

# 第4回音楽情報科学 国際会議既発表・萌芽・デモ・議論セッション

齋藤 大輔<sup>1</sup> 水野みか子<sup>2</sup> 山本紘征<sup>3</sup> 小椋裕太<sup>4</sup>  
松原正樹<sup>5</sup> 佃 洸撰<sup>6</sup> Tsung-Ping Chen<sup>7</sup> 高瀬遥矢<sup>8</sup> 平賀瑠美<sup>9</sup> 日下湧太<sup>10</sup> 植村あい子<sup>11</sup>  
渡辺みどり<sup>12</sup> 山口恭正<sup>13</sup> 田中瑞穂<sup>14</sup> 城田晃希<sup>14</sup> 能登楓<sup>14</sup>

概要：国際会議で先行して発表したが、日本の国内学会・研究会では未発表の研究について紹介する、国際会議既発表セッションと、まだ十分な成果の出していない萌芽期の研究やデモシステム、既発表ながら改めて参加者と議論したい研究などを披露する萌芽・デモ・議論セッションを行う。

## はじめに

本稿では2019年1月1日から2020年11月1日までに開催された音楽情報科学に関連する国際会議のうち、国内未発表の研究を紹介する「国際会議既発表セッション」における研究紹介と、十分な成果は出ていない萌芽的な研究やデモシステム、既発表ながら改めて参加者と議論したい完成度の高い研究などを一同に会する「萌芽・デモ・議論セッション」での発表を掲載する。オンライン開催に伴い、研究会のあり方が見直されていく中、今回は「国際会議既発表セッション」に10件、「萌芽・デモ・議論セッション」に5件の参加申し込みがあった。本セッションのような企画を通して、研究会における学術的活動の活性化とその活力の個々の研究者への還元がなされることを願っている。

(齋藤 大輔)

## 1. 国際会議既発表紹介

### 1.1

**The aesthetics of notation in Japanese Electroacoustic Music**, Mikako MIZUNO, *Musica/Tecnologia*, 2020, 13(1), pp.103-118, 2020

本研究は、日本の電子音響音楽を、記譜法とその根元にある美学との関係という視点から考察する。具体的には、現代のオリジナル記譜法が様々に試みられた1960年代から70年代に作曲されたライブ・エレクトロニック作品を取り上げ、西欧記譜法に適合されえない音像と時間構成を分析記述する。湯浅譲二、柴田南雄、藤田正典によって作曲された電子音響音楽作品には、日本的時間感覚や邦楽の新展開をめざす新邦楽との連動を指摘することができる。(紹介者: 水野みか子)

<sup>1</sup> 東京大学  
<sup>2</sup> 名古屋市立大学  
<sup>3</sup> 北陸先端科学技術大学院大学  
<sup>4</sup> 東京理科大学  
<sup>5</sup> 筑波大学  
<sup>6</sup> 産業技術総合研究所  
<sup>7</sup> Academia Sinica  
<sup>8</sup> 名古屋工業大学  
<sup>9</sup> 筑波技術大学  
<sup>10</sup> 東京工業大学  
<sup>11</sup> 日本大学  
<sup>12</sup> お茶の水女子大学  
<sup>13</sup> 東北大学  
<sup>14</sup> 公立はこだて未来大学

## 1.2

### **Jazz Harmony Analysis with $\epsilon$ -Transition and Cadential Shortcut**, Hiroyuki Yamamoto, Yui Uehara, Satoshi Tojo, SMC, 316, 2020

Tonal Pitch Space (TPS) は、二組の key と degree のペアに対して、聞き手にとっての自然さの度合いを反映した数値的な距離を与える理論である。これによって、コード列に対する最も自然な解釈列を求めることができる。本研究ではまず TPS を拡張し、jazz で重要となるマイナースケールの区別や四和音に対応できるようにする。次に、一つのコードに二つの解釈を与えることを可能にする  $\epsilon$ -transition を導入する。これによってピボットコードによる転調に対応する。さらに、明示的に cadence の成立を表現する cadential shortcut を導入する。これによって方向への依存性と tri-gram 以上の関係も表現される。これらの拡張により、jazz ハーモニー分析の表現力を向上し、同時に最短経路数を大幅に削減することができることを示す。(紹介者: 山本紘征)

## 1.3

### **Expectation-based Parsing for Jazz Chord Sequences**, Yuta Ogura, Hidefumi Ohmura, Yui Uehara, Satoshi Tojo, Kouichi Katsurada, Sound and Music Computing Conference, 350-356, 2020

