

# デモンストレーション：音楽情報処理の研究紹介 XVI

奥村 健太<sup>1</sup> 阪上 大地<sup>2</sup> 荒 哉太<sup>3</sup> 飯尾 英晃<sup>4</sup> 石垣 愛美<sup>3</sup> 齊藤 罌<sup>3</sup> 高橋 公太<sup>5</sup> 武田 郁弥<sup>3</sup>  
名畑 皓正<sup>3</sup> 能登 楓<sup>3</sup> 長谷川 麻美<sup>6</sup> 原 史也<sup>3</sup> 藤田 千尋<sup>6</sup> 柳田 拓郎<sup>3</sup> 類家 怜央<sup>3</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

奥村 健太（名古屋工業大学）、  
阪上 大地（株式会社コルグ）

本セッションでは、音楽情報処理研究分野における多様かつ最新の研究例を、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。近年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なっている。

2004年に始まり、今回で16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面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りうる。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。

## 発表内容の概要

今回のデモセッションでは、全13件の発表が予定されている。各発表者（敬称略）の発表内容は、概ね以下のよう  
に分類される：

<sup>1</sup> 名古屋工業大学  
<sup>2</sup> 株式会社コルグ  
<sup>3</sup> 公立はこだて未来大学  
<sup>4</sup> (株)UTSUWA  
<sup>5</sup> コータサウンド株式会社  
<sup>6</sup> 公立はこだて未来大学大学院

自動作曲： 荒  
自動演奏： 飯尾  
演奏支援： 石垣，齊藤  
音楽教育： 長谷川  
音響： 高橋  
楽曲分析： 武田，名畑  
楽曲推薦： 能登，類家  
歌声分析： 藤田

(演奏) 行動計測： 原，柳田

このように，各発表の研究内容は多岐にわたることから，発表会場では多様なディスカッションが期待される。

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。

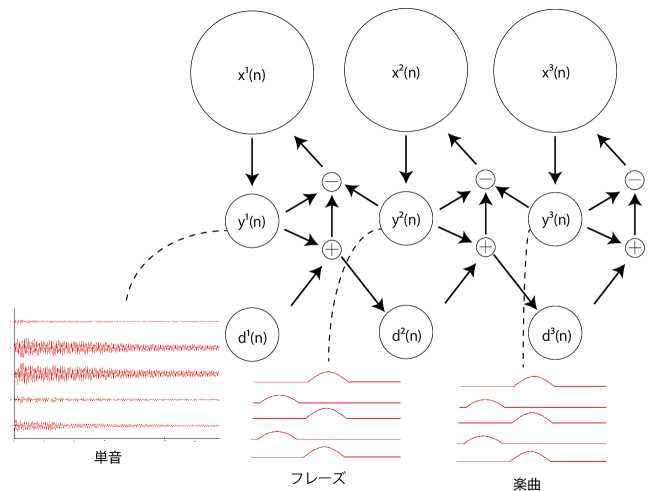


図 1 提案モデルの模式図

## 参考文献

- [1] Miwa Fukino, Yuichi Katori, Kazuyuki Aihara: A Computational Model for Pitch Pattern Perception with the Echo State Network. In Proceedings of 2016 International Symposium on Nonlinear Theory and its Applications (NOLTA2016), 2016.

## 動的プレディクティブコーディングを用いた聴覚情報処理モデルとそれを用いたフレーズ作成システム

荒 哉太，香取 勇一（公立はこだて未来大学）

### 概要

動的プレディクティブコーディング (PC) を用いた聴覚情報処理モデルを提案する。PC とは外界の情報を脳の高次野が予想する脳の情報処理である。聴覚情報処理における PC で説明可能なミスマッチ陰性電位の 1 つに，オドボール課題がある。オドボール課題で得られる脳の反応を説明できるモデルを，吹野ら (2016) が提案している [1]。このモデルから階層構造を持つ新しいモデルを構築し，フレーズ作成への応用を試みる。

### 提案モデル

図 1 は提案モデルの模式図である。提案モデルの教師信号は，楽曲から抽出したフレーズの音声波形である。提案モデルの低次の層では，単音の波形を学習し，波形の生成を行う。その生成結果と教師信号となる音声波形の誤差を高次の層に送る。この誤差の時系列は，音の出現タイミングになる。また，複数の単音の出現タイミングはフレーズになる。高次の層では，このフレーズを学習・生成する。また，高次の層からの予測を低次の層へフィードバックを行うことで次の単音の予測が可能になる。本提案では，この特性を用いてフレーズ作成システムの構築を試みる。

