

2 群(画像・音・言語) 9 編(音楽情報処理) 10 章(自動作/編曲)

10-1 自動作/編曲

(執筆者: 平田圭二)

10-1-1 位置づけと分類項目

作曲や編曲は、人が行う知的で創造的な活動の1つに数えられる。計算機に音楽を作/編曲させるという目標への取り組みは、人工知能(Artificial Intelligence)という言葉の誕生(1956年ダートマス会議)とほぼ同時期と言えるほど古い(1957年に作曲されたイリアック組曲(Illiack Suite)がその嚆矢)。当時は、音楽的な知性や創造性に対する科学的な興味、計算機を利用した新しいスタイルの音楽の開拓、作/編曲の技術や才能を持たない人に対する表現機会の提供などが目標であったと考えられる。その後の技術発展の経過や成果を振り返ると、現在これら目標の一部は実現されてはいるものの、まだ不十分なところも残っている。近年においても上述の目標はある程度有効だが、ゲーム操作中のBGM、公共空間での著作権フリーな楽曲、UGC/CGM向けの楽曲など、大量生産/大量消費される楽曲への需要が高まるにつれ、工学的な応用という側面も強く意識されるようになった。

自動作/編曲を研究対象として眺めてみよう。もし「自分が作るような(嗜好を反映したような)曲で、音楽的に一定のクオリティを満足するレベルを持ったものを自動作曲させたい」という研究目標を設定したとすると、次にそれを技術的な課題へとブレイクダウンしなければならない。例えば、自分が作るような曲とはどのようなもので、その情報をどうやって計算機に伝えるのか、計算機はその情報をどう処理するのか、どうやって音楽的に一定のクオリティを満足させるのか等である。ここで重要なことは、目標や音楽の対象が異なればそれらを適切に実現する技術も評価手法も異なってくるので、まず自動作/編曲を様々な観点から分類し、他の研究者による再現や追試が可能な程度に目的、条件、方法等を明確化することである。(表1.1)。表中、例えば最上段の「目標」の欄の意味は、自動生成された楽曲だけで出来上がるような作品を作るのか、それとも楽曲にビデオや絵や写真など他のメディアを組み合わせて作品とするのかのいずれかを選択する項目である。続く、長期間繰り返し聴くのかそれとも1回だけ聴くのかという選択肢は、先の選択肢とはまた独立の選択肢であり、実際の目標はこれら選択肢の組み合わせの数だけ存在する。

10-1-2 本来的に備える困難な性質

このように目的、条件、方法等を限定し明確化しても、自動作/編曲という課題が本来的に備える性質により、研究課題の設定と実験結果の評価法(測定・比較の方法)にはどうしても曖昧で不完全な部分が残ってしまう。その性質とは、(1)技術そのものの評価と生成された表現の評価が対応していない点、(2)その楽曲が音楽規則を守っているという意味で正しいか否かということとその人の意図を表現しているかという意味で好ましいか否かということが対応していない点、(3)楽曲は楽曲のみで存在しているわけではないという点である。

まず(1)に関して、これまで様々な自動作/編曲手法が提案され、中には、他分野で高い有用性を持つことが実証された技術も含まれているが、そのような技術が高い音楽性や選好性を保証するわけではない。技術の改良が直接的間接的に音楽性や選好性の改良に貢献するか否かは曖昧である。技術的に改良すべき点を同定するような評価を実現することも難しい。

表 1-1 自動作/編曲に関する分類項目 (の一部)

目標:	<ul style="list-style-type: none"> • 楽曲自体が作品 • 作品の一部(BGM) • 長期間繰り返し聴く • 1回だけ聴く • 自分だけ聴く • 他者に聴かせる • 専門家 • 非専門家
音楽の対象:	<ul style="list-style-type: none"> • 旋律, 和声, リズム(音楽の3要素) • パート(ソプラノ, アルト, テノール, バス) • 楽器(ピアノ, 管楽器, 弦楽器, 打楽器) • 曲長(4/8/16小節, あるいは数百小節) • バンドスコア/合奏譜, 即興演奏 • スタイル/ジャンル
生成方式:	<ul style="list-style-type: none"> • 決定的 <ul style="list-style-type: none"> - ルール - 事例 - 統計 • 非決定的 <ul style="list-style-type: none"> - 確率(乱数, カオス) - インタラクション(即興演奏, 進化計算)
制御方式:	<ul style="list-style-type: none"> • 数値を与える • 形容詞を与える • 例示曲を示す • 完全自動 • 半自動/支援

