

音響特徴量に着目した楽曲演奏の違い

Differences in Musical Performances Focusing on Acoustic Features

北野 愛†
Megumi Kitano

1. はじめに

音楽の演奏は、単に楽譜に書かれた音符を正確に再現するだけの作業ではなく、演奏者の持つ技術や表現、さらには使用する楽器の個体差や演奏や録音を行う空間の特性など、さまざまな要因によって大きく異なる音響の特徴を持つ。特にバイオリンのような弦楽器においては、演奏者の弓の動かし方や速度、圧力、微妙なビブラートやポジション移動のニュアンスなどが複雑に関与し、同じ楽曲であっても演奏ごとに全く異なる印象を与える場合が少なくない。これらの演奏の違いは、聴取者に与える音楽的な感動や評価、さらには動画配信サイト等での視聴行動にも大きな影響を及ぼすと考えられる。

従来、音楽の演奏における違いは、聴取者や専門家による主観的な評価や感性に頼ることが多く、その物理的側面は十分に議論されてこなかった。しかし、近年は音響信号処理や音楽情報処理技術の発展により、演奏の特徴を客観的に捉え、物理的な指標として表す試みが進められている。牧[1]は、年代区分の異なるバイオリンやアマチュアやプロのバイオリン奏者といった熟練度の異なる奏者による演奏の放射パターンを分析することで、その差異を物理的に明らかにした。枝ら[2]は、フルートを対象に、材質の違いによる演奏者の実感と音響特性を比較し、演奏感覚と客観的データの関係性を検討している。また、高橋ら[3]は、ピアノ演奏において熟達者と被熟達者の演奏表現の違いを比較し、熟達度が音響特徴量にどのように表れるかを明らかにした。さらに、小幡ら[4]はバイオリンの歴史と高価なストラディバリウスの音響的特徴を概説しており、楽器の選択が演奏に及ぼす影響について言及している。音響特徴量を用いた音楽理解の可能性については、山田ら[5]が整理を行い、スペクトル重心やゼロクロス、RMS などの特徴量が演奏分析に有効であることを指摘している。

しかし、これらの研究は主として演奏者や楽器の材質、名器と一般器の比較に主眼を置いたものが多く、聴取行動と演奏の音響的特徴を直接関連付けて分析した研究は少ない。現代では YouTube などの動画配信プラットフォーム上で多くの演奏動画が手軽に視聴されており、再生回数や高評価の数といった指標はある意味で聴取者がどのような演奏を魅力的に感じているかを反映するデータと考えられる。視聴者に多く聞かれる演奏とそうでない演奏との間には、演奏者の解釈や表現、使用する楽器や録音環境を含めた音響的な違いが存在し、それを客観的に分析することは、演奏者自身の改善や指導、さらには楽器選定の基準としても大きな意義を持つ。

そこで本研究では、バイオリンのソロ演奏に着目し、演奏者や楽器の個体差、演奏・録音環境など全てを含んだ「演奏の違い」を音響特徴量に基づいて定量的かつ定性的に分析・比較し、その違いを明らかにすることを目的とする。具体的には、動画配信サイトに公開されている同一楽曲の複数の演奏のうち、視聴回数が多い音源と少ない音源を対象に、複数の音響特徴量を算出し、比較を行う。これにより、演奏が

聴取者に与える印象や好まれる要素を客観的データとして捉え、演奏家や指導者が演奏を見直す際の参考や、楽器選び、録音制作における指標を提供することを目指す。

2. 関連研究

2.1 演奏者の熟練度・表現と音響特徴量

音楽演奏の物理的差異を明らかにするために、演奏者の習熟度や表現の違いを音響特徴量から分析する試みが多くなされている。牧[1]はバイオリンを対象に、奏者の熟練度および名器ストラディバリウスの空間放射特性を計測し、名器が奏でる音の周波数分布や放射指向性に明確な違いが現れることを示した。こうした物理的な差異は演奏技術や聴取印象に直結するものである。高橋ら[3]はピアノ演奏において熟達者と被熟達者の打鍵タイミングや強弱パターンの差異を音響データで分析し、熟練度が複数の音響特徴量に有意な影響を与えることを明らかにしている。

2.2 楽器の材質・構造と音響特性

楽器そのものの材質や構造が音響特性に与える影響についても多くの研究が存在する。枝[2]は異なる材質のフルートを用いた演奏を比較し、材質により奏者が演奏時に感じる吹奏感や音色には明確な違いがあることを示した。一方で、音響特徴量には有意な差が現れず、奏者が体感する違いと物理的に計測される音響特徴量との間には必ずしも対応がないことが示唆された。これは、奏者の感覚と聴取者に伝わる音響の違いの関係が単純ではないことを示す重要な知見である。

2.3 音響特徴量を用いた音楽理解・評価の枠組み

山田ら[5]はラウドネスやラフネスといった心理音響指標やスペクトル重心やゼロクロス、RMS などの音楽情報処理におけるパラメータが音楽情報処理や感性評価に広く利用されており、演奏の音響的特徴を把握するために有効であることを体系的に整理した。また、Sergio Giraldo ら[6]は機械学習によって演奏音のティンバー品質を評価するモデルを構築し、主観評価との相関も検証した。

2.4 動画配信プラットフォームにおける視聴行動と音響・視覚特徴

近年、音声・視聴行動の統合的分析に関して注目が集まっており、Yap Kah Yee ら[7]は Spotify の音響特徴と YouTube のソーシャルメディア指標を用いて楽曲の人気を予測し、特に再生回数やコメント数などのソーシャル指標が音響特徴単独よりも予測精度を 10~60%向上させることを示した。さらに、Cheng-Yuan Lee ら[8]は音響特徴に加え、音楽動画の視覚的特徴も人気予測に寄与することを報告している。

