

# Deep Augmented Performers: メロディモーフィングと身体機能の融合によるアンサンブル演奏システム

三浦 寛也<sup>1,a)</sup> 浜中 雅俊<sup>1</sup>

**概要:** 音楽情報処理と人間の身体機能との融合が未来の音楽表現を開拓する。本論文では、演奏者の未来の演奏のコントロールを可能とするインタラクティブなアンサンブル演奏システム「Deep Augmented Performers」について述べる。このシステムは、メロディモーフィング手法を用いて作編曲された楽曲を筋電気刺激 Electrical Muscle Stimulation (EMS) に変換し、手指や足に装着したデバイスによって複数の演奏者の身体動作を同期・制御し、アンサンブル演奏を達成するものである。演奏するメロディのバリエーションは、人が音楽を聴取る知能の仕組みを形式化した理論である Generative Theory of Tonal Music (GTTM) に基づいて作曲され、複数のメロディのセグメントを自由に操作することで、メロディの全体的な構造を保持したまま、装飾音のみを変化させることができる。提案システムを用いたユーザスタディによる評価から、その有効性を確認した。

**キーワード:** インタラクティブ演奏システム, メロディモーフィング, 筋電気刺激 (EMS)

## Deep Augmented Performers: A New Musical Experience by Fusion of Melody Morphing and Body Movements

**Abstract:** The fusion of music-information processing and human physical functions will enable new musical experiences to be created. We developed an interactive music system called “Deep Augmented Performers,” which provides users with the experience of conducting a musical performance. This system converts music arranged using melody morphing into electrical muscle stimulation to control the body movements of multiple performers (e.g., musicians and dancers) via devices attached to the performers’ hands and feet. The melodies used in the system are divided into segments, and each segment has multiple variations of melodies. The user can interactively control the performers, thus the actual performance. We confirmed the effectiveness as a new method of performance expression, from the evaluation of user studies using the proposed system.

**Keywords:** Interactive musical performance system, Melody morphing, Electrical muscle stimulation (EMS)

### 1. はじめに

深層学習の高度化に伴う様々な技術の発展や、Makerムーブメントによる楽器のフィジカルインタフェース構築の広がりによる新たな音楽表現の模索などを背景に、音楽情報科学研究は更なる進展を遂げている。とりわけ、機械に音楽を演奏させる自動演奏の試みは古くは16世紀から始まり、自動ピアノやパーカッションなどの楽器演奏の自

動化が進んできた。2010年代に発明された複数のオーケストラ楽器を同時に演奏させる装置 [14] や、ロボットによる高度な音楽パフォーマンス [16] の実現は、自動演奏システムのひとつの境地である。さらに近年では、人間拡張 (Human Augmentation) 研究の一環として、Electrical Muscle Stimulation (EMS) や外骨格デバイスを用いて手指の動きを制御し、運動や描画、そして楽器演奏を支援する研究も盛んに行われている [1], [2], [4], [13], [15], [17]。EMSとは、運動神経に電気刺激を与え、筋肉の収縮を促すための装置である。このEMSは、これまで理学療法の一環としてよく知られていたが、インタラクティブメディア研究では、従来の Virtual Reality (VR) や Augmented Reality (AR) のア

<sup>1</sup> 理化学研究所 革新知能統合研究センター  
RIKEN Center for Advanced Intelligence Project, Nihonbashi 1-chome Mitsui Building, 15th floor, 1-4-1 Nihonbashi, Chuo-ku, Tokyo, 103-0027, Japan

<sup>a)</sup> hiroya.miura@riken.jp

アプリケーションを超える触力覚インタフェースのための新しいメカニズムとして注目されている [12]. このような技術的潮流から、音楽情報処理技術と人間の身体機能との融合が未来の音楽表現を開拓すると考える。

本論文では、演奏者の未来の演奏のコントロールを可能とするインタラクティブなアンサンブル演奏システム「Deep Augmented Performers」を提案する。このシステムは、メロディモーフィング手法を用いて作編曲された楽曲を筋電気刺激 (EMS) に変換し、ステージ上のダンサーやミュージシャンなどの複数のパフォーマーの身体の動きを手指や足に装着したデバイスによって同期・制御し、アンサンブル演奏を達成するものである。演奏するメロディのバリエーションは、音楽理論 Generative Theory of Tonal Music (GTTM) [3] に基づいて作曲され、複数のメロディのセグメントを自由に操作することで、メロディの全体的な構造を保持したまま、装飾音のみを変化させることができる (図 1)。従来のアンサンブルによる音楽演奏では、指揮者はタイミングと演奏表現に関する指示を与え、演奏の統一性を図る。音楽の演奏は身振りを伴うことが自然であるが、このような身振りの獲得と理解には、多大な学習が必要である [18]。また、一般的なステージ演出の問題点として、演奏者は観客の動きや反応と同期することは難しい。本システムでは、あらかじめ作られたステージ演出を再生するだけでなく、その場にいる観客の反応を考慮しながら、演奏者やダンサーの未来の動作、そしてパフォーマンスそのものをインタラクティブかつジェネレイティブに作り上げることができる。

