲の嗜好性の評価要因として、歌手の好み、メロディの好み、伴奏音の好み、歌詞の好みに着目し、AHP を用いて調査を行なった。4 つの要素を考慮した AHP による楽曲の嗜好性に対する評価基準の有意な有効性が聴取者全員に対しては確認できなかった。その理由として、各重要度が AHP 分析の総合評価値で計算される加算結合でない可能性、さらには要素が十分でない可能性が考えられる。また、普段聴取する音楽ジャンルと重要視する要素との関係性があり得ることが示唆された。続いて、歌声の声質に着目した嗜好性評価について述べる。まず、歌声の声質に着目した音響パラメータとして、18 種類の音響パラメータを設計した。続いて、実験刺激と歌声の評価法について述べた。歌声の声質に好みの評価に対して、各聴取者の回答を目的変数、提案する音響パラメータを説明変数とし、ランダムフォレストによる分類を行なった結果、提案する倍音構造パラメータによる分類において、F-measure の値の平均はチャンスレベルである 50% より高い 0.58 であることが確認され、提案する音響パラメータの有効性が示唆された。一方、聴取者の判断基準のいくつかが音響パラメータに反映できていない可能性が高く、それらの音響パラメータを設計することでより F-measure の値が高い分類を行なうことが可能であることが示唆された。さらに、歌声の声質の好みについて、その判断基準を明文化できない聴取者も存在するので、SD 法などの印象評価の側面から好みの認知過程を明らかにすることによって音響パラメータの設計に活かすことができる可能性も示唆された。

最後に今後の課題及び展望について述べる。ヴァイオリン演奏、歌声の嗜好性評価についての今後の課題等については、4 章や 5 章で述べているので、ここでは、音響パラメータの展望という点を中心に述べる。現在はディープラーニングの台頭によって、音響パラメータなどを人間が設計せずとも、コンピュータが自動でなんらかのパラメータを設計、算出し、問題の解決を図るといった研究が増加している。これらの研究においては、人間がコンピュータによる推定結果などを評価するのが困難であるという側面がある。一方、人間が設計した音響パラメータを用いることで、その結果の解釈も前述のようなディープラーニング等を用いた手段よりも容易となる。その解釈から、人間の知覚認知研究に対する貢献、また逆も知覚認知研究からの音響パラメータ設計などの、それぞれの研究結果を活

用することで、音楽に関する分野横断型の交流ができるということが期待される。現在では、XAIも登場しているので、パラメータ等を用いた研究の解釈との比較も行なっていく、まだまだ未知の領域が多い音楽が科学的に明らかにする手法も検討していきたい。

謝辞

本論文の執筆，またそれに至るまでに，多くの方々の御支援と御指導を賜りました．皆様にここで感謝の意を示します．

学部生のころからこれまで温かいご指導，ご鞭撻をいただきました国立音楽大学音楽学部音楽文化教育学科音楽情報専修三浦雅展准教授に深く感謝を致します．三浦雅展准教授には，右も左もわからない学部生の頃からこれまでに，多くの御指導，ご鞭撻を賜りました．研究がどのようなことかをよくわかっていない学部生の頃から，研究におけるプログラミング等といった基礎技術だけではなく，研究倫理，研究者として振る舞いなど多岐に渡るご指導，ご鞭撻に感謝申し上げます．

講義等において御指導をいただいた国立音楽大学音楽学部音楽文化教育学科音楽情報専修神部智教授，友利修教授，横井雅子教授，早稲田みな子教授，中田朱美准教授に厚く御礼申し上げます．

また，本論文内で扱う感性という言葉においてディスカッションを行なってくださった国立音楽大学音楽学部音楽文化教育学科音楽情報専修瀬尾文子准教授に厚く御礼申し上げます．

さらに，音楽学分野と工学分野というような異なる分野間における研究の方針などについてアドバイスをいただいた国立音楽大学久保田慶一氏に厚く御礼申し上げます．

博士前期課程の在籍時に御指導，ご鞭撻，ご支援をいただいた八戸工業大学の嶋脇秀隆教授，藤岡与周教授，小久保温教授，武山泰教授，伊藤智也教授，小坂谷壽一教授，小玉成人教授，坂本禎智教授，石山武教授，高橋史郎教授，岩見一郎教授，山口広行准教授，笹原徹准教授，清水能理准教授，鮎川恵里准教授，本田和也技師，山日康之技師，教職員の皆様に厚く御礼申し上げます．

