

9軸慣性センサを搭載したグローブ型デバイスを用いた指揮動作の認識と動的な音響操作

坪田 泰知^{1,a)} 高田 峻介¹

概要：オーケストラの指揮者は、指揮ジェスチャを通じて演奏のテンポや音量といった音楽的要素を伝達している。指揮動作に基づく音響操作は、身体動作に内在する周期性や運動量を音楽制御へ反映可能な入力手法として検討されてきた。一方、既存手法の多くは、複数のセンサや設置環境に依存した構成を前提としており、使用環境や実装規模の面で課題が残されている。本研究にて、手の甲に9軸慣性センサを搭載した小型グローブ型デバイスを装着し、指揮ジェスチャを計測し、演奏のテンポや音量を制御する手法を提案する。検証実験では、一定テンポ下での指揮動作を用いた評価により、テンポ検出誤差が概ね1BPM以内に収まること、振りの大きさに応じて運動強度指標が段階的に変化することを示し、テンポへの適用可能性を確認した。

1. はじめに

オーケストラにおいて指揮者は、指揮棒を用いた指揮動作によって、拍の周期や振幅などを通じてテンポや音量に関する指示を奏者に伝達している。

これまでに、指揮動作を用いた音響制御に関する研究が提案されてきた [1], [2]。しかし、外部カメラや複数のセンサを必要とする構成が多く、設置や調整を要する点で装着性や可搬性に課題が残されている。

本研究では、9軸慣性センサのみを搭載した小型グローブ型デバイスを用い、指揮動作時のセンサ値から腕や手を振る拍と振幅を計測し、音楽のテンポと音量を制御する手法を提案する。図1に、本研究における指揮動作と音響パラメータの対応関係の概念図を示す。なお、本対応は指揮動作全体を網羅的にモデル化するものではなく、本研究で扱う音響操作（テンポ・音量）に必要な要素に焦点を当てた設計である。



図1 本研究における指揮動作と音響パラメータの対応関係

単一のウェアラブルデバイスのみで動作する構成とする

¹ 神戸市立工業高等専門学校 電子工学科

^{a)} r121320@g.kobe-kosen.ac.jp

ことで、外部機器に依存しない音響操作インタフェースの実現を目指す。

2. 関連研究

本章では、指揮ジェスチャを用いた音響制御手法と、それを実現するための計測・処理手法に着目し、関連研究を述べる。

2.1 指揮動作を用いた音響制御手法

指揮動作を音響制御に利用する試みは、音楽インタフェース分野においてこれまでに数多く報告されている [1-4]。また、音楽における身体動作は、単なる入力手段ではなく、音響的意味や表現と密接に結び付いたジェスチャとして捉えられることが指摘されている [5]。

一方で、これらの手法の多くは、特定の指揮動作や特徴的な運動パターンをあらかじめ定義し、それらを基にしきい値判定や状態遷移を用いて音響操作を実行する構成を採っている。このような手法では、ユーザごとの振幅や速度といった動作特性の違いが、操作の安定性や再現性に影響する可能性がある。

2.2 小型・ウェアラブルデバイスによる音楽ジェスチャ入力

Digital Baton [1] は、3軸加速度センサや圧電抵抗式トリップ、赤外線センサなどを用いた指揮型音楽インタフェースである。原田らは指揮棒を用いた演奏制御システムを提案しており、指揮動作に基づく音響制御の有効性

を報告している [2].

一方、これらの先行研究では、赤外線センサや外部設置機器などを併用する構成が多く、使用環境に応じた機器配置や事前調整を前提とする場合がある。このような構成では、即時的な利用や可搬性の面で制約が生じる可能性がある。

これに対し本研究では、外部カメラや固定設置機器を用いず、9軸慣性センサのみを搭載した小型ウェアラブルデバイスによって指揮動作を計測する構成を採用する。これにより、設置作業を必要としない高い即時利用性と可搬性を有する音響操作インタフェースの実現を目指す。

2.3 身体動作計測におけるサンプリング周波数

身体動作を計測する際のセンサのサンプリング周波数は、動作の周波数に合わせて決定する必要がある。

指揮動作を加速度データとして解析した研究では、拍の検出や動作評価において、時間分解能が計測精度に影響を与えることが示されている [6]。また、ジェスチャや行動認識を目的とした加速度信号処理に関する研究では、一般的に 100 Hz 程度のサンプリング周波数が用いられる一方で、打点付近の急峻な運動変化や短時間動作を安定して捉えるためには、より高いサンプリング周波数が有効であることが報告されている [7]。

