

旋律概形の手描き入力による日本語歌詞からの自動作曲*

大森 陽† 光本 大記† 濱田 康弘† 嵯峨山 茂樹†

概要: 自動作曲システム Orpheus は現在、ユーザが音高や音程を文字入力により指定することで旋律を制御する機能を有している。本研究ではこれを発展させ、より直感的な手法として、概形の手描き入力による旋律の制御と生成を行う。旋律の作曲を DP の経路問題として捉え、音楽理論による制約と日本語歌詞からの韻律の上下動による制約、概形の手描き入力による制約の三条件の下で、ユーザの創作意図を反映する旋律の生成を目指す。また既発表の内容に加えて追加実験を行った結果を報告する。

キーワード: 自動作曲, 旋律概形

1 はじめに

我々が開発を行っている自動作曲システム Orpheus[1]は、音楽理論と日本語歌詞の韻律の上下動を基にした伴奏つき歌唱曲の自動生成を目的としたシステムであり、主に作曲技能を持たない人の楽曲作成支援に有用であることが期待されている。

自動作曲システムにおいて重要であるのは、ユーザの意図を反映させることであるが、たとえば Orpheus では、旋律の音高が全て自動生成されてしまうために、ユーザが期待した楽曲が生成されづらいという問題がある。これに対してユーザが文字入力によって部分的に音高を指定する機能[2]が実現されているが、これは一音々々に対して指定を行うため、範囲の大きい音高の指定を容易に行うことができない。また、作曲技能を持たないユーザが音楽理論に即していない音高を入力する場合も考えられる。以上のことから、範囲の大きい音高の指定が可能である、具体的な音高の指定が不要であるなどの、ユーザの曖昧な意図を反映できるシステムへの発展が望まれる。

一方、旋律概形の手描き入力により音符列を編集する楽曲作成支援の研究が行われている[3]。これは既に作曲された楽曲を、ユーザが手描きで入力した概形により旋律を編集するものである。我々は、自動作曲においてユーザが手描き入力した旋律概形を制約として加えることで、従来の Orpheus システムでの音楽性を保持したまま、ユーザの意図を反映した楽曲を生成することができると期待している。

本研究では、旋律概形の手描き入力による旋律制御を用いて、音楽理論による制約、日本語歌詞の韻律の上下動による制約、そしてユーザが描いた旋律概形による制約の三条件の制約下での自動作曲手法を提案する。

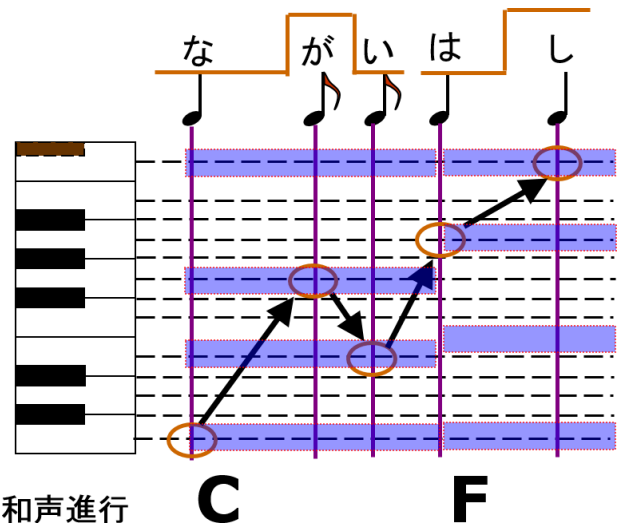


Fig. 1 和声進行の制約下の経路探索問題として
定式化した旋律生成の原理

2 手描き入力による旋律制御

2.1 日本語歌詞からの自動作曲での DP 経路

Orpheus システムでは、旋律の作曲の問題を、

- 旋律の音高が日本語歌詞の韻律の上下動に従う
- 与えられた和声進行と不協和が生じない
- 与えられた伴奏の低音との同時進行の禁則（平行 8 度など）を回避する
- 歌唱音域を守る
- 広すぎる跳躍を回避する
- 増音程を回避する

などの制約条件のもとで、確率最大の音高列を求める問題に帰着し、経路探索問題として旋律を自動生成している [1]。これは、時間を横軸、音高を縦軸とする 2 次元平面上の経路に確率が付与されているとき、その平面上で確率を最大化するような経路を探索する問題 (Fig. 1) を、動的計画法を用いて、音高系列の長さ に比例した計算量で解を探索している。本稿ではこの音高の組み合わせを 2 次元