音楽以外の視聴分析でも、YouTube における視覚情報の影響が指摘されており、例えば Chia-Jung Tsay[9]は音楽コンクールでは無音の映像から優勝者を選出することはできたが、音声のみや映像と音声を含む録音では優勝者を特定することができなかったと示した。これらは視覚・音響・ソ

† 関西学院大学, Kwansei Gakuin University

ーシャル要因が複合的に視聴行動に影響する可能性を示唆している。

2.5 本研究の位置づけと新規性

これらの研究を整理すると、演奏者熟練度や楽器材質、映像・音響。ソーシャル指標による人気予測といったそれぞれ異なる切り口があることが分かる。一方で、バイオリンのソロ演奏に対して、視聴回数と音響特徴量を包括的に紐づけて分析する研究はまだ存在しない。

本研究は、これまで個別に扱われてきた演奏者の違いと音響特徴、楽器構造と音響特性、さらに音響や映像による人気予測といったアプローチを総合し、同一の楽曲に血合いする複数の演奏動画について視聴回数という聴取行動の指標を用いて音響特徴量の影響を分析する点に独自性がある。この分析を通じて、演奏家や教育者、録音制作者がどのような演奏が多く聴取者に選ばれているのかを客観的に理解し、今後の演奏や指導、音源制作に役立てるための新たな知見を提供することを目的としている。

3. 研究手法

3.1 分析対象とデータ収集

本研究では、バイオリンのソロ演奏における音響的な違いを分析するために、動画配信サイト上に公開されている同一楽曲の演奏動画を対象とした。分析に際しては、視聴回数の多い動画と少ない動画を複数選定し、それらを比較することで演奏の音響特徴量と視聴行動(人気)の関係を調査した。

選定した動画は全て独奏(ピアノ伴奏を含まない)で、録音環境や撮影条件は様々である。これにより演奏者の表現、楽器の個体差、録音条件など演奏全体を含めた違いを総合的に評価することを目的とした。動画はmp3形式でダウンロードし、解析を行った。

3.2 使用ツールと分析環境

音響特徴量の計算にはPython3.10.11とlibrosa0.10.2を使用し、補助的にnumpy, scipy, matplotlibを利用した。ハードウェアはWindows10搭載のノートPC(Intel Core i5, メモリ 8GB)で処理を行った。

librosaにおける分析の主要なパラメータは以下の通りである。

- ・フレーム長(window length) : 2048 サンプル
- ・ホップ長(hop length) : 512 サンプル
- ・窓関数 : ハニング窓

3.3 抽出した音響特徴量と計算式

本研究では演奏音の持つ物理的特徴を多角的に把握するため、以下の音響特徴量を計算した。それぞれlibrosaを中心としたPythonライブラリを用いて求めた。

- ・スペクトル重心

スペクトル重心は、周波数成分のパワーに基づく加重平均であり、音の明るさや高域成分の重みを示す。各フレームについて高速フーリエ変換を行い、パワースペクトルの分布を計算した上で、各周波数ビンと振幅の積を総和し、全体の振幅和で割ることで算出している。librosaではlibrosa.feature.spectral_centroid(y=y, sr=sr)により計算した。フレームtにおけるスペクトル重心を C_t (Hz)、FFTで得られる周波数ビンの数をK、k番目の周波数ビンの中心周波数を f_k (Hz)、フレームtにおけるk番目の振幅スペクトル値を $X_t(k)$ とすると式(1)で算出される。

$$C_t = \frac{\sum_{k=1}^K f_k \cdot |X_t(k)|}{\sum_{k=1}^K |X_t(k)|} \quad (1)$$

- ・スペクトル平坦度

スペクトル平坦度は、スペクトルの形状がノイズ的に平坦か、あるいはピークを持つ(楽音的)かを示す指標である。各フレームのパワースペクトルに対して幾何平均を算出し、これを算術平均で割ることで計算しており、値が1に近いほどノイズ、0に近いほど純音的なスペクトルを示す。librosaではlibrosa.feature.spectral_flatness(y=y)により計算式(2)で計算される。

$$SF_t = \frac{(\prod_{k=1}^K |X_t(k)|)^{1/K}}{\frac{1}{K} \sum_{k=1}^K |X_t(k)|} \quad (2)$$

- ・ゼロクロス率

ゼロクロス率は、波形の時間的変化を捉える指標で、波形が0を跨ぐ回数をフレーム内のサンプル数で正規化して求める。粗い音やノイズ、高域成分が多いと高くなる。librosaではlibrosa.feature.zero_crossing_rate(y=y)を利用。フレームtにおけるゼロクロス率を ZCR_t 、フレームt内の時間インデックスnにおける信号振幅を $x_t(n)$ 、条件が成り立つときに1、それ以外0となる指示関数を $I\{\}$ 、フレーム内のサンプル数をNとすると式(3)で算出される。

$$ZCR_t = \frac{1}{N-1} \sum_{n=1}^{N-1} I\{x_t(n)x_t(n-1) < 0\} \quad (3)$$

- ・スペクトル帯域幅

スペクトル帯域幅は、スペクトルのばらつき(分散)を表し、中心周波数からどれだけ広くエネルギーが分布しているかを示す。重心周波数との差の二乗を重ね付けした平均を計算し、平方根をとって算出する。librosaではlibrosa.feature.spectral_bandwidth(y=y, sr=sr)で計算。BW_tをフレームtにおけるスペクトル帯域幅とすると式(4)で算出される。

$$BW_t = \sqrt{\frac{\sum_{k=1}^K (f_k - C_t)^2 \cdot |X_t(k)|}{\sum_{k=1}^K |X_t(k)|}} \quad (4)$$

