

第5回 MUS/EC研究会 萌芽・デモ・議論セッション

北原 鉄朗¹ 園部 和夫² 岩本 祐輝¹ 大下 沙偉¹ 次田 直樹¹ 島村 美羽¹ 阿久井 愛¹ 山川 峻¹
杉浦 磨矢¹ 円舘 健太¹ 渡邊 みさと¹ 森 鈴果³ 藤坂 亜南³ 菅野 翔平³ 小林 光⁴ 甲斐 大雅⁵

概要：2021年3月に開催される情報処理学会 音楽情報科学（MUS）研究会・エンターテインメントコンピューティング（EC）研究会 合同研究発表会では、第5回となる「萌芽・デモ・議論セッション」を実施する。本稿では、同セッションで発表される予定の14件の研究発表の概要を紹介する。

はじめに

国内で行われる研究会の1つの醍醐味は、発表者と聴講者がじっくりと議論してともに研究を育てていくことであろう。もちろん成果ががっかりとまとまっているハイレベルな研究発表を聴くのもいいのだが、まだまだ伸び代のある研究に対して意見や提案をしてその成長を見守るのも、研究会ならではの魅力がある。そのためには、長い時間をかけ、可能であればプロトタイプのデモを見ながらじっくりと議論できるのが望ましい。そのような観点から、MUS・EC両研究会では、定期的にデモセッションを実施してきた。特にMUS研究会では、2019年11月より「萌芽・デモ・議論セッション」と名前を変えて、デモを伴わない発表も積極的に取り入れてきた。

今回は、その「萌芽・デモ・議論セッション」の5回目である。MUS単独の研究会で実施した第3回、第4回と同様、現地での開催は叶わず、Zoomのブレイクアウトルーム機能を用いて実施することとなった。募集期間が短かったために、MUS研究会運営委員の研究室からの発表が主になってしまったが、それでも14件の発表が集まった。信号処理や音楽分析からダンス、デジタルファブ리케이션まで、ジャンルはかなり多岐に渡る。ぜひZoomのブレイクアウトルームを渡り歩いて様々な研究に触れていただければと思う。

以下では、本セッションで発表される研究内容の概要を紹介する。申込順を基本として所属機関ごとにまとまるように一部の順序を入れ替えて並べたため、研究内容を考慮した並び順にはなっていないことをご了承いただきたい。

¹ 日本大学
² 自営業
³ 関西学院大学
⁴ 関西大学
⁵ 熊本高等専門学校

1. 非線形性を積極的に利用する信号解析方法の提案

園部 和夫（自営）

従来とは異なる「非線形性を積極的に利用する」信号解析の方法を提案する。音声、音楽、その他、波形データに関連する各分野で役立つと考えている。信号源に非線形性が作用し各帯域間の独立性が無い場合（相互変調の状態）に、観測信号の局所的な位相の進み遅れ量を用いて原信号を回復する。局所的な位相を用いることから、信号源とセンサーの間に予め測定できない未知の（パワーに関する）伝達特性があってもバイパスされる。

現状では、シミュレーション実験により、基本的な正しさを明らかにしている [1][2]。また実測データ（例えば、DCASE2020-task2-fan）に対して、元々の正解が分からない状況の中で、ある程度説明可能な結果を得ている [3]。

- [1] <https://sigprocrandwalk.hatenadiary.org/entry/20170722/1500751760>
- [2] <https://sigprocrandwalk.hatenadiary.org/entry/2020/03/22/193447>
- [3] <https://sigprocrandwalk.hatenadiary.org/entry/2020/12/16/101552>

2. ユーザーの嗜好を学習する作曲ツールの実現に向けた予備検討

岩本 祐輝，北原 鉄朗（日本大学）

作曲自体には興味があるが、作曲の方法や定石を勉強する時間を確保するのが難しく、作曲に手を出すのを躊躇する人が少なからず存在すると思われる。そのため、専門知識がなくても計算機が作曲してくれるシステムの開発が行われているが、ユーザーの嗜好を十分に反映できるものに