これらの知見を踏まえ、本研究では、指揮動作に含まれる急峻な運動変化を余裕をもって取得するため、センサのサンプリング周波数を 200 Hz に設定した。

3. 提案手法

本研究では、このような観点に基づき、9軸慣性センサのみを搭載したグローブ型デバイスを用い、指揮動作に基づいて音楽再生のテンポおよび音量を制御する手法を提案する。

3.1 システム概要

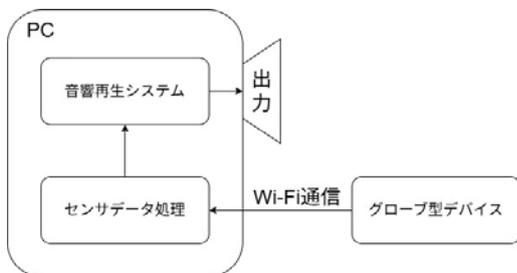


図2 システム構成

システム構成を図2に示す。グローブ型デバイスには、加速度・角速度・地磁気を計測可能な9軸慣性センサ(MPU9250)を搭載し、手首および前腕部の運動を計測する。サンプリング周波数 200Hz で取得されたセンサデータ

は無線通信により PC へ送信され、PC 側でリアルタイム処理を行う。

PC 側では、後述するように受信したセンサ値から指揮動作時の周期や振幅から、音楽プレーヤのテンポおよび音量を決定する。本手法では、明示的なジェスチャ分類は行わず、運動データから得られる連続的な物理量を音響制御量として直接用いる点に特徴がある。



図3 慣性センサを搭載したグローブ型デバイス

3.2 センサデータ処理

本研究では、加速度センサおよび角速度センサから得られる3軸データを用いて指揮動作の解析を行う。時刻 t における加速度ベクトル $\mathbf{a}(t) = (a_x(t), a_y(t), a_z(t))$ から、3軸成分の二乗和平方根として合成加速度 $A(t)$ を算出する。同様に、角速度ベクトル $\boldsymbol{\omega}(t) = (\omega_x(t), \omega_y(t), \omega_z(t))$ についても合成角速度 $\Omega(t)$ を算出する。

3.3 テンポ制御

本節では、実装した音響操作システムのテンポ制御について、算出手法と反映手法に分けて解説する。

3.3.1 テンポの算出手法

テンポ制御には、指揮動作に伴う周期的な運動成分に着目し、合成加速度 $A(t)$ の時間変化を用いる。指揮動作に伴う拍点は、加速度信号における局所的なピークとして現れる。このピークを安定して検出するため、 $A(t)$ に対して一次のローパスフィルタを適用する。平滑化後の信号 $\hat{A}(t)$ が所定のしきい値 A_{th} を下回った状態から上回った瞬間を拍点として検出する。連続する拍点の時刻差を ΔT_i 秒とすると、瞬時テンポは $BPM_i = 60/\Delta T_i$ により算出される。

誤検出を抑制するため、拍点検出後は一定時間の不応期を設け、極端に短い間隔での再検出を防止する。さらに、得られた複数のテンポ推定値に対して移動平均処理を行い、安定したテンポ推定値 $BPM(t)$ を得る。

3.3.2 テンポ制御量への反映

算出されたテンポ推定値 $BPM(t)$ は、音響再生に用いる制御テンポへ反映される。本研究では、過度に細かいテ

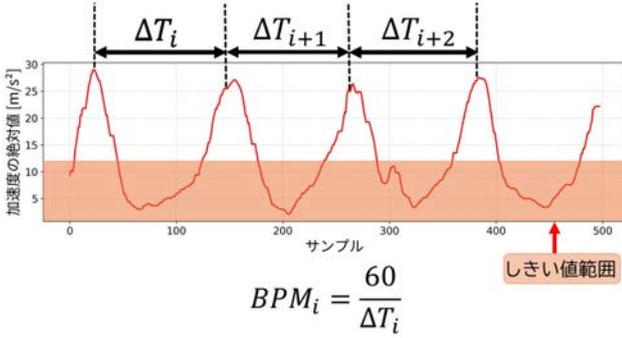


図 4 テンポ算出のイメージ

ンポ変動による聴感上の不安定さを抑制するため、テンポ値を一定幅で量子化する。基準テンポを BPM_0 、量子化幅を Δ_{BPM} とすると、制御テンポ $BPM_q(t)$ は次式で定義される。

$$BPM_q(t) = BPM_0 + \Delta_{BPM} \cdot \text{round} \left(\frac{BPM(t) - BPM_0}{\Delta_{BPM}} \right)$$

