

楽曲の再生履歴を利用した自動メロディ生成

松岡 優太† 酒向 慎司†

名古屋工業大学†

1 はじめに

自動作曲において、音楽的知識などの支援をすることはもちろん、ユーザの好みに合った曲の制作を支援することも重要である。しかし、人間の楽曲に対する嗜好は様々であいまいなものであるため、システムによって正確に捉えることは難しい。また、取得した嗜好をどのようにメロディ生成に反映させるかにも難しさがある。

よって本研究ではユーザの嗜好を反映した楽曲を生成するための自動作曲システムの方略として、ユーザの持つ楽曲の再生履歴情報を利用することを検討する。履歴情報には、ユーザの楽曲への嗜好が表れると考えられ、実際に楽曲推薦の分野でユーザの嗜好を得るために現在も広く用いられている情報である。作曲には従来手法でよく用いられている遺伝的アルゴリズム (GA) を利用し、履歴情報の取得には Spotify API[1] を利用する。

2 GA によるメロディ生成

本研究では、GA を用いて自動メロディ生成を行う。GA は、交叉・突然変異といった遺伝的操作と適応度関数を用いた自然淘汰の機構をモデル化した最適化アルゴリズムである。このアルゴリズムは局所的な解に陥りにくく、広い探索能力を持っている点から、大量の音符の組み合わせから最適解を探索する最適化問題である自動作曲において適した手法と考えられている。

3 楽曲の再生履歴情報によるユーザ嗜好の獲得

履歴情報は、ユーザが再生した曲の情報であり、その中には、ユーザの好む曲に関するさまざまな情報が含まれている。これらを有効に利用することで、メロディ生成にユーザ嗜好を反映させることが可能になると考える。以下に各曲から得た情報を示す。また、Spotify API のドキュメントにも詳細が示されている。

ジャンル: Spotify API では各曲単位でジャンル情報を提供していないため、iTunes Search API を用いた。ここでは 367 種類の小ジャンルが用いられているが、今

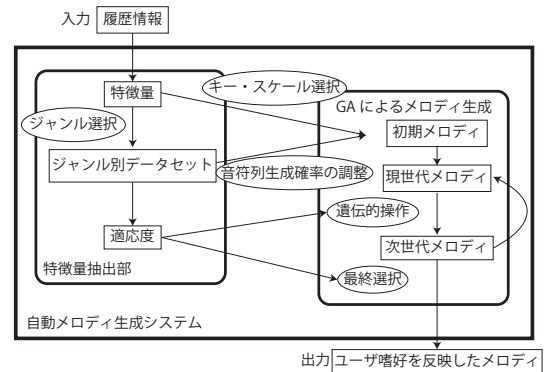


図 1: 提案手法の流れ

回はそれらの大ジャンルを参考にしながら、RWC 研究用データベースの音楽ジャンルデータベース [2] で用いられている 9 つの大ジャンルへと対応させて利用した。

キー・スケール: 曲全体の推定キー・スケールを取得できる。キーはピッチクラス記法の整数によって表現される (例: $0 = C, 1 = C\sharp/Db$)。スケールも同様に整数で表現される。(メジャー: 1, マイナー: 0)

テンポ: 曲全体の推定テンポを取得でき、1分あたりの拍 (BPM) によって表現されている。

Danceability: 曲のテンポとリズムの安定性、拍の強さなどの音楽的要素の組み合わせに基づいてどれだけ曲がダンスしやすいかを 0.0~1.0 の値で示す。

Energy: 曲の持つ熱気や活発さといった知覚的尺度を 0.0~1.0 で示したもので、これが高い曲は早く、大きく、騒々しく感じられる。

4 楽曲の再生履歴情報を用いたメロディ生成

本研究における提案手法の流れを図 1 に示す。メロディ生成における履歴情報の適用部は、初期メロディ生成、適応度計算、出力メロディの調整の 3 つの部分である。

4.1 初期メロディ生成

初期メロディ生成は、目標旋律からメロディをランダムに生成するという操作である。履歴情報から選択されたジャンルの既存楽曲データセットから目標旋律を選択し、その音符の位置によって音符生成確率を設定する。履歴から同様に得られるキー・スケールによって音高を定義し、音符生成確率に基づいて初期メロディ

Automatic melody composition using listening history
†Yuta Matsuoka, Shinji Sako · Nagoya Institute of Technology