計算論的音楽理論では、自然言語における文法理論を活用し、楽曲の階層的構造を明らかにしてきた。例えば、和声解析では、生成文法に基づく解析により和声構造を木構造によって表現する。これは楽曲聴取後の認知構造を表している。一般的に人間の音楽認知は楽曲の進行に基づき行われるが、これらの分析手法は楽曲の進行に伴う逐次的な処理を考慮していない。そこで、私たちは楽曲の進行に着目し、和声進行を漸進的に解析する枠組みを提案する。これにより、進行と関わりの深い、人間の音楽的期待に基づく解析が期待できる。本稿では、和声の漸進的解析法を詳述し、ジャズ楽曲における和声進行の解析結果から漸進的解析の有意性を論じる。(紹介者: 小椋裕太)

## 1.4

### **Auditory Gestalt Formation for Exploring Dynamic Triggering Earthquakes**, Masaki Matsubara, Yota Morimoto, Takahiko Uchide, The 14th International Symposium on Computer Music Multidisciplinary Research (CMMR2019), 983-987, 2019

本論文は動的誘発地震の探索のための地震波可聴化に関する地震学者とコンピューター音楽作曲家との共同研究成果を報告する。500 を超える観測点の地震波形信号から動的誘発地震の検知を支援するデータ可聴化手法を提案した。本手法は視覚による波形プロットを頼ることなく、またフィルタをかける前処理を行うことなく、人間の認知特性を活かして聴覚情報から検出できるようにすることが特長である。具体的にはゲシュタルト認知の類同の原理を用い、同じような動きしている時系列データが自然に浮かび上がってくるような可聴化マッピングを提案している。可聴化における音デザインの考案、およびインタラクティブな可聴化システムを構築し、2011 年東北地方太平洋沖地震のデータを元に有効性を検証した。(紹介者: 松原正樹)

## 1.5

### **Analysis of Song/Artist Latent Features and Its Application for Song Search**, Kosetsu Tsukuda, Masataka Goto, the 21st International Society for Music Information Retrieval Conference (ISMIR 2020), 2020

ユーザへの楽曲推薦の有効なアプローチのひとつに、アーティストと楽曲を潜在ベクトルで表現してユーザの楽曲に対する好みを推定する方法がある。本研究ではそれらの潜在ベクトルの位置関係に基づき、アーティストと楽曲の関係性として overall similarity と prominent affinity の二種類を提案する。前者はアーティストが持つ性質と楽曲との総合的な類似度であり、後者はアーティストが持つ特徴的な性質を楽曲が持つ度合いである。これらの関係性を用いることで、新しい観点から楽曲検索が可能になることを示す。具体的には、ユーザのアーティストへの親密性に基づく検索、アーティストにとっての楽曲の典型性に基づく検索、アナロジーに基づく検索の 3 種類の検索を提案する。(紹介者: 佃 洗撰)

## 1.6

**Chord Jazzification: Learning Jazz Interpretations of Chord Symbol**, Tsung-Ping Chen, Satoru Fukayama, Masataka Goto, Li Su, the 21st International Society for Music Information Retrieval Conference (ISMIR2020), 2020

コードネームは主にポピュラー音楽の和音の根音と種類を表示することに用いられるが、和音構成音の配置（どのオクターブで演奏するかなど）が演奏者に委ねられているなど、簡略化された和音の表現になっている。コードネームのみを生成する生成手法では、ジャズなどで重要な和音構成音の配置まで含めた自動生成ができず、幅広い音楽ジャンルを扱えなかった。そこで本研究では、和音の構成音とその配置をジャズのスタイルで生成できる Jazzification を提案する。Jazzification は Coloring（色付け）と Voicing（配置）の二段階のプロセスに分けられ、深層学習にもとづいて生成する。Coloring では colortones と呼ばれる和音構成音を生成し、Voicing では和音構成音の配置を行う。コードネームの解釈に重点を置いたポップス・ジャズ楽曲のデータベースを作成して実験を行なった結果、1) 和音構成音の配置では Coloring と Voicing の二段階処理による生成が end-to-end による生成より性能が良いこと、2) コード進行の生成では注意機構を用いた生成が再帰的ニューラルネットワークを用いた生成より性能が良いことがわかった。（紹介者: Tsung-Ping Chen）

## 1.7

**Support System for Improvisational Ensemble Based on Long Short-Term Memory Using Smartphone Sensor**, Haruya Takase, Shun Shiramatsu, Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 394–398., 2020

本研究では、コード進行の知識がなく楽器演奏の経験がない音楽初心者のための即興演奏支援システムを開発することを目的としている。体の動きを用いて直感的に旋律概形や発音タイミングを表現出来るように支援する演奏インタフェースの実現を目指す。本システムは、Android スマートフォンのセンサーや背景楽曲のコード進行を入力に Long Short-term Memory に基づいて発音タイミング、上下動、音高をそれぞれ推定するモデルの三つのモデルにより構成されている。テストデータを用いて精度評価を行った結果、発音タイミングの精度は十分であるが、上下動推定とピッチ推定の精度を向上させる必要があることが分かった。（紹介者: 高瀬遥矢）

