

# 俳句の情景に沿ったBGM自動生成システム

人間システム工学科 片寄研究室 B8626 池永考志

## 1. はじめに

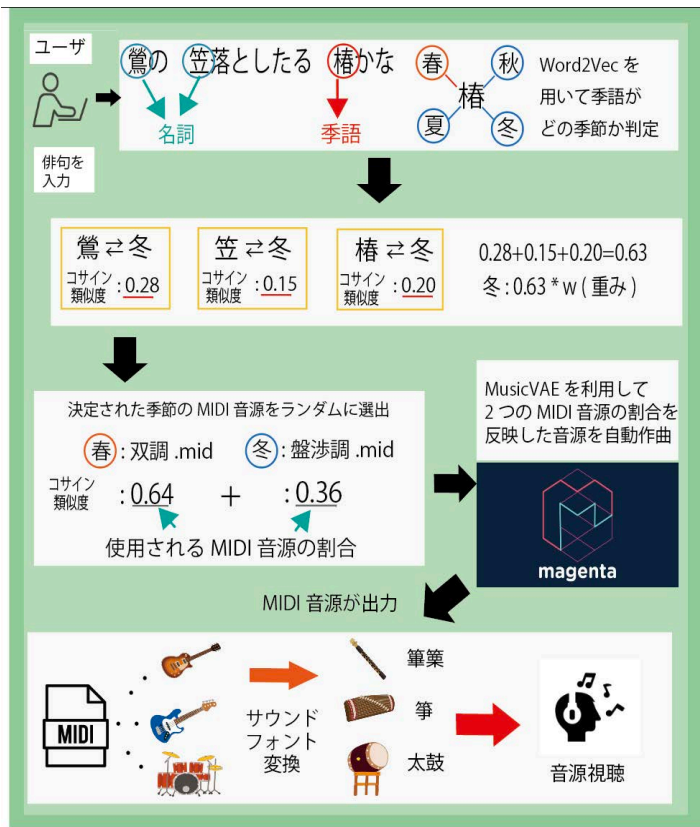
近年、機械学習による自動作曲が、動画や画像などの多方面に及んで発達し続けている。しかし、俳句に関する自動作曲の手法は存在しなかったため、自動作曲モデルの MusicVAE を利用した、俳句の情景に沿った BGM 自動生成システムの構築を行う。また、MubertAPI を用いたプロンプトエンジニアリングによる自動作曲 AI もシステムの一つとして利用する。よって本研究では、ユーザの入力した俳句から、MusicVAE と MubertAPI をそれぞれ利用した 2 つの BGM 自動作成システムを構築する。

## 2. 関連研究

ユーザの入力したプロンプトから自動作曲を行う MubertAPI のコード公開 [1] があり、こちらを使用する。また、平田ら [2] は、俳句中の名詞と季語の関連性の強さを Word2Vec によって算出し、俳句の定量的評価を行った。本研究では、Word2Vec を用いて、入力された俳句がどのような四季の要素を持っているかを単語の類似度等で定量的に判断するためにこのモデルを使用する。また、自動作曲モデルは MusicVAE[3] で発表されたモデルを使用する。

## 3. 提案手法

### 3.1 概要



初めに、ユーザが任意の俳句と季語を入力する。ユーザが入力した俳句の中に存在する名詞と季語から、Word2Vec を用いて各季節とのコサイン類似度を算出する。次に、俳句中の各名詞と四季との類似度を算出し、俳句の季節以外での類似性を検証する。この検証は、冬から春の時期は初春、春から夏の時期は晩春と言うように、春の前後での季節の要素を考慮するために行う。以上から、俳句中の名詞から得られる各季節の類似度の合計値を算出する。そして、俳句の前後の季節の合計値には固有の重み付けを行い、最大の合計値が算出された季節を俳句の考慮すべき季節として決定する。図では、春と冬の類似度が最も高かったため、両季節の類似度の比率を、生成される MIDI 音源の割合に反映させた。よって、双調(春)の音源と盤渉調(冬)の音源の特徴を有した BGM が生成される。なお、雅楽の理論では和歌での季節に応じて演奏する楽曲の調が決まっているため、実験で使用する音源は季節ごとに設定可能である。図の楽曲では、春らしさが約 6 割、冬らしさが約 4 割反映され、初春の特徴を持つ

とされる MIDI 音源を出力する。MusicVAE は、基本的にドラム、ベース、ギター の 3 つの楽器の MIDI 音源を 16 小節分出力する。雅楽では、箏、琴、太鼓が演奏される。よって、出力された MIDI 音源の楽器を、プログラムチェンジやサウンドフォントを使用し、雅楽で演奏される楽器の音源に変更する。次に、MubertAPI による自動作曲システムは、ユーザが俳句を入力し任意の雰囲気タグを選択すること(図 2)で、API は選択されたタグとユーザの入力した俳句を翻訳したプロンプトから生成されたタグから作曲する。俳句の翻訳には、OpenAI が作成した言語モデルである、GPT-3 の翻訳 API を使用している。この二つのシステムから得られた音源をユーザは試聴する。



図 2

### 3.2 使用する楽曲データと実行環境

雅楽の MIDI 再現 [5] というサイトに掲載されている MIDI 音源のデータセット 36 曲、3 時間ほどのデータを 16 小節ずつ分割し、教師データと評価データとして使用する。開発環境は Flask で、アプリケーションを作成した。

### 3.3 評価実験

第一の実験計画は被験者 10 名に構築したシステムで、各季節、計 4 句を入力してもらおう。入力された計 40 句から Musicvae と MubertAPI の両作曲システムが生成した BGM 計 80 曲を、評価用のデータとして用いる。被験者に生成 BGM を聴かせ、その評価の点数を季節別に分けて集計する。VAE による作曲では、季節や楽曲での調の比率、MubertAPI による作曲では、タグや入力する俳句によって、楽曲が持つ特徴の傾向や印象の変化を分析する。次に、アレンジでのユーザビリティを検証するため、MubertAPI による作曲の際に使用するタグの設定をジャンル分けし、被験者ごとに変更し、ユーザの期待する方向性に自動作曲が機能するかを評価してもらおう。また、著者が設定したタグが自動作曲に適切であるかの有効性を、ランダムなタグ設定による作曲での比較実験などで調査する。最後に、作曲のクオリティを検証するため、作曲に関与しない第三者として被験者を募り、どちらの作曲システムが任意の俳句の内容から得られるイメージや印象に沿った作曲を、より実現できているかなどの項目において評価データを集める、また、システムとしての改善点や優位性などについても検証する。

## 4. 今後の進め方

MusicVAE で使用する MIDI 音源が不足しているため、早急に増量し、本システムの評価実験に着手する。評価実験の項目や実施方法にも荒さがあるので、適宜改善していく。

## 参考文献

- [1] <https://github.com/MubertAI/Mubert-Text-to-Music/blob/main/Mubert-Text-to-Music.ipynb>, 2022/10/20
- [2] 単語の分散表現を用いた俳句の取り合わせの評価に関する研究 <https://www.slideshare.net/harmonylab/ss-245026576>, 2022-10-18
- [3] MusicVAE <https://magenta.tensorflow.org/music-vae>, 2022-10-14
- [4] 雅楽の MIDI 再現 <http://kaz3275.sitemix.jp/midi/gagaku/index.html>, 2022-10-17