様々な音楽操作を可能にする作曲支援システムの実現には、音楽の深層構造を理解可能な計算機の開発が必要であると考える。筆者らは、これまで 10 年以上に渡り、Lerdahl と Jackendoff が提案した音楽理論 GTTM に基づく楽曲構造分析およびメロディ操作に関する研究開発を進めてきた [5], [6], [8], [9]。計算機による GTTM 分析の性能は、深層学習を導入することで飛躍的に向上した [10]。この GTTM 分析の結果によって得られる楽曲の深層構造であるタイムスパン木は、重要な音符と装飾音符の関係を階層的に表現するものである。タイムスパン木を用いる最大の利点は、メロディの主要な構造を維持したまま、装飾音を段階的に減らしていくことで、メロディの音数を減らすことができることである。そしてこのタイムスパン木を用いたメロディ操作の一例として、2つの入力メロディから中間となるメロディを生成するメロディモーフィングがある。GTTM による楽曲構造分析の知見を実世界のアプリケーションや実パフォーマンスに応用・拡張する取り組みは未だ少ない。同様に、先述の通り EMS を用いた楽器演奏支援システムはこれまでも数多く存在するが、音楽情報処理との融合を試みた研究は、我々の知る限りでは本研究が初めてである。



図 1 研究概要

Fig. 1 Research Outline

## 2. メロディモーフィング手法

本章では、メロディ  $A$  と  $B$  を入力し、 $A$  と  $B$  それぞれの特徴を反映させる度合いを変化することで、メロディ  $A$  と  $B$  の間にある複数個のメロディ  $C$  をある尺度のもとで順序付けて生成するメロディモーフィング手法について述べる。この時、メロディ  $A, B, C$  は以下の条件を満たすものとする。

1.  $A$  と  $B$  より  $A$  と  $C$  のほうが類似している、かつ、 $A$  と  $B$  より  $B$  と  $C$  のほうが類似している。
2.  $B$  が  $A$  と同じ場合、 $C$  も  $A$  となる。
3.  $A$  と  $B$  それぞれの特徴を反映させる度合いを変化させることで、 $C$  が複数出力される。
4.  $A$  と  $B$  がいずれもモノフォニー (和音を含まない単旋律) の場合、 $C$  もモノフォニーとなる。

条件 1 と 2 は、メロディモーフィングに限らず、画像等のモーフィングにも一般にあてはまる条件である。一方、条件 3 と 4 は、メロディモーフィングの条件として我々が新たに考えたものである。

タイムスパン木は GTTM [3] による楽曲構造分析の結果求められるものである。GTTM は、人が音楽を聴取する知

能の仕組みを形式化した理論であり、グルーピング構造分析、拍節構造分析、タイムスパン簡約、プロロンゲーション簡約という4つのサブ理論から構成されている。このGTTMでは、楽譜に分析を加えることで、楽譜に内在する様々な階層構造を深層構造として顕在化させる。タイムスパン簡約は、あるメロディを簡約化することによって、そのメロディの装飾的な部分が削ぎ落とされ、本質的なメロディが抽出されるという直観を表したもので、構造的に重要な音が幹になるような2分木(タイムスパン木)を求める分析である(図2)。

## 2.1 タイムスパン木に基づくメロディの簡約と基本演算

図2は、タイムスパン木を用いたメロディの簡約の例である。図のメロディAの上にある木構造は、メロディAをタイムスパン簡約した結果得られたタイムスパン木である。タイムスパン木のレベルBより下にある枝の音符を簡約するとメロディBのようになる。さらに、レベルCより下にある枝の音符を簡約するとメロディCのようになる。この時、メロディBはメロディAとCの間のメロディであることから、メロディの簡約もメロディのモーフィングの一種と考えることができる。

