

第1回 音楽情報科学 萌芽・デモ・議論セッション

大村 英史^{1,a)} 沖 貴司^{2,b)} 峯村 開^{3,c)} 近藤 芽衣^{4,d)} 北原 鉄朗^{5,e)}

概要：音楽情報処理の研究分野のさらなる発展に向けて、同分野に関する研究を自由な形式で議論し合える「萌芽・デモ・議論セッション」を実施する。本稿では、このセッションで発表される予定の研究について、各発表者による概要を紹介する。

音楽情報科学 萌芽・デモ・議論セッションの開催によせて

大村 英史（東京理科大学）

萌芽・デモ・議論セッションは、音楽情報科学における十分な結果は出ていない萌芽研究から、他の研究会や国際会議や論文で既発表ではあるが本研究会で共有・議論したい完成度の高い研究まで、ポスターやデモンストレーションなど発表者が自らで選択した発表形式で紹介が行われる。自由度の高い発表形式であるため研究会参加者と発表者が直接議論できる点が、一般発表にはない本セッションの魅力である。情報処理分野における多くの研究集会では近年、ポスターセッションなどの形でカジュアルな議論の場を設けているが、それと比べて本セッションは、ポスターのみならず、システムの実演と体験のような自由な発表スタイルが特色の一つである。本セッションは、いままで毎年の夏開催の研究発表会での恒例イベントであったデモンストレーションセッションの発表形式をさらに自由度を高めたセッションである。

本セッションは一般発表とは異なり、原稿の提出は題名・著者名・アブストラクトのみであり、また原稿提出期限に近い日まで申し込みを受け付けている。このような形で発表申し込みの敷居を下げることで、できるだけ気軽に発表できるようにしている。また、発表会場の質疑や議論においても親しみやすい雰囲気づくりを心がけている。さらに、発表会場のスペースや発表者の希望次第では飛び入り

での発表も受け付けている。つまり、状況によっては原稿や説明資料がない状態でも発表に参加することができる。

本セッションでは、萌芽研究、デモンストレーション、参加者と議論したい研究など歓迎している。具体的には、

- 未完成だが現時点での内容について議論したい研究
- 研究的要素がはっきりしていない思い付きで作ったシステム
- 作ってみたものの一般発表に至らなかった研究
- 楽器演奏などパフォーマンスを伴う研究
- 実際に触れてみないと実感しにくい研究
- 同じ研究会にて一般発表もなされるデモンストレーション
- 他で既発表であるが、本研究会参加者と情報共有したい研究

などが上げられる。

また、一般発表に至らない研究は、高い将来性が期待される可能性を秘めている場合も多い。このような研究に対して発表の機会を得ることで、参加者との議論から研究の新しい展開を切り拓けるような効果も考えられる。このようなことから、本セッションの定期的な実施こそ、音楽情報科学における多様な研究の発展に寄与するものであると考える。

音楽情報科学研究会では、一般発表の動画中継およびアーカイブ化などを通して、発表会場に来場できない方々とも研究成果を共有する取り組みを進めてきた。しかし本セッションの発表者が実際にデモンストレーションを実施する場面は動画中継されない。よって、本発表に参加できるのは発表会場での参加者のみである。また、アーカイブ化された発表からは、発表会場の音場や残響を完全に再現することは難しい。つまり、研究成果を忠実に体験したければ発表会場に足を運ぶのが一番の手段ということになる。本セッションは同じ空間を共有する発表者と参加者による体験型の発表を実現する機会である。研究成果を体験しな

¹ 東京理科大学
² 高知工科大学
³ 慶應義塾大学
⁴ お茶の水女子大学
⁵ 日本大学
a) hidefumi.ohmura@gmail.com
b) 235057g@gs.kochi-tech.ac.jp
c) minekai374@gmail.com
d) g1620517@is.ocha.ac.jp
e) kitahara@chs.nihon-u.ac.jp

がらのインタラクティブな議論を通じて、一般発表とは異なる体験や意見交換の場となることを期待する。それと同時に、本セッションが本研究会への参加のモチベーションの一つになることを期待する。

以下、各発表者による概要を紹介する。

Coconet を用いたポピュラー音楽におけるメロディからのコード進行生成の検討

沖 貴司, 高田 喜朗 (高知工科大学)

本研究では, Huang らによる Coconet を用いて, ポピュラー音楽におけるメロディから適切なコード進行の自動生成が行えるかどうかを検討した。Coconet は, バッハの賛美歌の様式の曲を自動生成する畳み込みニューラルネットワークである。Coconet をポピュラー音楽のコード進行生成に用いるために, ポピュラー音楽のデータを収集してデータセットを作成し, また, コード進行のデータ表現を検討した。コード進行のデータ表現については, コンヴェールらの先行研究で検討されている 4 種類のデータ表現を用いた。Coconet を用いて適切なコード進行が生成可能かどうかを調査するための予備実験として, データセットから Bm(b5) を除く C メジャーのダイアトニック・コードのみを使用した 8 小節の長さのデータを用いて学習を行い, Lim らによるコード進行生成の先行研究と精度 (accuracy) の比較を行った。その結果, 2 種類のデータ表現に関して, 先行研究を上回る精度を達成した。

