

研究概要報告書【サウンド技術振興部門】

(/)

研究題目	人の感性に基づく人工知能による進化的楽曲自動生成システムの提案	報告書作成者	森直樹
研究従事者	大阪府立大学 森直樹		

研究目的

近年、計算機による音楽の自動生成に関する研究が積極的になされており、新しい文化や産業の創造に繋がるとして多方面から強く期待されている。

しかしながら、音楽のような芸術分野において作品の評価は個人の嗜好や感性に大きく依存するため、これを定量的に数値化することは非常に困難とされている。このような問題に対して、進化型計算 (Evolutionary Computation: EC) を応用した対話型進化型計算 (Interactive Evolutionary Computation: IEC) が提案されている。

IEC は人間の評価系そのものを評価関数として最適化システムに導入することで、ユーザの感性を反映しつつ定量的な評価が可能となるという特徴を有しており、音楽の自動生成に関しても IEC の適用例が報告されてきた。

一方で、IEC にはユーザ負荷の観点から解を評価できる回数に大きな制約があり、探索の際に必要な適応度評価の回数が十分に確保できず、本来 EC が持つ探索能力を發揮しにくくなるという問題もある。特に音楽の自動生成の分野においては、絵や文字のように並列に解を提示して評価することが困難であり、各々の解を逐次的に評価する必要があることから、評価回数の制約が顕在化しやすい。そこで、ユーザに対する適応度評価の負荷を小さくするための手法として、適応度景観を学習し評価関数を近似的に表現する surrogate model が提案されており、IEC においてもその有用性が報告されている。

また、深層学習を用いた音楽の自動生成の手法も提案されている。特に、Variational Autoencoder (VAE) および Generative Adversarial Network (GAN) などの生成モデルを活用した手法は、入力データから獲得した確率分布を用いることで新規にデータを生成するという特徴を有しており、幅広い応用が期待されている。更に、これらの手法は入力データが潜在的に内包している意味を抽象化する表現学習とも類され、抽象化によって得られたベクトル (分散表現) についても様々な研究がなされている。深層学習を用いた手法は計算機が特徴量を自動的に抽出するために非常に定量的であるとされている一方で、主には事前に学習した分布の再現になるため生成物に対してユーザの感性を直接的に反映させることは困難であるという課題も存在する。

本研究では以上の点を背景として、surrogate model を応用した IEC による対話型音楽自動生成システムを提案する。対話型音楽自動生成システムでは、音の高さおよび長さの概念を反映した木構造によって楽曲を表現した。また、Genetic Programming (GP) の拡張手法である Genetic Programming with Multi-layered Population Structure (MLPS-GP) を音楽の探索アルゴリズムとして用いた。更に、surrogate model として既存の楽曲データから音楽的特徴を抽出、学習し、個体に対する近似評価をする評価モデルを導入した。評価モデルの学習に当たっては、楽曲における数学的特徴を定義し学習させた手法および深層学習を用いた手法の 2 つを提案する。深層学習を用いた手法では、楽曲学習時に生成される潜在空間に着目しつつ IEC と組み合わせることでより定量的かつユーザの嗜好や感性を反映した音楽の自動生成を目的とする。

研究概要報告書【サウンド技術振興部門】

(/)

研究内容	<p>本研究では人の感性に基づく人工知能による進化的楽曲自動生成システムの構築を目的としている。しかしながら、音楽のような芸術分野において、作品の評価は個人の嗜好や感性に大きく依存するため、これを定量的に数値化することは非常に困難とされている。そこで、本研究では進化型計算とディープラーニングを組み合わせることで、定量的にユーザの感性を楽曲に取り入れることを可能としている。具体的な人の感性の推定方法としては、潜在変数と呼ばれる楽曲潜在的な特徴を反映した値に着目し、これに基づいて人の感性と楽曲の特徴を比較するというものである。潜在変数は楽曲だけでなく、文章や画像など様々なメディアの潜在的な特徴を有することができ、シームレスに扱うことが可能である。このため、例えば本システムを通じて SNS の文章や画像などから人の感性を推定し、それに基づいた音楽の自動生成が可能である。このような進化型計算とディープラーニングの併用手法は未だ報告されておらず、非常に新しい手法となっている。</p>
------	---