表 1: 式 (1) 中の i の値に対応する各評価項目

i	評価項目に対応する特徴量	i	評価項目に対応する特徴量
1	総音数	8	打鍵間隔 (最短幅)
2	最高音高推移	9	打鍵間隔 (平均幅)
3	平均音高推移	10	音高平均
4	音高の最大連続上昇	11	最高音
5	音高の最大連続下降	12	最低音
6	音高の最大連続平行	13	拍でのオンセット数
7	打鍵間隔 (最長幅)		

を生成する。

4.2 適応度計算

適応度は既存楽曲のデータセットから取得された音楽理論的な特徴量と、生成メロディとの距離を求める評価関数 (1) によって定義される。

$$f(i) = \sum_{n=1}^{13} ((s_i - v_i) / \sigma_i)^2 \quad (1)$$

i は各評価項目の特徴量を示し、表 1 のように対応している。 s_i は生成メロディの評価項目 i の特徴量の値、 v_i は既存楽曲の評価項目 i の特徴量の平均値を指す。 σ_i は既存楽曲の評価項目 i の特徴量の分散で、この値で除算を行うことにより各評価項目を正規化する。使用する既存楽曲データセットは履歴情報から得たジャンルによって選択される。また履歴情報から得られた Danceability, Energy の値によって特徴量を調整する。具体的には、Danceability によって拍でのオンセット数を、Energy によって総音数を調整する。調整に用いる計算式を (2) に示す。

$$v_i = x_i[E + (1 - E_m)] \quad (2)$$

x は既存楽曲の特徴量の平均値を指し、 E は履歴情報から取得した Energy や Danceability の値の平均値を指す。そして E_m は Spotify で取得可能な全ジャンルからそれぞれ 1000 曲ずつ取得した Energy, Danceability の平均値である。

4.3 出力メロディの調整

GA による最終選択によって出力されたメロディに対して、音の強さ、テンポの部分で調整を行う。Danceability は拍の強さを示す数値であるため、拍の部分の音を強く調整し、Energy は曲全体の騒々しさを表すため、曲全体において基準となる音の強さを強く調整することでこれらを反映させた。テンポは、履歴から得られた推定テンポに基づいて調整した。

5 実験

履歴情報を用いた GA によって生成されたメロディに、それを用いていない GA によって生成されたメロ

ディと比べて変化が表れているかを確認するための主観評価による実験を行った。

5.1 実験条件

ジャンルを Pop, キー・スケールを C メジャー, Danceability と Energy を 4.2 節で述べた平均値に固定して生成したメロディを履歴情報を用いていないメロディとし、比較対象として 3 節で述べた利用する履歴情報からジャンルを除いたものそれぞれの要素の内 1 つを変化させたメロディとする。

被験者 7 名は履歴情報を用いずに生成したメロディと 1 要素だけ用いて生成したメロディを聴き比べ、用意した質問に対して 5 段階評価をする。質問はそれぞれキー・スケールを変化させた楽曲では「雰囲気が変わったか」、テンポでは「曲の速さは変化したと感じられたか」、Danceability では「リズムを強く感じられるようになったか」、そして Energy では「曲の雰囲気が明るくなったか」というものを用意した。

5.2 結果と考察

キー、スケール、テンポでは評価点の平均がそれぞれ 3.71, 3.43, 4.36 となった。これらから、この 3 つの要素は曲に変化を与えることができているということが示唆された。一方で Danceability, Energy はそれぞれ 2.86, 3.04 の値をとった。これらの要素は平均が低いが、メロディによって評価点に偏りが見られ、良い点を取るメロディも存在した。よってこれらは目標旋律の違いによる影響と考えられる。目標旋律もこれらの特徴を持ったものを選択することで改善されると考えられる。

6 むすび

本研究では、ユーザ嗜好を捉えたメロディ生成の方策として履歴情報の利用を検討した。今後は、目標旋律の選択方法や、特徴量の調整法など検討していきたい。

参考文献

- [1] Web API | Spotify for Developers. <https://developer.spotify.com/documentation/web-api/>. (Accessed on 11/27/2019).
- [2] 後藤ら. RWC 研究用音楽データベース: 音楽ジャンルデータベースと楽器音データベース. 情報処理学会研究報告音楽情報科学 (MUS).