この処理により、指揮動作に応じたテンポ変化を保ちつつ、再生テンポの急激な揺らぎを抑制する。

3.3.3 テンポ制御への反映手法

音響再生においては、制御テンポ $BPM_q(t)$ に基づき、再生テンポが逐次更新される。基準テンポ BPM_0 に対する再生速度倍率 $r(t)$ は、 $r(t) = BPM_q(t)/BPM_0$ で定義される。

再生速度倍率 $r(t)$ は時間的に変動するため、音響出力の安定性を確保する目的で、更新されたテンポ値が直前の制御値と異なる場合にのみ再生テンポを更新する。

音声信号を離散時間信号 $x[n]$ とすると、テンポ制御後の出力信号 $y[n]$ は、時間伸縮処理によって $y[n] = x(n/r(t))$ で表される。

ここで、 $r(t) > 1$ の場合には再生速度が増加し、 $r(t) < 1$ の場合には再生速度が低下する。時間伸縮後の出力信号 $y[n]$ は、元の音声信号 $x[n]$ を時間軸方向に再配置することで生成される。このとき、出力サンプル $y[n]$ に対応する入力信号上の参照時刻 $\frac{n}{r(t)}$ は、一般に整数値とならないため、実際の実装では補間処理を用いて音声信号を再構成する。

以上の処理により、指揮動作に含まれる周期的な運動が音楽再生テンポとして反映され、ユーザの動作に応じた連続的なテンポ制御が実現される。

3.4 音量制御

本節では、実装した音響操作システムの音量制御について、算出手法と反映手法に分けての解説する。

3.4.1 音量の算出手法

音量制御には、指揮動作の運動強度を表す指標として、

加速度および角速度を統合した運動強度指標 $I(t)$ を、合成加速度および合成角速度の線形結合として定義する。

ここで g は重力加速度、 w_a および w_w はそれぞれ並進成分と回転成分の寄与度を調整する重みである。一定時間幅 W における運動強度指標 $I(t)$ の二乗平均平方根 (RMS) を運動強度指標 $E(t)$ として算出する。この指標は、一定時間内における指揮動作の振幅や勢いを反映した量となる。さらに、瞬間的な変動による音量の不安定化を抑制するため、算出された運動強度指標 $E(t)$ に対して一次遅れ系による平滑化処理を施す。

$$E_{LPDF}(t) = E_{LPDF}(t-1) + \lambda(t)(E(t) - E_{LPDF}(t-1))$$

$$\lambda(t) = \begin{cases} \alpha, & E(t) > E_{LPDF}(t-1) \\ \gamma, & E(t) \leq E_{LPDF}(t-1) \end{cases}$$

ここで、 $E_{LPDF}(t)$ は平滑化後の運動強度指標、 α および γ はそれぞれ上昇時および下降時の追従性を決定する係数である。この処理により、指揮動作の開始に対しては素早く反応し、動作の減衰に対しては緩やかに追従する運動強度指標が得られる。

3.4.2 音量制御の反映手法

算出された運動強度指標 $E_{LPDF}(t)$ は、音響再生ソフトウェアにおける音量制御パラメータへ反映される。人間の聴覚は音圧の比に対して対数的な感度を持つため、音量は一般にデシベル (dB) スケールで表現される。このため、音声信号の振幅を線形に変化させた場合でも、知覚される音量変化は必ずしも線形とはならない。

そこで提案手法では、運動強度指標 $E_{LPDF}(t)$ を直接振幅倍率に写像するのではなく、音圧レベル $L(t)$ (dB) として扱うことで、聴覚特性を考慮した音量制御を行う。