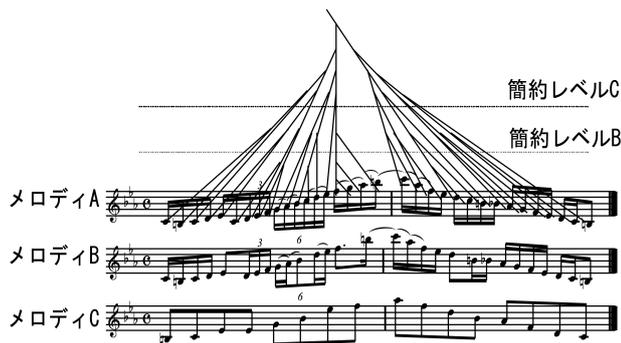


図2 タイムスパン木  
Fig. 2 Time-Span Tree

メロディのモーフィングを実現するため、文献 [5], [6], [8], [9] で定義されている演算である、包摂関係  $\sqsubseteq$ , meet (最大下界)  $\sqcap$  と join (最小上界)  $\sqcup$  を用いる。包摂関係  $\sqsubseteq$  は、 $F_1$  を下位の構造、 $F_2$  を上位の構造(下位の構造を含んでそれ以上の構造を持つ)とした場合、 $F_1 \sqsubseteq F_2$  と表記し、 $F_2$  は  $F_1$  を包摂すると言う。例えば、図2のメロディA, B, Cのタイムスパン木(簡約されたタイムスパン木)、 $T_A, T_B, T_C$ の包摂関係は、以下のように表せる。

$$T_C \sqsubseteq T_B \sqsubseteq T_A$$

meet (最大下界) は、 $T_A, T_B$  の共通部分のタイムスパン木  $T_A \sqcap T_B$  である。join (最小上界) は、メロディA, Bのタイムスパン木  $T_A, T_B$  が矛盾を起こさない限り統合したタイムスパン木  $T_A \sqcup T_B$  である。

## 2.2 GTTMに基づくメロディモーフィング手法

モーフィングという用語は、通常、2つの画像がある場合に、片方の画像からもう一つの画像へ滑らかに変化していくよう、その間を補うための画像を作成することをいう。2次元の顔画像のモーフィングの場合、例えば以下のような操作で中間的な画像の生成が実現できる；(1) 目や鼻など2つの画像の特徴点の対応づけ、(2) 各画像の形状(位置)や色の重み付け、(3) 2つの画像の重ね合わせ。一方、本研究で提案するメロディモーフィングでは、以下のような操作で中間的なメロディの生成を実現する；(1) 2つのメロディの共通部分の対応づけ、(2) 各メロディについてメロディの部分簡約、(3) 両方のメロディの重ね合わせ。

### (1) メロディの共通部分の対応づけ

2つのメロディA, Bのタイムスパン木  $T_A, T_B$  を求め、その共通部分(最大下界)  $T_A \sqcap T_B$  を求める。これにより、タイムスパン木  $T_A, T_B$  は、それぞれ共通部分と非共通部分に分けることができる。 $T_A \sqcap T_B$  は、タイムスパン木  $T_A, T_B$  をトップダウンに見て、最も大きく共通する部分を取り出すが、その際、オクターブの異なる2音(例えばC4とC3)を異なる音とみなす場合と、同じ音とみなす場合で結果が異なる。異なる音とみなす場合には、 $C4 \sqcap C3$  の解は空疎  $\perp$  となる。一方、同じ音とみなす場合には、オクターブの情報捨棄され解はCとなる。オクターブ情報が未定義の場合、(2)以降の処理が困難となることから、オクターブの異なる2音は、異なる音として扱うこととした。

### (2) メロディの部分簡約とその合成

(1)で求めたメロディAとBのタイムスパン木  $T_A, T_B$  の非共通部分には、それぞれ相手のメロディにはない特徴が表われていると考えられる。したがって、メロディのモーフィングを実現するためには、それら非共通部分の特徴をなめらかに増減させ、中間的なメロディを生成するような手法が必要となる。

本研究では2.1節で述べたメロディの簡約を応用し、メロディの非共通部分についてのみ簡約を実行するメロディ部分簡約法を提案する。メロディ部分簡約法では、メロディAのタイムスパン木  $T_A$  と、メロディA, Bのタイムスパン木の共通部分、 $T_A \sqcap T_B$  から、次のアルゴリズムでメロディ  $C_m$  ( $m=1, 2, \dots, n$ ) を生成する。 $C_m$  の添え字  $m$  は、 $C_m$  に含まれるが、 $T_A \sqcap T_B$  には含まれない部分(メロディの非共通部分)に表われる音符の数である。図3の場合では、 $T_A$  には、 $T_A \sqcap T_B$  にはない音符が9個含まれているため、 $T_A$  と  $T_A \sqcap T_B$  の間のメロディが8種類得られることになる。

上記のようにして求めたメロディ  $C$  は、メロディBにはないメロディAのみが持つ特徴の一部を減衰させたものと考えられる。同様に、 $T_B$  と  $T_A \sqcap T_B$  から下記

