

吹奏楽のための音形と音程可視化による管楽器練習支援システム

夜久 佳月 Kazuki Yaku

法政大学情報科学部デジタルメディア学科

kazuki.yaku.8p@stu.hosei.ac.jp

Abstract

Uniformity in performance expression within wind ensembles is crucial for enhancing the overall quality of ensemble playing. However, it is often difficult for individual performers to objectively grasp the differences between their own expression and that of skilled performers. To address this challenge, this study develops a system that visualizes note shapes and pitch from recorded performances, aiming to support the unification of performance expression during practice. The proposed system extracts amplitude envelopes and fundamental frequencies trajectories from recorded audio data. It provides visual feedback on errors in tone shape and pitch by displaying comparison results against teacher data after performance. This visualization enables performers to quantitatively and intuitively understand discrepancies in their own performance. In experiments, performance data using two types of instruments was collected from seven participants (two experts and five novices). The effectiveness was evaluated by comparing performances with and without the system. Results showed improvements in note pattern matching rate from 3.4% to 12% and in average pitch error from 0.7 cent to 4.7 cent. Furthermore, subjective evaluations revealed that users recognized differences between their own performance and that of the experts by checking the system's output, reporting the system as useful for practice. These findings suggest the proposed system has potential as an effective method for supporting the unification of musical expression.

1 はじめに

吹奏楽における合奏では、音程や音量だけでなく、アーティキュレーションやフレーズの流れといった演奏表現を団体内で統一することが求められる。しかし、これらの表現は数値化が難しく、指導者の言葉や模範演奏によって感覚的に伝えられることが多い。また、合奏練習においては、パートあるいは複数パート間の音量や音色のバランス調整に重点が置かれることが多く、個々の演奏者に着目した表現の違いまでは十分に指導することが難しい。そのため、演奏者自身が自分の演奏を客観的に把握することは容易ではなく、表現のばらつきが生じる要因となっている。

特に中学校・高等学校の吹奏楽部においては、音色や発音、音量、フレーズを「揃える」ことが重視され、個人の自由な表現よりも全体としての統一感が美德とされる文化がある。このような傾向は、吹奏楽が学校の部活動として行われることが多く、合奏を通じて協調性や一体感を育成するという教育的側面が重視されていることに起因すると考えられる。

近年、学校の吹奏楽部において、熟練した指導者の不足が課題として指摘されている。吹奏楽経験を有する教員が在籍しない場合、経験の浅い、あるいは管楽器の演奏経験を持たない音楽教員が顧問を務めることも少なくない。そのような環境では、金管楽器や木管楽器に特有の発音や奏法に関する技術的指導を十分に行うことが難しい。

また、指導者が練習指導に十分な時間を割けない状況や、個人練習

が中心となる場面では、演奏表現の統一を図ることはさらに困難となる。このような背景から、演奏者自身が自律的に自らの演奏を振り返り、改善点を把握できる支援手法が求められている。

これまで、音楽演奏を対象とした可視化や支援に関する研究は数多く行われており、音程や音量、タイミングなどを視覚的に提示する手法が提案されてきた。しかし、吹奏楽における演奏表現の統一を目的として、音形と音程の両方に着目した可視化研究は十分に検討されていない。

そこで本研究では、吹奏楽における演奏表現の統一を支援することを目的とし、演奏音の音形および音程を可視化するシステムを開発する。提案システムでは、教師データ（模範演奏）と演奏者の演奏を比較して提示することにより、理想的な表現との差異を視覚的に把握できるようにする。これにより、従来は感覚的にしか伝えられなかった演奏表現を、客観的かつ具体的に評価し、演奏者自身が改善点を認識できる環境の構築を目指す。

2 関連研究

これまで、楽器演奏の練習支援を目的として、音の可視化に関する研究がさまざまな角度から行われてきた。

Park らは演奏中の音楽ダイナミクスを視覚的に提示する音楽ダイナミクス可視化手法（MDVM）を提案している [1]。この研究では、録音済みの演奏音と演奏者がリアルタイムで演奏する音の振幅を同一時間軸上に重ねて表示するインターフェースを開発し、演奏者が自身のダイナミクス変化を直感的に把握できるようにしている。実験の結果、MDVM を用いることで音楽ダイナミクスに関する誤差が大幅に低減し、システムを使用しない条件と比較して約 256% の改善が確認されている。一方で、同手法は個人演奏におけるダイナミクス表現の向上を目的としたものであり、合奏における演奏表現の統一や、音程や音形を含めた総合的な演奏表現を対象としてはいない。

金子らは吹奏楽器演奏者の音程および音形を支援するロングトーン練習支援システム「Pik-kun」を提案している [2]。本システムは、演奏中の音程および音形をリアルタイムに可視化する Web アプリケーションであり、初心者演奏者が自身の演奏を客観的に把握しながらロングトーン練習を行うことを目的としている。理想的な音形や問題のある音形を示したリーフレットを併用することで、熟練指導者が不在の状況においても自己評価を促進する工夫がなされている。実験では、音程や音形に対する意識の向上や演奏の安定化が報告されており、ロングトーン練習支援における有効性が示されている。一方で、本システムはロングトーンという基礎練習に着目した支援手法であり、単音を持続的に演奏する状況を主な対象としている。

