



## 《第59回》音とITで生活を調和させる

植村 あい子

### 1. はじめに

私は、「ハーモニー」をテーマに音楽情報処理、音響信号処理の研究に取り組んでいます。これまで音楽を対象とするコンテンツの視聴支援、演奏のスキル獲得支援やコンピュータによる創作支援の技術開発に携わってきました。たとえば、テレビ音楽番組におけるアーティストのトークと楽曲場面の検出、ライブバージョンの楽曲検索やピアノ譜面データのバラード編曲などです。IT (Information Technology) を駆使することで、ユーザを支援・補助する機能を付けたり、レベルに合わせて段階的な調整をしたりもできるので、ユーザの特性に合わせた支援を行えると考えています。

大学3年のときに音楽情報処理という分野を知り、コンピュータで音楽を人間のように解釈したり生成したりする技術があることを知りました。幼少のころから中学生ぐらいまでピアノと歌を習っていて、大好きな楽曲を分析したいという思いがあったこと、ピアノがあまり上達せず演奏や作編曲もうまくできないもどかしさを抱えていたこともあり、音楽情報処理の世界に飛び込みました。同じような思いを抱えている方を技術で支援したいという思いが今日まで繋がっています。

本稿では、現在取り組んでいる生成系 AI (Artificial Intelligence) の研究で、曲の雰囲気を変えてくれるハーモニー生成の手法について説明します。

### 2. 曲の雰囲気に変化を付けるハーモニー生成

ハーモニーのニュアンスを変化させる手段の1つにリハーモナイゼーションがあります。リハーモナイゼーションとは、メロディは変更せずに元のコードを別のコードに置換し、曲の雰囲気に変化を付ける操作のことをいいます。コンテンツ産業において、クリエイタの参入形態

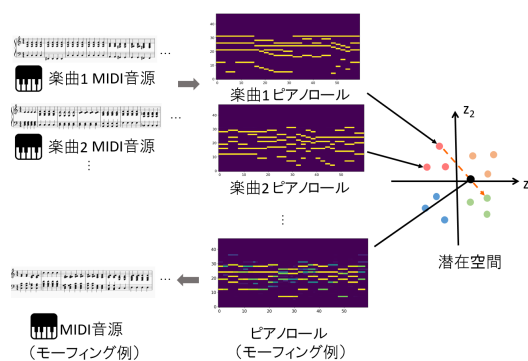


図1 モーフィングによるコード進行生成のイメージ

も変化してきている中、コンテンツ創作に取り組む方が増加しました。自分の好みに合わせて作編曲できる技術の需要も増加傾向にあります。一方で、コンテンツ創作には専門知識やソフトウェアを使いこなす技能が必要になり、初心者にとってハードルがあるのが現状です。

それを解決するため、本研究ではユーザがハーモニーの響き（ポップスっぽい、ジャズっぽいなど）を手掛かりに好みのハーモニーを探索して生成する手法を検討しています。ここでは、図1のようにコントローラとして、Variational Autoencoder (VAE) の潜在空間におけるモーフィングに基づくコード進行の生成に着目しました。モーフィングとは、与えられた2つのインスタンス間の内挿データを生成する手法のことです。従来の深層学習を用いた手法では、リハーモナイズされたデータが少ないこともあり、潜在空間内の周囲の適当なデータで補完されてメロディに合わない不協和音や、時間方向の接続が合わない和音が生成されてしまう問題がありました。そこで本研究は、図2のようにVAEとLong Short-Term Memory (LSTM) をベースにメロディ情報や和音構成音から計算される2音の共起性をネットワークに追加し、不協和を抑制するモデルを設計しました<sup>1)</sup>。ここでは和音

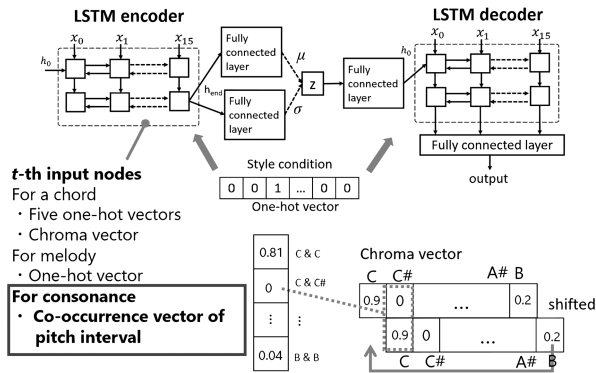


図2 不協和を抑制する提案モデル<sup>1)</sup>

認知に関する心理物理モデルの不協和度<sup>2)</sup>から協和度を定義して、学習時に各2音の共起ベクトルと理想的な協和度との損失が少なくなるように設計しました。各和音構成音が調和するように学習されることが期待されます。