運動強度指標 $E_{LPDF}(t)$ に基づき、音圧レベル $L(t)$ を区分線形写像として定義する。 $E_{LPDF}(t) \leq E_{\text{mid}}$ の場合、 $L(t)$ は

$$L(t) = L_{\text{min}} + (L_{\text{mid}} - L_{\text{min}}) \frac{E_{LPDF}(t) - E_{\text{min}}}{E_{\text{mid}} - E_{\text{min}}}$$

とし、 $E_{LPDF}(t) > E_{\text{mid}}$ の場合には

$$L(t) = L_{\text{mid}} + (L_{\text{max}} - L_{\text{mid}}) \frac{E_{LPDF}(t) - E_{\text{mid}}}{E_{\text{max}} - E_{\text{mid}}}$$

と定義する。ここで、 $L_{\text{min}}, L_{\text{mid}}, L_{\text{max}}$ はそれぞれ最小音量、通常動作時の音量、および強調動作時の最大音量を表す音圧レベル [dB] であり、 $E_{\text{min}}, E_{\text{mid}}, E_{\text{max}}$ は対応する運動強度指標の代表値である。この区分線形写像により、通常の指揮動作に対しては安定した音量変化を保ちつつ、大きな動作に対しては音量変化を強調することが可能となる。

この写像により、小さな指揮動作に対しては過度な音量変化を抑制しつつ、大きな動作に対しては明確な音量変化を与えることが可能となる。音圧レベル $L(t)$ から、振幅倍率は $V(t) = 10^{L(t)/20}$ により求められる。音量制御後の出力信号は、音声信号に振幅倍率 $V(t)$ を乗算することで得られる。この処理は、音声信号の振幅を時間的に変化させる操作に相当し、指揮動作の運動強度が音の大きさとして直接出力に反映される。運動強度指標は時間的に変動するため、振幅倍率 $V(t)$ も逐次更新される。音量の急激な変化による聴感上の違和感を抑制するため、振幅倍率に対して一次遅れ系による平滑化処理を施す。

以上の処理により、指揮動作から算出された運動強度指標は、dB スケールでの音圧レベル表現を経て振幅倍率へ変換され、音声信号の振幅として出力に反映される。この結果、指揮動作の大きさや勢いに応じた連続的かつ聴覚特性を考慮した音量制御が実現される。

4. 検証

本章では、提案した指揮動作に基づく音響制御手法について、テンポ制御および音量制御のそれぞれに関して検証実験の概要を説明する。

4.1 テンポ検出精度の検証

本節では、テンポ検出精度の測定について、目的、実験条件、評価指標について解説する。

4.1.1 目的

テンポ制御においては、指揮動作から算出されるテンポ推定値が、実際の指示テンポをどの程度正確に反映しているかが重要である。本実験では、既知のテンポに対する指揮動作を入力とし、提案手法によるテンポ検出精度を評価する。

4.1.2 実験条件

実験では、メトロノームを用いて $BPM = 60, 120, 180$ の3種類の一定テンポを提示し、各テンポに同期するよう著者自身が指揮動作を行った。テンポ推定における安定化処理として用いる移動平均回数を1回、2回、4回の3条件に設定し、各テンポ条件と組み合わせ、合計9条件の実験を行った。各条件において、20回分の指揮動作を行い、システムが出力したテンポ推定値 $BPM(t)$ を記録した。

4.1.3 評価指標

テンポ検出精度の評価には、提示テンポ BPM_{ref} と検出テンポ $BPM(t)$ の差に基づく平均値および標準偏差を用いる。これにより、移動平均回数の違いによるテンポ推定の安定性および応答性の変化を比較する。

4.2 音量調整用の運動強度指数検出精度の検証

本節では、検出精度の測定について、目的、実験条件、評価指標について解説する。

4.2.1 目的

音量制御では前述したとおり、指揮動作の振り幅や勢いが運動強度指標として算出され、音量制御量へ反映される。本実験では、振りの大きさの違いが算出される運動強度指標にどのように反映されるかを検証する。

4.2.2 実験条件

実験では、メトロノームのテンポを $BPM = 120$ に固定し、図5に示すように、小さめの振り（半径10cm程度）、中程度の振り（半径20cm程度）、大きめの振り（半径30cm程度）の3種類の振り動作を行った。

