

# J-POP のボーカル音高別発声時間の分析と可視化

理学専攻 情報科学コース 2140671 松本 留奈 (指導教員: 伊藤 貴之)

## 1 はじめに

ヒット曲の歌唱難易度の評価は歌唱愛好者にとって有益な情報である。我々は歌唱難易度を評価する要素としてメロディの声域と跳躍に着目し、その分布から歌唱難易度を評価することを検討している。

その初期段階として本手法では、J-POP 曲を対象として、音高別の発生時間の分布を可視化し、そのパターンについて考察する。本研究では J-POP 曲に対して音源分離を適用してボーカル音源を抽出し、その単位時間ごとの音高を推定して、音高別の出現頻度をヒストグラムとして可視化する。このヒストグラムから音高分布のパターンについて考察し、歌唱難易度の評価への適用を検討する。

## 2 処理手順

本章で説明する処理手順は以下の通りである。

1. J-POP ヒット曲の音源から無伴奏ボーカル音源を抽出し、MP3 形式で保存する。
2. 無伴奏ボーカル音源の音高別の発声時間を算出する。
3. 算出結果を CSV ファイルとして出力し、ヒストグラムとして可視化する。

### 2.1 ボーカル音源の抽出

J-POP ヒット曲の音源から、MP3 形式の無伴奏ボーカル音源を抽出する。本研究では、フランスの音楽配信サイト Deezer が開発した Python ライブラリの Spleeter[1] を適用し、Google Colaboratory 上で実行した。

### 2.2 音高と発声時間の算出

抽出した無伴奏ボーカル音源の基本周波数を推定する。本研究では、Python ライブラリ LibROSA[2] が提供する音高推定アルゴリズム pYIN を適用している。そして、各曲を 1/30 秒ごとに音階に変換して CSV 形式で出力する。無音となる時刻の音階は「NaN」とする。

### 2.3 ヒストグラムで可視化

2.2 節で出力した CSV ファイルから各音階の出現回数を集計し、結果をヒストグラムで可視化する。横軸を音高、縦軸をその音高が出現する回数とすると、一曲を通して音高の分布を知ることができる。

## 3 実行例

出力された音高分布のヒストグラムを図 1~3 に示す。いずれの図も、音高分布が類似する 2 曲を左右に並べている。多くの曲においてヒストグラムに複数のピークが見られるが、これは音域の低い方から地声、

ミックスボイス、裏声というように音域ごとに声の使い方が切り替わっていることが考えられる。

図 1 に示す 2 曲および図 2 に示す 2 曲は、いずれもテンポや雰囲気の異なる曲であるが、ヒストグラムから音高分布が類似していることがわかる。

図 1 は、LiSA の「炎」と DA PUMP の「U・S・A」の音高分布を示す。「炎」は BPM76 でゆったりと壮大な雰囲気のある曲である一方、「U・S・A」は BPM140 でダンスビートの軽快な曲でかけ離れたイメージがある。しかし頻出する音が低音部、高音部ともに同じ高さのため、音高面での歌唱難易度は近いと考えられる。

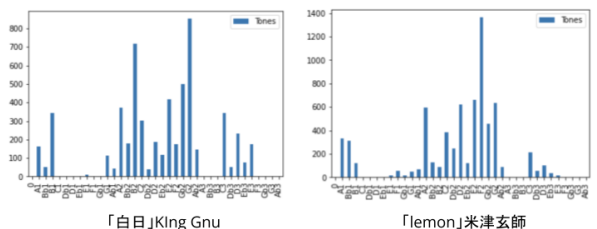


図 1: 「白日」(左), 「lemon」(右) の音高分布をヒストグラムで出力した結果。

図 2 は、King Gnu の「白日」と米津玄師の「lemon」の音高分布を示す。「白日」は男性ツインボーカルで力強い曲である一方、「lemon」は哀愁があるバラード曲である。しかし最低音、最高音付近の分布が似ているため、音高面での歌唱難易度は近いと考えられる。