- ・RMS(Root Mean Square)

RMSは音の時間的エネルギーを示し、フレーム内の振幅を二乗して平均し、平方根をとることで求める。これにより演奏中のダイナミクスや音圧レベルの変化を定量的に評価する。librosaではlibrosa.feature.rms(y=y)で計算。フレームtにおけるRMSの値を RMS_t とすると式(5)で算出される。

$$RMS_t = \sqrt{\frac{1}{N} \sum_{n=1}^N x_t(n)^2} \quad (5)$$

- ・HNR(Harmonics-to-Noise Ratio)

HNRは音の周期成分と非周期成分の比を示す指標であり、楽音の明瞭さやノイズ性の評価に用いる。自己相関関数を計算し、最も高いピークとラグ0の値を比較することで推定し、対数比としてデシベルスケールで表現した。フレームtにおけるHNR(dB)を HNR_t 、フレームtにおける遅延時間 τ に対する自己相関関数、自己相関が最大となる遅延を τ_{max} 、自己相関の最大値を $R_t(0)$ とすると式(6)で算出される。

$$HNR_t = 10 \log_{10} \left(\frac{R_t(\tau_{max})}{R_t(0) - R_t(\tau_{max})} \right) \quad (6)$$

- ・時間的重心

時間的重心は、音のエネルギーが時間軸上のどの位置に分布しているかを示す指標であり、信号の重心を意味する。これは演奏全体の構造や力点の違いを捉える上で重要な手がかりとなる。時間的重心 T_c は、全フレーム数を N 、フレーム n に対応する時間を t_n (秒)、フレーム n におけるエネルギーを $E(n)$ とすると式(7)で算出される。

$$T_c = \frac{\sum_{n=1}^N t_n \cdot E(n)}{\sum_{n=1}^N E(n)} \quad (7)$$

- ・テンポ

テンポは、楽曲の拍の速さ (BPM) を示す指標であり、演奏の速さやリズムの印象を定量化する上で重要である。

- ・フーリエテンポグラム

フーリエテンポグラムは、音声信号のオンセットストレングス (onset strength envelope) に対して短時間フーリエ変換 (STFT) を適用することで、時間軸上のテンポ成分を周波数領域で表現する手法である。

- ・自己相関テンポグラム

自己相関テンポグラムは、音楽信号中の周期的なリズム構造を時間的に可視化・定量化する手法である。短時間自己相関を計算して、時間ごとの周期性を表現する。

3.4 特徴量の統計処理と可視化

全ての音響特徴量はフレーム単位で算出し、演奏全体について平均値および標準偏差を求めた。これにより単一の値として動画ごとの代表値を比較可能とした。

4. 実験結果および考察

本研究では、複数の楽曲についてそれぞれ異なる演奏者によるバイオリンソロ演奏動画を分析対象とした。演奏者 A, B, C は比較的視聴回数が多い動画の演奏者であり、一方演奏者 D, E, F は同様の楽曲において視聴回数が少ない動画の演奏者である。これにより、視聴回数が多い群と少ない群を明確に分け、音響特徴量の違いを視聴行動の指標である視聴回数との関連から分析した。本研究で扱った楽曲は2曲ある。楽曲1は、バッハの無伴奏ヴァイオリンのためのパルティータから「シャコンヌ」ニ短調 BWV 1004、楽曲2は、パガニーニの24のカプリース Op.1 第24番である。

4.1 スペクトル重心の比較

まず一曲目について、演奏者ごとのスペクトル重心の平均値および標準偏差を算出し、その結果を図1,2に示す。全体的に視聴回数が多い演奏の方がスペクトル重心の平均が低くなっている。スペクトル重心の標準偏差は視聴回数が多い演奏の方が高い値を示した。この結果から、視聴回数が多い音源の方が、スペクトル重心の時間的な変動が大きい、すなわち音色により多くの揺らぎが含まれていたことがわかる。

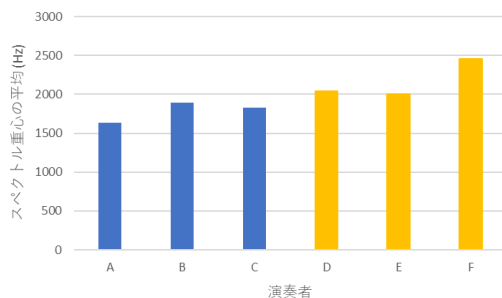


図1:楽曲1のスペクトル重心の平均

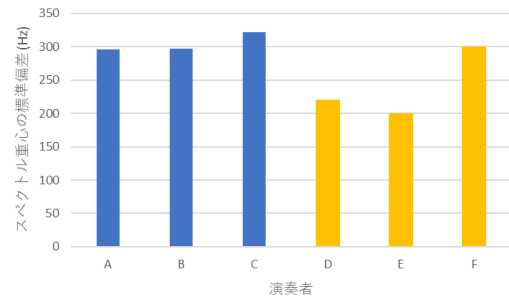


図2:楽曲1のスペクトル重心の標準偏差

次に、楽曲2について演奏者ごとのスペクトル重心の平均値および標準偏差を算出し、その結果を図3,4に示す。図3から、こちらの楽曲においても視聴回数が多い演奏の方がスペクトル重心の平均が低いことが分かる。またスペクトル重心の標準偏差は1つ大きな値をとり、分布の偏りに強く影響を与えていることが分かる。実際、他2つの数値は、視聴回数が多い演奏の値よりも低く、全体としては一貫した傾向が見られない。従って、この演奏はスペクトル重心の時間的な変動が特に大きかった特殊な例であり、演奏法や録音条件などが他の音源とは大きく異なっていた可能性がある。

