

デモンストレーション：音楽情報処理の研究紹介 XIV

阪上 大地¹ 竹川 佳成² 浜中 雅俊³ 大野 涼平⁴ 菅野 沙也⁶ 清川 隼矢⁷ 栗原 拓也⁴
黒田 元気⁷ 小池 宏幸⁸ 鈴木 潤一⁵ 土橋 彩香³ 長村 佳祐⁴ 橋田 光代⁹ 平井 辰典¹⁰

概要：本デモセッションでは、音楽情報処理の研究分野における若手研究者のさらなる発展に向けて、若手による研究事例をデモンストレーション形式で紹介する。

はじめに

阪上 大地, 竹川 佳成

デモセッションは音楽情報科学研究会（音情研）の恒例企画のひとつであり、研究者が各々に音楽情報処理システムを持ち寄り、実地でのデモを行いながら、議論と交流を深める場となっている。本セッションでは一般発表の内容を補う形で展示されるシステムから、思いつきで作ってしまったシステム、開発中のシステム、お蔵入りになってしまったがどうしても披露したいシステムなどが幅広く発表される。デモ発表では投稿プロセスが簡略化されているのも特徴で、一般発表よりもメッセを遅く設定し、原稿の分量も最小限とするなど、誰もが気軽に発表できるようとなっている。また、原稿メッセに間に合わなかった場合でも、会場のスペース次第で当日の飛び入り参加も可能である。

デモセッションの面白さを挙げ始めるときりがないが、ここで強調しておきたいのは発表の多様さである。音楽情報科学には様々なバリエーションがある。たとえば、工学の一分野である「情報処理」と芸術のひとつである「音楽」をどうまとめるか。言語処理や機械学習の手法を音楽に応用する seed-based な取り組みでもよいし、ピアノ演奏などの音楽的なタスクを支援する need-based な取り組みでもよい。評価についても、音楽の専門家による多面的な評価を行うもの、Signal-to-Noise Ratio (SNR) など客観的な評価軸を活用するもの、デモを通して新しいタスクや視点

を提案するもの、など様々である。また、参加者の顔ぶれも幅広い。本企画は当初、若手研究者の研究を紹介する場として設けられたが、現在は若い学生、キャリアのある研究者、企業で製品開発に携わる発表者などが肩をならべてデモを行う場となっている。さらには多様なバックグラウンドを持つ一般参加者と、個々のシステムが発するデモ音が加わり、会場はいつも独特の盛り上がりとなる。

さて、今年で 14 回目となるデモセッションには 11 件の応募があった。中でも声にかかわる発表が 6 件と多く、歌声情報処理の人気を感じさせる。声以外の発表の内訳は、演奏支援が 3 件、編曲支援が 1 件、楽曲推薦が 1 件であり、いずれも工夫に富んだ独創的な発表である。以降のページには、これらのデモ発表の概要を掲載した。

デモセッションで展示されるシステムのほとんどは、作曲・演奏・鑑賞の各要素に複合的にかかわっている。デモ発表としてひとまとまりのシステムを展示するため、という理由ももちろんあるが、同時に音楽情報科学の無数のタスクが相互に関連しあい、単純な細分化ができないからでもある。音楽情報科学の研究では、「演奏支援」を目指すふたつの研究が全く異なるシステムに結実したり、「演奏支援」と「楽曲解析」のシステムが同じ理論的背景を共有している、ということが数多くある。ここにも本研究分野独特の奥深さがあるといえる。

このように、小型のデモ環境を気軽に持ち込み、展示やディスカッションができることは音楽情報科学の顕著な特色である。ぜひ会場に足を運び、議論に参加していただければ幸いである。

¹ 株式会社コルグ
² 公立はこだて未来大学
³ 京都大学
⁴ 日本大学
⁵ 日本大学大学院
⁶ お茶の水女子大学
⁷ 東京工科大学
⁸ プラスアド株式会社
⁹ 相愛大学
¹⁰ 早稲田大学

韻律変換実現のための一試行：高橋みなみ風の音声を小嶋陽菜風に変えてみた

大野 涼平, 北原 鉄朗 (日本大学)