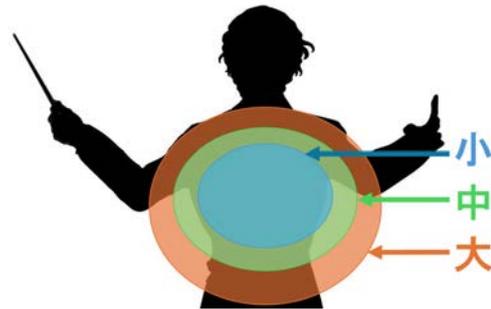


図5 それぞれの振りの大きさでの腕の移動範囲の目安

それぞれの振り動作における腕および手首の移動範囲については、事前に設定した範囲を図5に示す。

各条件において指揮動作を行い、運動強度指標 $E_{LPPF}(t)$ の算出値を記録した。測定は各振り条件につき1回ずつ行い、合計3条件の測定を行った。

4.2.3 評価指標

各振り条件において得られた運動強度指標の平均値および時間変化を比較することで、振りの大きさの違いが運動強度指標に段階的に反映されているかを評価する。この結果をもとに、提案した音量制御手法が指揮動作の強弱を適切に捉えているかを検証する。

5. 結果

本章では、検証実験の結果をテンポ検出精度、運動強度指標検出精度の二つに分けて説明する。

5.1 テンポ検出精度

3種のBPMと3種の移動平均回数の組み合わせによる計9条件の測定データ（平均値と標準偏差）を図6に示す。

基準テンポと検出テンポの平均の差分は、基本的に小さく、約1BPM程度に収まっていることが分かる。BPM = 180での値は他BPMの時と比べて少し大きい傾向が見られた。

標準偏差に関しては平均回数が少なく、基準BPMが大きい方が値が大きくなる傾向が見られた。BPM = 180の測定では、他BPMでは見られた平均回数が1回と2回での差がほとんど見られなかった。標準偏差で見られたこの傾向は、BPMの大小による、1BPMの時間範囲の違いが

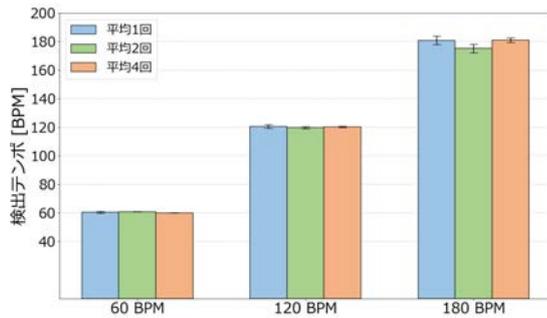


図 6 各 BPM ごとの測定値の平均値と標準偏差

関係していると考えられる。BPM が 1 変化するために必要な運動周期変化量は変化前の BPM に依存する。変化前 BPM が小さければより大きな、BPM が大きければより小さな BPM が変動する。同じ周期変化量でも BPM, 基準 BPM が高 BPM である方が BPM の値の変動量が大きくなるという現象が、高 BPM での標準偏差の値の増加に寄与していると考えられる。

5.2 運動強度指標の測定

図 7 にそれぞれの振りの大きさでの運動強度指標の測定値と標準偏差を示す。

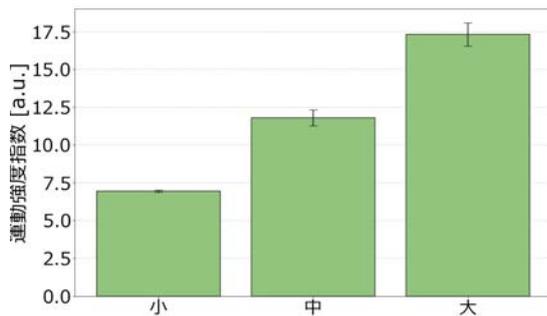


図 7 運動強度指標の測定値の平均値と標準偏差

運動強度指標の値は、振りの大きさに合わせてそれぞれおよそ、小：7、中：12、大：17となった。この値は、運動強度指標の代表値 E_{min} , E_{mid} , E_{max} の決定に用いることができると考えられる。

標準偏差は振りの大きさに合わせて大きくなっていった。これは振りが大きいほど指揮動作に緩急がつき、 $E(t)$ の 1 周期内での変動が激しくなるためと考えられる。