以上の関連研究より、楽器演奏の可視化に関する研究は、個人演奏や基礎練習を対象とし、特定の演奏要素の改善を目的としたものが中心であることが分かる。一方、吹奏楽において重視されるフレーズ全体の演奏表現を対象とし、演奏者が自身の表現を振り返ることで、結果としてバンド全体の演奏表現の統一につなげることを目的とした支援手法は十分に検討されていない。本研究では、音形および音程を可視化し、教師データとの比較を通じて演奏者が表現の違いを客観的に把握できる環境を構築することで、個人練習を通じた演奏表現の統一支援を目指す。

3 演奏音可視化システム

吹奏楽における演奏表現の統一を支援することを目的として、演奏音可視化システムを提案する。本システムは、合奏中に使用することを想定したものではなく、個人練習や演奏後の振り返りの場面での活用を想定して設計されている。演奏後に自身の演奏を可視化結果として確認することで、演奏中には意識しきれなかった音量変化や音程のずれ、フレーズ全体の表現の流れを客観的に捉えることを可能とする。

演奏表現の統一というのは、単に音量を大きくする／小さくするといった指示を共有するだけでは達成できない。例えば、クレッシェンドという同じ表現であっても、「どのタイミングから音量を上げ始めるか」「どの程度の幅で増加させるか」「ピークをどこに置かか」などは演奏者によって差が生じやすい。図1は、クレッシェンドの音形の違いを例として示したものである。青色の例ではアタック後に一度音量を下げてから大きくしているのに対し、赤色の例ではアタック後の音量をほとんど変化させず、小さい幅で音量を増加させている。このように、同じ意図を持って演奏している場合でも音量変化の形状は一致しないことがあり、演奏表現の統一が容易ではないことが分かる。

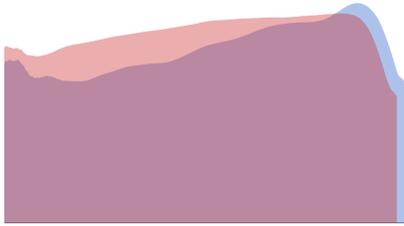


図1. クレッシェンドの幅の違いの例

このように表現の統一には多くの困難が伴うが、その重要性は大きい。例えば、30人中29人が同じように演奏していても、1人だけが音を長く演奏してしまったり、音程がずれてしまったりすると、観客には全体がずれて聴こえてしまう。また、音量変化についても、各自がバラバラの音量で演奏すれば、曲の表現は不明瞭になってしまうが、逆に全員で音量変化を揃えることで、迫力や緊張感を観客に鮮明に伝えることができる。さらに、ひとりひとりの表現を統一させることで、まとまりのあるサウンドになり、主旋律が聞き取りやすくなる。本システムでは、個人練習の質を高めることで、最終的には合奏全体の演奏表現の統一に寄与することを期待している。

音は音量・音程・音色の三要素に分けられるが、本研究では音量および音程の二要素を対象とする。これらは定量的に測定可能であり、演奏者間で客観的に比較・評価できる利点がある。一方、音色は感覚的側面が強く、可視化を行っても演奏者間で共通理解を得ることが難しい。また、音色は表現の最終段階で調整される要素であり、練習の初期段階では音量や音程の調整がより重要である。

一方で、可視化システムを使用するとそのシステムに頼ってしまい、耳による判断力が育たなくなるという懸念点がある。例えば、聴覚に基づいて音程のずれを判断する能力が低下する可能性がある。ただし、今回はリアルタイムの表示ではないため、演奏中に意識することはシステムを使用しない場合と変わらない。演奏後に、まず自身の聴覚による振り返りを行った上で可視化結果を確認することで、耳で感じた印象と可視化結果を照合でき、自己評価の精度向上にもつながると考えられる。

以下では、本システムにおける音形および音程の可視化方法について述べ、教師データとユーザーデータを比較提示する表示形式について説明する。

3.1 音形

吹奏楽における演奏表現の統一では、フレーズ全体の強弱だけでなく、各音の立ち上がりの鋭さや音の長さ、減衰の仕方などが揃っていることが重要となる。しかし、これらの要素は演奏者自身が演奏中に客観的に把握することが難しく、指導者の主観的な指摘に依存しやすい。そこで本システムでは、演奏後に音形を可視化することで、演奏者が自分の発音の特徴を振り返りやすくすることを目的とする。

音形の可視化にあたり、教師データおよびユーザーデータの演奏音をそれぞれ録音し、同一時間軸上に重ねて表示する。本システムでは、教師データを青色、ユーザーデータを赤色として表示することで、両者の違いを直感的に比較できるようにした。

一般に、録音された音声波形は高周波成分を含むため瞬間的な振幅変化が大きく、波形そのものをそのまま表示しても音量変化の傾向を把握しにくい。特に管楽器演奏では、波形の細かな周期的変動（音の周期構造）よりも、音量が時間的にどのように変化しているかが演奏表現の理解に直結する。そのため、音声信号の包絡線（エンベロープ）を算出し、音形として表示する。

