

# デモンストレーション：音楽情報処理の研究紹介 XVII

糸山 克寿<sup>1</sup> 飯島 祥<sup>2</sup> 梅村 祥之<sup>3</sup> 尾形 正泰<sup>4</sup> 加藤 淳<sup>4</sup> 柴田 健太郎<sup>5</sup> 津島 啓晃<sup>5</sup> 佃 洗撰<sup>4</sup>  
出口 幸子<sup>6</sup> 錦見 亮<sup>5</sup> 橋田 光代<sup>7</sup> 濱崎 雅弘<sup>4</sup> 廣瀬 均<sup>8</sup> Junichi Yamagishi<sup>9</sup> 吉久 怜子<sup>10</sup>  
和田 雄介<sup>5</sup>

概要：音楽情報処理の研究分野のさらなる発展に向けて、同分野に関する最先端の研究を紹介するセッションを、デモンストレーション形式で実施する。本稿では、このセッションで発表される予定のデモシステムについて、各発表者による概要の紹介を行う。

## デモセッションの開催によせて

糸山 克寿（東京工業大学）

本セッションでは、音楽情報科学研究分野における多様かつ最新の研究事例をデモンストレーション形式で紹介する。発表者が自らの研究成果であるシステムを実演、参加者がそれらを体験し、両者が直接ディスカッションを行えることは一般発表にはない特色である。デモセッションは主に音楽情報科学研究会の夏の研究発表会に合わせて開催されており、本研究会における恒例イベントの一つとして定着しつつある。

2004年に始まり、今回で17回目となる本セッションでは、毎回多彩な年代や顔ぶれによる発表が行われてきている。本セッションは一般発表とは異なり、申し込みの締め切りになるべく余裕を設けるなど、できるだけ気軽に発表できるよう配慮している。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飛び入りでの発表も受け入れており、この場合は原稿や説明資料がない状態でも発表を行うことができる。

例年、本セッションでは

- 新しい形態の楽器
- 新しい音楽体験のためのインタフェース

など、実際に触れてみないとわかりにくいものだけでなく、

- 既存の楽器を用い、演奏を伴うもの
- 開発中のもの（未完成のものも含む）
- 以前に発表したが、それ以降デモンストレーションの機会に恵まれていないもの
- 本セッションと同回に一般発表がなされるものなどの発表も積極的に受け入れている。また、過去の開催回では一般発表では発表しにくいような、
- 思いつきで作ってしまったもの
- 作ってみたものの、発表されることなくお蔵入りとなってしまったもの

なども少なからず披露されている。そのようなものも発表の機会を得ることで、参加者とのディスカッションを通じて新たな気づきが生じ、これを契機としてのちに日の目をみる研究へと発展する可能性がある。本セッションを継続的に実施することは、本研究分野における多様な研究の発展に資するものと考えている。今後もより多くの研究に発表の機会を提供できるよう努めていく所存である。

音楽情報科学研究会では、ニコニコ生放送を通じて一般発表を動画中継・アーカイブする取り組みを進めてきた。しかし本セッションの各発表者が実際にデモを行う場合は動画中継されないため、発表内容を体験できるのは発表会場に居合わせた参加者のみである。研究の成果を同じ時間と場所で共有しながらディスカッションを行える体験型の発表形態は、時系列メディアである音楽を扱う研究発表において、一般発表とは異なる側面からの意見を得られる有益な機会となる。また参加者にとっても、本研究分野で扱われる多様かつ最新の話題に最も近いところで触れられる機会となる。本セッションを通じ、多くの方に何らかの新たな気づきを得ていただければ幸いである。以降、各発表者による概要の紹介を、代表発表者の五十音順で掲載する。

<sup>1</sup> 東京工業大学

<sup>2</sup> IAMAS

<sup>3</sup> 広島工業大学

<sup>4</sup> 産業技術総合研究所

<sup>5</sup> 京都大学

<sup>6</sup> 近畿大学

<sup>7</sup> 相愛大学

<sup>8</sup> アップビョ・ミュージック・スタジオ事務所

<sup>9</sup> National Institute of Informatics

<sup>10</sup> レコチョク

## 観客に向けた、 ループペダルの演奏メカニズムの視覚化による、 ライブ演奏の拡張

飯島 祥 (IAMAS)