既存の機械作曲の多くが教師信号として演奏情報を使用している。しかし，音声波形から，演奏情報に変換する際に多くの情報が失われる。本システムでは音声波形を教師信号として使用するので，既存の機械作曲よりも元の楽曲を考慮したフレーズの作成が期待できる。

## 音源解析技術を活用した自動カホン演奏ロボット

飯尾 英晃，門原 公子 ((株)UTSUWA)，  
辻本 友樹，宮木 孝輔，三宅 正人，三熊 智哉，吉井 和佳  
(京都大学)

### 概要

打楽器カホンを演奏するロボットを開発しました。このロボットはカホンに取り付けて使用します。ロボットはアームでカホンを叩くことにより演奏します。ロボットはスマートフォンアプリと連携しており，アプリを通じてスマートフォンから演奏の指示をします。

ロボットにどのように演奏させるかを規定する演奏データは，スマートフォン内にある楽曲データをアプリで解析し生成します。これで，どんな楽曲でも，カホンロボットとセッションを楽しめます！

### 謝辞

アプリ開発にあたってご協力をいただいた，株式会社 Modelor 各務 将士氏にお礼申し上げます。

## ピアノ学習支援システムを用いた 演奏習得におけるチャンク分析 - GTA を用いて -

石垣 愛美, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

ピアノ演奏では、譜読み、正確な打鍵、リズム感覚など、さまざまな技術が求められる。そのため、学習コストの高さに利用を断念したり、習熟効率の低さから挫折してしまう演奏者が多い。これらの問題を解決するために、竹川らはピアノ初心者を対象とした、鍵盤上に演奏補助情報を投影するピアノ学習支援システムを構築した。また、田村らは竹川らのピアノ学習支援システムを用いて実験を行った。その結果、このピアノ学習支援システムを利用した学習者は、課題曲をうまく弾けるようになるにつれて、一連の複数の音符をまとまりとして認識するチャンクを形成することを明らかにした。しかし、チャンクの形成要因に関しては、被験者が記入したチャンクの形成理由を実験者が推測しているといった理由で、結論に客観的妥当性を見出せていない。

そこで本研究では、GTA にもとづく統計的手法による、チャンクの変容およびその要因を明らかにすることを目的とする。我々は被験者間で、チャンク形成の変化に傾向があることを発見した。本シンポジウムでは、これまでのチャンクに関する分析とその応用について議論したい。

## バイオリン初級者のための マルチモーダル情報提示による ポジショニング学習支援システムを目指して

齊藤 壘, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

本研究では、バイオリン初級者のためのマルチモーダル情報提示によるポジショニング学習支援システムを目指している。従来のポジショニング学習支援の多くは、学習者が弾いた演奏音が音高の正しい理想音からどの程度ずれているのか、その音高差のフィードバックを視覚的に提示しているが、その手法だけでは学習者が音高差フィードバックに依存してしまい、視覚的補助情報が無い時に正しいポジショニングで演奏できない可能性がある。

そこで本研究では、視覚と聴覚へのマルチモーダル情報提示を活用したポジショニング学習支援を行う。従来行われている音高差フィードバック提示のほか、聴覚的な手法として、学習中のガイド音提示や、演奏音の音高正誤のチェックなどを実施する。聴覚的な手法により学習者の音感を向上させることで、演奏音と理想音の音高差の認知力が上がり、より精度の高いポジショニングが可能となる。また、視覚的補助情報が無い場合でも、正しいポジショニングができるようになり、システムからの離脱を期待できる。本発表では、本システムが提供する練習モードおよびテストモード、そしてそれらに含まれる各機能についてデモンストレーションを行う。

## 音を力の変化の信号を使用して再生

高橋 公太 (コータサウンド株式会社)

### 概要

私は、従来の音再生システムによる再生音全般に不満を感じ、20年以上に渡って新しい再生方式について研究を行ってきました。研究開始時に、スピーカの振動板には重さがあるためこれを駆動するためにはどうしても力が必要となり、従来の再生方式では空気の振動した振幅信号をそのまま増幅してスピーカを駆動していたため、正しく振動を再現することができないことを発見しました。その後、理論を実証するために様々な装置を製作してきましたが、新しく音の信号から力の変化の信号を計算して求め、その信号を使用して音を再生することに成功しました。また、昨年基本となる理論に必須となる処理を追加したところ、この研究を始める理由となった当初の不満を払しょくする効果を得ることができました。以上の成果について、今回の音楽情報科学研究会でデモを行い、報告させていただきます。