CNN・CRNN による音楽ジャンル分類の実装と LINE Bot アプリによるデモ

峯村 開 (慶應義塾大学)

① 先行研究の実装

音楽のジャンル分類を CNN と CRNN の 2 つのモデルで実装し, 精度の測定・ジャンル予測・ハイパーパラメータのチューニングを行った。GTZAN Genre Collection というデータセット (正解ラベルと 1 曲 30 秒の音楽ファイルのセット) を用いた。ジャンルは Blues, Classical, Country, Disco, Hiphop, Jazz, Metal, Pop, Reggae, Rock の 10 個で, 各ジャンル 100 曲, 計 1000 曲のデータを用いた (訓練で 800 曲, テストで 200 曲)。音楽ファイルをメルスペクトログラムに変換し Numpy 配列として入力し, 精度を出力した。また, 特定の音楽ファイルに対するジャンル予測も行った。

② デモ

ジャンル分類を行う LINE Bot を作成した。ユーザが音声を録音または音声ファイルをアップロードすると, 上位

3 つのジャンルの候補とその確率を返す。①の学習済みモデルとその重みを使って, ユーザからの入力に対するジャンルの予測を行っている。

*デモでは, 実際に LINE Bot を皆様に使っていただきたいと考えております。

<http://nav.cx/oK0WoD0>

同一楽曲に対する多数の歌唱の可視化に向けて

近藤 芽衣, 伊藤 貴之 (お茶の水女子大学)

中野 倫靖, 深山 覚, 濱崎 雅弘, 後藤 真孝

(産業総合技術研究所)

「多数の歌唱者による同一楽曲」が近年 Web 上で多く公開されるようになり, 鑑賞する機会が増えた。このような歌唱群の音響データから同一楽曲に対する各歌唱者の癖や個性の違いを理解するために, 我々は以前, 同一楽曲を歌った多数の歌唱音響データからそれぞれの音高 (基本周波数) の推移を抽出し, その分布を可視化する手法を提案した。本発表ではこの研究の今後の展開について議論する。まず, 音高以外の特徴量の可視化について議論したい。例えば, 音響データから得られるスペクトル包絡, あるいは歌唱技法に関する数値指標 (ビブラートの深さ, 等), さらに各歌唱者の人気や評価を表す数値指標 (動画共有サイトでの再生数, 等) などの可視化が考えられる。続いて本研究における望ましいビジュアル表現やインタラクションについて議論したい。さらに, 歌唱群の可視化によってユーザがどんな知識を得られるか, あるいはどのような実用が考えられるか, といった点についても議論したい。

HMM を用いたジャズのウォーキングベースラインの生成

北原 鉄朗, 志賀 あゆみ (日本大学)

本研究では, 与えられたコード進行からジャズのウォーキングベースラインを隠れマルコフモデルで生成する手法を提案する。コード進行を観測系列, 生成するベースラインの各ベースノートを隠れ状態とし, 隠れ状態の設計方法として 3 つの手法 (音名を用いるもの, 音高を用いるもの, 音名と拍節位置の組み合わせを用いるもの) を提案した。客観評価および主観評価の結果, 3 つ目のものが最も高評価であった。なお, 本発表は次の発表内容を SIGMUS 参加者向けに改めて発表するものである。

Ayumi Shiga and Tetsuro Kitahara: “Generating Walking Bass Lines with HMM”, *Proc. of CMMR 2019*, pp. 83-90, Oct. 2019.

<http://www.kthrlab.jp/members/shiga/>

印象語を使って操作可能なグラフィックイコライザー実現のための一調査

北原 鉄朗, 増田 誠也 (日本大学)

饗庭 絵里子 (電気通信大学)

本研究のゴールは、アマチュア音楽家が直感的にグラフィックイコライザー (GEQ) を使えるようになることである。このゴールを実現するため、我々は、いくつかの印象語 (明るさ, 温かさなど) に対する強度を入力すると, GEQ のパラメータ (各バンドのゲイン) に自動的に変換するシステムの実現を目的とする。本発表では, その実現に向けた聴取実験の結果と簡易的なプロトタイプシステムの作成について述べる。なお, 本発表は次の発表内容を SIGMUS 参加者向けに改めて発表するものである。

Seiya Masuda, Eriko Aiba, and Tetsuro Kitahara: “An Investigation towards Verbally Controllable Graphic Equalizer for Singing Voices”, *Proc. WIMP 2019*, Paper ID 7, Sept. 2019.

<http://www.kthrlab.jp/members/masuda/web/masuda.html>