また，学会において研究に関するディスカッションを行なっていただいた先生方に御礼申し上げます．

また，筑波大学水谷孝一名誉教授，若槻尚人准教授，金沢工業大学山田真司教授，九州大学岩宮眞一郎名誉教授，京都芸術大学津崎実教授，東京藝術大学亀川徹教授，東京情報大学西村明教授，長野看護大学松本じゅん子准教授，青山学院大学西宮康治朗助教，立命

館大学正田悠助教，埼玉大学安井希子助教，防衛大学校黒山喬允講師には，研究に関するディスカッションだけではなく，研究者としての振る舞い等，多岐に渡りお話をさせていただき，今後の研究人生としての糧となりました．厚く御礼申し上げます．

リタルダンド演奏に関する印象評価に関する研究で，共同研究を行なっていたいでいる川上央教授，三戸勇氣教授には，様々な御指導，御鞭撻，また御支援をいただきました．熱く御礼申し上げます．

日本語の指導など多岐に渡りご指導いただきました八戸工業大学三浦研究室 OB の勝呂明央氏，山口翔也氏に厚く御礼を申し上げます．

そして，本論文内の実験にご協力いただきましたすべての皆様に深く感謝申し上げます．

最後に，研究だけではなく様々な面でご支援をしていただきました，龍谷大学，八戸工業大学，国立音楽大学三浦研究室の諸氏に厚く御礼申し上げます．

参考文献

秋吉康晴, “フォノグラフ、あるいは「音を書くこと」の来歴——録音再生技術の着想をめぐる考察——”, 京都精華大学紀要, 第五十一号, pp4-30 (2017).

安藤由典, “楽器の音色を探る,” 音講論集, pp. 799– 802 (2003.1).

荒木雅弘, “フリーソフトではじめる機械学習入門: Python/Weka で実践する理論とアルゴリズム 第2版”, 森北出版 (2018) .

Billboard Japan, “Billboard Japan 2021 年上半期チャート発表、優里が【HOT 100】／SixTONES が【HOT Albums】／YOASOBI が【TOP Artists】首位に”, <https://www.billboard-japan.com/special/detail/3189> (2021.10 参照).

Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., ... & Serra, X. (2013, October). Essentia: an open-source library for sound and music analysis. In Proceedings of the 21st ACM international conference on Multimedia (pp. 855-858).

Breiman, L., “Random forests”, Machine learning, 45(1), pp.5-32(2001).

Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. Mathematics of computation, 19(90), 297-301.

De Cheveigné, A., & Kawahara, H. “YIN, a fundamental frequency estimator for speech and music”, The Journal of the Acoustical Society of America, 111(4), pp.1917-1930 (2002).

Fastl, H., “The hearing sensation roughness and neuronal responses to AM-tones”, Hearing research, 46(3), pp.293-295 (1990).

藤原弘将, 後藤真孝, “Vocalfinder: 声質の類似度に基づく楽曲検索システム”, 情報処理学会研究報告音楽情報科学 (MUS), 2007(81 (2007-MUS-071)), pp.27-32 (2007).

古井貞熙, “電子・情報工学入門シリーズ 2 音響・音声工学”, 近代科学社, p.25 (1992).

Goto, M., “PreFEst: A predominant-F0 estimation method for polyphonic musical audio signals”, Proceedings of the 2nd Music Information Retrieval Evaluation eXchange (2005).

波部斉, “ランダムフォレストの基礎と最近の動向”, 映像情報メディア学会誌, 70(9), pp.788-791 (2016).

Hedger, S. C., Heald, S. L., & Nusbaum, H. C., “Absolute pitch may not be so absolute”,

Psychological science, 24(8), pp.1496-1502 (2013).

平江遼, 西隆司, “感性に基づくクラシック音楽の分類”, 日本音響学会誌, 64(10), pp.607-615 (2008).