本報告の基本理論は、第109回音楽情報科学研究会 [1] と、第79回情報処理学会全国大会 [2] で発表させていただいています。上記必須となる追加した処理については、別途研究会などで報告する機会をいただきたいと思います。従来の方式とははるかに再生精度が異なる、驚愕の再生音をご確認いただきたいと思います。ハイレゾなどの方式を実施する前に、この方式の対応をされることを強く希望します。

### 参考文献

- [1] 高橋公太: 音の再生信号に関する研究 (第1.3版). 情報処理学会 研究報告, 音楽情報科学 (MUS), 2015.
- [2] 高橋公太: 音の再生信号の情報処理に関する研究. 情報処理学会 第79回全国大会講演論文集, 2017.

## 三次元に拡張したメロディ概形を用いた ギター特殊奏法の自動検出法について

武田 郁弥, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

本発表では、特徴次元数を拡張した新たなメロディ概形を用いて、楽曲中からギターのバンドやスライドといった特殊奏法を音響信号処理によって自動検出するシステムについて紹介する。

ピアノやドラムなどの伴奏に起因するものや、ギターの押弦時に発生するノイズに頑健な、メロディ概形を用いることで実際の録音音源中からギターの特殊奏法部の時刻とその種類を自動検出するシステムを実装する。また、従来のメロディ概形に音響信号の持つパワーや周波数成分などの特徴量を付加し次元拡張を行うことで、従来では検出できなかった特殊奏法を検出する。

本システムの実現によって、実際の録音音源中から特殊奏法を自動検出することで楽譜への書きおこしの自動化が可能になる。ある程度の知識と経験を必要とするコストの高い書きおこしという行為を自動化することで、楽器演奏者の練習支援への貢献を目指す。

今回は次元拡張した新たなメロディ概形の可視化と、実際に音源中から特殊奏法がどのように検出されるのかをデモンストレーションする。

## スパイラルアレイモデルを用いた 階層的調性構造を抽出する試みについて

名畑 皓正, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

認知的音楽理論である GTTM が提唱するタイムスパン木は、調性が階層構造を持つことを示唆している。そこで本研究では、調性の階層構造（以下、階層的調性構造）を特徴量として楽曲間の類似度の算出を行い、階層的調性構造が今までにない音楽情報を持つことを検証する。

階層的調性構造が具体的にどのような音楽構造なのかは、本研究の中で解明していく。本研究では階層構造を得るための方法として、Spiral Array Model を採用する。Spiral Array Model では、調性が3次元空間中の点として表されるため、階層的調性構造は複数の点の配置で表現される。

本研究の有用性の例としては、従来にはない楽曲推薦の実装が挙げられる。同主調転調などの部分転調や、楽曲途中から調性が+2 上がるなどの一般的な転調といった、高度な音楽構造に着目した楽曲推薦が実装可能となる。音楽理論的な楽曲推薦が一般的になることで、楽曲を構造的に捉えて楽しむ人が増え、音楽理論文化が今まで以上に広がり豊かになる。

本発表では、楽曲の階層的調性構造の現段階でのイメージと、調性同定の方法として採用した Spiral Array Model の仕組みを説明する。

## 暗意—実現モデルに基づく 作曲家らしさ識別へ向けての予備検討

能登 楓, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

従来、旋律を対象とし、作曲家らしさを取り出す方法としては1小節内音符数、最高音、最低音、旋律の傾き、平均音高などがある。しかし、このような統計的な指標では作曲家らしさを取り出すには弱く、旋律の構造を反映した特徴量が必要である。

本研究では、Eugene Narmor が提案した暗意—実現モデルによる旋律のパターン分けを利用することによって、作曲家らしさに関する特徴量を取り出すことを目指す。暗意—実現モデルは旋律のゲシュタルト構造に着目したモデルであるため、作曲家らしさに関する特徴量として適切であると考えた。

作曲家らしさの特徴量を獲得することによって、新しい曲推薦システムを作成することができる。推薦システムではお気に入りのフレーズを基に、新しい曲を推薦することができるようになる。