なっているとは限らない。そこで、本研究では、ユーザーの嗜好を学習して、ユーザーが気に入しやすい曲を自動で作成する作曲ツールの実現を目指す。

本研究室では、以前より、ユーザーが盛り上がり度の曲線を描くと、それに合わせて音素材を自動で挿入するルーブシーケンサーの開発を行ってきた。このシステムでは、ユーザーが盛り上がり度曲線というものを手動で描くと、一定間隔毎に平均の盛り上がり度を計算して、「Sequence」「Synth」「Bass」「Drums」の4パートのそれぞれに音素材が入るかどうかを決めた後に、事前に各音素材に割り振られた盛り上がり度と平均の盛り上がり度を比較して、自動で音素材を割り振られるように出来ている。本研究では、このシステムにユーザーが手動で音素材を挿入する機能を追加し、音素材挿入ログを収集・分析することで、そのユーザーの音楽的嗜好を学習することを目指している。

3. フルートの練習支援のための音響分析

大下 沙偉, 北原 鉄朗 (日本大学)

フルートの吹き方を教則本などで調べてみても、息の方向や口の形などは必ずしも具体的に書いておらず、自分で吹いてみて少しずつ調整して慣れていくという方法しかなかった。そこで、本研究では、自分が吹いた音からどのように修正すれば良い音になるかを示してくれるシステムの実現を目指している。これによって初心者は吹けるようになるまでの時間が減り、中級者は自分に変な癖がついていた時にそれを修正するのに役に立つと思われる。

現在、様々な奏者による演奏音を収集した上で、スペクトログラムや基本周波数を分析し、どんな時にどのような特徴が得られるのかを調べている。その結果、今のところ倍音が多く、基本周波数が揺れず、音量が一定であるときが良い音であると思われる。今後は、その得られた特徴から音の良し悪しを自動で判断するプログラムを作成するとともに、良い音の条件をさらに詳しく調べていく。まず、単音を判断できるプログラムを作成し、その後、音階や旋律を吹いたものでも判断できる拡張していく予定である。

4. PC用キーボードを使った演奏ソフトウェアの予備検討

次田 直樹, 北原 鉄朗 (日本大学)

PC用キーボードで音楽を演奏をする場合、ピアノやMIDIキーボードに比べて、キーが小さい、キーの押しが小さい、キーを押す強さが認識されない、横一列に並ぶキー数が少ないといった特徴がある。ピアノの演奏では、右手でメロディーを、左手で伴奏を弾くことが多い。しかし、PC用キーボードは、横一列に並ぶキー数が少ない。そのため、右手でメロディーを、左手で伴奏を弾くことを考えると、複数の列にまたがって演奏しなければならない。

複数の列にまたがる方法では、手が重なってしまうため演奏しにくい。このような理由から、ピアノと同じように演奏することは難しい。本研究では、このような制約のあるPC用キーボードを使って演奏できるようにするために必要な工夫について検討する。

5. 総譜から連弾譜への自動編曲に向けて

島村 美羽, 阿久井 愛, 北原 鉄朗 (日本大学)

オーケストラ楽曲などは、人数が必要な編成になっていることが多いため、演奏にとりかかる壁が高いことが多い。一方、ピアノソロは独りで手軽に演奏できるものの、身体制約から演奏の幅は制限される。そこで、本研究では、オーケストラ楽曲などの総譜を連弾譜に自動編曲するシステムの実現を目指す。

連弾譜に編曲するための主たる課題は、いくつもの楽器が同時に演奏する多声の旋律を、いかに4本の腕で演奏できる複雑さに落とし込むかである。その1つの手段として、音高やリズムが共通する旋律を省略することが考えられる。現在は、その第一段階として、旋律のクラスタリングを検討中である。総譜の各パートの旋律を特徴ベクトルに変換し、k-means法などを用いて旋律をクラスタリングする。