* Melody Control by Hand-Drawn Input in Automatic Composition from Japanese Lyrics

†Yo Ohmori, Daiki Mitsumoto, Yasuhiro Hamada, Shigeki Sagayama

‡Meiji University

平面上の経路としてとらえたものを、DP(Dynamic Programming)経路と呼ぶ。

$$\hat{Q} \approx \underset{s}{\operatorname{argmax}} p(s_1, s_2, \dots, s_N)$$

$$\propto \underset{s}{\operatorname{argmax}} p(s_1)p(s_1|s_{1:N})p(s_2)p(s_2|s_{3:N})$$

$$\dots p(s_N)$$

経路探索における計算量において tri-gram 以上は現実的ではないため、言語モデルでよく用いられるような bi-gram まででの近似を行う。

$$\approx \underset{s}{\operatorname{argmax}} p(s_1)p(s_1|s_2)p(s_2)p(s_2|s_3)$$

$$\dots p(s_N)$$

uni-gram 確率としては、増、減和音を考慮した調性内音に限定する制約と、日本語歌詞の韻律の上下動により生成旋律が音域の上限及び下限に偏る問題を防ぐような音域制限確率分布[2]を与える。

bi-gram 確率としては、和声連結についての禁則等の制約、日本語歌詞の韻律の上下動についての制約、そして順次進行と跳躍進行の度合いを考慮するような跳躍度数分布[2]を与える。

以上の議論により、Orpheus システムにおける uni-gram 確率を $p(s_i)$ 、bi-gram 確率を $p(s_i|s_{i+1})$ とすると、経路 $Q = \{s_i\}_{i=1}^N$ を最大化する問題として、以下のように定式化できる。

$$= \underset{s}{\operatorname{argmax}} \prod_{i=1}^{N-1} p(s_i) p(s_i|s_{i+1}) \quad (1)$$

2.2 旋律制御方法

ユーザが描いた旋律概形による制約を DP 経路に反映させる問題は、旋律概形に近い音高にはそれだけ高い確率を与える確率分布を uni-gram 確率に付加することにより解決できる。

今回、確率分布として Gauss 分布を用いた。パラメータ σ が ∞ の場合は、旋律概形による制約はなく、従来手法により旋律が生成され、 σ が 0 に近い場合は、旋律概形の制約のみ反映して旋律が生成される。このように σ を変えることによって、旋律概形の制約の強さを連続的に操作することができる。これはそれぞれの確率における結合係数を操作することに相当し、ユーザの意図の具体性に合わせて調整することが可能である。

以上の議論により、新たに加える旋律概形の手描き入力による制約を $\mathcal{N}(\mu_i, \sigma_i^2)$ とすると、経路 $Q = \{s_i\}_{i=1}^N$ を最大化する問題として、(1)式を以下のように再定式化できる。

$$\hat{Q} = \underset{s}{\operatorname{argmax}} \prod_{i=1}^N \mathcal{N}(\mu_i, \sigma_i^2) p(s_i) p(s_i|s_{i+1})$$

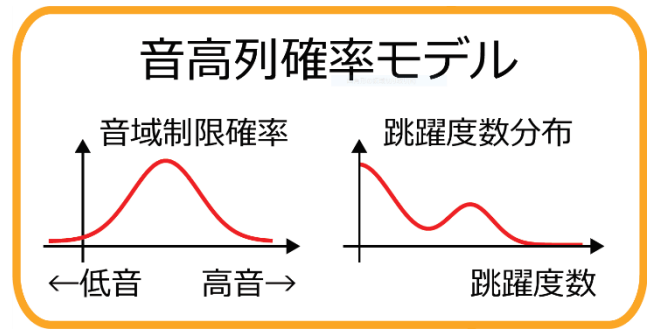


Fig. 2 音域制限確率と跳躍度数分布

3 客観評価

3.1 目的と評価方法

客観評価実験によって評価したのは次の 3 項目である。

- ① 旋律概形による制約についての評価
- ② 日本語歌詞の韻律の上下動の制約についての評価
- ③ 音楽理論の制約についての評価

①は提案手法により生成旋律が旋律概形を反映しているか、またパラメータ σ の値によって連続的に制約の強さを操作できるかを、旋律概形と生成旋律とのユークリッド距離、そして相関係数を用いて評価する。