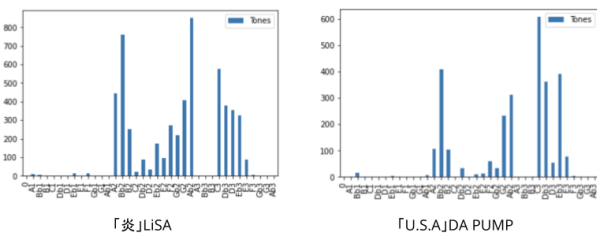


図 2: 「炎」(左), 「U.S.A」(右) の音高分布をヒストグラムで出力した結果。

このようにヒストグラムの似ている曲はただ音域が似ているだけでなく、頻出する音高が似ているため、出現する音域が似ていると考えられる。このような楽曲ペアについて、どちらか 1 曲が歌える人は、もう 1 曲も歌いやすい、という可能性がある。

図 3 は、DISH// の「猫」とあいみょんの「マリーゴールド」の音高分布を示す。同じあいみょんの作曲による曲のペアでも、男性が歌唱者の場合は音高分布が低い方に広く、音域に幅が見られた。歌唱に使用する

る音域が広い曲ほど歌唱難易度が高いため、歌手の性別によって音高面での歌唱難易度が違う傾向にある可能性が考えられる。

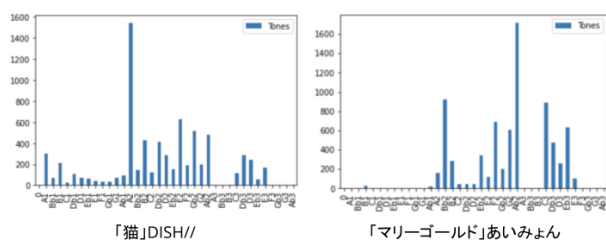


図 3: 「猫」(左), 「マリーゴールド」(右) の音高分布をヒストグラムで出力した結果。

## 4 議論

前章で示したヒストグラムからわかるように、ビートや曲の雰囲気が似ているように感じられない曲の中にも、音高分布が類似する楽曲があることがわかる。音高分布と歌いやすさに関連があるとすると、ビートや雰囲気が異なっても歌いやすい楽曲がある場合があると考えられる。

本研究では、今後以下のように楽曲の類似についての検証を進める予定である。

### 4.1 音の跳躍の検討

現状の可視化では横軸に音高、縦軸に各音高の出現回数を割り当てている。一方で、フレーズの類似度や歌いやすさには跳躍幅も関係あると著者らは考えている。そこで前後の音階の跳躍幅をヒストグラムに横軸に割り当て、縦軸にその跳躍幅の出現回数を割り当てる案を検討している。

一曲を通しての跳躍を集計する方法として、ヒストグラム以外のグラフを併用する案も検討している。前後の音の変化を1曲で集計することで、跳躍が多さや幅の大きさの可視化を考えている。

これらの方法で、音程調節の観点で跳躍回数が多く大きな跳躍がある楽曲ほど歌いにくい楽曲である直感が適切であるかを検証したい。

### 4.2 ユーザ実験による実証

音高別発声時間を表すヒストグラムと跳躍を表すグラフの2つから難易度が難しいと考えられる曲について、ユーザ評価実験を実施し、本当に歌いにくい曲と評価されるのかをアンケートや歌唱実験で実証することを検討している。

また、ヒストグラムとグラフから歌唱難易度が似ていると推測される曲について、片方の曲が歌いやすい場合に他方の曲も歌いやすいかについても、あわせて検証を進めることで、カラオケでの歌唱曲推薦に役立てられると考えている。

## 5 まとめ

本報告では、ボーカルの音高を抽出して音高分布をヒストグラムで可視化した結果を示し、楽曲間の音高分布の特徴を比較し、歌いやすい曲と歌いにくい曲の違いを議論した。本研究は初期段階であり、前章で示した今後の方針をもとに検討を進めたい。

謝辞：本研究にあたり、研究方針や手法についてアドバイスをいただいた伊藤貴之教授、伊藤研究室メンバーに感謝の意を表す。

## 参考文献

- [1] Romain Hennequin, Anis Khelif, Felix Voituret, Manuel Moussallam, "Spleeter: a fast and efficient music source separation tool with pre-trained models," The Journal of Open Source Software, 5(50):2154, pp. 1-4 (June. 2020).
- [2] Brian McFee, Colin Raffel, Dawen Liang, Daniel P.W. Ellis, Matt McVicar, Eric Battenberg, Oriol Nieto, "librosa: Audio and Music Signal Analysis in Python," Proc. of the 14th Python in Science Conference (SciPy 2015), pp. 18-24 (July. 2015).