6. J-popのラテン風ピアノアレンジに向けた予備検討

阿久井 愛, 島村 美羽, 北原 鉄朗 (日本大学)

J-popなどの人気曲をピアノで演奏するのは代表的な音楽の楽しみ方であるため、様々な譜面が市販されている。なかでも、ジャズやバラードなどの他のジャンルへアレンジした譜面は、音楽の楽しみの幅を広げる意味でも、有用である。しかし、音楽には様々なジャンルがあるため、自分が演奏したい楽曲を自分が演奏したいジャンル用にアレンジした譜面を見つけるのは、容易ではない。本研究では、ジャズの一つであるラテンジャズに着目し、与えられた楽曲をラテンジャズ風にアレンジしたピアノ演奏用譜面を生成することを目指す。現在は、ラテンジャズに出現するリズムパターンに着目し、どのような条件でどのようなリズムパターンが出現するかの分析を進めている。本発表では、その途中経過を報告するとともに、その途中経過に基づいて発表者が人手で行ったアレンジの例について紹介する。

7. GTTMを用いたj-popのメロディ分析のための部分タイムスパン類似度の実装

山川 峻, 北原 鉄朗 (日本大学)

J-POPは、日本国内で最も人気のある音楽ジャンルであるが、その楽曲的特徴は時代を追うごとに変化している可能性がある。本研究では、そのような楽曲的特徴の変化を分析することを目的としている。

我々は、J-POP のメロディの構造が時代を追うごとに複雑になっているとの仮説を立てた。この仮説を実証するため、音楽理論 Generative Theory of Tonal Music (以下、GTTM という) を使用する。GTTM は、音楽の認知的構造をタイムスパン木として木構造で表すことができ、本研究に適用している。

現在は、GTTM のタイムスパン木を用いたメロディの複雑さの分析に向け、浜中らが提案した部分タイムスパン類似度の実装を進めている。本発表では、その途中経過について報告する。

8. 仮想スタジオを用いた遠隔ダンスレッスンシステムの実現に向けて

円舘 健太, 杉浦 磨矢, 北原 鉄朗 (日本大学)

本研究では、仮想スタジオを利用したダンスレッスンを家や遠隔地でも楽しめるシステムの実現を目指している。現在、ダンスは体育の必修科目となっており、メジャーなスポーツになりつつある。しかし、新型コロナウイルスの影響により外出ができない、都市部から離れていて有名な先生のレッスンを受ける事ができないという問題がある。そこで、仮想空間を用いて目の前に受講したい先生がいるかのようにレッスンを受けられるようにできれば、場所を問わずに本格的なダンスレッスンを受けることができる。また、ダンスにおいて音と動きのタイミングを合わせる事はとても重要な要素になるため、レッスン後にタイミングに関してのフィードバックを行う事ができれば、音と振り付けのタイミングの違いを自分自身で気づくことができる。

本研究で目指しているシステムは、講師が踊る動画を撮影し、動画を OpenPose を用いて動きを推定し CG モデルを作成する。その CG を仮想スタジオで再現する。生徒は、レッスン受講の際にはカメラで自分自身の姿を写せる状態でレッスンを受ける。生徒の動きと先生の動きの音とのズレをシステムの FB として返す。現在は、このシステムの実装に向けて、Unity を用いた仮想スタジオの実装、および、OpenPose を用いた講師役のダンスに見立てた動画の分析を進めている。本発表では、現時点での成果物について報告する。

9. 管楽器を対象とした基礎的演奏技術向上のための音響特徴可視化アプリの試作

渡邊 みさと, 尾上 洋介, 植村 あい子,
北原 鉄朗 (日本大学)