エンベロープとは、信号振幅の外形を滑らかに近似した曲線であり、演奏音の強弱変化を時間的に表現するために用いられる。これにより、演奏者が意識的に制御していると考えられる音量変化を視覚的に捉えやすくなる。また、アタックの鋭さ、コアの安定性、リリースの長さといった発音の特徴も、エンベロープの形状として表れやすい。

エンベロープの算出手順としては、まず録音波形 $y[n]$ に対して絶対値処理を行い、負方向の振幅を正方向に変換する。これは、演奏者が実際に制御しているのは波形の正負ではなく振幅の大きさであり、音形の比較においては正負を区別する必要が小さいためである。次に、絶対値波形に対して移動平均を適用し、時間方向に平滑化することでエンベロープを得る。さらに、平滑化後の値に対して対数変換を行った。移動平均の窓幅（フレーム長）を W とすると、エンベロープ $e[n]$ は式(1)で表される。

$$e[n] = \log \left(\frac{1}{W} \sum_{k=0}^{W-1} |y[n-k]| + \varepsilon \right) \quad (1)$$

ここで窓幅 W は、音形の滑らかさを決定する重要なパラメータである。窓幅を小さくすると細かな振幅変化まで追従できる一方で、ノイズや局所的な揺れが強調される可能性がある。逆に窓幅を大きくすると大域的な音量変化を捉えやすくなるが、アタックなどの急激な変化が丸められる可能性がある。本研究では、アタック・コア・リリースの特徴を損なわずにフレーズ全体の流れを比較できるよう、窓幅を 80ms に設定した。

図2にフレーズ全体の音形可視化結果を示す。結果を見ると、アタックのタイミングは概ね一致している一方で、教師データはマルカートに近い明瞭な発音となっており、ユーザーデータはテヌートのように音を保った形状になっていることが確認できる。このように、フレーズ全体としては同じリズムで演奏していても、各音の音形の違いが積み重なることで、演奏表現の印象が変化する。本システムでは、演奏後にこのずれを確認し、意図した表現に合わせて再度演奏するという練習サイクルを支援する。

3.2 音程

音程の表示には周波数そのものではなく、セント（cent）を用いる。セントは吹奏楽の現場で一般的に用いられている単位であるほか、周波数の比を対数で表す尺度であり、人間の音高知覚に近い表現となる。基準周波数を $f_{ref} = 442 \text{ Hz}$ (A4) とすると、周波数 f をセントへ変換する式は式(2)で表される。

$$cents = 1200 \cdot \log_2 \left(\frac{f}{f_{ref}} \right) \quad (2)$$

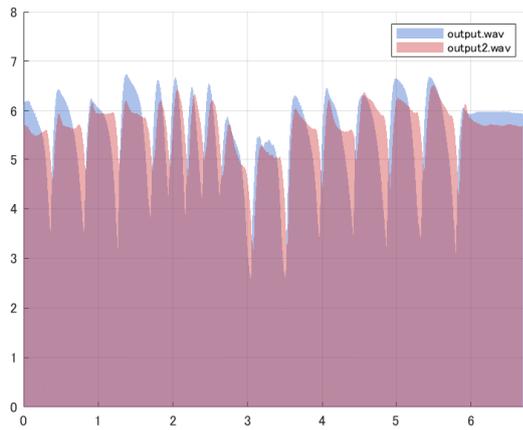


図 2. フレーズの音形可視化

ここで、1200 セントは 1 オクターブ（周波数が 2 倍）に対応し、100 セントは半音に相当する。この尺度を用いることで、異なる音高における音程誤差を統一的に扱うことができ、教師データとユーザーデータの比較が容易となる。

音程の可視化は、目的に応じて 2 種類の表示方式を採用した。1 つ目は、図 3 に示すフレーズ全体での音程表示である。この表示では、演奏全体の音程の推移を時間軸上に示し、音程が大きく外れている箇所を一目で確認できる。個々の音を詳細に把握することは難しいものの、フレーズ全体の傾向を大まかに捉えることが可能であり、重点的に練習すべき音や不安定な区間を把握する手がかりとなる。

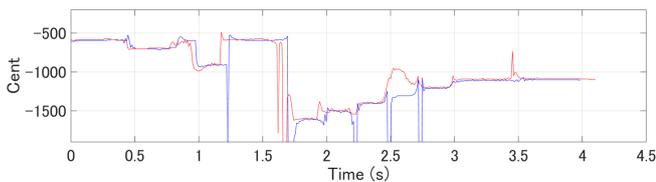


図 3. フレーズ全体の音程表示 (赤: ユーザ, 青: 教師)

2 つ目は、時間の進行に合わせて表示が追従するアニメーション表示である。フレーズ全体の表示では見落としやすい短時間の音程変動や、発音直後の揺れ、音の終止における下がりを確認しやすくするため、時間の進行に合わせて表示が追従する形式とした。各音を追従するように表示することで、音程のずれがどの程度生じているかを直感的に把握でき、演奏中の音程制御の安定性を確認できる。図 4 に、アニメーション表示の一部を示す。