<https://shoijima.github.io/AfterImage.html>

本研究では、ギターを演奏する動画や、小規模のライブ会場、演奏者がディスプレイに映る大きな会場などの、演奏者がギターを持って演奏している様子がはっきりと分かる状況において、ループペダルを用いたギター演奏における、視覚的に不明瞭なプロセスの可視化を行うことで、観客が、演奏者の精神表現に対する解釈を深めることを目指す。

デモ展示では、視覚的な情報を通じて、ループペダルにおける、「パターン化された音型の反復」と、「リアルタイムで個人の演奏を重ねていく多重録音」という情報を伝え、かつその際に、ループペダルを操作する動きが、映像の変化との対応関係を持つようにするための表現についての構想案を展示する。

## DNN による楽曲データからの情動の推定

梅村 祥之 (広島工業大学)

機械が人の感じる曲の情動を推定できれば、自動作曲で生成した曲から良い曲を取捨選択するなどの応用につながるなどの考えから、情動の推定の研究を行っている。今回、世界の民謡を集めた楽曲コーパス The Essen folksong collection に収集された楽曲の中から、予備実験により選定した 373 曲を 10 数名の評価者に聴取してもらい、もう一度聴きたくなる曲かどうかを判断してもらった。このデータから、曲ごとに、もう一度聴きたいと答えた評価者数のパーセンテージを計算し、曲の好ましさに関する数値を得た。

この値を機械推定するディープニューラルネットワークを開発すべく、旋律情報をピアノロールで表現し、1次元畳み込み層を重ねたネットワークを構成した。実験結果として、高い性能を得るに至っておらず、改良を継続中である。そこで、本デモセッションにて、参加者のご意見をお聞きしたい。会場では、評価者の数値と機械推定の数値のついた曲リストを提示し、その場で聴取できるデモを行う。

## Songle Sync:

音楽に連動させて多様なデバイスを  
大規模に制御できるプラットフォーム

尾形 正泰, 井上 隆広, 加藤 淳, 後藤 真孝  
(産業技術総合研究所)

<http://api.songle.jp/sync>

音楽に連動してインターネット経由で 100 台以上のス

マートフォン・パソコン・IoT デバイスを制御する Songle Sync (<http://api.songle.jp/sync>) を開発した。音楽理解技術で得たビートやサビなどの情報に応じて表示・動きが変化する多様な演出を手軽に体験できる。CG を用いた映像・ロボット・照明などを音楽に連動させて新たな演出を生み出せるデモンストレーションを開発したほか、音楽連動アプリケーションを手軽に作成するための開発キットも公開した。音楽イベントでの実証実験を経て、音楽連動制御技術で同期した映像を来場者のスマートフォンに映して音楽体験を拡張できること、また混雑した通信環境でも利用可能なことを確認した。デモ会場では、スピーカーから流れる音楽に合わせてノート PC のほかに iPad, iPod touch といったモバイル端末の画面が一斉に同期するデモンストレーションを展示する。また、来場者がブースに表示された QR コードを読み取って、各自のスマートフォンでも Songle Sync のアプリケーションを体験できるデモンストレーションを提供する。

## TextAlive:

楽曲歌詞の Kinetic Typography 動画の  
ための統合制作環境

加藤 淳, 中野 倫靖, 後藤 真孝 (産業技術総合研究所)