研究のポイント	本研究では人の感性に基づく人工知能による進化的楽曲自動生成システムの構築を目的としている。ここで、自動生成とは音楽理論に基づくものではなく、既存の音楽データから統計的な特徴を学習し再現するものを対象としている。このため、本研究を通じて得られた知見や楽曲を用いることで、既存の音楽理論や楽曲を統計的に解析し、定量的に評価することや新しい理論の構築なども可能であると考えており、これらの考え方を人工知能や機械学習によって実現する点が本研究のポイントである。
研究結果	本研究の完成に伴い、やや敷居が高いという風に考えられている作曲について、様々なユーザが簡単にその場面に応じた楽曲を作成できるようになると考える。例えば、作曲初学者の場合は本システムの GUI を通じて簡単に楽曲を作成することができる。また、作曲に習熟した者もアイデア出しの方法の 1 つとして活用可能である。更に、本システムでは生成過程においてランダムな要素を含むため、ユーザが想像しない、既存の音楽からは得られないような新しい楽曲の獲得が期待される。この結果、本システムを通じて幅広いユーザによる新しい音楽文化の創造が期待される。具体的に Vibrational Recurrent Auto-encoder と遺伝的プログラミングを併用するシステムで人工知能による楽曲が狩野であることを示した。
今後の課題	音楽の自動生成に関する研究は国内外問わず幅広く為されている。特に注目を集めているものとして Google の Project Magenta が挙げられる。この研究は本研究と同様にディープラーニングを用いた音楽の自動生成に取り組むものである。しかしながら、Project Magenta の場合、自動生成の際に事前に学習した統計的な分布の再現になるため、生成物に対し、人の感性を直接的に反映させることは困難であるという課題も存在する。一方で、本研究は進化型計算およびディープラーニングを組み合わせた手法であるため、楽曲な統計的な特徴を反映しつつも、ユーザの嗜好を反映したフレーズを保存し、再利用する事が可能である。このように他の研究手法との比較および Web による成果の公表が重要な今後の課題である。

近年、計算機による楽曲の自動生成に関する研究が積極的になされており、新しい文化や産業の創造に繋がるとして多方面から強く期待されている。しかしながら、作品の評価は個人の嗜好や感性に大きく依存するため、これを定量的に評価することは非常に困難とされている。このような問題に対して、人間の評価系そのものを評価関数として最適化システムに導入した対話型進化型計算 (Interactive Evolutionary Computation: IEC) が注目をされている。IEC は定量化が難しい人の感性を評価できるという利点がある。一方で、ユーザ負荷の観点から解を評価できる回数に大きな制約があり、特に楽曲の自動生成の分野においては各々の解を逐次的に評価する必要があることから、評価回数の制約が顕在化しやすいという問題がある。

また、深層学習を用いた楽曲の自動生成の手法も提案されている。特に、Variational Auto-Encoder (VAE) などの生成モデルを活用した方法は、入力データが潜在的に内包している意味を反映し写像した潜在空間を作り上げることから幅広い応用が期待されている。しかしながら、深層学習を用いた手法は主に事前に学習した分布の再現になるため、生成物に対してユーザの感性を直接的に反映させることは困難であるという課題も存在する。加えて、深層学習に基づくモデルの場合、楽曲のような可変長のデータを適切に扱うのは難しいという問題もある。

本研究では以上の点を背景として、2 点の提案した。1 点目としては深層学習および Genetic Algorithm (GA) に基づく楽曲の分割手法の提案である。これにより、可変長の楽曲を定量的に分割することが可能になり、深層学習モデルが扱いやすい形式に変換することが可能となった。2 点目としては segmented VRAE と呼ばれる VAE の拡張モデルを提案し、IEC による対話型楽曲自動生成システムにおける surrogate model の構成要素として適用した。図 1 に segmented VRAE の構造を示す。segmented VRAE は楽曲データを部分構造に分割して扱うことができるもので、これにより可変長の楽曲をシステム内でより適切な定量評価を目的としたものである。定量評価の方法としては、segmented VRAE から生成される潜在空間の情報およびユーザからの入力を組み合わせることで達成した。また、この surrogate model の導入により IEC の課題であるユーザ評価の負荷を軽減した。提案システムでは音の高さおよび長さの概念を反映した木構造によって楽曲を表現し、Genetic Programming (GP) の拡張手法である Genetic Programming with Multi-layered Population Structure (MLPS-GP) を楽曲の探索アルゴリズムとして用いた。最後に専門知識を持たないユーザでも本システムを用いて楽曲生成ができるように、アプリケーションの開発をした。図 2 にアプリケーションの画面を示す。

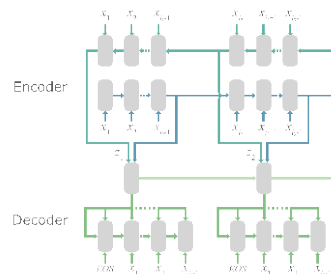


図 1:segmented VRAE の構造

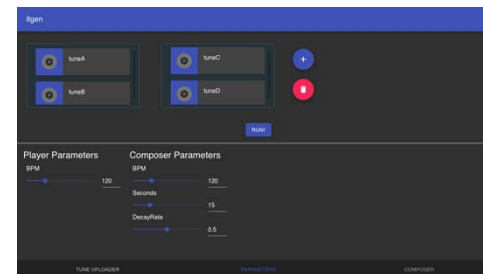


図 2:Web アプリケーション