本研究では、任意の音声の韻律をユーザが指定した音声の韻律に似せる方法の実現を目指している。これまで声質を変換する研究は数多くなされてきたが、韻律の変換については、あまり研究が行われていない。我々は、韻律変換実現のための第一歩として、韻律の特徴が大きく異なる小嶋陽菜と高橋みなみ（ともに尾木プロ）を取り上げ、次のことを試みた。

- 1) アマチュア声優に同一文（ATR 音素バランス文）を、小嶋陽菜および高橋みなみの韻律にできるだけ似せて発話してもらう（以下、「小嶋風音声」「高橋風音声」と呼ぶ）。
- 2) 韻律のうち、基本周波数（F0）と話速が重要であると考え、音節ごとに F0（または F0 の変化量）と話速が小嶋風音声と等しくなるよう、高橋風音声の F0 と話速を変換する。
- 3) 被験者に変換後の音声を聞いてもらい、評価する。この結果、F0 と話速を小嶋風音声と等しくなるように変換するより、少し大きさに変換した方が、より小嶋に似ているとの結果が得られた。本デモでは、我々が収録した音声および変換後の音声を聴いていただき、より精度のよい変換の実現に向けて議論を行いたい。

入力文書の印象と感情に基づく楽曲提供の一手法

菅野 沙也, 伊藤 貴之 (お茶の水女子大学), 高村 大也

我々は文書に印象の合った音楽を生成・提供することで、文書の内容に直接言及することなくその印象を理解・伝達することが容易になり、文書を鑑賞・共有する楽しみを拡げられるのではと考えている。そこで、文書の印象や感情に基づいた楽曲生成の一手法を提案する。本手法では前処理として、日本語を構成する各単語に対する感性極性値を記録した辞書を作成する。またコード進行とリズム進行をユーザーに聞かせてその印象を数値入力させることで、コード進行やリズム進行とその印象との関係を学習させる。続いてユーザーが文書を入力すると、まず文書の形態素解析結果に対して感性極性辞書を参照することで文書の印象値を求める。続いて文書の印象値に近い印象をユーザーが有するコード進行とリズム進行を、文書の場面の前後関係も考慮しながら選出する。このようにして選出されたコード進行とリズム進行を合成することで楽曲を生成し、ユーザーごとに提供する。

初音ミクを使用した n 平均律の音響生成と音階アルゴリズム

清川 隼矢, 伊藤 彰教, 伊藤 謙一郎 (東京工科大学)

n 平均律の音高を発生させる初音ミクに、本研究者の考える音階アルゴリズムを用いて歌わせ、新たな音楽的響きを生み出す試みである。本研究者は NICOGRAPH International 2015 で 14 や 25, 100 といった様々な平均律（これを n 平均律と定義する）で音源を正確に鳴らすシステムを発表した。初音ミクにこのシステムを適用し、n 平均律の音高を発生させ、歌わせるようにする。ここで 1 つの疑問が生じる。それはどのような音階で音楽を作るべきなのか、ということである。本研究者は音階を計算するアルゴリズムと、それを計算するシステムを開発した。作曲家は n 平均律が生み出す何らかの音程の響きを重視すると本研究者は考えた。だとするならば、理想的な音階とは、その音程が発生しやすい音階なのではないか、という仮説が立つ。本アルゴリズムは、ピッチクラス集合論の考えを応用し、音高を低い方から n 平均律の音を数値化し、計算することで音程の発生頻度を求める。結果として、ある音階の発生しやすい音程と発生しにくい音程を導くことができる。この音階から音を抜き出し、初音ミクに歌わせ響きの変化を聴く、というのが本研究発表である。

大型ディスプレイを用いたカラオケとヘッドマウントディスプレイを用いたカラオケの比較実験

栗原 拓也, 北原 鉄朗 (日本大学)

今日カラオケは娯楽の一つとして根付いていて、複数人でカラオケに行く機会が多くある。しかし、複数人でカラオケに行った際に歌を聴いている人たちはカラオケの画面を見ている事が多く、歌い手は孤独を感じカラオケで一体感を得ることができない。そこで、通常の大ディスプレイを用いたカラオケとヘッドマウントディスプレイを用いたカラオケの両方を試行することで、両者において聴き手の視線がどう変化するかを検証した。

タッチインタラクションシステム上での非和声音を含んだフレーズ生成

黒田 元気, 伊藤 彰教 (東京工科大学),
渡邊 賢悟 (渡辺電気株式会社),
伊藤 謙一郎 (東京工科大学)

本研究では, Objective-C[1], Swift[2] それぞれの言語で, iOS 上で SuperCollider をサウンドエンジンとして利用するためのライブラリである「iSCKit」について発表した. 音楽を再生しながらスケールを変更できることは確認したが, 和声課題のような, 非和声音を含むフレーズを生成するまでには至らなかった.