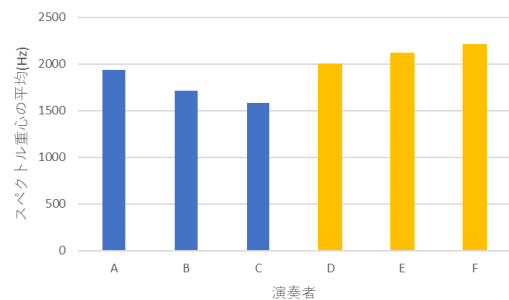


図3:楽曲2のスペクトル重心の平均

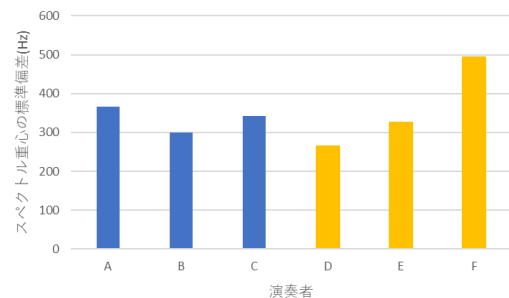


図4:楽曲2のスペクトル重心の標準偏差

これらの結果から、2つの楽曲に共通して、視聴回数が多い演奏はスペクトル重心が低めであることが分かる。すなわち、高周波成分が過度に強調されていない音源が、より多くの聴取者に好まれていることが分かる。特に、バイオリン独奏などのクラシック楽曲では、高周波成分が強すぎると耳に刺さるような印象を与える場合があり、長時間の聴取や繰り返しの再生において聞き疲れを引き起こすことがあると指摘されている。そのため、スペクトル重心の平均値が比較的低く、より滑らかで落ち着いた音色を有する演奏が好まれる傾向にあると考えられる。また、スペクトル重心の標準偏差では2つの楽曲において、スペクトル重心の標準偏差に関する傾向は一樣ではなかった。楽曲1では、視聴回数が多い音源の方が標準偏差が大きく、音色の時間的な変動が大きい傾向が見られた。一方、楽曲2では平均値として視聴回数が少ない音源の方

が高いものの、その中の1つが極端に高い値を示しており、全体の傾向として視聴回数との明確な関係は見出しにくかった。

4.2 スペクトル平坦度の比較

同様に、楽曲 1 について演奏者ごとのスペクトル平坦度の平均値を図 5 に示す。図 5 より視聴回数が多い演奏の方がスペクトル平坦度が低い値を示した。この結果から、視聴回数が多い演奏ではスペクトル平坦度の値が一貫して低く、倍音構造が明瞭な演奏であることが示唆される。視聴回数が少ない演奏の中には、特に高い値を示す演奏もあり、よりノイズ的な性質を持つ可能性がある。

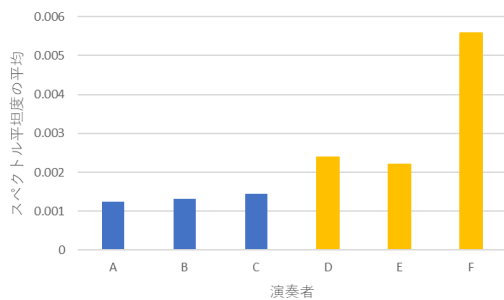


図 5: 楽曲 1 のスペクトル平坦度の平均

また、楽曲 2 について演奏者ごとのスペクトル平坦度の平均値を図 6 に示す。図 6 より視聴回数が多い演奏のスペクトル平坦度の平均が低いことが分かる。特に視聴回数が少ない演奏の内 2 つの値が極めて高い値を示していることから、ノイズの傾向が顕著な演奏が存在していたことが分かる。

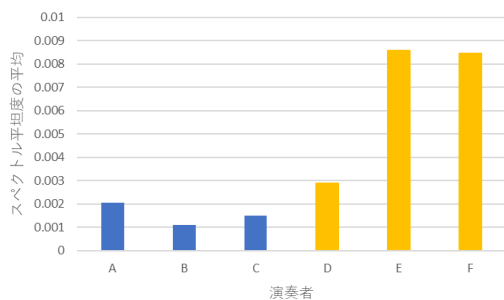


図 6: 楽曲 2 のスペクトル平坦度の平均

2 つの楽曲において、いずれも視聴回数が多い演奏ではスペクトル平坦度の平均値が顕著に低く、少ない演奏では高い値を示す傾向が確認された。このことから、スペクトル平坦度が低く、倍音構造が明瞭な音源の方が聴取者に好まれる傾向があることが示唆される。特に、クラシック音楽やソロ楽器演奏においては、楽器固有の音色や音の輪郭が明確であることが演奏の質や表現力に直結するため、スペクトル平坦度が低い、つまり音の構造が明確な音源が好まれるという関係性は自然な結果といえる。また、視聴回数が少ない演奏の中には、極端に高い値を示すものもあり、こうしたノイズの傾向の強い音源は、音の明瞭さや演奏の表現が損なわれ、結果的に聴取者の支持を得にくくなった可能性がある。

4.3 ゼロクロス率の比較

次に楽曲 1 について演奏者ごとのゼロクロス率を図 7 に示す。図 7 から視聴回数が多い演奏の方がゼロクロス率が低いことが分かる。この結果から、視聴回数が多い演奏の方がゼロクロス率が低く、より周期的で滑らかな音の性質を有している可能性が示唆される。