人は、その表現が自分の意図に照らし合わせて適切なか/好ましいかどうかよく分からない場合があるが、これは、自動作/編曲における汎用的な評価尺度の構築が難しいことを示唆している。

次に(2)に関して、一般に、人の意図を効果的に表現したり、人が求めているような質の高い音楽を生成するには、構造的な普遍性を規定する音楽的な規則に従って楽曲を生成するだけでは不十分で、そこから多少逸脱することで音楽性や嗜好性を高める必要があると考えられている*。しかしその逸脱に規則性を認めるのは難しく、また統計的に最尤な振る舞いが常に妥当とも限らない。音楽性を高める逸脱と嗜好性を高める逸脱の区別も曖昧である。さらにユーザの意図を計算機に伝達するには、ユーザの意図を何らかの記号の形式で表現しなければならない。ユーザの意図を最も具体的に指示する方法は、楽譜エディタ上で生成する楽曲の1音1音をユーザが操作することであろう。逆に最も抽象的な指示方法は、ユーザが幾つかの大域的なパラメータや指示を与えて自動作曲するものである。前者は精密な操作が可能だがユーザに高度な音楽的スキルを求める。一方後者の指示方法は簡便であるがユーザの意図を正確に伝達するのが難しい。この意図指示に関する抽象度と操作性、あるいは簡便さと意図の伝達度は一般にトレードオフの関係にある。

最後に(3)に関して、システムが楽曲を自動生成して五線譜上の音符として記述したとしよう。しかし人がその楽曲を聴取する時は、五線譜上の音符の情報だけを聴きとっているわけではない。編曲、演奏、音色、その楽曲に関連した文章(例えばアルバムのライナーノート)、楽曲がビデオ作品のBGMとして利用されるような時は映像に関するような付帯的な情報も同時に鑑賞している。鑑賞には、聴取環境や以前聴取した楽曲との関連性も影響を及ぼ

* そもそも音楽には正しい楽曲か否かを判別する明確な規則が存在しない。対して、例えば自然言語では、ある文法規則に関して正しい文と非文の区別は容易である。

す。従って、自動作/編曲の技術は、自動作/編曲された楽曲を鑑賞する環境まで考慮することが望まれる。

前述の(1)や(2)で述べた曖昧さが音楽自身に由来するとすれば、(3)は外部的な要因に由来する曖昧さに関連している。

自動作/編曲には上に挙げた本来的に備える困難な性質があるため、従来なら効果を発揮してきたような方法論が、うまく使えない場面が出てくる。例えば、音楽情報処理の分野では(他のメディア処理の分野でも)、他の研究者による実験の再現や新手法の検証を可能とするために、標準的な正解集(コーパス)を構築することが一般的である。しかし、自動作/編曲の場合、もしシステムがそのコーパスに含まれないような楽曲を生成してもそれを単純に不正解とすることはできない。なぜなら、コーパスに含まれないような新しい表現を生成する自動作/編曲システムの実現というのを目標に掲げることのできるからである。

音楽情報処理において(他のメディア処理の分野でも)、個人の選好を適切に扱える技術の1つにソーシャルフィルタリングがある。この技術は、すでに大量に存在している作品やコンテンツに対する個人の選好(振る舞い)を処理の対象とする。一方、自動作/編曲は、基本的に、この世に存在しない未知の楽曲を生成する技術である。もし自動作/編曲で生成された楽曲の評価にソーシャルフィルタリング技術を応用する場合には、楽曲の類似度を定義する(楽曲のモデル化をする)必要がある。しかし一般的に、研究コミュニティが芸術性や選好性の要素を含む類似度に関して合意に到達するのは難しいだろう。

10-1-3 研究事例紹介

自動作/編曲分野を網羅的に紹介した文献には¹⁾や Wikipedia の Algorithmic Composition のページ²⁾があるので参考にされたい。以下代表的な自動作/編曲システム研究を紹介する。

音楽学者である David Cope は 1981 年より作曲システム Experiments in Musical Intelligence (EMI) の開発を開始した^{3,4)}。EMI はある類似度をもって内部データベースから適切なメロディ断片を検索し、SPEAC という音楽文法に従って断片を接続して楽曲を生成する。EMI は数多い自動作曲システムの中でも質の高い楽曲を創作することで有名である。

コンピュータ音楽の作曲家である Robert Rowe は Cypher というインタラクティブな自動作曲システムを製作した⁵⁾。Cypher に入力された音楽は、listener モジュールによって特徴量空間にマップされ、フレーズが検出されて音楽を理解する。player モジュールはその情報を元に実時間でユーザに応答を返す。インタラクティブにすることで、ユーザの意図の曖昧さを低減させることに成功した。