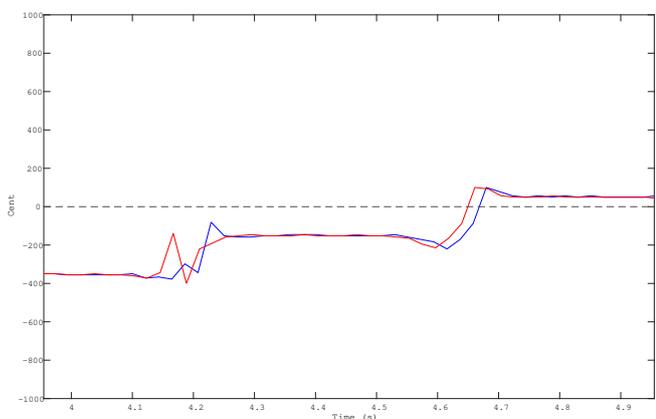


図 4. 音程のアニメーション表示 (赤: ユーザ, 青: 教師)

以上の 2 種類の表示を組み合わせることで、演奏者は「フレーズ全体の傾向」と「瞬間的なずれ」をそれぞれの観点から確認できる。これにより、演奏中には把握しづらい音程の課題を可視化し、次の演奏に向けた改善点を明確化できると考えられる。

3.3 再生機能

本システムでは、可視化結果と聴感の対応づけを支援するため、録音音声の再生機能を実装した。再生方式は「教師のみ」「ユーザーのみ」「教師とユーザー同時」の 3 種類とし、それぞれ再生ボタンによって切り替え可能とした。教師のみの再生では模範演奏の表現を確認でき、ユーザーのみの再生では自身の演奏を客観的に振り返ることができる。また、教師とユーザーの同時再生では、音量変化や音程のずれ、アーティキュレーションの違いを聴感上でも比較しやすくなり、可視化結果と実際の聴こえ方の対応を理解する助けとなる。

4 評価

4.1 内容

提案システムが演奏者の表現統一に効果的であるかを検証するため、実験を行った。被験者は吹奏楽経験を有する大学生 5 名とした。使用楽器はユーフォニアムおよびトロンボーンであり、ユーフォニアム奏者 3 名、トロンボーン奏者 2 名である。また、比較の基準となる演奏データとして、ユーフォニアム奏者 1 名およびトロンボーン奏者 1 名が教師データを収録した。ここで、教師とは基準となる演奏データを提供する奏者を指し、ユーザーとはシステムの評価対象者を指す。

実験は、システムを使用しない条件と使用する条件の 2 条件で実施した。各条件において、被験者は異なる 5 種類のフレーズをそれぞれ演奏し、各フレーズにつき 5 回の反復練習を行った。システム使用条件では、被験者は可視化結果を確認しながら練習を行った。一方、システム未使用条件では、可視化結果は提示せず、録音した演奏音のみを被験者に提供した。全被験者は、本研究の目的および実験内容について事前に説明を受け、同意を得た上で実験に参加した。

演奏音はコンデンサマイクを用いて収録し、オーディオインターフェースを介して PC に入力した。サンプリング周波数は 48 kHz とし、モノラルの WAV 形式で録音した。使用したオーディオインターフェースは Roland 製 UA-25EX である。マイクは楽器のベルから約 15 cm の位置に設置し、すべての被験者において同一条件とした。なお、演奏時にはテンポを一定に保つため、被験者に片耳のみイヤホン装着させ、メトロノーム音を聴取させた。

4.2 評価方法

音形の評価は、フレーズ区間におけるエンベロープ差の平均値を算出した。また、音形一致率はエンベロープの差が 0.1 以下となるフレームの割合を求め、表示した。

音程に関しては、平均音程差 (cent) と ± 20 cent 以内の割合 (%) を表示した。音程差は、教師データおよびユーザーデータの基本周波数から算出したセント値の差として求めた。教師のセント値を $cents_{teacher}$ 、ユーザーのセント値を $cents_{user}$ とすると、音程差は式 (3) で表される。

$$\Delta cents = |cents_{user} - cents_{teacher}| \quad (3)$$

この際、基本周波数推定の誤りや音ミス（別の音の発音）による外れ値が含まれる可能性があるため、音程差が ± 200 cent を超える（1 音以上）となる区間は評価から除外した。

また、実験後にアンケートを実施し、音程や音形の違いの認識、および表現統一への効果について主観評価を収集した。内容は表 1 に示す。

4.3 結果

4.3.1 システム使用例

トロンボーン奏者の一例を図 5 から図 8 に示す。青線が教師データ、赤線がユーザーデータであり、1 回目と 5 回目の音形と音程の可視化結果をそれぞれ示している。

まず音形について、図 5 では教師データと比較してユーザーデータのエンベロープ形状に差が見られ、フレーズ全体の音量変化の傾向が一

表 1: アンケート項目

| 番号 | 内容 |
|----|------------------------------|
| Q1 | 出力結果を見て熟練者とあなたの演奏の違いは分かりましたか |
| Q2 | 違いが分かった場合、それはどのような点でしたか |
| Q3 | このシステムは自分の演奏の改善に役立ちましたか |
| Q4 | 練習に取り入れるとしたらどのように活用したいと思いますか |