本発表では、暗意—実現モデルに基づく特徴量を基に作成した推薦システムのデモンストレーションを行う。

## 3 視点演奏動画を用いた ピアノ演奏者の指使い分析ツールの提案

長谷川 麻美 (公立はこだて未来大学大学院),  
竹川 佳成, 平田 圭二 (公立はこだて未来大学),  
兼重 直文 (三重大学)

### 概要

本研究の目的は、ピアノ教師のための指使い分析支援である。指使いは、繊細な打鍵の強弱や速弾きなど、高度な

演奏技術の基礎として重要である。しかし、関節の柔らかさや、腱の結合の強さなど、手の構造は人に依存して異なる。手の構造によっては、打鍵時に力みがあったり、鍵盤から指が外れていたり、指使いに悪癖が見られる。演奏中の生徒の指使いについて、ピアノ教師がすべての悪癖を指摘することは難しい。多くの対面レッスンの場合、ピアノ教師は一方向からしか指使いを見ることができないうえ、じっくり観察することはできない。また、悪癖があっても、僅かな打鍵強度の歪みとなって表れることもあり、一聴しただけでは悪癖の有無を判断できない。

そこで、生徒の演奏を3視点から撮影し、複数視点の演奏動画から、一音ずつ生徒の指使いを分析できるツールの提案を行う。提案ツールは、複数視点の演奏動画を同期再生する機能や、1音ごとに再生できる機能、コマ送り機能をもつ。映像の見せ方や、和音を含む楽曲への応用についての議論を行いたい。

## 音量制御のための音声フィードバック器具の設計と実装

原 史也, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

近年、Google Glass と Apple Watch などのウェアラブル機器が社会に普及している。それらのデバイスは、ユーザに特定の選択肢を与える情報を提示している。しかし、例えば以下のような状況などにはそのような情報を利用できない：

- 音量が大きくなっている周りの子供を静かにしたいとき
- 緊張しているにも関わらずプレゼンテーションで明確に話すことができないとき

そこで、本研究ではユーザのストレスを感じさせないことを目指し、ユーザの行動を意識せずに制御するウェアラブルデバイスの提案を行う。

研究の第1段階は音声制御のための音声フィードバックシステムの設計と実装である。デバイスを実装するために、ここではロンバード効果を適用する。このデバイスの特徴は以下の通りである：

- 環境音を取得するマイク、発話音量を取得するマイク、スイッチ、Arduino、音量増幅回路で構成される
- ユーザの音量によってイヤホンからフィードバックするユーザの音量を変化させる
- ユーザの音量を意識的に高めるためにホワイトノイズを出力する機能を有する

## 歌唱から歌詞の発声タイミング抽出と発声タイミングモデルの提案

藤田 千尋 (公立はこだて未来大学大学院),  
竹川 佳成, 平田 圭二 (公立はこだて未来大学)

### 概要

人による歌唱は必ずしも楽譜上のタイミング通りに発声されるとは限らず、あえて楽譜上の発声タイミングよりも早めに発声する部分や遅らせて発声する部分が存在する。特にあえて遅らせて発声する歌唱表現をタメ表現と呼ぶ。本研究の目的は以下の2つである：

- (1) 歌唱における発声時間のずらしを楽曲中から抽出すること
- (2) 音高や歌詞の遷移を隠れ状態とし、発声タイミングのずらし時間を出力記号系列とした隠れマルコフモデルを構築すること

また、本研究では次のような手法を採用する：

- (1) オーディオ編集ソフトである Melodyne によって楽曲から実際の歌唱部分を標準 MIDI ファイル (SMF) として抽出する
- (2) 市販の楽譜からずらしのない歌唱の発声時間を取得する
- (3) それらを比較して、各母音と子音の発声時間のずらしを算出する

### DTW を用いた歌唱における発声のずらし時間の算出

楽曲歌唱の発声時間の抽出とずらし時間の算出は、次のような流れで行う：

- (1) Melodyne によって楽曲中の歌唱部分を抽出し、SMF に変換する
- (2) Melodyne によって変換した SMF と市販の SMF から発声時間や音高などの情報を抜き出してそれぞれ CSV に変換する
- (3) それぞれの譜面中の音符同士の対応を取り、対応された音符同士の発声時間の差を求めそれをずらし時間とする