Hrimaly, J. 作曲, 岩船雅一 編, “フリマリーヴァイオリン音階教本”, 東京電機大学出版局, pp. 165-167 (2007).

IFPI, “Global Music Report”, https://www.ifpi.org/wp-content/uploads/2022/04/IFPI_Global_Music_Report_2022-State_of_the_Industry.pdf (2022.07.19 参照).

一般社団法人日本ディープラーニング協会監修, “深層学習教科書 ディープラーニング G 検定 (ジェネラリスト) 公式テキスト 第2版”, 翔泳社 (2018).

市原清志, “バイオサイエンスの統計学 正しく活用するための実践理論 3 相関係数検定表の項”, 皆江堂, p. 363 (1995).

市原茂, “布の好みの個人差の因果分析的研究”, 人間工学, 32(1), pp.21-27 (1996).

伊藤雄哉, 山西良典, 加藤昇平, “音楽ゆらぎ特徴を用いた楽曲印象の推定”, 日本音響学会誌, 68(1), pp. 11-18 (2011).

加納政芳, 山田雅之, 遠藤守, “コンピュータサイエンス教科書シリーズ 12 人工知能原理”, コロナ社, p.1 (2017).

川上央, “フランス国立音楽音響研究所 (IRCAM) における音のデザイン活動の紹介 (< 小特集> 音のデザイン: 感性に訴える音が豊かな社会を生み出す)”, 日本音響学会誌, 68(1), pp.43-48 (2011).

城戸健一, “音響入門シリーズ B-1 デジタルフーリエ解析 (I) 基礎編”, コロナ社 (2007).

桐本篤, 佐々木史織, 清木康, “風景画像とサンプル楽曲を用いた環境状況コンテキスト対応型音楽推薦システムの実現”, 情報処理学会研究報告データベースシステム (DBS), 2008(88 (2008-DBS-146)), pp.157-162 (2008).

小西夕貴, 岩見直樹, 三浦雅展, “練習支援を目的としたドラム基礎演奏における熟達度の自動評価”, 信学論, J94, pp.549-559 (2011).

高校新聞 ONLINE, “高校生・中学生が選んだ「好きなアーティスト」ランキング”, <https://www.koukouseishinbun.jp/articles/-/6818> (2021.9 参照).

小島寛之, “完全独習 統計学入門”, ダイヤモンド社 (2006).

Lartillot, O., & Toiviainen, P., “A Matlab toolbox for musical feature extraction from audio”, International conference on digital audio effects, Vol. 237, p. 244(2007).

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O., “librosa: Audio and music signal analysis in python”, In Proceedings of the 14th python in science conference, Vol. 8, pp. 18-25(2015).

Moorer, J. A., “Signal processing aspects of computer music”, A survey. Proceedings of the IEEE, 65(8), pp.1108-1137(1977).

長町三生, “感性工学: 感性をデザインに活かすテクノロジー”, 海文堂出版, p22-27 (1989).

中山功一, 乗富喜子, 大島千佳, “顔の好みの分析”, 人工知能学会全国大会論文集 第26回全国大会, pp. 3B2R27-3B2R27). 一般社団法人 人工知能学会 (2012).

難波精一郎, 桑野園子, “音響テクノロジーシリーズ 4 音の評価のための心理学的測定法”, 日本音響学会編, コロナ社, pp. 76-77 (1998).

中野倫靖, 後藤真孝 & 平賀譲, “楽譜情報を用いない歌唱力自動評価手法”, 情報処理学会論文誌, 48(1), pp.227-236 (2007).

日本音響学会 (編), 鈴木陽一, 赤木正人, 伊藤彰則, 佐藤洋, 荳木禎史, 中村健太郎 (共著), “音響入門シリーズ A-1 音響学入門”, コロナ社, p.40 (2011).

日本音響学会監修, “新版 音響用語辞典”, コロナ社, p. 297 (1988).

西川直毅, 糸山克寿, 藤原弘将, 後藤真孝, 尾形哲也, 奥乃博, “歌詞と音響特徴量を用いた楽曲印象軌跡推定法の設計と評価”, 研究報告音楽情報科学 (MUS) 2011.7, pp.1-8 (2011).