致していない。一方、図 6 では、1 回目と比較して教師データとの重なりが増加しており、音量変化の推移が教師データに近づく傾向が見られた。また、5 回の練習を経て、2 小節目 4 拍目と 3 小節目 1 拍目の間をつなげる表現やテンポのずれが改善し、教師データに近づいた。

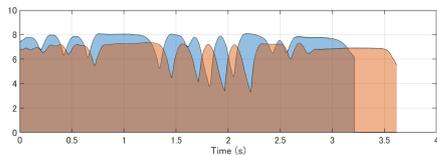


図 5. 1 回目の音形比較 (赤: ユーザ、青: 教師)

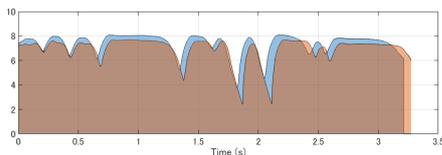


図 6. 5 回目の音形比較 (赤: ユーザ、青: 教師)

次に音程について、図 7 および図 8 に音程推移の比較結果を示す。1 回目では Eb および Ab において教師データより低い傾向が見られた。5 回目では、1 回目と比較して一部の音における教師データとの差が縮小していることが確認できる。

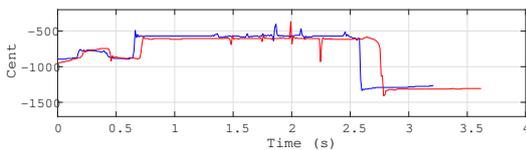


図 7. 1 回目の音程比較 (赤: ユーザ、青: 教師)

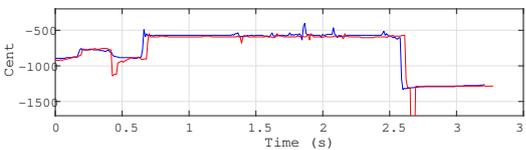


図 8. 5 回目の音程比較 (赤: ユーザ、青: 教師)

4.3.2 音形指標の推移

トロンボーン奏者の 1 回目から 5 回目までの音形指標の推移を図 9 および図 10 に示す。平均エンベロープ差は 1 回目から 2 回目にかけて変化が見られた一方で、2 回目以降は変化量が小さく、大きな変化が見られない傾向が確認された。同様に、音形一致率についても初期段階で改善があったが、2 回目以降は大きな変化が見られなかった。

4.3.3 各評価指標結果

表 2 に各評価指標の結果を示す。

平均エンベロープ差はシステムを使用しない条件では -0.006 、使用した条件では -0.0367 であった。音形一致率はシステムを使用しない条件で 3.4% 、使用した条件で 12.2% であった。

平均音程差はシステムを使用しない条件で -0.7cent 、使用した条件

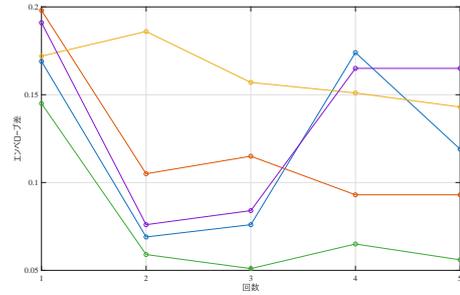


図 9. トロンボーン奏者における平均エンベロープ差の推移 (1 回目~5 回目)

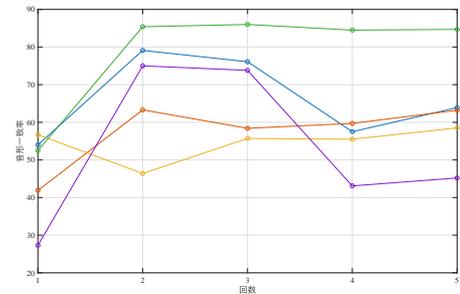


図 10. トロンボーン奏者における音形一致率の推移 (1 回目~5 回目)

表 2: 可視化システム使用有無による演奏評価結果の比較

| | 指標 | 1 回目 | 5 回目 | 改善度 |
|-------|-----------------------------|-------|--------|----------|
| システム無 | 平均エンベロープ差 | 0.123 | 0.117 | -0.006 |
| | 音形一致率 (%) | 61.4 | 64.8 | 3.4 |
| | 平均音程差 (cent) | 35.2 | 34.5 | -0.7 |
| | $\pm 20\text{ cent}$ 以内 (%) | 48.7 | 48.7 | 0.0 |
| システム有 | 平均エンベロープ差 | 0.134 | 0.0974 | -0.037 |
| | 音形一致率 (%) | 57.0 | 69.2 | 12.2 |
| | 平均音程差 (cent) | 34.8 | 30.1 | -4.7 |
| | $\pm 20\text{ cent}$ 以内 (%) | 48.3 | 50.6 | 2.3 |

で -4.7cent であった。また、 $\pm 20\text{ cent}$ 以内に収まる割合の改善量は、システムを使用しない条件で 0.0% 、使用した条件で 2.3% であった。