従来手法から用いられている、日本語歌詞の韻律の上下動による制約と音楽理論による制約が提案手法においても守られているか、②については、生成旋律の上下動が歌詞の韻律に反した数の割合を用いて評価し、③については、Orpheus が考慮した音楽理論の中で禁則を犯した数の割合によって評価する。

研究者の意図が干渉することを防ぐため、評価に用いる日本語歌詞は、自動作詞システム[4]により楽曲ごとに異なる歌詞を作詞した。作曲条件は Orpheus に用意された旋律リズム、和声進行、伴奏音型に従って、それぞれ『ポップス』『バラード』『ロック』の 3 種類を用意した。

3.2 客観評価実験

実験条件について、入力は前述の 3 種類の楽曲条件に対し、3 種類の旋律概形、そして付加する Gauss 分布の標準偏差として予備実験より以下の 5 種類を与え自動作曲を行った。

$$\sigma = \infty(\text{制約なし}), 12, 6, 3, 1$$

3.3 旋律概形による制約についての評価

手描き入力した旋律概形の例と自動作曲結果例を Fig.3, 4, 5 に示す。

旋律概形と生成旋律とのユークリッド距離を、楽曲ごとに 3 つの旋律概形において平均化したものを Fig. 6 に示した。グラフより、ユークリッド距離がパラメータ σ の値に従う傾向にあり、楽曲によっては、『制約なしの生成旋律』と比較すると、『制約ありの生成旋律』の距離が全て 1/2 以

下となるような結果も得られた。

ある楽曲条件を例に、各旋律概形と生成旋律との相関係数を Fig. 7 に示す。グラフよりパラメータの値に従って相関が強くなるという結果が得られた。これは、前述したユークリッド距離の測定結果と相関関係において矛盾しない。

以上により、生成旋律が旋律概形を反映していることと、 σ の値を変えることによって旋律概形の制約を連続的に操作できることが示された。

3.4 日本語歌詞の韻律の上下動による制約についての評価

生成旋律の上下動が日本語歌詞の韻律の上下動に反した数の割合を Fig. 8 に示す。本実験では、その割合は 14% 未満であり、制約なしの生成旋律を下回る結果も得られた。

3.5 音楽理論についての評価

音楽理論における禁則のうち、2.1 に挙げた Orpheus の制約を違反した結果は無かった。

4 主観評価

4.1 目的と評価方法

主観評価実験によって評価したのは次の 2 項目である。

- ① ユーザの意図の反映についての評価
- ② 音楽性の保持についての評価

主観評価実験により、①は自動作曲された楽曲がユーザの意図に沿っているか、②は従来手法により自動作曲された楽曲の音楽性が提案手法によって損なわれていないかについて評価を行った。

実験条件について、楽曲条件は客観評価と同じものを用い、先行研究[3]を参考に、被験者に楽曲ごとに 2 回ずつ旋律概形の手描き入力を行ってもらった。1 回目は、3 小節目にあたる歌詞の音を高くするように指示し、2 回目は、被験者自身の自由な発想で作曲してもらった。その後、自動作曲結果を聴いてもらい、次の項目をパラメータごとに 7 段階評価で回答してもらった。

Q1 旋律概形を手描きで入力し、出力された作曲結果は自分の意図が反映されたものだったか。

Q2 出力された楽曲は良い曲であったか。被験者は大学生 5 名 (内 2 人は音楽経験あり) とした。

4.2 意図の反映に関する評価

被験者ごとの平均評価値を Fig. 9 に示す。パラメータ σ が小さくなるほど、より意図が反映されるという評価だった。

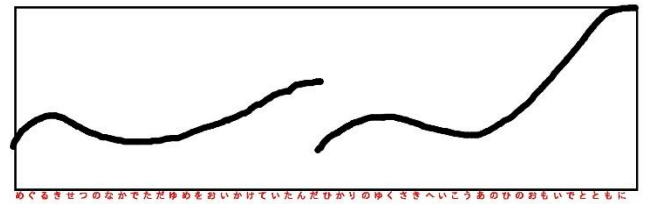


Fig. 3 旋律概形の手描き入力の実験例

sample_01 Orpheus

Music engraving by LilyPond 2.18.2 - www.lilypond.org

(制約なし)

sample_01 Orpheus

Music engraving by LilyPond 2.18.2 - www.lilypond.org

(制約あり $\sigma=1$)

Fig. 4 手描き入力(Fig. 3)による自動作曲結果例

4.3 音楽性の評価

被験者ごとの平均評価値を Fig. 10 に示す. 評価値は σ の値に従って増加し, $\sigma=6$ を最大値として次第に減少する傾向が示された. 制約なしの生成結果と比較して制約ありの生成結果は, 平均値は高いが分散も大きく顕著な例として $\sigma=1$ が挙げられる.