Nonogaki, A., Shimazu, S., Emura, N., Miura, M., Akinaga, S. & Yanagida, M., “Use of spline curve to evaluate performance proficiency of a Czerny piano piece,” Proc. Int. Symp. Performance Science, pp. 68–74 (2011).

三浦雅展, 江村伯夫, 秋永晴子, 柳田益造, “ピアノによる 1 オクターブの上下行長音階に対する熟達度の自動評価,” 音響学会誌, 66, 203-212 (2010).

Ohishi, Y., Goto, M., Ito, K., & Takeda, K., “Discrimination between singing and speaking voices”, Ninth European Conference on Speech Communication and Technology (Eurospeech 2005),

pp.1141-1144(2005).

岡田創太, 山口翔也, 三浦雅展, “女性アイドルポピュラ音楽を対象とした動的パラメータによる年代推定システムの構築”, 日本音響学会誌, 74(7), pp.363-371 (2018).

Okemoto, M., & Miura, M., “A Proficiency estimation system for violin performance”, 15th International Conference on Music Perception and Cognition Abstract Book D4G: Demonstration 4, p. 501 (2018).

Okemoto, M., & Miura, M., “Relation between subjective evaluation for proficiency, expression or technique and acoustic feature on violin performance” Proc. of ISMA(International Symposium on Musical Acoustics) 2019, pp.83-85 (2019).

桶本まどか, 三浦雅展, “ヴァイオリン演奏に対する熟達・技術・表現の評価と音響パラメータの関係”, 音楽音響研究会資料 MA2019-34, pp.31-36 (2019).

桶本まどか, 三浦雅展, “ヴァイオリン演奏の音響特徴を考慮した1オクターブ上下行音階演奏に対する熟達度推定手法”, 日本音響学会誌, 76(12), pp.673-683 (2020).

桶本まどか, 三浦雅展, “日本語歌唱を伴う楽曲に対する嗜好性の評価要因に関する検討”, 日本音楽知覚認知学会 2021 年度秋季研究発表会資料 JSMPC2021-(2)-5, pp.15-18(2021).

桶本まどか, 三浦雅展, “J-POP 楽曲の歌声を対象とした嗜好性の推定のための音響パラメータの提案”, 国立音楽大学 大学院研究年報第 34 輯, pp.87-102 (2022).

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. “The measurement of meaning”, University of Illinois press (1957).

Peeters, G. “A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project”, Technical Report; IRCAM: Paris, France, (2004).

Prétet, L., Hennequin, R., Royo-Letelier, J., & Vaglio, A, “Singing voice separation: A study on training data”, ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 506-510, (2019).

Quinlan, J. R., "Learning with continuous classes." 5th Australian joint conference on artificial intelligence. Vol. 92, pp.343-348 (1992).

Russell, J. A, “A circumplex model of affect”, Journal of personality and social psychology, 39(6), pp.1161–1178 (1980).

Saaty, T. L., “Decision making with the analytic hierarchy process”, International journal of

services sciences, 1(1), pp.83-98(2008).

坂本真樹, “感性情報学—オノマトペから人工知能まで—”, コロナ社, pp.1-4 (2018).

佐藤保, “回帰分析の方法：主成分分析の応用,” 三田学会雑誌, 64, 819–824 (1971).

Seashore, C. E., “The Psychology of Music”, Music Educators Journal, 23(4), pp. 30–33 (1937).

Serrà, J., Corral, Á., Bogaña, M., Haro, M., & Arcos, J. L., “Measuring the evolution of contemporary western popular music,” Sci. Rep., 2, pp.1–6 (2012).

渋谷恒司, 菅野重樹, “バイオリン演奏における感性表現としての音色と運動の関係,” 計測自動制御学会論文誌, 32, pp.1259–1266 (1996).

総務省, “平成28年版情報通信白書 | 人工知能 (AI) とは”, <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/html/nc142110.html> (2022.7.11 参照).

杉山孝男, 椎塚久雄., “階層的ファジィ積分による意思決定法 (< 特集論文> 人文・社会科学へのファジィ理論の応用)”, 日本ファジィ学会誌, 5(4), pp.772-782 (1993) .