管楽器は、発音そのものが簡単ではないため、音の強さ・高さ・音色が安定した音を出せるようになるための基礎的な反復トレーニングが不可欠である。しかし、初心者にとって、これらが安定しているかを音から判断するのは難しく、初心者が独力で練習する 1 つの障壁になってい

た。本研究では、管楽器を練習したいユーザが気軽に録音し、音の強さ・高さ・音色の安定度を視覚的に確かめられる Web アプリ [1] の開発を進めている。これらの音響特徴をグラフ化するだけでなく、継続的に録音を行うことで、これらの音響特徴の安定度がどのように改善しているかを可視化する機能の実装も進めている。本発表では、本アプリのプロトタイプについてデモを交えて報告する。

[1] <https://musicvis.vdslab.jp/>

10. 恋ダンスの達人：深層学習を利用したダンスのキレ採点システム

森 鈴果, 田中 哉汰 (関西学院大学),
橋田 光代 (福知山公立大学),
片寄 晴弘 (関西学院大学)

“踊ってみた動画” が動画共有サイトで上位ランキングの常連となっているように、ダンスを踊ること、共有、閲覧することが身近なエンタテインメントとして成立している。一方、もう一つの自己表現系の身近なエンタテインメントである歌唱では、カラオケでの採点機構が提供され、楽しみ方に新たな付加価値が生み出されている。そこで、ダンス動画にカラオケと同様の採点機構を構成することで、ダンスエンタテインメントにおける新たな付加価値を提供できると考える。

一般的にダンスの採点対象には振付の正確さ、リズム表現、オリジナリティ等が存在する。ここでは踊ってみた動画に対するコメントとして、上手さを表現する代表的なもの 1 つ “キレ” の採点を行う。採点機構の構成においては、踊ってみた動画とそのコメントに着目することで、採点に必要な大量のデータを用意する。

我々は、踊ってみた動画とコメント中の “キレ” というキーワードの有無を基準として、キレの判別器を製作した。デモセッションでは、キレ判別器を応用したダンス採点システム “恋ダンスの達人” を紹介する。

11. フレーズ表現の意図伝達に関する Web 聴取実験システム

藤坂 亜南 (関西学院大学), 橋田 光代 (福知山公立大学),
片寄 晴弘 (関西学院大学)

演奏を聴く人々は、聴こえてきた音から感情、表現といった演奏者の意図を感じ取っている。同じ曲でも演奏者によって聴こえ方が違うのは、各演奏者の表現意図が異なるからであり、聴取者はその違いを聴き分けるのも楽しみの一つである。しかし、演奏者の意図が聴取者に正確に伝わっているかはよく分からない。本研究では意図の中でも「フレーズ構造」に焦点を当てて、演奏者の意図が聴取者に伝わる様相に関する検討を実施する。この際、聴覚情報に加えて、演奏者の動きといった視覚情報がどのような影響

を与えるのかについての分析も実施する。「フレーズ構造」に関する意図伝達の度合いについては、演奏者および聴取者に自身が解するフレーズ構造の境界および頂点音位置の聞き取りを実施し、それらの調和平均で算出する。聴取者には、実際の演奏を聴きながらフレーズ構造を記述してもらう必要があるが、演奏を見てフレーズ構造を書くというタスクを同時に行わなければならない。またコロナ禍の影響で実験形態をオンライン化に対応しなければならない。ここでは、この目的に向けて作成した被験者実験用 Web システムを紹介する。

12. J-POP におけるメロディの音高上昇と発声法の関係の定量的分析

菅野 翔平, 片寄 晴弘 (関西学院大学)

最近のポップス音楽が歌いにくくなっているという声をよく耳にする。その主要因にボーカルのメロディの音高上昇が考えられているが、学術的な視点から調査したものは少ない。高音域の発声法の一つにミックスボイスが挙げられるが、曖昧な要素の多い技術であり、データは不足しているのが現状である。