現在の手法では、実際の歌唱パートの発声（モーラ、音素）と楽譜上の歌唱の発声に対応しておらず、素朴な DTW ではマッチングが困難である。

実際に抽出した Melodyne による SMF と市販の SMF の発声タイミングの例を図 2 に示す。

### 発声タイミングのずらし時間のモデル構築

本研究では、隠れマルコフモデルを用いた発声タイミングのずらし時間のモデル構築を行う。発声タイミングのずらしの大きさが音高や歌詞と関係していると考え、隠れ状態を音高と歌詞とし、出力記号系列をずらし時間とする。

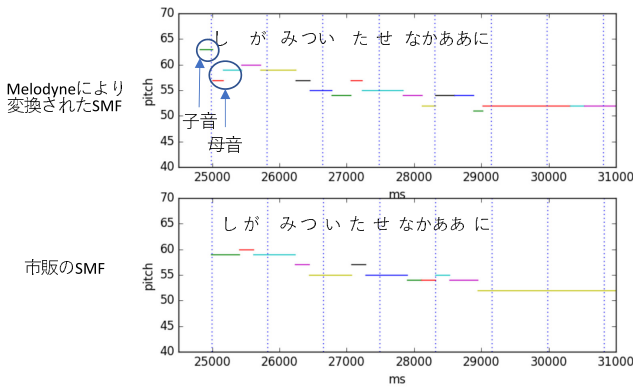


図 2 楽曲から抽出した歌唱の発声タイミング

また、歌詞をそのまま隠れ状態としてしまうと状態数が多くなってしまうため、母音、半母音、閉鎖音、摩擦音、破擦音、鼻音、流音の7つに分類する。

## 筋電位と MIDI 情報を用いた ドラム演奏における疲労度の 簡易計測法の提案

柳田 拓郎, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

ドラム演奏などの動的筋収縮を伴う運動において、疲労は必然的に起こり得るものである。本研究ではドラム演奏時に用いる筋肉から取得した筋電位信号の分析を行い、疲労度の定量的な表現を目指す。

具体的な手法として、得られた筋電位信号の周波数成分から、パワーの総面積を縦に二分できる周波数中央値 (Median Power Frequency, 以下 MdPF) を経時的に観察する。筋疲労は周波数成分の徐波化によって確認できることがこれまで多くの文献で報告されている。MdPF の低値推移から筋疲労を確認する手法は様々な研究で用いられているため、本研究では叩打時刻ごとに区切った筋電位信号それぞれの MdPF を求め、経時的にどのような変化があるかを観察する。他にも、屈筋と伸筋の共収縮にも着目し、筋疲労との関連性について考察を行う。

実験内容として、電子ドラムパッドの叩打を被験者に行わせる。その間の筋運動を筋電位信号で、また演奏情報を MIDI 情報で取得し、分析を行う。本発表では筋電位信号および MIDI 情報の簡易計測手法の提案を行う。この簡易計測手法によって筋疲労情報の取得が可能かを確認するため、ある一定の音量幅を維持した叩打を演奏者に行わせ、筋疲労の指標の一つである MdPF の推移を観察する。

## 音楽音響信号の分散表現法の提案と その楽曲推薦システムへの応用について

類家 怜央, 竹川 佳成, 平田 圭二  
(公立はこだて未来大学)

### 概要

現在多くの楽曲配信サービスが定額で聴き放題のサービスを取り入れている。しかし、そのサービスで多くの楽曲を聴けるにも関わらず検索方法は古くから存在している方法が主流である。

そこで本研究は、自然言語分野で用いられる分散表現を音楽音響信号に適用することで、多くの楽曲の中からユーザが聴取したい楽曲を「○○という曲に△△らしさを足し××らしさを引く」といった方法で検索できる楽曲推薦システムを提案する。

分散表現を音楽音響信号に適用する方法として、音楽音響信号から得たテンポや調などを単語に変換することで文章化する方法と歌詞を使用する方法を想定している。単語のボキャブラリは、Hevner が楽曲の旋律やリズムが感情価に及ぼす実験で用いた 67 の形容詞とする。

この研究の有用性として、楽曲内の要素の加減算で別の楽曲を探し出せるようになる。また、楽曲の要素のベクトルを比較することにより、流行の起点となった楽曲やアーティストを特定できるようになるなど、更に別の楽曲推薦方法を提案できる。

本発表では、Word2Vec を用いた歌詞内の要素の加減算を行う。