その後本研究では, 刺繍音, 倚音などの非和声音を含むフレーズを生成するアルゴリズムを構築した. また, アルゴリズムを iOS アプリケーションとして実装し, オブジェクトに対するタップ, スワイプなどのタッチインタラクションと, スケール, リズムパターン, 非和声音の出現頻度などの音楽的要素の変化を結びつけた.

関連 Web ページ URL:

<https://github.com/wdck/iSuperColliderKit/>

参考文献

- [1] 黒田他, : タッチインタラクションを想定したリアルタイム・モーダルチェンジ・インタラクションシステムの構築, 研究報告音楽情報科学 (MUS), 2015-MUS-106 (4) (2015).
- [2] Ito, A. et al.: iSuperColliderKit: A Toolkit for iOS using an internal SuperCollider Server as a Sound Engine, in Proc. ICMC 2015 (2015) (in press).

スマート電子楽譜 piaScore

小池 宏幸 (プラスアド株式会社)

スマート電子楽譜「piaScore」は, スマートフォンやタブレットを活用した, 音楽演奏者向け楽譜リーダーである. 本製品の主要開発者が, 長年, 研究開発を主体としたソフトウェア開発でありながら, アマチュアピアノ演奏者でもあることから, ユーザ目線の機能が随所に盛り込まれており, 主要各国 (日中米英独) のユーザレビューで5点満点中4.5点を頂くなど, 世界的にユーザの評価が高い. また, 2015年7月現在, 全世界で累計130万ダウンロード (うち65%は海外, 1万人以上の学校へ導入済) を記録し, 日本発の電子書籍リーダーとして成功した事例であると考え.

また, 主要開発者は IPA 未踏スーパークリエイターであり, 今後は, そのプロジェクト (ソーシャル音楽院 <http://wikivatoire.org/>) で開発した技術を, 楽譜リーダーをプラットフォームとしたネットレッスンや演奏 SNS に展開していく予定である.

本デモセッションでは, アプリのご紹介, 並びに, スタートアップとアカデミックの接点について議論したいと考えている.

関連 Web ページ URL: <http://piascor.com/>

無声ボイスパーカッションの音響分析

鈴木 潤一 (日本大学大学院), 北原 鉄朗 (日本大学)

ヒューマンビートボックスの一技巧として「無声ボイスパーカッション」がある. これは唇の破裂音のみでパーカッションを演奏するため, 鼻歌と組み合わせることで二重奏が可能になる. しかし, 有声ボイスパーカッションに比べ習得が難しい. 我々は, 信号処理技術などを使って無声ボイスパーカッションの習得支援を行うことをめざし, その第一歩として経験者と未経験者の無声ボイスパーカッションを録音し, 音響的な分析を行ったので, その結果について報告する.

歌声・伴奏音・打楽器音分離に基づく音楽演奏支援システム

土橋 彩香, 池宮 由楽, 糸山 克寿, 吉井 和佳 (京都大学)

音楽の楽しみ方には, 単純に音楽を聴くだけではなく, 音楽に合わせて歌ったり, 楽器を演奏したりすることも含まれる. この際, 市販 CD に録音されている音楽音響信号と同等の品質の伴奏音が利用できれば, 音楽演奏が一層楽しめるようになると考えられる. 本研究では, 市販 CD に収録されている複雑な音楽音響信号を歌声・伴奏音・打楽器音に分離することで, 演奏したいパートを除去した音響信号を自動生成し, 音楽の進行に合わせて演奏に役立つ様々な音楽内容 (ビート時刻やコード進行など) をスクロール提示するシステムを提案する. 本システムでは, MIDI を用いずに元の音楽音響信号から直接伴奏音を生成するため, もとの豊かな音色が保持できる利点がある. また, 音楽内容の自動推定には誤りが不可避であるという前提にたち, 能動的音楽鑑賞サービス Songle からそのつど音楽内容を取得することにした. 本サービスでは, 多くのユーザの協力で誤りが少しずつ訂正される仕組みを持つため, 音楽内容を逐次アップデートすることができる.