主観評価では、システムの出力結果を確認することで、熟練者との演奏の違いを理解できたとの意見が得られた。特に、アクセントの強さや音の長さ、音程がずれるタイミングを視覚的に把握できた点が有用であると評価された。また、指導者が不在となる自主練習時において、本システムを演奏改善の支援ツールとして活用したいという意見を得た。

4.3.4 演奏経験別の結果

演奏経験の違いによる改善傾向を確認するため、初心者群 (ユーフォニアム 2 名) と経験者群 (ユーフォニアム 1 名、トロンボーン 2 名) で比較を行った。結果を表 3 および表 4 に示す。システム使用時における音形指標では、経験者群の方が初心者群より改善量が大きい傾向が確認された。音形一致率の改善度は、初心者群で 8.1% 、経験者群で 14.9% であった。また、平均エンベロープ差の改善度は、初心者群で -0.015 、経験者群で -0.052 であった。

一方、音程指標では、平均音程差の改善度は初心者群で -3.4cent 、経験者群で -5.6cent であり、いずれも改善傾向が見られた。さらに、 $\pm 20\text{ cent}$ 以内に収まる割合の改善度は、初心者群で 0.9% 、経験者群で 3.3% であった。

4.3.5 楽器別の結果

表 5 および表 6 に、楽器別の評価結果を示す。ユーフォニアムでは、システム無条件において音形一致率は 64.4% から 66.5% へ増加した一方、システム有条件では 60.6% から 69.2% へ増加し、改善度は 8.6 ポイントであった。また、平均エンベロープ差はシステム有条件において 0.115 から 0.0987 へ減少した。

表3: 初心者(ユーフォニアム2名)の評価結果

| | 指標 | 1回目 | 5回目 | 改善度 |
|-------|--------------|-------|-------|--------|
| システム無 | 平均エンベロープ差 | 0.122 | 0.122 | 0.000 |
| | 音形一致率(%) | 61.3 | 61.7 | 0.4 |
| | 平均音程差(cent) | 30.3 | 31.1 | 0.8 |
| | ±20cent以内(%) | 60.6 | 57.9 | -2.7 |
| システム有 | 平均エンベロープ差 | 0.116 | 0.101 | -0.015 |
| | 音形一致率(%) | 59.8 | 67.9 | 8.1 |
| | 平均音程差(cent) | 29.7 | 26.3 | -3.4 |
| | ±20cent以内(%) | 57.1 | 58.0 | 0.9 |

表4: 経験者(ユーフォニアム1名, トロンボーン2名)の評価結果

| | 指標 | 1回目 | 5回目 | 改善度 |
|-------|--------------|-------|--------|--------|
| システム無 | 平均エンベロープ差 | 0.123 | 0.113 | -0.010 |
| | 音形一致率(%) | 61.6 | 66.9 | 5.3 |
| | 平均音程差(cent) | 38.5 | 36.8 | -1.7 |
| | ±20cent以内(%) | 40.7 | 42.6 | 1.9 |
| システム有 | 平均エンベロープ差 | 0.147 | 0.0949 | -0.052 |
| | 音形一致率(%) | 55.2 | 70.1 | 14.9 |
| | 平均音程差(cent) | 38.3 | 32.7 | -5.6 |
| | ±20cent以内(%) | 42.4 | 45.7 | 3.3 |

トロンボーンでは、システム無条件において音形一致率は57.0%から62.2%へ増加した一方、システム有条件では51.6%から69.3%へ増加し、改善度は17.7ポイントであった。さらに、平均エンベロープ差はシステム有条件において0.163から0.0954へ減少した。

音程に関しては、ユーフォニアムの平均音程差はシステム有条件で33.3centから28.8centへ減少し、±20cent以内の割合は54.6%から56.5%へ増加した。トロンボーンにおいても、平均音程差はシステム有条件で37.1centから32.1centへ減少し、±20cent以内の割合は38.9%から41.2%へ増加した。

表5: ユーフォニアムの評価結果

| | 指標 | 1回目 | 5回目 | 改善度 |
|-------|--------------|-------|--------|--------|
| システム無 | 平均エンベロープ差 | 0.112 | 0.110 | -0.002 |
| | 音形一致率(%) | 64.4 | 66.5 | 2.1 |
| | 平均音程差(cent) | 33.4 | 33.0 | -0.4 |
| | ±20cent以内(%) | 54.4 | 55.0 | 0.6 |
| システム有 | 平均エンベロープ差 | 0.115 | 0.0987 | -0.016 |
| | 音形一致率(%) | 60.6 | 69.2 | 8.6 |
| | 平均音程差(cent) | 33.3 | 28.8 | -4.5 |
| | ±20cent以内(%) | 54.6 | 56.5 | 1.9 |

表6: トロンボーンの評価結果

| | 指標 | 1回目 | 5回目 | 改善度 |
|-------|--------------|-------|--------|--------|
| システム無 | 平均エンベロープ差 | 0.139 | 0.127 | -0.012 |
| | 音形一致率(%) | 57.0 | 62.2 | 5.2 |
| | 平均音程差(cent) | 37.9 | 36.8 | -1.1 |
| | ±20cent以内(%) | 40.1 | 39.3 | -0.8 |
| システム有 | 平均エンベロープ差 | 0.163 | 0.0954 | -0.068 |
| | 音形一致率(%) | 51.6 | 69.3 | 17.7 |
| | 平均音程差(cent) | 37.1 | 32.1 | -5.0 |
| | ±20cent以内(%) | 38.9 | 41.2 | 2.3 |