6. 今後の課題

提案手法の課題を述べる。

6.1 テンポ検出精度の評価指標の改良

本研究では、検出テンポと基準テンポとの差分を BPM 単位で評価指標として用いた。しかし、BPM は非線形な指標であり、低 BPM と高 BPM では、同一の BPM 誤差が異なる時間的ずれを意味する。そのため、拍間隔におけ

る時間誤差や、基準テンポに対する再生速度倍率といった指標を併用することで、テンポ制御精度をより適切に評価できる可能性がある。

6.2 ユーザ実験による音楽的応答性の評価

本研究で実施した検証実験は、主に数値的な精度評価を目的としたものであり、操作に対する応答の速さや、テンポ変化の自然さといった音楽的観点での評価は十分に行えていない。今後は、実際の音楽再生を伴うユーザ実験を通じて、音響制御の応答性や操作感について定性的な評価を行う必要がある。

6.3 ユーザ動作に応じた音響制御パラメータの最適化

本研究では、テンポ推定値の安定化のために移動平均回数を固定値として設定した。しかし、移動平均回数は応答性と安定性のに大きく影響する。検証結果から、平均回数によってテンポ推定値の分散特性が変化することが示されており、ユーザの動作特性や演奏場面に応じて制御パラメータを動的に調整する適応的な音響制御手法の検討が今後の課題である。

6.4 テンポ変動時の音響制御手法の改良

現状の実装では、基準テンポより極端に低いテンポが指示された場合、音楽の時間伸縮処理に起因して、音が断続的に再生される問題がある。この問題に対して、補間手法の改良や、音楽フレーズ単位での制御など、聴感上の自然さを考慮した音源操作手法の確立が求められる。

6.5 音楽表現要素の拡張

本研究では、テンポおよび音量の制御に着目したが、指揮動作にはアクセントや強拍・弱拍といったさらなる表現情報が含まれている。これらの要素を運動データから抽出し、音響制御へ反映することで、より表現豊かな音楽インタラクションが実現できると考えられる。

7. まとめ

本研究では、9 軸慣性センサを搭載した小型グローブ型デバイスを用い、指揮動作に基づいて音楽再生のテンポおよび音量を制御する音響操作手法を提案した。提案手法では、明示的なジェスチャ分類を行わず、加速度および角速度から得られる連続的な運動量を音響制御パラメータへ直接対応付ける構成を採用した。

検証実験を通じて、指揮動作に含まれる周期性および運動強度が、計測できることを確認した。

今後は、テンポ検出精度の評価指標の改良や、ユーザ実験による応答性の定性的評価を行うとともに、テンポ変動時の音響操作手法や、アクセントなどのさらなる音楽表現要素への拡張について検討する予定である。

参考文献

- [1] Nakra, T. M. and Paradiso, J. A.: The Digital Baton: a Versatile Performance Instrument, *International Conference on Mathematics and Computing*, pp. 313–316 (1997).
- [2] 原田勉, 森田秀之, 大照完, 橋本周司: 指揮棒を用いた演奏制御, 第 42 回情報処理学会全国大会講演論文集, 情報処理学会, pp. 313–314 (1991).
- [3] 小尾正和, 鈴木健嗣, 橋本周司: 音楽音響信号を対象としたジェスチャによる音楽指揮システム, 第 68 回情報処理学会全国大会講演論文集, 情報処理学会, pp. 161–162 (2006).
- [4] 小尾正和, 鈴木健嗣, 橋本周司: ジェスチャによる音楽の再生制御, 第 67 回情報処理学会全国大会講演論文集, 情報処理学会, pp. 363–364 (2005).
- [5] Schacher, J. C.: Moving Music: Exploring Movement-to-Sound Relationships, *Proceedings of the 3rd International Symposium on Movement and Computing*, MOCO '16, Association for Computing Machinery (2016).
- [6] 安田隆, 葉田善章: 3 軸加速度データを用いた指揮法の評価手法の開発, 研究報告音楽情報科学, Vol. 2021-MUS-132, No. 6, pp. 1–6 (2021).
- [7] 吉村直也, 前川卓也, 天方大地, 原隆浩: ジェスチャ・行動認識のための加速度信号アップサンプリング手法に関する検討, 研究報告ユビキタスコンピューティングシステム, Vol. 2018-UBI-58, No. 10, 情報処理学会, pp. 1–8 (2018).