5 考察

客観評価から, 旋律概形の手描き入力による制約, 日本語歌詞からの韻律の上下動による制約, そして音楽理論による制約の3条件を満たした自動作曲を行うこと, パラメータ σ の値を変えることで旋律概形の制約を操作できることが定量的に示された.

また主観評価から, ユーザの意図を反映し, 従来の Orpheus システムの音楽性を保持したうえで自動作曲ができることが分かった.

制約なしの自動作曲結果が制約ありの結果と比較して, 日本語歌詞の韻律を違反した数が多い楽曲が見られる. これは bi-gram 確率として与えられている歌詞からの韻律の上下動による制約と, uni-gram 確率として与えられている音域制限確率の競合によるものであり, 旋律概形が低い音高に uni-gram 確率を付加したため, 歌詞からの韻律を違反した音高列の数が減少したと考えられる.

主観評価実験の際に, 被験者からの「評価が難しいものがあった」という意見と, 客観評価における旋律概形の制約についての実験結果を加味すると, 条件によっては2つの異なるパラメータの下でほとんど同じ旋律が生成されてしまうことがあると分かった. これはパラメータの設定が偏っていたことが原因であると考えられる. しかし, 「複数の異なる候補を出してくれることが良い」という被験者からの意見も併せると, パラメータの選択によって結果の向上が期待できることと, ユーザの意図の具体性に合わせたパラメータの調整が好ましいことが分かった.

その他にも, 「試行回数を重ねれば, より良い楽曲が作れる」という被験者の意見があった. 被験者5人中3人が音楽未経験者であったが, 自動作曲結果の楽曲において良い評価を付けていたため, 作曲技能を持たないユーザの楽曲作成支援における, 有用性を期待することができる. 音楽性の保持について, Fig. 10 のような結果が得られるのは, 前述のようにパラメータ σ が大きいと現 Orpheus システムをより考慮した結果が得られ, パラメータが小さくなり過ぎると旋律概形で入力した通りの単調な旋律が生成されてしまうため, 音楽性が保持されなくなると考えられる.