<http://textalive.jp>

Kinetic Typography は、文字を動かして見せることで意味をより深く伝えられる手法として、楽曲のプロモーション動画などで頻りに利用されている動画演出手法である。本発表では、楽曲に応じて歌詞の Kinetic Typography 動画を自動生成して自由に編集できる動画制作環境 TextAlive (<http://textalive.jp>) を紹介する。従来は、まず動画制作ツールの使い方に慣れる必要があった。さらに、歌詞を文字ごとに時刻同期させて動きを手付けするため、多くの手間がかかっていた。そこで、TextAlive は Web 上に公開されている楽曲と歌詞の時間的対応付けを自動で推定し、Kinetic Typography を用いた動画を自動生成する。また、生成された動画をインタラクティブに編集でき、動画制作の手間を従来より大幅に削減できる。さらに、文字の動きを決めるアルゴリズムをライブプログラミングでき、プリセットのテンプレートに縛られない自由な表現が可能である。デモ参加者は TextAlive で制作された動画を視聴したり、好みの楽曲から動画を自動生成したり、既存の動画を編集したりできる。

## Factorial HSMM に基づく

音楽音響信号に対するカバー譜生成

柴田 健太郎, 錦見 亮, 中村 栄太 (京都大学),  
深山 覚, 後藤 真孝 (産業技術総合研究所),  
糸山 克寿 (東京工業大学), 吉井 和佳 (京都大学)

本研究ではポピュラー音楽の音響信号からカバー譜を生成する手法を提案します。これは自動採譜タスクの派生であり、原曲を典型的なバンド編成での演奏で近似することを目指します。通常バンド譜はボーカル、ドラム、その他の伴奏パートで構成されます。歌声・ドラムの音源分離および採譜についてはすでに様々な研究がなされているので、本研究ではその他の伴奏パートをリズムギター・リードギター・ベースギターの3つのパートで近似的に再構成する楽譜を推定します。そのために提案モデルでは各ギターに対応する3つのマルコフ連鎖を持つ階乗隠れセミマルコフモデルに基づくベイジアンアプローチを行います。1つの連鎖はリズムギターのコード列を表現し、残りの2つの連鎖はリード・ベースギターの単旋律を表現します。これらの連鎖から生成される3つの低ランクスペクトログラムの和で入力音響信号のスペクトログラムを近似します。入力信号を受け取ったとき、全てのパラメータと潜在系列をギブスサンプリングで推定します。実験により提案手法の有効性を示しました。

## コードとメロディの階層的生成モデルに基づくインタラクティブ作曲システム

津島 啓晃, 中村 栄太, 吉井 和佳 (京都大学)  
<http://sap.ist.i.kyoto-u.ac.jp/members/tsushima/ismir2018/>

本稿では、ユーザが初めにメロディを与えたのちに、メロディに対してコード進行を自動生成する操作（和声付け）及び、コード進行に対してメロディを自動生成する操作をユーザが繰り返すことによって、ユーザが楽曲を洗練させることができるインタラクティブ作曲システムを提案する。従来の自動作曲システムの多くは、メロディやコード進行の完全な自動生成が目標となっているため、生成結果の修正は手動で行う必要があるという課題があった。また、従来のコードのマルコフ性を仮定した和声付け手法では、コード進行に内在するコードの繰り返し構造が考慮されないという課題もあった。この問題を解決するため、メロディとコードの相補的な自動生成操作によって、コードとメロディのそれぞれを少しずつ洗練させていくことができる自動作曲システムを提案する。そのために、木構造モデルである確率的文脈自由文法(PCFG)に基づく和音系列生成モデル、マルコフモデルに基づくコードとメロディそれぞれに関するリズム生成モデル、コードの条件付きマルコフモデルに基づく音高系列生成モデルからなる階層的生成モデルを定式化する。さらに、このメロディとコード進行に関する統一的な生成モデルを予め学習しておくことで、現在得られているメロディやコード進行のうちで、ユーザが指定した一部のみを事後分布に従ってサンプリングする手法を提案する。また、被験者実験による主観評価