巢籠悠輔, “詳解ディープラーニング 第2版 ~TensorFlow/Keras・PyTorchによる時系列データ処理~ (Compass Books シリーズ)”, マイナビ出版, pp.141-142 (2019).

Sundberg, J. 著, 榊原健一 監訳, 歌声の科学, 全音楽譜出版社 (2008).

Terhardt, E., “On the perceptions of periodic sound fluctuation (Roughness)” Acustica 30, pp.201-213, (1974).

徳田恵一, 小林隆夫, 今井聖, “メル一般化ケプストラムの再帰的計算法”, 電子情報通信学会論文誌 A, 71(1), pp.128-131 (1988).

東京大学教養学部統計学教室, “統計学入門 (Vol. 1)”, 東京大学出版会 (1991).

追木智明, 櫻惇志, 宮崎純, “物体の色や表情情報を利用した画像の印象にあった音楽推薦手法の提案”, 研究報告情報基礎とアクセス技術 (IFAT), 2018(25), pp.1-6 (2018).

Turumeso, “vocal-remover v4.0.0”, <https://github.com/tsurumeso/vocal-remover> (2021.9 参照).

堤盛人, 清水英範, 井出裕史, “多重共線性に対する適切化手法とその実証的比較研究,” 応用力学論文集 1, pp. 137–145 (1998).

上野山努, 櫻村雅章, 小沢慎治, “ドラム音の音色における感性情報と工学的パラメータとの対応付け”, 日本音響学会誌, 49(10), pp.671-681 (1993).

梅本堯夫, “音楽心理学 ヴァイオリン演奏の特徴の項”, 誠信書房, p. 332 (1996).

山田真司, 三浦雅展, “音楽情報処理で用いられる音響パラメータによる音楽理解の可能性”, 日本音響学会誌, 70(8), pp.440-445 (2014).

Yamasaki, M., & Miura, M., “Proficiency estimation for audio of cello performances,” Proc. Forum Acusticum 2014, R12-1 (2014).

山内勝也, “ハイブリッド車・電気自動車などの静音性対策の動向: 車両接近報知音のデザイン (< 小特集> 音のデザイン: 感性に訴える音が豊かな社会を生み出す)”, 日本音響学会誌, 68(1), pp.31-36 (2011).

吉井和佳, 後藤真孝, “音楽情報処理技術の最前線: 7. 音楽推薦システム”, 情報処理 50.8 pp.751-755 (2009).

吉村季織, 福原亘治, 三ツ木健一郎, 高柳正夫, “Microsoft Excel を用いたケモメトリクス計算 (4)—主成分回帰,” J. Comput. Chem. Jpn., 10, 33–43 (2011).

関連発表論文

I. 査読付き原著論文（2件，国内誌：2件）

- 1 桶本まどか，三浦雅展，“ヴァイオリン演奏の音響特徴を考慮した 1 オクターブ上下行音階演奏に対する熟達度推定手法”，日本音響学会誌 76 巻 12 号, pp.673 - 683 (2020).
2020.12
- 2 桶本まどか，三浦雅展，“J-POP 楽曲の歌声を対象とした嗜好性の推定のための音響パラメータの提案”，国立音楽大学 大学院研究年報第 34 輯, pp.87-102 (2022). 2022.3

II. 国際会議論文（6件，abstract 査読付き：6件）

- 1 Madoka Okemoto, Masanobu Miura, “Proficiency estimation for violin performance using pitch and vibrato parameters”, Proceedings of the 6th Conference of the Asia-Pacific Society for the Cognitive Sciences of Music, p.28 (2017) . 2017.08
- 2 Madoka Okemoto, Masanobu Miura, “A Proficiency estimation system for violin performance”, 15th International Conference on Music Perception and Cognition Abstract Book D4G: Demonstration 4, p. 501 (2018). 2018.7
- 3 Madoka Okemoto, Masanobu Miura, Yuki Mito and Hiroshi Kawakami, “How to Play Ritardando?: Factor analysis for impression to the piece of “Pomp and Circumstance”, The International Symposium on Performance Science 2019 (2019). 2019.7
- 4 Madoka Okemoto, Masanobu Miura, “Tempo estimation for musical audio of tempo-variated music”, 23th International Congress on Acoustics, pp.6491-6493 (2019). 2019.9
- 5 Madoka Okemoto, Masanobu Miura, “Relation between subjective evaluation for proficiency, expression or technique and acoustic feature on violin performance” Proc. of ISMA(International Symposium on Musical Acoustics) 2019, pp.83-85 (2019) . 2019.9
- 6 Madoka Okemoto, Masanobu Miura, ”Investigation of relationship between beat-timing and acoustics-parameter for method of beat-timing and tempo estimation”, 24th International Congress on Acoustics, (2022). 2022.10