を満たすメロディ D を生成する。

$$T_A \cap T_B \subseteq T_D \subseteq T_B$$

メロディ A と B のそれぞれのタイムスパン木の一部を簡約したメロディ C とメロディ D を統合 (最小上界) し、合成したメロディ E を生成する。

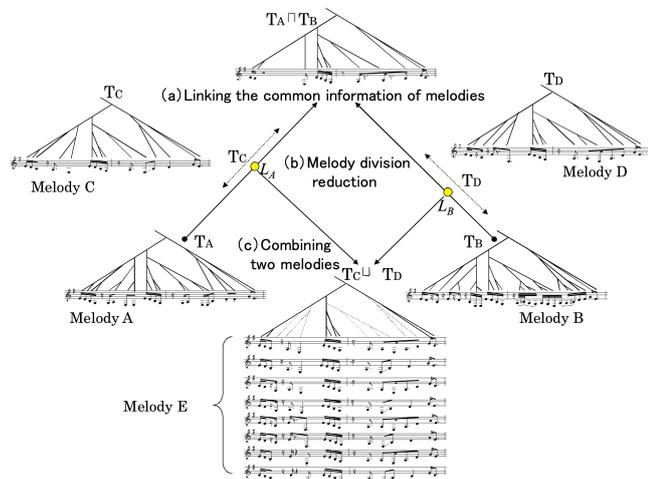


図 3 メロディモーフィング手法  
Fig. 3 Melody Morphing Method

### (3) 音楽家による調整

メロディモーフィング手法は自動化を目指しているが、現時点では音楽家による作業なしでは実現できない。本研究における実験では、音楽家は以下の 3 項目の作業を行った。第 1 の作業では、GTTM に基づきメロディの構造分析を行い適切なタイムスパン木を取得する。第 2 の作業では、2.2 節 (2) のメロディの部分簡約において音符を抽象化する順序を決定する。第 1 と第 2 の作業の後、[N1, N2] のようになっている箇所では、どちらかを選択していくことで、モーフィングメロディは自動的に取得可能である。しかし、この方法でメロディをモーフィングすると、メロディのバリエーションごとの変化が極めて少なくなることがある。そこで、メロディの変化が少なかった場合、第 3 の作業として、音楽家が手でメロディの調整を行った。

本研究では、3 名の音楽家にモーフィングメロディの作成を依頼した。1 名の音楽家は、GTTM 分析について 10 年以上の経験があり、モーフィング手法についても深い理解がある。残りの 2 名は GTTM 分析の経験は 2 年である。上記 3 名は分析結果のクロスチェックを行っており、ほぼ同等の能力であることを確認している。モーフィングの元となるメロディの長さや曲名は指定せず、音楽家が自身で考え制作を行い、最終的なモーフィングメロディを作成した。

## 3. EMS システム

提案システムの主幹となる人間の身体動作を制御するための方法として、本研究では、筋電気刺激 Electrical Muscle

Stimulation (EMS) を用いて、多人数 (2~8 人) の手指や足などの身体動作を同期・制御するための装置 (以下、EMS システム) を開発した。本システムは、Arduino、電気刺激回路、PC、直流安定化電源、電極パッドで構成され、電気刺激を発生させている (図 4, 5)。電極パッドには、市販の低周波治療器に用いられている湿式パッド (OMRON 社製) を採用している。また電気刺激回路には、パルスを発生させるために 555 タイマーを使用し、電圧を上げるために変圧器を用いている。刺激パルスの周波数は 40~70Hz で、パルス幅は 0.8ms である。このパルス電圧はユーザ自身が制御でき、制御したい筋肉の箇所や種類に応じて動的に変更できる。直流安定化電源の電圧を上げることで、電気刺激が強くなり、人体への反応を高めることができる。なお、安全のため自己復帰型ヒューズを電源直後に挿入し、人体に過剰な電流が流れることを防いでいる。



図 4 システム構成  
Fig. 4 System Configuration

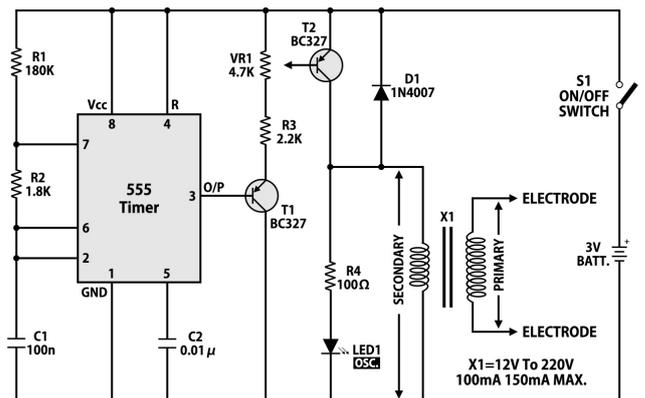


図 5 EMS システムの回路図  
Fig. 5 Schematic of EMS System

本システムでは、複数の EMS 発生器を統合的に制御することで、多人数の身体動作の同期が実現できる。併せて、本システムを用いた複数の楽器演奏に関するフィージビリティスタディを実施した (図 6)。本検証から、EMS システムの装着者は、自身の意思とは無関係に身体の一部が