によって提案システムのユーザの楽曲制作を支援するシステムとしての高い有用性があることが示された。

## Lyric Jumper:

### アーティストごとの歌詞の傾向を考慮したトピックモデルに基づく歌詞探索サービス

佃 洸撰, 石田 啓介, 後藤 真孝 (産業技術総合研究所)  
<https://lyric-jumper.petitlyrics.com/>

各アーティストには、「恋愛」や「友情」といったトピックの歌詞が書かれやすい、といった傾向が存在する。本発表ではそうしたアーティストごとの歌詞トピックの傾向に基づいて様々な歌詞に出会える歌詞探索サービス Lyric Jumper (<https://lyric-jumper.petitlyrics.com/>) を提案する。Lyric Jumper では、アーティストごとのトピックの傾向の可視化や、トピックの類似度に基づくアーティストの推薦などの機能を提供する。そうした機能を利用することで、Lyric Jumper によって様々なアーティストや楽曲の歌詞へのユーザの理解が深まることを目指している。Lyric Jumper の実現にあたり、我々はアーティストごとの歌詞トピックの傾向を推定するためのモデルを提案する。提案モデルでは、各アーティストがトピックの分布を持ち、その分布に応じて各歌詞にひとつのトピックが割り当てられる。歌詞データセットを用いた実験により、従来研究において歌詞トピックを推定する主要な手法である latent Dirichlet allocation (LDA) よりも提案モデルが優れていることを定量的に示した。

## 数字譜を用いた演奏システムの試用と楽譜・楽曲に関する検討

出口 幸子, 篠原 瑠花, 佐々木 敦規, 相賀 将志  
(近畿大学工学部)

本研究では五線譜に慣れていない人でも容易に演奏できるように、音名や数字等を用いた楽譜表示およびそれに対応した演奏システムの UI を開発している。今までの研究で、演奏経験の少ない人の場合、演奏しながら歌うには五線譜より数字譜の方が容易であることが示された。

そこで、数字譜とそれに対応した UI を、公開講座において一般の人に使って頂き、講座の最後にアンケートを実施した。60・70代の19名のアンケート結果および受講状況から、高齢者でも演奏できることが確認できた。またアンケート結果より、演奏経験が少ない高齢者は、演奏しながら歌うことが難しいと示唆された。

一方、本研究では楽譜 DB から、音名・数字・色・図形の楽譜を生成している。生田流箏曲譜(縦譜)に基づき横譜にした形式、および山田流箏曲譜(横譜)に基づいた形式を比較した。その結果、両者とも利点があるため、両方を生成することにした。

また、著作権が消滅した楽曲から楽譜 DB を作成しており、それを基に、俳句に既存楽曲のメロディを付けるシステムを開発中である。また楽譜 DB のデータ形式を Humdrum から MusicXML に変更中である。

## 調とリズムを考慮した 階層隠れセミマルコフモデルに基づく 歌声 F0 軌跡に対する音符推定

錦見 亮, 中村 栄太 (京都大学),  
後藤 真孝 (産業技術総合研究所),  
糸山 克寿 (東京工業大学), 吉井 和佳 (京都大学)

本稿では歌声 F0 軌跡から音楽的に自然な音符系列を推定する統計的手法を示す。歌声の発音時刻や F0 は楽譜に示されたビート時刻や音符の音高からの大きな逸脱を含むため、歌声 F0 軌跡の時間・周波数方向への離散化による音符推定の精度を向上するためには、楽譜の音楽的な自然さを表現する楽譜モデルが重要である。我々は調とリズムに依存する音符の音高を表現する楽譜モデルと楽譜 (音符系列) から時間・周波数方向に逸脱する歌声 F0 軌跡を表現する F0 モデルとを統合した準ビート同期階層隠れセミマルコフモデル (HHSMM: hierarchical hidden semi-Markov model) を提案する。楽譜モデルでは、確率的に生成された調に従って音符の音高が生成される。さらに、音符の開始位置はビートの 1 次元格子以上に定義されたマルコフ過程に従って生成される。F0 モデルでは、歌声の発音時刻の時間方向の逸脱、音符間における F0 の滑らかな遷移、F0 の周波数方向の逸脱が確率的に生成され、楽譜に付与される。提案法では、楽譜モデルと F0 モデルが音符推定に与える影響を考慮しながら、入力 of 歌声 F0 軌跡から尤もらしい音符系列を推定する。実験結果から調やリズムを考慮しない場合と比較して、提案法による音符系列の推定精度が向上することを示した。