歌唱におけるピッチの正しさとピッチ知覚の関係性に関する予備調査

長村 佳祐, 岩崎 順, 小泉 遼, 北原 鉄朗 (日本大学)

本研究の最終的な目標は、音痴を矯正するシステムを開発することで、誰もが恥ずかしながらカラオケを楽しめる環境を実現することである。「音を外す」現象については、音を外していることを自覚できていないから正しいピッチが出せないのか、音を外していることを自覚しているにも関わらず正しいピッチが出せないのか、1つの疑問として言われている。我々もこの疑問に着目し、よく知っているメロディと初めて聴くメロディを歌ってもらう実験と、別のメロディの数箇所のみピッチを10~100centずらしたものを聴いて、ずれているところを回答する実験を行った。本発表では、両者の実験結果の関係性を調べることで、この疑問に迫ってみたい。

聴取を通じた演奏表現の把握と楽譜アノテーション

橋田 光代 (相愛大学)

一般に、楽器演奏においては、用意された楽譜を用いて旋律を奏でていくが、発音における強弱・緩急表現については、楽譜に記述されているものだけでは情報不足に陥るか絶対的な指示効力が弱いために、奏者の判断に委ねられる側面が強い。そのため奏者は、各フレーズに対する表現内容を、練習過程を通じて明確に定め、必要に応じて楽譜に書き込み(アノテーション)を行っていく。ところがそのアノテーション手法は個人によってまちまちで、確立されていない。強弱・テンポ表現に焦点を当てた表情付け支援システムを構築する上では、フレーズ範囲の指定およびその表現内容を指示するインタフェースデザインが不可欠である。システムにおける直感的な入力手法を定めるにあたり、演奏表現を聴取して楽譜に書き込む様子を考察する必要がある。本デモでは、楽器演奏経験の有無によって(1)聴取した演奏からどのような強弱・テンポ表現を想起するか、(2)それをどのような形で楽譜に書き表そうとするか、事例を示す。合わせて、参加者に対する公開実験も兼ねて、演奏表現の聴取分析についての議論を試みる。

関連 Web ページ URL: <http://mixtract.m-use.net/>

MusicMixer: ビート及び潜在トピックの類似度を用いた DJ システム

平井 辰典 (早稲田大学), 土井 啓成 (株式会社ドワンゴ),
森島 繁生 (早稲田大学理工学術院総合研究所 / JST CREST)

本デモは、本研究会にて口頭発表を行う「楽曲のビート類似度及び潜在トピックの類似度に基づく DJ プレイの自動化」の技術を用いた DJ システム MusicMixer のデモである。

MusicMixer は、ユーザが入力した楽曲群から、あらかじめ算出されたビートと潜在トピックの類似度に基づいて自動/半自動で DJ プレイを行うシステムである。DJ プレイとは、楽曲と楽曲の繋ぎ目をなるべく自然に保ちながら次々と曲を流す音楽のパフォーマンススタイルの一つであり、主にクラブやパーティなどで行われている。DJ は場を盛り上げることに優れており、近年その機材の普及と共にプライベートなパーティ等での活躍も目立ってきている。しかし、よりプライベートな空間である人々の日常生活にまで DJ が踏み込むことはなく、音楽のリスナーは音楽プレイヤーのランダム再生やプレイリスト再生を楽しむことしかなかった。

本システムは DJ プレイ (ミックスと呼ばれる) を自動化するもので、これによりユーザのプライベートな音楽鑑賞においても DJ がミックスしたかのような音楽を楽しむことができる。また、本システムはユーザ自身はその嗜好に基づいてインタラクティブに DJ プレイに参加できるインタフェースを備えており、リスナーと DJ としての両視点からミックスを楽しむことができる。