Ⅲ. 国内全国大会 (11 件)

- 1 桶本まどか, 三浦雅展, “テンポ変動楽曲の事例分析に基づくリタルダンドの印象評価～「威風堂々」第一番の場合～”, 日本音楽知覚認知学会 2018 年春季研究発表会・日本音響学会音楽音響研究会 2018 年 5 月研究会共催研究発表会, 日本音楽知覚認知学会平成 30 年度春季研究発表会資料 JSMPC2018(1) (音楽音響研究会資料 Vol.37 No.2 MA2018-08), pp.19-24 (2018). 2018.5
- 2 桶本まどか, 三浦雅展, 三戸勇氣, 川上央, “リタルダンドにおける時間遅延量の形状とその印象の関係”, 日本音響学会平成 30 年度春季研究発表会資料 3-1-11, pp.1231-1234, (2019). 2019.3
- 3 桶本まどか, 三浦雅展, 嶋脇秀隆, 三戸勇氣, 川上央, “リタルダンドの印象における音楽経験の影響”, 日本音楽知覚認知学会令和元年度春季研究発表会資料 JSMPC2019(1)-3, pp.11-16 (2019). 2019.6
- 4 桶本まどか, 三浦雅展, “ローカルテンポ推定の性能評価および求められる精度に関する考察”, 日本音楽知覚認知学会令和元年度秋季研究発表会資料 JSMPC2019(2)-6, pp.17-22 (2019). 2019.11
- 5 桶本まどか, 三浦雅展, “拍時刻推定を目的としたオンセット時刻の認知に対する個人差の基礎的検討”, 日本音楽知覚認知学会 2020 年度秋季研究発表会 JSMPC2020-8 (共催: 日本音響学会音楽音響研究会 2020 年 10 月研究会 MA2020-38), pp.33-38 (2020). 2020.10
- 6 高倉怜花, 桶本まどか, 三浦雅展, “ポピュラ音楽を対象とした各セクションにおける付与残響の自動付加手法に関する提案”, 日本音楽知覚認知学会 2020 年度秋季研究発表会資料 JSMPC2020-18 (共催: 日本音響学会音楽音響研究会 2020 年 10 月 MA2020-48), pp.83-88 (2020). 2020.10
- 7 桶本まどか, 高倉怜花, 三浦雅展, “印象の変化に基づいた音楽音響信号の自動分割に対する検討”, 日本音響学会 2021 年春季研究発表会資料, 1-9-8, pp. 1267-1270 (2021). 2021.3
- 8 桶本まどか, 三浦雅展, “拍時刻推定のためのオンセット時刻の知覚に対する検討～演奏音を構成する楽器数に関する調査～”, 日本音響学会 2021 年秋季研究発表会 2-9-3, pp. 1159-1162 (2021). 2021.9

- 9 桶本まどか, 三浦雅展, ”日本語歌唱を伴う楽曲に対する嗜好性の評価要因に関する検討”, 日本音楽知覚認知学会 2021 年度秋季研究発表会資料 JSMPC2021-(2)-5, pp.15-18 (2021). 2021.11
- 10 平田友花, 桶本まどか, 三浦雅展, ”なぜオンラインライブは感動が薄いのか? ~心拍変動による音楽聴取時のストレス評価による検討~”, 日本音楽知覚認知学会 2021 年度秋季研究発表会資料 JSMPC2021-(2)-7, pp.25-30 (2021). 2021.11
- 11 桶本まどか, 三浦雅展, “歌唱言語の違いが拍のタイミング知覚に与える影響~Let It Go の場合~”, 日本音楽知覚認知学会 2022 年度春季研究発表会資料 JSMPC2022-(1)-14, pp.53-56(2022). 2022.6