## フレーズ表現を多角的に俯瞰するための 演奏表現支援システム

橋田 光代 (相愛大学)  
<http://mixtract.m-use.net/>

フレーズングはいわゆる発声や呼吸に関連した音楽的表現であり、言語に依らない意図の伝達手段として直感的に取り組みやすいものである。その表現に関わる方法論は、伝統的に往年の演奏家らの言を通じて経験的に語り継がれてきた知見が演奏解釈理論としていくつか取りまとめられているが、その使用者が優れたリスナーであることを前提としており、解決すべき多義性や曖昧性が残されている。本発表では、フレーズングに焦点を当てたユーザ主導型の演奏デザインにおいて、煩雑な作業の一部を自動化技術で代替するという方式の演奏デザインシステム Mixtract を

介して、音楽・音声表現生成に共通する課題について多角的に議論する。

## Songrium: Web-native music の俯瞰的な鑑賞を支援する 音楽視聴支援サービス

濱崎 雅弘, 石田 啓介, 佃 洗撰, 深山 覚,  
中野 倫靖, 後藤 真孝 (産業技術総合研究所)  
<http://songrium.jp>

Web で公開され、Web で視聴されてフィードバックを受け、Web で新しい派生作品が公開される、そうしたオープンコラボレーションが創発する音楽を、我々は Web-native music と呼んでいる。本発表では、そうした Web-native music の俯瞰的な鑑賞を支援する音楽視聴支援サービス Songrium (<http://songrium.jp>) を紹介する。Songrium では自動的に Web 上の楽曲を収集・分類し、様々な観点に基づく可視化インタフェースを提供する。音楽星図および Songrium3D では空間的な広がりを用いた可視化によりコンテンツの膨大さを感じられるインタフェースを提供する。バブルプレーヤおよび超歴史プレーヤではコンテンツの時間的な変化を感じられるインタフェースを提供する。さらに生成力ランキングでは、公開された派生作品数に基づく新しいランキングを提供する。Songrium の全ての機能は Web 公開中であり、デモンストレーションでは実際に触って各機能を参加者に体験していただく。

## 作曲用コンパイラの紹介

廣瀬 均 (アッピー・ミュージック・スタジオ事務所)  
<http://appyo.jp>

作曲のための道具として Web ブラウザから音楽をテキストで表現して快適な操作で音楽が作成できる作曲用のコンパイラの紹介を行います。少ない文字入力力で音楽用のテキスト・ソース・ファイルが生成されコンパイル、リンクを行なうことでプログラミング感覚で MIDI データが作成されます。

設計方針は、できるだけ作曲者の手間を省き創造力の妨げにならず目的となる音楽作成に集中して優れた楽曲が作成できるソフトウェアの提供をすること目的にしています。

今回のデモは、Raspberry Pi を使った LINUX サーバに音楽用のコンパイラ Web サーバ、MIDI 出力、ステータスを表示して楽曲作成の進行を助ける LCD、7 セグの数値表示器、各種ステータス LED、電源コントロールユニットが備わったサーバーを使って行います。

## Sneak Preview of the 2nd Voice Conversion Challenge 2018

**Junichi Yamagishi, Jaime Lorenzo-Trueba**  
(National Institute of Informatics),  
**Tomoki Toda (Nagoya University),**  
**Daisuke Saito (The University of Tokyo),**  
**Fernando Villavicencio (ObEN),**  
**Tomi Kinnunen (University of Eastern Finland),**  
**Zhenhua Ling**  
(University of Science and Technology of China)  
<http://www.vc-challenge.org>

現在、声質変換の性能を共通のコーパスで競う「2nd Voice Conversion Challenge」が進行中であり、国際会議 Speaker Odyssey 2018 にて特別セッションが開かれる予定である。本デモでは、この 2nd Voice Conversion Challenge の要約と提出された 25 システムの結果を説明する。声質変換された音声の品質の評価では最大 MOS 値 4.1 が出るなど、これまでの声質変換技術を大きく上回る結果となっている。

