

時系列感性パラメーターによるハーモナイゼーションシステムの開発に向けて

関西学院大学大学院 理工学研究科 人間システム工学専攻 M1 47022825 中村友紀

1. はじめに

メロディーに自動でハーモナイゼーションを行う技術は自動音楽生成において、不可欠なテーマである。またハーモナイゼーションを行うためには、音楽的構造、和声の知識が必要とされるため、オートハーモナイゼーションが可能になることによって、音楽理論等の知識を持っていない人が作曲を始めるハードルを下げることが出来る。さらに、人間のハーモナイゼーション方法を解明することで、人間の作曲理解にも繋がる。

現在までのオートハーモナイゼーション技術では、LSTM を用いたモデルによって、長期的な関係性を考慮できるようになり、Transformer を用いたモデルによって、音楽的に意味のある構造を捉える事が可能になった。しかし、これまでのモデルの多くは、モデルパラメーターが固定されており、静的にハーモナイゼーションを行っていた。よって、自分好みの音楽を作りたいユーザーにとっては満足出来ないものかもしれない。加えて、現実の音楽では同じメロディーに対して、音楽構造や編曲者によって、様々な解釈を持つことが出来る。よって、同じメロディーに対して、多様性を持ったハーモナイゼーションを生成することは重要な課題である。

そこで、本研究では自動ハーモナイゼーションで生成したコード進行を直感的に制御出来るシステムの構築を行う。

2. 関連研究

2.1 TaleBrush[1]

時系列パラメーターで物語の生成を制御したモデルに「TaleBrush」があります。UI 部分はこれを参考にします。

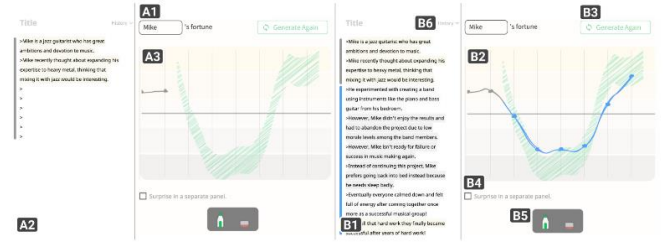


図 1: TaleBrush の曲線によるストーリー制御^[1]

2.2 Translating Melody to Chord: Structured and Flexible Harmonization of Melody With Transformer[2]

Transformer を用いたオートハーモナイゼーション技術である。VAE を導入することで、ハーモナイゼーションの制御を可能にしている。具体的に論文内では使用するコードの種類を制御している。

2.3 SurpriseNet[3]

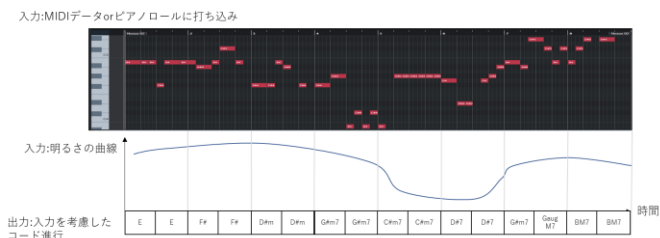
LSTM モデルに CVAE を導入することで、時系列パラメーターを用いてオートハーモナイゼーションを制御したモデルに SurpriseNet があります。論文内ではコード進行の意外性を数値化した意外性値について制御を可能にしている。

3. 研究概要

本研究では、Chang らの論文[1]からアイデアを得て、メロディーを MIDI 等で入力し、それに対して横軸時間、縦軸コード進行の明るさ等のグラフに曲線を書き込むことで、曲線によって制御されたコード進行が生成される。

3.1 システム構成

まず 8 小節のメロディーを MIDI ファイルもしくはピアノロールに直接入力する。次に、縦軸がコード進行の明るさ、横軸が時間を表す二次元グラフに、コード進行の明るさを制御する曲線を描画する。同様にして、楽曲の雰囲気制御する曲線を描画する。値としては各小節の左端の値と真ん中の値を考慮したいと考えている。上記入力を考慮して、Lee らの論文[2]を基にしたモデルでコード進行を生成する。



3.2 コード進行の制御方法

Lee らの論文[2]のモデルを使用して、Hooktheory Lead Sheet Dataset 内のモードの値を用いる。具体的には、[2]内で潜在変数に重みづけを行っており、使用されるコードの固有コード数について重みづけを行っている。その部分にモード値を入力して学習を行うことで、コード進行の雰囲気制御出来るのではないかと考えた。

3.3 学習結果

図2は学習済みモデルに対して、メロディとコード進行を入力して潜在変数の次元目を予測した結果です。青い点がメジャーにアノテーションされた楽曲で、赤い点がマイナーにアノテーションされた楽曲です。X軸が潜在変数の次元目で、Y軸がその他の潜在変数の次元に対して tSNE を用いて圧縮したものである。

図3は同じメロディに対して、潜在変数の次元目を指定して生成した結果です。-1がメジャーより、+1がマイナーよりで0はその間になると考えられる。

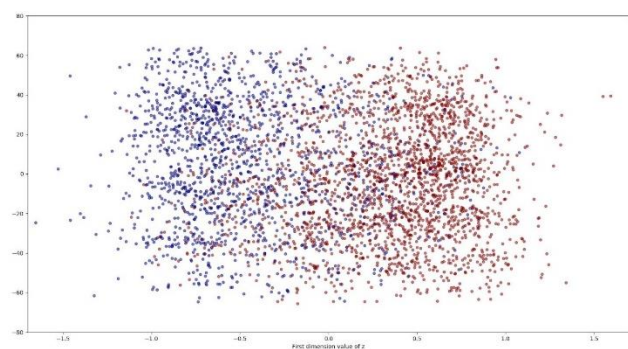


図2: 学習済みモデルの潜在変数の予測



図3: 潜在変数の次元目を指定してコード生成した結果

3.4 結果と考察

図2よりおおよそのモードの値でクラスタリング出来ていると考えられる。

図3より+1に潜在変数を指定して生成されたコード進行は Bm⁵ や E7 といった暗い響きを持つコードが生成されており、0に関しては、Dのコードがブルーノート由来だと考えられる為、メジャーとマイナーの間としては妥当であると考えられる。また-1に関しても殆どが C,F,G から生成されており、メジャーの響きとして妥当である。

4 まとめと今後の課題

8小節を通してなら、ある程度の雰囲気の制御は可能であると考えられる。しかし、部分的に見れば暗い部分や明るい部分がある。よって、今後の方針は時系列で制御可能にすることである。しかし、Transformerではデータを一括入力するため、[3]を用いたいと考えている。

加えて、明暗の最大値は古典派クラシックのルールを適用し、コードの雰囲気を変化させるため、提案したコードに沿ったメロディに変換する機能も加えたい。

参考文献

- [1] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching Stories with Generative Pretrained Language Models. In Proceedings of the ACM Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, to appear.
- [2] S. Rhyu, H. Choi, S. Kim and K. Lee, "Translating Melody to Chord: Structured and Flexible Harmonization of Melody With Transformer," in IEEE Access, vol. 10, pp. 28261-28273, 2022, doi: 10.1109/ACCESS.2022.3155467.
- [3] Chen, Yi-Wei and Lee, Hung-Shin and Chen, Yen-Hsing and Wang, Hsin-Min, "SurpriseNet: Melody Harmonization Conditioning on User-controlled Surprise Contours", in Proc. of the 22nd Int. Society for Music Information Retrieval Conf., Online, 2021..
- [4] C. Anderson, D. Carlton, R. Miyakawa and D. Schwachhofer, Hooktheory, Sep. 2021, [online] Available: <https://www.hooktheory.com>.