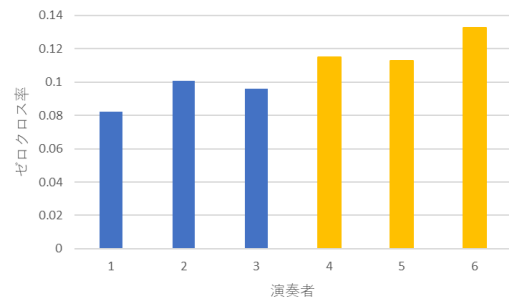


図 7: 楽曲 1 のゼロクロス率

楽曲 2 における演奏者ごとのゼロクロス率を図 8 に示す。図 8 からここでも視聴回数が多い演奏のゼロクロス率が低いことが分かる。この結果から、視聴回数が多い演奏のゼロクロス率が一貫して低く、視聴回数が少ない演奏に比べてノイズ的成分が少ない傾向が確認された。

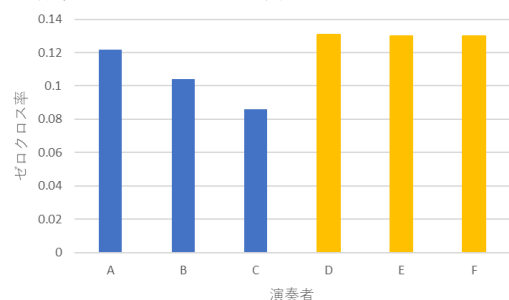


図 8: 楽曲 2 のゼロクロス率

以上の結果から、2 つの楽曲に共通して、視聴回数が多い演奏ではゼロクロス率が有意に低い傾向が観察された。ゼロクロス率が低いということは、信号内での極性の変化が少なく、音がより滑らかで周期的であることを意味しており、聴取者にとって聴きやすく、快適に感じられる可能性がある。特にクラシック音楽や独奏楽器の演奏においては、楽音の音色の滑らかさや明瞭さが重視されることから、ゼロクロス率の低い演奏が好まれる傾向にあることは自然な結果である。一方、ゼロクロス率の高い演奏では、音の粗さや摩擦音に近い成分が含まれており、これが音の印象に影響を与え、結果として聴取者の支持を得にくくなる可能性がある。

4.4 スペクトル帯域幅の比較

次に楽曲 1 について演奏者ごとのスペクトル帯域幅の平均値を図 9 に示す。図 9 から若干視聴回数が少ない演奏のスペクトル帯域幅の平均値が高い傾向が見られるが、値の差はあまり大きくないことが分かる。

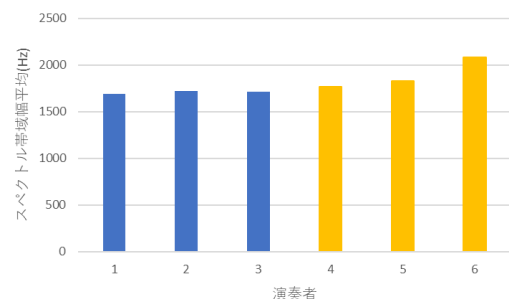


図 9: 楽曲 1 のスペクトル帯域幅の平均値

楽曲 2 について演奏者ごとのスペクトル帯域幅の平均値を図 10 に示す。図 10 から楽曲 1 と同様に視聴回数が少ない演奏にやや高い値が見られたが、全体的な違いは僅かであることが分かる。

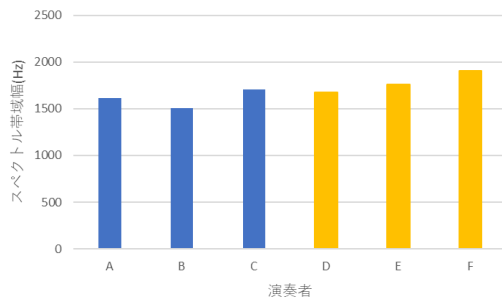


図 10: 楽曲 2 のスペクトル帯域幅の平均値

以上の結果より、スペクトル帯域幅の平均値においては、視聴回数が多寡による顕著な違いは確認できなかった。特定の演奏では高い値を示す場合もあるが、それは個別の演奏や録音環境の影響による可能性も考えられるため、人気の要因としてスペクトル帯域幅が大きく寄与しているとは言い難い。

4.5 RMS の比較

楽曲 1 について演奏者ごとの RMS の標準偏差を図 11 に示す。図 11 から視聴回数が多い演奏の方が RMS の標準偏差が大きいことが分かる。この結果から、視聴回数が多い演奏では一貫して RMS の標準偏差が高く、音量の変化がより豊かである傾向が認められた。

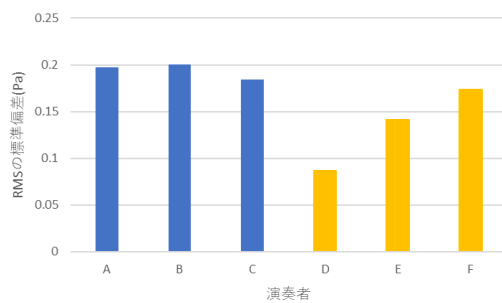


図 11: 楽曲 1 の RMS の標準偏差

楽曲 2 について演奏者ごとの RMS の標準偏差を図 12 に示す。図 12 から視聴回数が多い演奏の方がやや RMS の標準偏差が高いことが分かる。しかし、視聴回数が少ない演奏のうち比較的高い値を示す例もあり、全体としての差は楽曲 1 ほど顕著ではない。