## 1.8

**Computer-Based Music Training with Hearing Impairments: Lessons From an Experiment**, Rumi Hiraga, Kjetil Falkenberg, SMC2020, 426-433, 2020

補聴器具の性能向上は音声に関して著しいが、非言語音（音楽や環境音）聴取に関する研究はあまりなされていない。本研究では、以前開発した音楽ゲームを聴覚障害者が一定期間使用する実験を行った。実験の前後、途中で聴取に関するテストを行い、音楽ゲームを使用することで音聴取能力の向上の可能性があるか、また、トレーニングとして音楽ゲームを使用する時の問題点は何があるか、について述べた。（紹介者: 平賀瑠美）

## 1.9

### ONSET-INFORMED SOURCE SEPARATION USING NON-NEGATIVE MATRIX FACTORIZATION WITH BINARY MASKS,

Yuta Kusaka, Katsutoshi Itoyama, Kenji Nishida and Kazuhiro Nakadai, The 23rd International Conference on Digital Audio Effects, 2020

本稿ではバイナリマスク付き非負値行列因子分解(NMF)に基づく、分離したい楽器音の発音時刻を利用する新しい音源分離手法について説明する。提案法で利用する発音時刻は、ユーザが楽曲を聴取しながらデバイスをタッピングするなどして作成可能な準備に手間がかからない有用な情報である。既存のNMFベースの音源分離モデルにおいて発音時刻を利用するため、楽器音のオン/オフを表現するバイナリマスクを導入して拡張を行い、マスク上で表現された発音時刻を含めてモデルを推定するアルゴリズムを導出した。ユーザが発音時刻を作成する際に予想されるずれなどのノイズも考慮した分離実験を行い、提案法の分離精度と安定性の評価を行った。(紹介者: 日下湧太)

## 1.10

### Morphing-Based Reharmonization using LSTM-VAE, Aiko Uemura, Tetsuro Kitahara, The 2020 Joint Conference on AI Music Creativity, 2020

本研究では、Long Short-Term Memory (LSTM) と Variational Autoencoder (VAE) を用いて、コード進行間のモーフィングに基づくリハーモナイゼーションを行う方法を提案する。リハーモナイゼーションとは、ある楽曲のコード進行を変更することで楽曲の雰囲気などに変化をつけることをいう。リハーモナイゼーションが自動化されれば、初心者でもハーモニーの印象を簡単に変えることが可能になる。本研究では、メロディの条件付けを行った LSTM-VAE の潜在空間において、2つのコード進行間のモーフィングを行った。実験の結果、メロディに適合したコード進行を生成することができた。(紹介者: 植村あい子)

## 2. 萌芽・デモ・議論セッション

### 2.1

メタ情報と特徴量の関係把握のための楽曲群の可視化  
渡辺 みどり, 黒子 なるみ (お茶の水女子大学), 大矢 隼士 (株式会社レコチョク), 伊藤 貴之 (お茶の水女子大学)

近年、音楽の自動分類や推薦といった研究およびサービスが盛んである。しかし、分類や推薦の基準とするものにはメタ情報や特徴量など様々存在し、それによって結果は変化する。また、分類後に何をもって成功とするのかを判断することは困難である。分類結果の判断をするためには、可視化することが有用であると考えられる。本研究では、メタ情報(タグ)を付与した楽曲群について、音楽解析ツールや機械学習を用いて特徴量を算出して、得られた特徴量に基づいて分布を可視化し、結果の観察を行う。これを通して、音楽分類タスクに効果的なメタ情報、特徴量、機械学習手法、可視化手法を検討すること、音響特徴量とメタ情報の新たな関係を発見することを目指す。用いるメタ情報や特徴量の算出方法、次元削減や可視化の手法を変えることによって、同一のタグを持つ楽曲が同じクラスタに分類されるか、より良い性能を有する特徴量や機械学習はどれか、などを観察する。現時点では、年代や作曲家のメタ情報、テンポや音量平均などの特徴量、機械学習などを用いた実験を行った。本発表では、既に行った可視化結果を紹介し、より有効なメタ情報や特徴量について議論したい。

### 2.2

#### 演奏会レパートリーデータベースの構築に向けて

山口恭正 (東北大学)