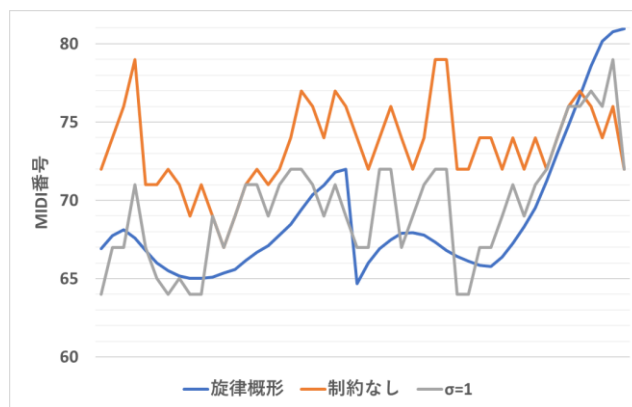


Fig.5 Fig.3 での旋律概形と生成旋律

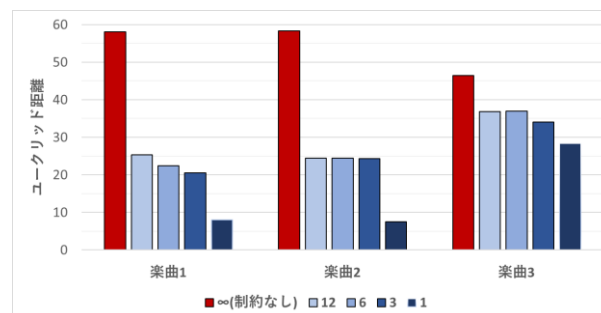


Fig. 6 ユークリッド距離による類似度評価

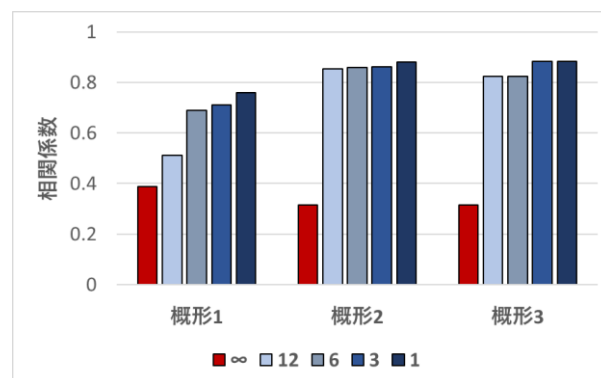


Fig. 7 相関係数による類似度評価例

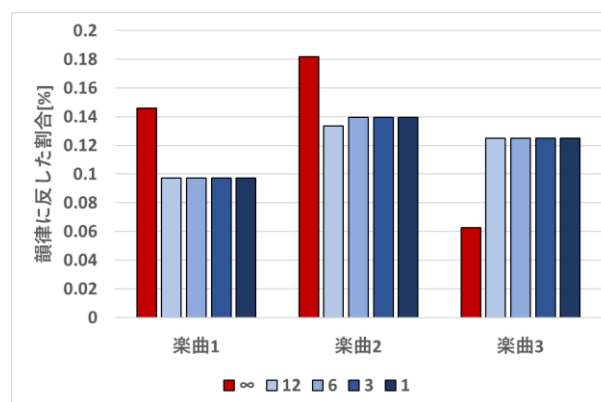


Fig. 8 日本語歌詞の韻律に反した数の割合

6 おわりに

本稿では、旋律概形の手描き入力による制約、日本語歌詞の上下動による制約、音楽理論による制約を満たす自動作曲を行った。3条件の制約を統合した確率的再定式化と動的計画法による最適化問題によって旋律を自動生成する枠組みにおいて、客観評価実験により、動的計画法によって探索する経路へ制約を課すことで、目的が実現できることを示した。また主観評価実験によって、自動作曲におけるユーザの意図の反映と従来手法により生成される楽曲における音楽性の保持を示した。

今後は、付加する確率分布のパラメータの検討、そして本アルゴリズムの、自動作曲システム Orpheus への実装を検討する予定である。

参考文献

- [1] 嵯峨山, 酒向, 堀, 深山, “確率的手法による歌唱曲の自動作曲 (<特集>音楽制作と情報処理の友好関係),” システム/制御/情報, Vol. 56, No. 5, pp. 219-225, 2012.
- [2] 深山, 齋藤, 嵯峨山, “日本語歌詞からの自動作曲における DP 経路制御による旋律制御,” 情報処理学会第 75 回全国大会講演論文集, pp. 65-66, Mar. 2013.
- [3] 土屋, 北原, “音符を単位としない旋律編集のための旋律概形抽出手法,” 情報処理学会論文誌 Vol. 54 No. 4 1302-1303 Apr. 2013.
- [4] 堀, 嵯峨山 “分散意味表現を利用した自動作詞,” 計量国語学 Vol. 29, No. 3, pp. 104-106, 2013.

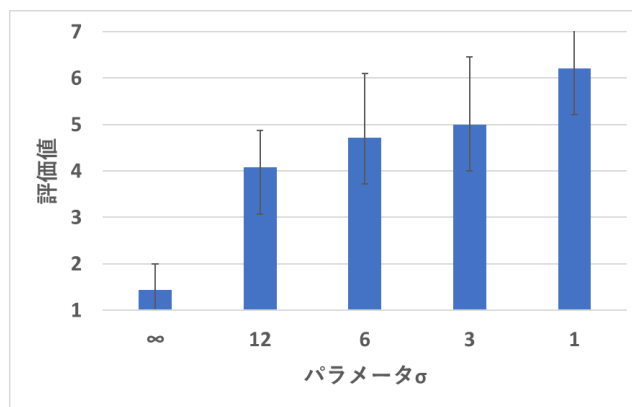


Fig. 9 ユーザ意図の評価値

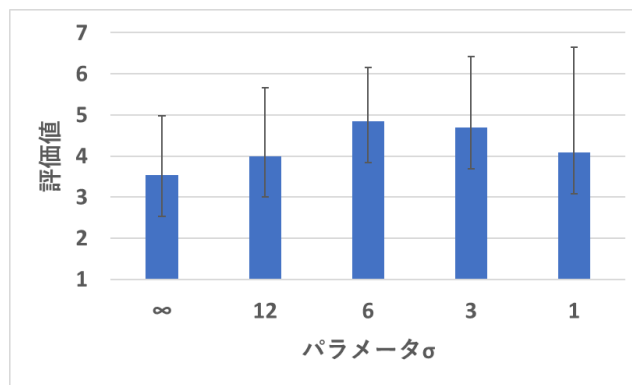


Fig. 10 音楽性についての評価値