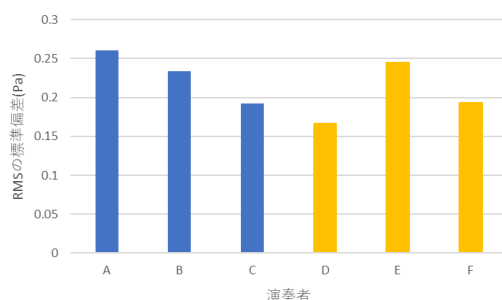


図 12: 楽曲 2 の RMS の標準偏差

これらの結果から、楽曲 1、楽曲 2 ともに、視聴回数が多い演奏は音量の時間的変化が大きい、すなわち演奏表現におけるダイナミクスが豊かである傾向が認められた。これは、聴取者が平坦で単調な演奏よりも、強弱のコントラストを伴う表情豊かな演奏を好む可能性を示唆している。また、楽曲 1 では差が顕著であり、このことから、ダイナミクスの幅が大きい演奏が、より高い視聴回数を得る要因の一つである可能性が高いと考えられる。一方で楽曲 2 では、視聴回数が少ない演奏

にも比較的高い標準偏差を持つ例が含まれており、音量変化の大きさだけでは聴取者の好みに直結しない可能性を示された。演奏の音量変化は、曲調やフレージング、録音環境などの他の要素とも複雑に絡み合うため、今後は RMS の標準偏差を他の音響特徴量と組み合わせて検討することが重要である。

4.6 HNR の比較

楽曲 1 について演奏者ごとの HNR を図 13 に示す。図 13 から視聴回数が多い演奏は、少ない演奏と比較してわずかに高い値を示しており、音の明瞭さに関して顕著な差は見られなかった。

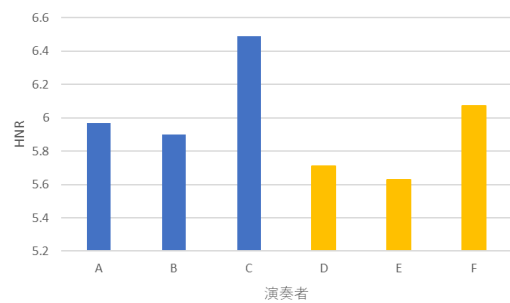


図 13: 楽曲 1 の HNR

楽曲 1 について演奏者ごとの HNR を図 14 に示す。楽曲 2 では視聴回数が多い演奏が、視聴回数が少ない演奏よりも明らかに高い HNR 値を示しており、視聴回数との関連性が明確に表れている。特に、全体的に視聴回数の少ない演奏では HNR が 5 未満に留まっており、音の明瞭度が相対的に低い可能性が考えられる。

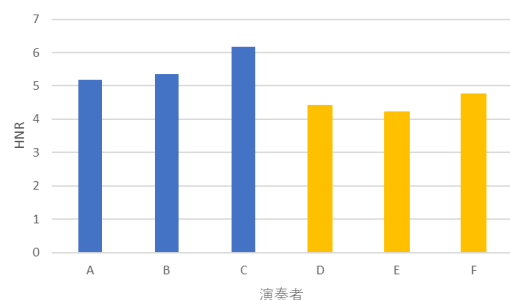


図 14: 楽曲 2 の HNR

以上の結果から、HNR は楽曲の印象や聴きやすさに影響を与え、それが視聴回数の違いにも一定の関係性を持つ可能性があると考えられる。特に、楽曲 2 の結果からは、ノイズの多い演奏が敬遠されている傾向が見て取れ、HNR が聴取者の好みに影響する一つの要因となりうることを示唆された。

4.7 時間的重心の比較

楽曲 1 について演奏者ごとの時間的重心を図 15 に示す。図 15 から視聴回数が多い演奏の方が後ろよりのエネルギー分布であることが分かる。これは、音楽の構成において後半に盛り上がりや印象的なパッセージが集中していた可能性を示唆しており、聴取者の印象に残りやすい構成が視聴回数に影響している可能性がある。

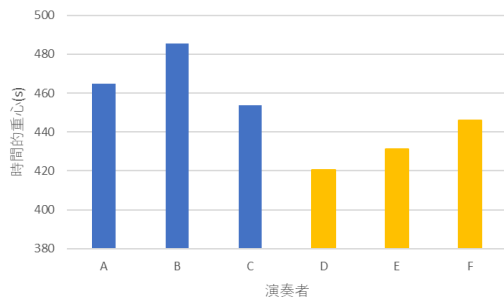


図 15:楽曲 1 の時間的重心

楽曲 2 について演奏者ごとの時間的重心を図 16 に示す。図 16 から視聴回数が多い演奏がおおよそ後半よりであることが分かる。楽曲 1 ほどの差は見られないものの、視聴回数が多い演奏の方が一貫してエネルギーの重心が後方にある傾向が見られた。

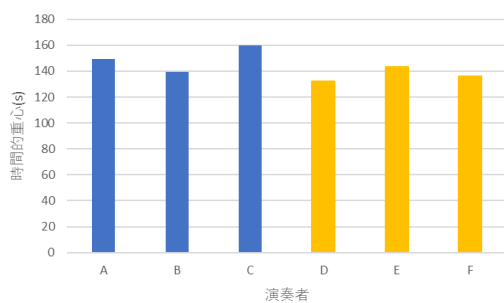


図 16:楽曲 2 の時間的重心

両楽曲において共通して、視聴回数が多い演奏は時間的重心が後半寄りである傾向が見られた。これは、楽曲の広範に向かって盛り上がるような構成や、印象的な要素が配置されていることが、聴取者の興味を持続させ、結果として視聴回数の向上に寄与している可能性を示唆している。