動き、結果として、手首や肘の手首の関節を曲げるといった単純な動作を完全に制御可能であることを確認した。また、背中に電極パッドを装着し、微弱な電気刺激を与えることで、メトロノームのようにリズムを直接体感できることが確認された。一方で、ギターやピアノを弾くような細かい手の動きを制御することは、現時点では困難であることが確認されたため、将来的な課題として認識している。



図 6 EMS システムを用いた楽器演奏の試行

Fig. 6 Feasibility Study of Performance by EMS System

#### 4. Deep Augmented Performers の実装

Deep Augmented Performers は、演奏者の未来の演奏のコントロールを可能とするインタラクティブなアンサンブル演奏システムである。加えて、演奏者やダンサーの動作、そしてパフォーマンスそのものをリアルタイムに作り上げることができる。本システムは、第2章で詳述したメロディモーフィング手法によって生成された楽曲を EMS に変換し、複数の演奏者の手足に装着したデバイスを介して身体動作を同期・制御することで、アンサンブル演奏を達成する。第3章で述べたフィージビリティスタディの結果から、本システムは特に打楽器演奏に適していることが確認された。本システムで用られるメロディは、メロディモーフィングの分析結果をもとに複数のセグメントに分かれており、各セグメントには複数のメロディのバリエーションが存在する。このメロディのバリエーションを誰でも簡単に操作することができるグラフィカルユーザーインターフェース (GUI) を開発することで、演奏者の動きの指揮が実現できる (図 7)。

GUI の操作方法は以下の 3 手順である；(a) 異なる 2 つの曲を選択し、“GENERATE MUSIC” ボタンの押下により、メロディモーフィングの分析結果から得られた複数のメロディセグメントが生成される (例：1~6)、(b) 生成されたメロディから演奏する曲を選択する、(c) EMS システムにより電気刺激が生成され、各音に対応する電極パッドに送電される (例：C4 ~ C5)。以上のプロセスを繰り返すことにより、システム利用者はメロディを自由に操作することで演奏者の未来の演奏を指揮することができる。

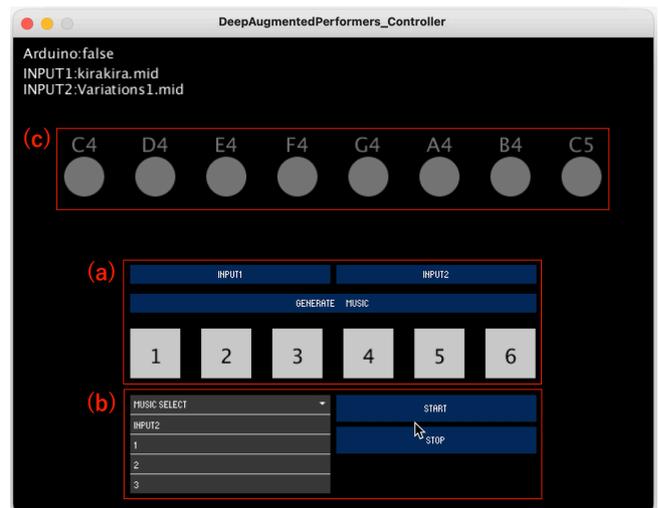


図 7 GUI 画面

Fig. 7 System Graphical User Interface

#### 5. ユーザスタディによる評価

本章では、提案システムの評価について述べる。楽器演奏支援システムの評価では、その行為における人間の諸感覚をどのように扱うべきかに関する課題に取り組むこととなる場合が多いため、目標設定や環境構築、被験者の収集、分析に掛かるコストなどの多様な困難が伴う。奥村らは、演奏支援システムと評価手法の設計について、定量的研究と定性的研究に分類可能であり、注目する研究対象の側面によりそれぞれのアプローチは対をなす関係にあることを指摘している [7], [11]。一般に演奏支援システムの多くは計算機だけで完結することは少なく、人がシステムを利用する。従って、システム利用者が提案システムをどのように使っていたかについての定性的な評価は有益である。また、文献 [7] において、竹川らは、目的ではなく手法を評価することの重要性を提言している。提案システムでは、音楽情報処理と人間の身体機能との融合による新たな音楽表現を提案している。本システムを音楽的かつ芸術的な観点でその斬新さを評価することは難しい。一方で、本研究では EMS システムによる動作制御手法を提案しており、その精度は定量的な評価が可能である。以上を踏まえ、本論文では、EMS システムの時間応答性評価 (5.1 節)、実利用による評価 (5.2 節)、さらにインタビューによる評価 (5.3 節) を実施した。