そこで本研究では、J-POP のアーティストが実際にミックスボイスで歌唱することが増加しているのか確認し、発声法の変化の視点を含めて、メロディの音高上昇の状況を定量的に分析することを目的とする。

本研究のアプローチとして、ヒットチャートに入る人気楽曲を対象に、ラベル付けしたデータから発声法の推移の確認、楽譜データより抽出した音楽特微量から、年代ごとのメロディの音高上昇を分析した。

結果として、時代の流れに合わせて男性ボーカルのミックスボイスの使用率が上昇し、地声がミックスボイスに切り替わりつつあることを確認した。

また、楽曲の最高音の上昇、最低音の下降、標準偏差の値から男性ボーカルの音域が上下に広がっていることを確認した。

発声法の推移を示したデータと、メロディの音高上昇を示したデータは強い正の相関があり、J-POP において両者の相関関係が確認された。

13. ものづくりにおける基礎技能の獲得と意欲維持を目的としたデジタルファブリケーションプロセスの設計

小林 光, 松下 光範 (関西大学)

本研究の目的は、デジタルファブリケーションにおけるものづくりに対して初学者が学習を始める段階へ到達することを目指し、そのために必要な基礎技能の獲得と創作意欲の維持を目的とした体験プロセスを設計することにある。デジタルファブリケーションとはデジタル工作機

器によってデジタルデータをもとにものづくりを行う技術である。近年ではデジタル工作機器が安価になり一般家庭での利用が可能になっている。しかし、デジタル工作機器を利用したものづくりへの新規参入は難しいとされている。これはもととなるデータ作成の困難さ、作品の構想づくりへの先入観が原因として挙げられる。この原因となるデータ作成の困難さや作品の構想づくりはものづくりに対する創作の意欲に影響し、この意欲の欠如が初学者の新規参入の難しい要因となっている。そのため、初学者におけるデジタルファブリケーションプロセスを設計し、基礎技能の獲得や意欲維持が可能な物づくり体験を企図する必要がある。そこで本稿では、デジタルファブリケーションに利用される 3D プリンタを対象に、初学者が一通りの創作活動を行うことが可能な導入体験を企図することで基礎技能の獲得と意欲維持を目指す。先行研究から初学者の技術的障壁、必要な技能、作品の構想に必要な要素を用いて体験の企図を行った。本稿の主な体験としてゲームコンテンツを活用した空間認識能力の獲得、創造性のルーブリックを基にしたアイデアノートによる作品の構想づくり、作品の構想から簡易 3D モデリングによる制作の 3 つの体験を企図している。

14. 貪欲法を用いた電子楽器の音色制御パラメータの最適化

甲斐 大雅, 藤井 慶 (熊本高等専門学校)

現代の音楽制作においてシンセサイザは欠かせない電子楽器の一つである。シンセサイザの特長は様々な音色を作成できることだが、音作りの自由度が高い分操作が複雑であるため、特に初心者にとって制御が難しいという問題がある。またシンセサイザを用いる理由の一つとして、生楽器とは違う独特の質感が得られるという点も挙げられる。仮にシンセサイザで生楽器と同一の音色が得られる場合、独特の質感はその音色の近傍で得られる可能性がある。これらのことを踏まえ、本研究ではこの問題を最適化問題の 1 つとして捉え、音色制御パラメータを自動的に最適化する乗算方式シンセサイザを実現することを目的とする。

最適化には MFCC のユークリッド距離の最小化を指標とした貪欲法を用いた。乗算方式シンセサイザの音色空間は多峰性であるため、遺伝的アルゴリズムを適用する研究例が複数あるが、今回用いる貪欲法は単峰性の最適化手法である。貪欲法は距離が単調減少し比較的少ない回数で収束するため、最適化に近づく過程を再現しやすい。このことから、見本音に似てかつ独特の質感のある解が得られることを期待し、貪欲法を採用した。局所解に陥る問題については初期値候補群をランダムに 50000 個用意することで対応した。