4.8 テンポの比較

楽曲 1 について演奏者ごとの推定テンポを図 17 に示す。図 17 より視聴回数が多い演奏のテンポは、視聴回数が少ない演奏よりやや遅くなっていることが分かる。

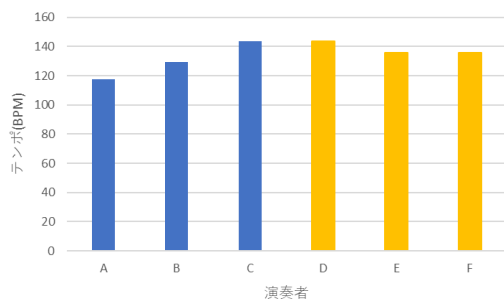


図 17:楽曲 1 の推定テンポ

楽曲 2 について演奏者ごとの推定テンポを図 18 に示す。図 18 より視聴回数が多い演奏は一定のテンポを保っており、視聴回数が少ないはテンポにばらつきがあることが分かる。

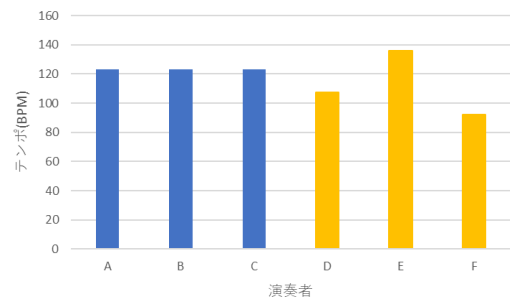


図 18:楽曲 2 の推定テンポ

以上の結果から、視聴回数が多い演奏ではテンポが極端に速すぎず、また一定している傾向があり、テンポの適度な速さや一貫性が聴取者に好まれる要因のひとつであると考えられる。

4.9 フーリエテンポグラムの比較

楽曲 1 について演奏者ごとのフーリエテンポグラムを図 19 に示す。図 19 より視聴回数が多い演奏は、100~150BPM の間に明るい帯が比較的一定に見られる。また、100~150BPM 付近に強調される箇所が多く、全体的に BPM ごとの成分が鮮明である、一方、視聴回数が少ない演奏は明るい帯がややバラバラかつ弱めであり、ブレが大きい結果もある。また、一部で 32~64BPM 付近が強調されており、非常に遅いテンポで表現されている。加えて、やや曖昧でぼやけている部分も含まれている。

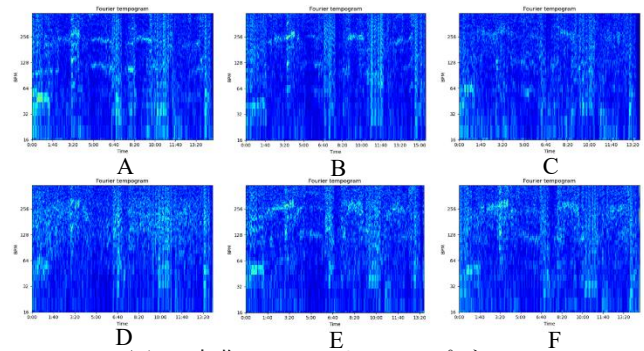


図 19:楽曲 1 のフーリエテンポグラム

楽曲 2 について演奏者ごとのフーリエテンポグラムを図 20 に示す。図 20 より視聴回数が多い演奏は、全体的に明瞭なテンポ帯が見られる。また、3 人とも比較的共通したテンポ傾向があり、時間軸に沿ってテンポの変化が滑らかである。一方で、視聴回数が少ない演奏は、テンポが不安定な傾向があり、演奏者間でばらつきがみられる。また、テンポ構造が複雑または曖昧になっている。

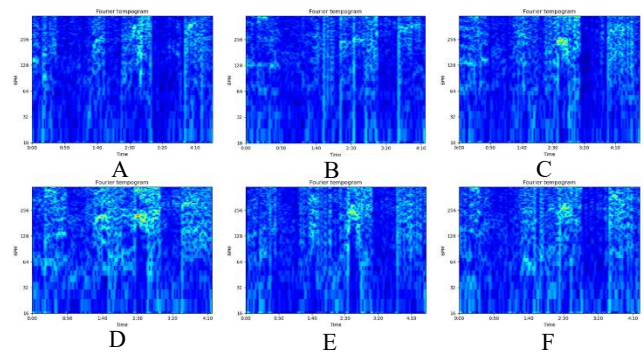


図 20:楽曲 2 のフーリエテンポグラム

以上の結果から、視聴回数が多い演奏では明瞭なテンポ帯が見られ、ピークが鮮明である。これらはテンポが安定し、明

瞭で統一感があるため、聴取者にとって心地よく分かりやすい演奏であった可能性が高い。一方で、視聴回数が少ない演奏はテンポ構造が複雑または曖昧であり、BPM 帯のエネルギー分布がばらけており、明確なピークが薄い。これらは聴取者にとって分かりづらく、魅力が伝わりにくくなる可能性があると考えられる。

4. 10 自己相関テンポグラムの比較

楽曲 1 について演奏者ごとの自己相関テンポグラムを図 21 に示す。図 21 より視聴回数が多い演奏は 3 人とも赤の帯が高い BPM 帯に広く安定して分布していることが分かる。また、BPM の縦方向の広がり比較的狭く、縦縞が細かく均一である。一方で、視聴回数が少ない演奏は BPM の分布のばらつきが多く、縦縞の乱れや途切れがあり、周期性の明瞭さに欠けている。