##### 5.1 EMS システムの時間応答性評価

提案システムの主幹となる EMS システムによる動作制御手法の定量評価として、その時間応答性評価を実施した。本実験では、被験者にアイマスクと耳栓を装着し、視覚・聴覚を遮断した状態で、提案システムを用いて筋電気刺激のみを同時に与えた環境下で、手首の周期的なリズム動作が同期されるかを観察した。手首関節の屈折を機能する長

橈骨筋伸筋と短橈骨筋を刺激する箇所に電極パッドを装着し、刺激周波数 60Hz の刺激を律動的に与えた(図 8)。全 10 名の被験者(男性 6 名, 女性 4 名)に対して計 20 回実施し、その伝達時間遅延を計測した。

実験の結果、全被験者による伝達遅延時間の平均は 242.8ms (標準偏差 86.4ms) であることが分かった。また全被験者を通して、20~45° 程度の手首の掌屈が確認された。伝達時間の遅延が最も大きかった被験者 1 のコメントを収集した結果、電極パッド貼付位置などの事前キャリブレーションが不十分であり、適切な刺激を与えられていなかったことが分かった。実験結果より、視覚・聴覚入力が遮断された状態であっても、EMS システムによる運動覚同調のみで筋発揮のタイミングを同調できることが分かった。これにより、システム利用者が互いに視認していない場合や、体の向きといった空間的整合性が確保されていない状態であっても身体動作を同期可能であることが示唆された。一方で、電極パッド貼付位置の最適化や刺激関数の最適化について今後検討する必要があると考えられる。



図 8 EMS システムによる伝達時間遅延の検証実験  
Fig. 8 Verification Experiment on Transmission Time Delay

## 5.2 実利用による評価

提案システムの有効性を検証するため、実利用による定性的評価を目的としたケーススタディを実施した。本ケーススタディでは、ハンドベル奏者 4 名, ドラム奏者 1 名, ダンサー 1 名による計 6 名のパフォーマンスを対象とした(図 9)。演奏曲は、きらきら星変奏曲 (12 Variationen über ein französisches Lied “Ah, vous dirai-je, maman”) の主題と第 1 変奏の 2 曲を入力とし、メロディモーフィングの分析結果から得られた原曲を含む 8 項目のメロディセグメントを設定した。本システムの操作(指揮)は、提案システムおよび GTTM の機構を熟知した本著者の 1 名が一貫して操作した。各奏者の動きを適切に制御するため、約 5 分間の事前動作テストとキャリブレーションによって、筋肉の動きを考慮した適切な電極パッドの装着位置を決定した。電極パッドの装着位置と電気刺激の提示方法は、以下の通りである。

・ハンドベル奏者は、ベルを鳴らす際に手首を振って音を鳴らす必要がある。そのため、手首関節の屈折を機能する長橈骨筋伸筋と短橈骨筋を刺激する箇所に電極パッドを装着した。ハンドベル奏者には、メロディモーフィングから生成された主旋律を情報として与え、完全に動作を制御した状態で演奏可能である。

・ドラム奏者には、上腕二頭筋、腕橈骨筋、足関節の底屈を制御する腓腹筋を刺激する箇所に電極パッドを装着した。これにより、前腕と手首が動くため、ドラム奏者がスティックでビートを刻むことができる。また、踵を床につけ、足を上に向けた状態で腓腹筋を刺激することで、足のリズムカルな動きが制御できる。ドラム奏者には、演奏曲のテンポ情報を読み取ることで、対応する 4 ビートを自由自在に演奏可能である。

・ダンサーには、曲のテンポ情報を入力とし、背中に装着した電極パッドから弱い電気刺激を与えた。これにより、メトロノームのようにリズムを感知して、他の奏者とリズムを合わせながら任意のダンスを踊ることができる。

本システムを用いた実際の演奏は約 5 分間行われた。メロディセグメントの切り替えは、合計 16 回実施された。演奏における顕著なミスとして、ハンドベル奏者とドラム奏者の動作において、0.5 秒以上の遅延箇所は 8 回であった。これは、主に連符を演奏する際に、連続した細かい動作が強いられる場面で生じたものであった。

## 5.3 インタビューによる評価

前節で述べたケーススタディの実施後、全被験者の内 4 名(ハンドベル奏者 3 名, ドラム奏者 1 名)を対象にインタビュー調査を実施した。本調査では、主に被験者が「提案システムをどのように使っていたか」、「どのように態度が変わったか」、「どのような新しい気づきがあったか」といった項目についてのヒアリングをおこない、提案システムによる有効性の検証を図ることを目的としている。インタビュー結果の要旨は以下の通りである。