### リスナー間の楽曲嗜好傾向の可視化の一手法

吉久 怜子, 大矢 隼士 (レコチョク),  
伊藤 貴之 (お茶の水女子大学),  
山内 和樹 (レコチョク)

リスナーが保有している楽曲のリスト群を用い、リスナー間での楽曲の視聴傾向の可視化する一手法を提案する。具体的には、リスナーと楽曲をノード、リスナー-楽曲間の視聴関係をエッジとし、リスナーはリスナー同士、楽曲は楽曲同士でクラスタリングする。このエッジがより束化するようにクラスタを配置することを目指す。

### 楽曲中の歌声とユーザ歌唱の リアルタイムアラインメントに基づく 伴奏追従型カラオケシステム

和田 雄介, 坂東 宣昭, 中村 栄太,  
糸山 克寿, 吉井 和佳 (京都大学)

<http://sap.ist.i.kyoto-u.ac.jp/members/wada/smc2017/index.html>

本稿では、入力された音楽音響信号から伴奏音を抽出し、ユーザ歌唱のテンポ変化に自動で追従して再生するカラオケシステムを提案する。このシステムによって、ユーザは任意の楽曲を、テンポを自由にアレンジしながら歌うことが可能になる。このシステムの主な利点は、ユーザが楽譜 (MIDI ファイル) を用意する必要がないことと、システムを起動した後すぐにカラオケを楽しめることである。これらを実現するために、このシステムでは音源分離手法および audio-to-audio アラインメント手法をオンラインで並列に実行する。まず、入力された音楽音響信号が、ロバスト非負値行列因子分解 (RNMF) のオンライン版を用いて歌声と伴奏音に分解される。その後、分離された歌声信号と

ユーザ歌唱が、動的時間伸縮 (DTW) によって時間方向に同期される。最後に、DTW によって推定されたワーピングパスを用いて伴奏音が伸縮され、再生される。被験者実験により、このシステムの有効性が確認され、このシステムは新しい歌唱の楽しみ方を提示しうることが示された。

## 正誤表

下記の箇所に誤りがございました。お詫びして訂正いたします。

訂正箇所	誤	正
1 ページ 著者	糸山 克寿 (東京工業大学) 飯島 祥 (IAMAS) 梅村 祥之 (広島工業大学) 尾形 正泰 (産業技術総合研究所) 加藤 淳 (産業技術総合研究所) 柴田 健太郎 (京都大学) 津島 啓晃 (京都大学) 佃 洗撰 (産業技術総合研究所) 出口 幸子 (近畿大学) 錦見 亮 (京都大学) 橋田 光代 (相愛大学) 濱崎 雅弘 (産業技術総合研究所) 廣瀬 均(アッピョ・ミュージック・スタジオ事務所) Junichi Yamagishi (National Institute of Informatics) 吉久 怜子 (レコチョク) 和田 雄介 (京都大学)	糸山 克寿 (東京工業大学) 飯島 祥 (IAMAS) 梅村 祥之 (広島工業大学) 柴田 健太郎 (京都大学) 津島 啓晃 (京都大学) 出口 幸子 (近畿大学) 廣瀬 均 (アッピョ・ミュージック・スタジオ事務所)

本文：以下の各セクションの掲載は誤りです。

- Songle Sync: 音楽に連動させて多様なデバイスを大規模に制御できるプラットフォーム
- TextAlive: 楽曲歌詞の Kinetic Typography 動画のための統合制作環境
- Lyric Jumper: アーティストごとの歌詞の傾向を考慮したトピックモデルに基づく歌詞探索サービス
- 調とリズムを考慮した階層隠れセミマルコフモデルに基づく歌声 F0 軌跡に対する音符推定
- フレーズ表現を多角的に俯瞰するための演奏表現支援システム
- Songrium: Web-native music の俯瞰的な鑑賞を支援する音楽視聴支援サービス
- Sneak Preview of the 2nd Voice Conversion Challenge 2018
- リスナー間の楽曲嗜好傾向の可視化の一手法
- 楽曲中の歌声とユーザ歌唱のリアルタイムアラインメントに基づく伴奏追従型カラオケシステム