4.3.6 長期的な練習

トロンボーン奏者1名を対象にシステムを用いて、20回の練習を行った。表7に、1,5,10,15,20回目の評価指標の実データを示す。ま

た、表8に1回目からの改善度をそれぞれ示す。ここで、1は平均エンベロープ差、2は音形一致率(%), 3は平均音程差(cent), 4は±20cent以内(%)とする。

平均エンベロープ差は1回目の0.170から20回目には0.099へと減少し、練習を重ねることで音形の差が縮まったことが示されている。音形一致率は1回目の50.7%から20回目の68.4%へと改善した。また、±20cent以内の割合も49.0%から61.0%へと向上している。

音程の平均音程差は全体としては緩やかな改善傾向が見られ、31.3centから27.3centへ減少した。

表8に示す改善度では、特に音形一致率が最大18.4ポイント向上するなど、練習回数の増加に伴い着実な改善が確認できる。

表7: 長期的な練習の評価結果

| 指標 | 1回目 | 5回目 | 10回目 | 15回目 | 20回目 |
|----|-------|-------|-------|-------|-------|
| 1 | 0.170 | 0.121 | 0.123 | 0.089 | 0.099 |
| 2 | 50.7 | 65.8 | 63.2 | 69.1 | 68.4 |
| 3 | 31.3 | 31.5 | 32.6 | 27.9 | 27.3 |
| 4 | 49.0 | 51.7 | 52.9 | 58.7 | 61.0 |

表8: 長期的な練習の改善度

| 指標 | 5回目 | 10回目 | 15回目 | 20回目 |
|----|--------|--------|--------|--------|
| 1 | -0.049 | -0.047 | -0.081 | -0.071 |
| 2 | 15.1 | 12.5 | 18.4 | 17.7 |
| 3 | 0.2 | 1.3 | -3.4 | -4.0 |
| 4 | 2.7 | 3.9 | 9.7 | 12.0 |

また、音程の2つの指標に関して、1回目から20回目の散布図および近似曲線を図11および図12に示す。どちらも徐々に改善傾向にはあるが、わずかな差であり、本システムでは20回の練習を重ねても有効であるとは言えない。

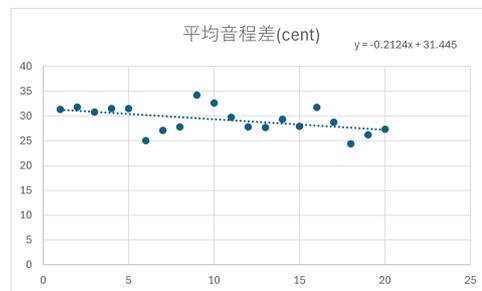


図11. 平均音程差の散布図および近似曲線

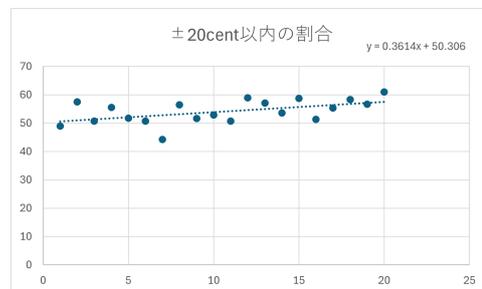


図12. ±20cent以内の割合の散布図および近似曲線

4.4 考察

吹奏楽の練習において求められる到達目標は、熟練者と完全に同一の演奏を実現することではなく、演奏者自身が自分の演奏表現のずれを認識し、それを修正できる状態に到達することであると考えられる。実際の吹奏楽の現場では、音程や音量、フレーズといった表現要素は常に変動しており、熟練者であっても毎回全く同じように演奏す

ることは困難である。本研究で用いた評価指標の数値がどの程度変動し得るかを把握するため、ユーフォニアムの熟練奏者が複数のフレーズをそれぞれ2回ずつ演奏した際の比較結果を算出した。これらの結果を平均したところ、平均エンベロープ差は約0.0440、音形一致率は約90.0%、平均音程差は約15.1cent、±20cent以内の割合は約84.4%であった。このことから、熟練奏者であっても、演奏フレーズの違いにより演奏表現には一定のばらつきが生じることが確認された。そのため、表現統一の本質は、演奏者が理想的な演奏との違いを理解し、合奏の中で適切に調整できる能力を身につけることにある。

本研究の結果からは、熟練奏者のずれと比較すると、音形の改善には大きな効果が見られた一方で、音程については改善傾向は見られたものの、明確な改善効果を示すには至らなかった。音形は、息の量やリズムなど比較的意識的に調整しやすい要素であるため、視覚的なフィードバックによって演奏者が自己修正を行いやすく、大きな改善につながったと考えられる。