### 被験者 A (ハンドベル奏者) による回答

電気刺激による不快感を感じることなく、すぐにシステムに慣れることができた。実際に筋肉が震え、その反動で手首が動かされる感覚が印象的で興味深かった。未知のメロディを演奏する際に、自分の予期していないタイミングで動作することに敏感になるため、そこに安定性があると尚良い。知らない曲をみんなと演奏すること自体は楽しく、演奏者同士での一体感が感じられた。演奏のテンポや強弱を可変できるようになると表現の幅が広がると考える。

### 被験者 B (ハンドベル奏者) による回答

電気刺激の感覚はもっとビリビリとする感じを想



図9 Deep Augmented Performers による演奏の様子

Fig. 9 Performance by Deep Augmented Performers

像していたが、痛さなどはなくドクンドクンという感じで、システムに慣れると楽しく演奏することができた。演奏した曲は知らないメロディだったが、どこかで聴いたことがあるような自然な曲調だった。楽器演奏の経験は皆無で不得意ではあったが、自分の感覚とは裏腹に勝手に演奏が達成された感覚はこれまでにない体験だった。例えば、指などの細かい動きも操作できれば、より豊かな演奏体験が実現できるという期待がある。

#### 被験者 C (ハンドベル奏者) による回答

電気刺激の伝わり方や動作の感じ方は、実際に腕がどのように動いているのかが分かった。最初はピクッと反応の方に気を取られていたが、システム慣れてくると電気がどのように通って、筋肉が動くのかということが分かるようになり、それが面白く感じた。何かを習得する際には、まずは頭で理解することが必要だが、身体に直接的に無意識な部分に働きかけるという意味で多方面に活用できるだろう。

#### 被験者 D (ドラム奏者) による回答

システムを使うことで、体の動かし方やバランス感覚の捉え方の参考になった。直接筋肉を刺激することで、自分の意思とは無関係に動かされることになるが、人によっては苦手な人もいるのかもしれない。電極パッドはより小さく複数箇所に着装できると、刺激が伝わりやすく、細かい動きにも対応でき、使いやすくなると思った。システム利用者が演奏するため技能を平等かつ自由に取り入れることができることに多大な可能性があると感じた。

ユーザスタディによる評価結果から、メロディモーフィング手法を用いたインタラクティブなメロディの操作および、EMS システムを用いた演奏者間の動作の同期による、新たな演奏表現法としての有効性を確認した。加えて、電気刺激の強さやタイミング、電極パッドの装着位置を考慮することで、様々な動作やパフォーマンスに応用可能であることを確認した。また、制御する部位やパターンの多様化を進め、より細かい動作の制御を実現することで、演奏表現のバリエーションが拡張可能であると考えられる。一方で、EMS システムによる制御に関して、タイミングの遅延などが生じ、完全なタイミングの同期が難しいことが分かった。そのため、複数箇所に着用パッドを装着し反射速度を向上させるという方法や、ビートに合わせていくことで、これらの諸問題に対処していく必要がある。

## 6. おわりに

本稿では、音楽情報処理と身体機能の融合による新たな音楽表現として、演奏者の未来の演奏のコントロールを可能とするインタラクティブなアンサンブル演奏システム「Deep Augmented Performers」の構築について説明した。概要映像は以下に公開している；<https://youtu.be/djaexG7so14>。本システムは、GTTM のタイムスパン木を利用したメロディモーフィングにより、ユーザーが自由に音楽を操ることができることに加え、複数のパフォーマンスの体の動きを制御し、インタラクティブかつジェネレイティブなパフォーマンスを実現する。ユーザスタディとして、EMS システムによる動作制御手法の定量評価、実利用およびそのインタビューによる定性評価を

実施し、提案システムの有効性を確認した。

応用シナリオとして、システム利用者が互いに視認していない場合や、体の向きといった空間的整合性が確保されていない状態であっても身体動作を同期可能であるという提案システムの特徴を活かし、暗闇演奏会や360°ステージア라운드オーケストラ、オーディエンスの状態をフィードバックし一緒にコンサートを作り上げる演奏会などが挙げられる。また、今後の課題として、ピアノやギターなど、より複雑な楽器演奏を可能にするシステムの開発を目指す。加えて、電極パッド貼付位置および刺激関数の最適化、制御する部位やパターンの多様化、動作方法のモデル化を進め、より細かい動作の制御に取り組んでいく。

本研究で得られた知見は、人間とシステムとの関わり、それによって生まれる演奏体験のあり方に変化をもたらす可能性を秘めているだろう。日々進化を遂げる情報処理および人工知能技術は、我々の仕事や知性を置き換えるものではなく、自身の能力や特性を拡張してくれるものであると考える。我々は、楽器演奏能力やより高度な技能の自由自在なインプットが可能な未来の実現を目指すとともに、音楽情報処理技術と人間が融合したシステムが新たな技術発展を推し進めるものであると強く認識している。本論文が、当該分野の発展の一助となれば幸いである。