IV. 国内研究会 (10 件)

- 1 桶本まどか, 岡田創太, 三浦雅展, “ヴァイオリン演奏におけるヴィブラートを考慮した熟達度推定の試み”, 日本音響学会音楽音響研究会研究会資料 MA2016-68, pp.25-31 (2017). 2017.03
- 2 桶本まどか, 三浦雅展, “心拍変動パラメータを用いた知覚感情推定手法の検討”, 音楽音響研究会資料 MA2017-28, pp.43-48 (2017) . 2017.10
- 3 桶本まどか, 新川芳行, “リタルダンド演奏を対象とした CPN によるモデル化の試み”, 日本音響学会音楽音響研究会資料 MA2017-47, pp. 1-4 (2017). 2017.11
- 4 桶本まどか, 三浦雅展, ”クラシック及びジャズ楽曲に対するローカルテンポ推定手法の精度向上への試み”, 音楽音響研究会資料 MA2018-36, pp.1-6 (2018) . 2018.10
- 5 桶本まどか, 三浦雅展, ”ローカルテンポ推定の評価法に関する一考察”, 音楽音響研究会資料 MA2018-83, pp.53-58 (2019) . 2019.02
- 6 桶本まどか, 三浦雅展, “ヴァイオリン演奏に対する熟達・技術・表現の評価と音響パラメータの関係”, 音楽音響研究会資料 MA2019-34, pp.31-36 (2019). 2019.10
- 7 桶本まどか, 三浦雅展, 嶋脇 秀隆, “ローカルテンポ推定のための短区間ビート数判定に関する検討”, 音楽音響研究会資料 MA2019-54, pp.25-30 (2019). 2019.11
- 8 高倉怜花, 桶本まどか, 三浦雅展, “ヘッドフォンを用いた音楽聴取における環境音の影響に関する基礎的検討”, 日本音響学会音楽音響研究資料 MA2020-72, pp.43-48, (2021). 2021.2

- 9 桶本まどか, 三浦雅展, “演奏音の立ち上がり形状がオンセット知覚に与える影響”, 日本音響学会音楽音響研究資料 MA2020-73, pp.49-52 (2021). 2021.2
- 10 桶本まどか, 三浦雅展, “オンライン会議システムを介した演奏音に対する印象評価”, 日本音響学会音楽音響研究会資料 Vol.41, No.1, pp.13-16, (2022). 2022.4

V. 国内支部大会 (3件)

- 1 桶本まどか, 三浦雅展, “音楽音響信号のローカルテンポ推定手法におけるモジュール群の有効性検証”, 平成30年度電気関係学会東北支部連合大会, 2D12(2018). 2018.9
- 2 桶本まどか, 三浦雅展, “ローカルテンポ推定手法の精度向上に関する一検討”, 第1回東北地区音響学研究会 1-15 (2018) . 2018.11
- 3 桶本まどか, 高倉怜花, 三浦雅展, “音楽音響信号における印象切り替わり時刻の自動推定に関する基礎的検討”, 第3回東北地区音響学研究会, 3-19 (2020). 2020.11

IV. 受賞 (6件)

- 1 SEMPRE Conference Award at ICMPC15/ESCOM10 (Madoka Okemoto, Masanobu Miura) (2018.7)
- 2 日本音響学会第19回学生優秀発表賞 (桶本まどか, 三浦雅展, 三戸勇氣, 川上央) (2019.9)
- 3 ICA-ASA Young Scientist Conference Attendance (Madoka Okemoto, Masanobu Miura) (2019.9)
- 4 日本音響学会音楽音響研究会奨励賞 (桶本まどか, 三浦雅展) (2019.10)
- 5 八戸工業大学大学院工学研究科電気電子・情報工学専攻学位論文優秀賞 (2020.3)
- 6 日本音響学会音楽音響研究会奨励賞 (桶本まどか, 三浦雅展) (2020.10)