Generate & Test 手法を使うと、自動生成の問題の一部を認識の問題に帰着できる。Gerhard Widmer は、対位法の楽曲事例から作曲ルールを帰納推論するシステムを Prolog 言語を用いて構築した⁶⁾。帰納推論する際、既存の音楽理論 (Generative Theory of Tonal Music (GTTM) と Implication-Realization Model (IRM)) を背景知識とすることで、音楽知識を表現するための基本概念が与えられ、学習の効率が高まった。こうして学習されたルールを Generate & Test の Test に用いることは、上述(1)、(2)の曖昧さ克服に効果的であろう。

Francois Pachet は、ユーザの演奏スタイルを実時間で学習する Continuator という即興演奏器を作成した⁷⁾。入力されるフレーズのピッチを可変次数マルコフモデルによって学習し、その獲得されたモデルに基づいて応答の旋律を生成する。和声やリズムも学習、模擬できる

ように学習モデルに修正を加えた。入力された旋律の学習と応答の生成が同時に実時間オンラインで実行されるようアルゴリズムに工夫を加えた。この工夫により、例示による意図指示が可能となり、上述の意図指示に関する抽象度と操作性のトレードオフの課題に対処している。

浜中らは、既存の音楽理論 (GTTM と Tonal Pitch Space (TPS)) が定義する旋律のタイムスパン簡約構造に基づいて、ユーザが弾く可能性の高い音列を予測する予測ピアノを制作した⁸⁾。GTTM と TPS では、タイムスパン簡約構造の安定度を算出する手順が与えられており、安定度の高い構造ほど音楽的に正しい解釈を与えていると考えられている。音楽的に正しい旋律を生成するという曖昧な課題を、音楽的に妥当な解釈を持つ旋律を生成するという課題に置き換え、音楽に内包されている曖昧さの問題を軽減させた。

深山らは、ユーザが与えた歌詞の韻律を反映したような旋律を自動作曲するシステム Orpheus を構築している⁹⁾。様々な旋律候補の中から、テンプレートとして与えられている和音パターンやリズムパターンを最もよく満足する旋律を出力する。自動作/編曲に、歌詞の韻律と旋律の対応という新しい視点を持ち込んだ点は興味深い。さらに Web 上では、人工音声による歌声トラックや伴奏トラックを付加する編曲のサービスが提供されており、生成した旋律をどのように聴かせるかという環境まで考慮した例の 1 つである。

安藤の構築した作曲支援システムでは、意図指示に関する抽象度と操作性のトレードオフの課題に対処するため、クラシック音楽の作曲手法を模倣するような木構造型遺伝子と進化プロセスを用いたインタラクティブな遺伝アルゴリズム (Genetic Algorithm) を採用している¹⁰⁾。楽曲プールを世代更新する際、人が楽曲プール中の候補楽曲を直接評価し淘汰するかどうかを決定する。評価作業を行うユーザの負担という課題はあるものの、human-based computation の観点からも興味深い方法論である。

参考文献

- 1) Gerhard Nierhaus, *Algorithmic Composition: Paradigms of Automated Music Generation*, Springer (2009).
- 2) http://en.wikipedia.org/wiki/Algorithmic_composition.
- 3) David Cope, *Experiments in Musical Intelligence*, A-R Editions, Inc. (1996).
- 4) David Cope, A Musical Learning Algorithm, *Computer Music Journal*, Vol.28, No.3, pp.12-27 (2004)
- 5) Robert Rowe, *Interactive Music Systems – Machine Listening and Composing*, The MIT Press (1993).
- 6) Gerhard Widmer, Qualitative Perception Modeling and Intelligent Musical Learning, *Computer Music Journal*, Vol.16, No.2, pp.51-68 (1992).
- 7) Francois Pachet, The Continuator: Musical Interaction with Style, In *Proceedings of ICMC 2002*, pp.211-218.
- 8) Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo, Melody Expectation Method Based on GTTM and TPS, In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR2008)*, pp.107-112 (2008).
- 9) 深山覚, 中妻啓, 米林裕一郎, 酒向慎司, 西本卓也, 小野順貴, 嵯峨山茂樹, Orpheus 歌詞の韻律に基づいた自動作曲システム, *情報処理学会 音楽情報処理科学研究会 研究報告*, 2008-MUS-76, No.30, pp.179-184 (2008).
- 10) 安藤大地, 対話型進化論的計算による作曲支援に関する研究, 博士論文 東京大学大学院新領域創成科学研究科 (2009). あるいは, *人工知能学会誌 特集「人工知能分野における博士論文」*, Vol.25, No.1 (2010).