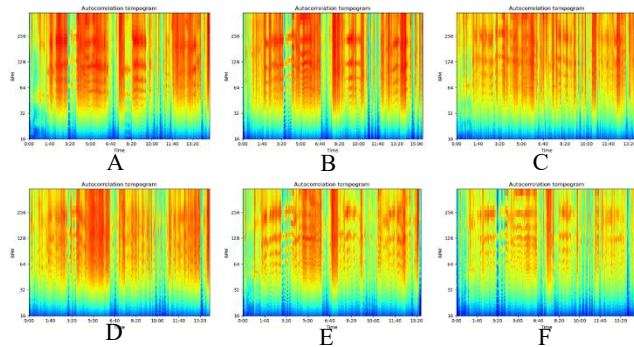


図 21: 楽曲 1 の自己相関テンポグラム

楽曲 2 について演奏者ごとの自己相関テンポグラムを図 22 に示す。図 22 より視聴回数が多い演奏は中高 BPM 帯にわたり赤〜黄の帯が広がっていることが分かる。また縦方向のパターンが整っている。一方で、視聴回数が少ない演奏は、テンポの帯が短く、不連続である。また、BPM 帯ごとのエネルギーがばらついている。

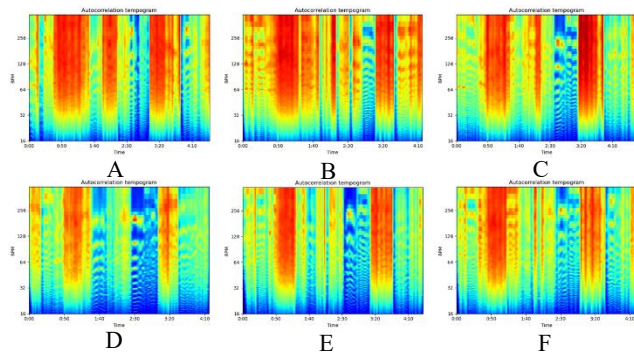


図 22: 楽曲 2 の自己相関テンポグラム

以上の結果から、視聴回数が多い演奏は、BPM の分布が均一であり、テンポに一貫性がある。このことから安定した演奏で、リズム感が良い傾向がある。一方で、視聴回数が少ない演奏はテンポが不安定であり、ばらつきが見られる。このことは聴取者がリズムを感じづらくなると考えられる。

5. まとめ

本研究では、バイオリン独奏における複数の演奏者による同一楽曲の演奏動画を対象とし、動画の視聴回数を聴取行動の指標と位置づけ、音響特徴量との関連性を分析した。具体的には、スペクトル重心、スペクトル平坦度、ゼロクロス率、スペクトル帯域幅、RMS、HNR、時間的重心、テンポ情報など、多様な音響特徴量を用いて定性的・定量的な比較を行った。

その結果、スペクトル重心やスペクトル平坦度、ゼロクロス率、RMS、HNR などの特徴量において、視聴回数の多い演奏と少ない演奏の間に傾向の違いが確認された。特に、視聴回数の多い演奏ではスペクトル重心が低く、倍音構造が明瞭であり、ゼロクロス率が低く滑らかな音質である傾向が見られた。また、RMS の標準偏差が大きく、ダイナミクスの変化が豊かな演奏が多く、聴取者にとって魅力的である可能性が示唆された。さらに、時間的重心やテンポグラムの結果から、演奏後半に印象的な構造を持ち、テンポに一貫性がある演奏が好まれている傾向が明らかになった。

これらの結果は、演奏の音響的な特性と聴取者の選好に一定の関係があることを示すものであり、今後、演奏家が自らの表現を見直す際の客観的な指標としての活用や、教育現場における演奏評価の補助、録音・配信における音源制作の参考としての応用が期待される。

今後の課題としては、分析対象となる楽曲や演奏者の数をさらに増やすことで、得られた傾向の一般化やジャンル間の比較が可能となると考えられる。また、視聴回数のような客観的な行動指標に加え、実際の聴取者に対するアンケート調査を実施し、演奏に対する印象や好みに関する主観の評価との対応関係を検討することで、より包括的な理解につながると期待される。

参考文献

- [1] 牧勝弘, 楽器演奏音の空間放射特性の計測: 奏者の熟練度と名器の特徴. 自動制御連合講演論文集 第 60 回自動制御連合講演会, 8-14, 2017.
- [2] 枝美穂子, フルートの材質による演奏者実感と音響特性の差異に関する研究, Ph.D. thesis, 九州大学, 2019.
- [3] 高橋範行, and 大串健吾, ピアノ演奏における熟達者と非熟達者の演奏表現の比較, 音楽教育学, 34(1), 1-11, 2004.
- [4] 小幡哲史, et al, ストラディバリウスはなぜ高い? バイオリンの歴史とその値段, 研究報告音楽情報科学 (MUS), 2016(12), 1-4, 2016.
- [5] 山田真司, and 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), 440-445, 2014.
- [6] Sergio Giraldo, George Waddell, Ignasi Nou, Ariadna Ortega, Oscar Mayor, Alfonso Perez, Aaron Williamson, and Rafael Ramirez, Automatic Assessment of Tone Quality in Violin Music Performance, frontiers, Volume 10, 2019.
- [7] Yap Kah Yee, and Mafas Raheem, Predicting Music Popularity Using Spotify and YouTube Features, Indian Journal of Science and Technology, Volume 15, 1786-1799, 2022.
- [8] Chaeng-Yuan Lee, and Yi-Ning Tu, Predicting Hit Songs Using Audio and Visual Features, IEEE ICKII 2024, 2025.
- [9] Chia-Jung Tsay, Sight over sound in the judgment of music performance, 2013