**謝辞** 本研究の一部は、科研費(基盤研究(B)21H03572)および一般社団法人未踏(Mitou Foundation)による第2期AIフロンティアプログラムの助成を受けたものである。北野宏明氏、吉崎航氏、清水亮氏、暦本純一氏、中原啓貴氏には研究の構想段階から多くの助言を頂戴した。ここに記して感謝する。

## 参考文献

- [1] Ayaka Ebisu, Satoshi Hashizume, and Yoichi Ochiai, Building a feedback loop between electrical stimulation and percussion learning, In Proceedings of ACM SIGGRAPH 2018 Studio (SIGGRAPH '18), ACM, pp.1-2 (2018).
- [2] Emi Tamaki, Miyaki Takashi, and Jun Rekimoto, Possessed-hand: Techniques for controlling human hands using electrical muscles stimuli, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), ACM, pp.543-552 (2011).
- [3] Fred Lerdahl and linguist Ray Jackendoff, A Generative Theory of Tonal Music, MIT Press (1983).
- [4] Jun Nishida and Kenji Suzuki, bioSync: A Paired Wearable Device for Blending Kinesthetic Experience, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '17), ACM, pp.3316-3327 (2017).
- [5] Keiji Hirata and Shu Matsuda, Interactive Music Summarization based on Generative Theory of Tonal Music, Journal of New Music Research, Vol.32, No.2, pp.165-177 (2003).
- [6] Keiji Hirata and Rumi Hiraga, Ha-Hi-Hun plays Chopin's Etude, In Working Notes of IJCAI-03 Workshop on methods for automatic music performance and their applications in a public rendering contest, pp.72-73 (2003).
- [7] 奥村健太, 竹川佳成, 堀内靖雄, 橋田光代, 評価のための問題設定: 演奏支援システムの事例から, (社) 情報処理学会 音楽情報科学研究会, 2014-MUS-102, No.20, pp.1-6 (2014).
- [8] Masataka Goto, Isao Hidaka, Hideaki Matsumoto, Yosuke Kuroda, and Yoichi Muraoka, A Jazz Session System for Interplay among All Players - VirJa Session (Virtual Jazz Session System), In Proceedings of the 1996 International Computer Music Conference (ICMC '96), pp.346-349 (1996).
- [9] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, Melody Expectation Method Based on GTTM and TPS, In Proceedings of the 9th International Conference on Music Information Retrieval conference (ISMIR '08), pp.107-112 (2008).
- [10] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, deepGTTM-III: Multi-task Learning with Grouping and Metrical Structures, In Proceedings of the 2017 International Symposium on Computer Music Multidisciplinary Research (CMMR '17), pp.238-251 (2017).
- [11] 橋田光代, 鈴木泰山, 奥村健太, 馬場隆, 柴崎正浩, 生成音楽評価の20年, (社) 情報処理学会 音楽情報科学研究会, 2014-MUS-102, No.18, pp.1-4 (2014).
- [12] Niels Henze and Kai Kunze, A Dagstuhl Seminar Looks beyond Virtual and Augmented Reality, Journal of IEEE Multi-Media, IEEE, Vol.24, Np.2, pp.14-17 (2017).
- [13] Nobuhiro Takahashi, Hayato Takahashi, and Hideki Koike, SoftExoskeleton Glove Enabling Force Feedback for Human-Like Finger Posture Control with 20 Degrees of Freedom, In Proceedings of IEEE World Haptics Conference (WHC '19), IEEE, pp.217-222 (2019).
- [14] Pat Metheny, Orchestration - An Excerpt from The Orchestration Project, Retrieved August 21, 2021 from <https://youtu.be/evHVh4bqa0Q> (2013).
- [15] Pedro Lopes, Max Pfeiffer, Michael Rohs, and Patrick Baudisch, Hands-on Introduction to Interactive Electrical Muscle Stimulation, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '16), ACM, pp.944-947 (2016).
- [16] Yoichiro Kawaguchi, Squarepusher × Z-MACHINES, Retrieved August 21, 2021 from <https://youtu.be/VkUq4s04LQM> (2013).
- [17] Yoichi Nagashima, Bio-sensing systems and bio-feedback systems for interactive media arts, In Proceedings of the Conference on New Interfaces for Musical Expression (NIME '03), ACM, pp.48-53 (2003).
- [18] 寺澤洋子, 星-柴玲子, 柴山拓郎, 大村英史, 古川聖, 牧野昭二, 岡ノ谷一夫, 身体機能の統合による音楽情動コミュニケーションモデル, 日本認知科学会論文誌「認知科学」, Vol.21, No.1, pp.112-129 (2013).