音楽の持つ意味や役割は時代とともに変化し、オーケストラの演奏曲目も時代や社会とともに変化している。しかし、それらを実証的に分析した研究というものは少ない。近年の行政の文化事業に関する法律や制度の改正、情報通信技術の発展、音楽と聴衆の社会的な役割や性質の変化を背景として、アートマネジメントおよび社会学、データ科学の観点から、日本におけるオーケストラ作品の演奏記録およびその軌跡を見直す必要があり、昨年度の研究では日本国内のプロオーケストラの活動のメインである定期演奏会の演奏記録に着目し、その特質や傾向を追った。研究を進める過程で、オーケストラ業界全体としての情報社会への対応の遅れや、電子データそのものを保持していないケースが目立った。しかし各オーケストラの演奏記録を適切な県家データベースとして構築すれば、その可視化や分析を通してプロモーションやマーケティングにつながると考えられる。特に、Web 上での可視化と相性が良く、レパートリーデータベースは多くのアマチュア音楽愛好家に興味を持ってもらえるコンテンツになる可能性を秘めている。本研究は、オーケストラの演奏会記録に適した関係データベースの構築を模索するものである。

## 2.3

### 合成音声歌唱に固有なヴィブラートの楽譜レベルでの学習について

田中 瑞穂, 竹川 佳成, 平田 圭二 (公立はこだて未来大学)

近年, 合成音声歌唱を用いた楽曲が増加している。また, 合成音声歌唱に人間らしい歌唱を付与するシステムも提案されている。しかし, 合成音声歌唱には, 人間の歌唱では表現できない独特な歌唱技術が存在する。例として, 8分音符以下の短い音に対するヴィブラートの付与, オクターブ単位の極端なピッチ変化が挙げられる。本研究では, 機械学習を用いて合成音声歌唱特有の歌唱技術の自動付与を目指す。今回のデモ発表では, 以下3点について述べる。1つ目は, 合成音声歌唱特有の歌唱技術の統計である。ヴィブラートや極端なピッチ変化がそれぞれの音価や音高でどの程度出現しているか, 合成音声歌唱を用いた楽曲を集めて統計をとった結果を報告する。2つ目は, LSTMを用いたヴィブラート付与モデルの実装である。本研究では, LSTMを用いた基本的な学習モデルに対して, 2種類の入力データ表現形式を導入し比較した。3つ目は, さらに正確なヴィブラート付与を実現するため, Attentionを導入したモデルを提案する。Attentionとは特定の箇所に着目することであり, 導入することで精度の上昇が見込める。今回のセッションでは, どのような箇所に着目すべきか議論したいと考えている。

## 2.4

### 音楽的協和性を加味した条件付き変分オートエンコーダによる非和声抑制方法の検討

城田晃希, 平田圭二, 竹川佳成 (公立はこだて未来大学)

近年, 深層学習を用いて音楽生成を行う研究が隆盛である。変分自己符号化器 (Variational Autoencoder) を用いた旋律生成法として, 訓練データに含まれる特徴を埋め込んだ潜在空間を学習したのち, その潜在空間上の適切な座標を選択して旋律を生成する方法が考えられる。この時, 出力の旋律は確率的に生成されるため, 非和声を含む可能性があり, それが自然聞こえる場合もあれば, 不自然に聞こえる場合もある。そこで本研究では, 条件付き変分自己符号化器 (ConditionalVAE) を用いて, 音楽的に自然・不自然な非和声を制御する方法を提案する。具体的には, CVAEが旋律を学習する際に, 旋律の和音に対する音楽的協和度をCVAEの条件信号として与える。旋律生成時, CVAEの条件信号を変化させることで, 生成される旋律に含まれる非和声を制御する。本デモでは, 提案モデルを用いて楽曲生成のデモンストレーションを行い, いくつかの実験結果を示し, 考察を加える。

## 2.5

### 暗意実現モデルに基づき音どうしの抽象的な関係性に着目した旋律特徴量の検討

能登楓, 平田圭二, 竹川佳成 (公立はこだて未来大学)

本研究の目的は, 音の参照関係に着目した旋律の特徴量設計である。これまでの旋律の学習では, 旋律の期待感や終止感といった, 音どうしの抽象的な関係性を学習することは困難であった。そこで, 本研究では認知的根拠を持つ音楽理論の一つである暗意実現モデル (IRM) に着目する。IRMでは, 先行する2音と後続する1音の参照関係を, 音程と音方向に関する2つの原理によって暗意の実現, 否定に基づく8つのシンボルへと分類する。IRM分析を旋律に適用することで得られるシンボル列は, 音の参照関係を抽象化の特徴量であることが期待される。本研究では, 旋律をシンボル列へと変換するIRM分析器の作成を行い, 旋律に対するIR分析結果に深層学習を適用し, 新しい旋律特徴量を検討する。発表では, IRM分析器の提案および楽曲に対する分析結果の提示, IR分析結果を旋律特徴量とした際の, Variational Recurrent Auto-Encoder (VRAE) における潜在空間の可視化によって, 提案する特徴量の効果を検討する。