一方で、音程の改善は一音ごとに口腔内の大きさや形、息のスピードを微細に調整する必要があり、演奏者が十分に習得・改善するのは難しかった可能性がある。20回の繰り返し練習においては、平均音程差はやや変動を繰り返しつつも徐々に減少し、±20cent以内に収まる割合も緩やかに増加している。しかし、複数回の練習を行ったとしてもわずかな改善しか見られなかったため、システムの機能を見直す必要がある。本研究では演奏後に音程推移を可視化し、次の演奏で修正を試みる形式であったが、音程は瞬間的な調整に依存する要素であるため、演奏後に結果を確認しても演奏時の感覚を再現しながら修正することが難しい。このことから、音程の改善には演奏中にずれを即時に把握できるリアルタイム可視化が有効であった可能性がある。

初心者群では、システム使用時に改善傾向は確認されたものの、経験者群と比較すると改善量が小さい傾向が見られた。これは初心者の場合、ある要素（音量やリズムなど）を意識して改善しようとする、他の要素（音程や音のつながりなど）が崩れてしまうことが多く、複数の表現要素を同時に安定させることが難しいためであると考えられる。

経験者群では、初心者群と比較してシステム使用時の改善度が大きい傾向が見られた。これは経験者の場合、熟練者の演奏に近づけるためにどのような点を調整すべきかある程度理解しており、可視化結果から得られる差分情報を演奏に反映しやすいためであると考えられる。

また、録音のみを提示したシステム未使用条件においても一定の改善は確認されたが、システム使用条件では改善度がより大きくなる傾向が見られた。このことから、本システムによる視覚的フィードバックは、経験者に対しても演奏改善を促進する有効な支援手段となる可能性が示された。

さらに、楽器による違いについても検討したところ、トロンボーンではユーフォニアムと比較して、システム使用時の音形一致率の改善度が大きい傾向が確認された。これは、トロンボーンがスライド操作によって音高や音の立ち上がりのタイミングが変化しやすく、演奏中のわずかな息の使い方やアタックの差が音形に反映されやすいためであると考えられる。その結果、可視化によって得られる「教師データとの差分」が演奏者にとって理解しやすく、修正方針を立てやすかった可能性がある。

一方、ユーフォニアムでは音形一致率の改善は見られたものの、トロンボーンほど大きな改善には至らなかった。これは、ユーフォニアムは比較的滑らかで安定した音量変化を出しやすく、初期段階から音形が大きく崩れにくいことが影響した可能性がある。そのため、音形可視化による差分が小さくなりやすく、改善量としては相対的に小さく表れたと考えられる。

また、音程指標に関しても両楽器で改善傾向が見られたが、いずれも短期間で大きな改善を得ることは難しかった。特にトロンボーンはスライド位置の微小なずれが音程に直結するため、視覚的フィードバ

ックのみでは安定化に時間を要する可能性がある。

なお、ユーフォニアム奏者には初心者者が2名含まれていたため、経験差の影響が結果に含まれる可能性がある。

本システムの有用性として、従来の吹奏楽指導で多用される言語的・感覚的な表現を、演奏者が直感的かつ定量的に理解できる形へ変換できた点が挙げられる。例えば、「小節線を超えるとときにエネルギーをもつ」「停滞しない」といった指導は、演奏者の解釈に依存しやすく、改善すべき箇所を具体的に特定することが難しい場合がある。一方、本システムでは、フレーズのつながりや音量の落ち込みがエンベロープ形状として可視化されるため、演奏者は「2小節目4拍目から3小節目1拍目をつなげる」「最後の音が小さくならないよう維持する」など、具体的な修正方針として捉えやすくなる。このことから、本システムは演奏者の自己認識を促し、自主練習における改善点の発見を支援する可能性が示された。

以上より、本研究で提案した可視化システムは、音形を中心とした演奏表現の改善を支援する有効な手段であると考えられるが、音程のような細かな表現の統一を支援するためには、さらなる機能拡充が必要である。

5 おわりに

本研究では、吹奏楽演奏における表現の統一を支援することを目的として、演奏音を可視化するシステムを開発した。実験結果から、音形についてはシステム使用時に改善傾向が確認されたものの、改善は主に初期段階に見られ、2回目以降は大きな変化が見られなかった。一方、音程に関してはシステム使用による明確な効果は見られなかった。

本システムは、指導者が不在となる自主練習の場面においても、演奏者の自己認識を促し、演奏改善を支援するツールとして有効である可能性が示された。

今後の展望としては、音程の許容範囲（帯）を可視化する手法の導入や、リアルタイムでの教師演奏とのcent差表示を追加することで、より即時かつ具体的なフィードバックを実現し、演奏技術の向上を支援することが期待される。

参考文献

- [1] Eun Ji Park. Music dynamics visualization for music practice and education. *Multimedia Tools and Applications*, January 2025. Published online: 31 January 2025.
- [2] D Kaneko, H Kunimune, M Kurayama, T Morishita, T Yamamoto, and H Oguchi. The effects of a long-tone exercise support system on wind instrument players' pitch and tone shape. *Proceedings of EdMedia + Innovate Learning*, pp. 1182–1187, 6 2019.