評価実験では、84曲のメロディ（5アレンジ以上ついているデータ）のMIDI（Musical Instruments Digital Interface）データを使用し、生成された合計11400個のコード進行を分析しました。コード進行は多変量正規分布に従う潜在変数ベクトルを用いて表現され、曲 $z_1$ とその他のスタイルの曲 $z_2$ のモーフィングを行うことで得られます。生成されるコード進行 $z$ は以下のように表現されます。

$$z = (1 - \alpha)z_1 + \alpha z_2 \quad (1)$$

実験では $0 < \alpha < 1$ で $\alpha$ を0.25ずつ変化させ、スタイルについても3種類（ $z_1$ に対してスタイル $s_1$ で生成、 $z_2$ に対してスタイル $s_2$ で生成、 $s_3 = (s_1 + s_2)/2$ のスタイル $s_3$ ）で生成しました。

図3に非標準的な和音や垂直方向の不協和音音程の比率を示します。提案モデルを使用した場合、非標準的な和音や不協和音音程が減少することが確認できました。共起ベクトルと協和度を使用することで、モデルに影響を与えることが示唆されました<sup>(注1)</sup>。

### 3. おわりに

これまでずっと私の研究の中心にあるのは「ハーモニー」という言葉です。最近では共同研究の機会も頂戴し、フェムテック（女性が抱える健康の課題をテクノロジーで解決すること）の一環として腹部表面体温の分析にも取り組

<sup>(注1)</sup>生成されたコード進行例は以下よりご参照いただけます。 <http://uemaik.org/demo/reharmodemo/index.html>

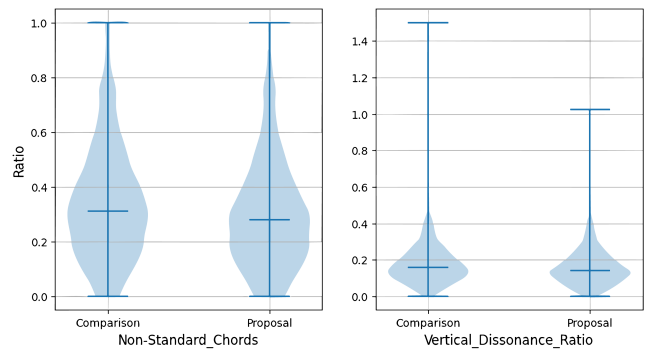


図3 非標準的な和音，垂直方向の不協和音音程の比率

んでおります。「ハーモニー」には音楽用語で「和声」という意味がありますが、ほかにも「調和」という意味もあります。「ITを使って生活の調和を目指す」というのが、研究の根本にあるように感じています。

近年の生成系AIの進歩は目覚ましく、とてもおもしろい時代に研究できることをうれしく思います。今後は、微妙なニュアンスをもコントロールできるようになり、さまざまな背景をもつユーザが参入して一期一会のライブ感あるコンテンツが生成されていくのではないかと期待しています。私もこの状況を楽しみながら、AIやITを駆使して、だれもが楽しんでより生活が調和できるような支援の研究に励んでいきたいと思ひます。

(2023年8月3日受付)

#### 参考文献

- 1) A. Uemura and T. Kitahara: Morphing-based Reharmonization with VAE: Reducing Dissonance with Consonance-based Loss Function, *Proceedings of the 3rd Conference on AI Music Creativity* (2022)
- 2) N. D. Cook and T. X. Fujisawa: The Psychophysics of Harmony Perception: Harmony is a Three-tone Phenomenon, *Empirical Musicology Review*, 1-2, 106/126 (2006)

#### [著者紹介]

植村 あい子 君

2016年早稲田大学大学院基幹理工学研究科博士課程修了。博士（工学）。同大基幹理工学部通信情報工学科助手を経て、2017年より日本大学文理学部情報科学科助手を務める。2021年から、同大生産工学部マネジメント工学科助教。音楽情報処理、音響信号処理の研究に従事。

E-mail: [uemura.aiko@nihon-u.ac.jp](mailto:uemura.aiko@nihon-u.ac.jp)

所属：日本大学生産工学部 千葉県習志野市泉町1丁目2-1