

カラオケ歌唱曲選定のためのユーザの適性を表現する 楽曲特徴の検討とその可視化

宮地 沙綾[†] 岡田 龍太郎[†] 峰松 彩子[†] 中西 崇文[†]

[†]武蔵野大学データサイエンス学部

1. はじめに

近年、カラオケは、日本発祥の音楽文化として、今や世界的に広まり、ユーザが気軽に歌唱演奏する機会が増加しており、カラオケに関する研究も世界中で行われている[1][2]。しかしながら、一般的にユーザが自身の歌唱スキルに合致した楽曲を選定することは容易ではない。この課題に対処するために、本研究ではユーザが歌唱演奏のレパートリーを容易に増やすために、楽曲の難易度を表す楽曲特徴の検討と可視化を行う。

本研究の目的は、カラオケ歌唱曲選定においてユーザの適性を表現する楽曲特徴の検討とその可視化である。本研究においては、歌唱演奏において相対的な難易度は、リズムの複雑度、音程の変化度、音域、およびロングトーンの要素によって構成されると仮定する。これらの楽曲特徴をもとに、ユーザが自身に適した歌唱曲を容易に選定できるように楽曲メディアコンテンツの可視化を実現する。

2. 関連研究

He らの研究[1]では、オンラインカラオケプラットフォームにおけるユーザーの歌唱行動に基づく曲の推薦モデルを提案しており、このモデルは、ユーザーの好みと歌唱能力のバランスをどのように取るかを中心に研究がされており、我々が提案するユーザの適性を表現する楽曲特徴を組み合わせることが可能である。

Wang ら[2]は、PCRP(Personalized Competence-based Rating Prediction)モデルを提案することで、ユーザの歌唱力に合わせて適切なキーを推薦するシステムを提案している。

Tanaka ら[3]は、カラオケ来店の際の機会や気分の変化を考慮するために、セッションベースの多変量リカレントニューラルネットワーク(RNN)を提案し、過去に再生された曲のメタデータ、現

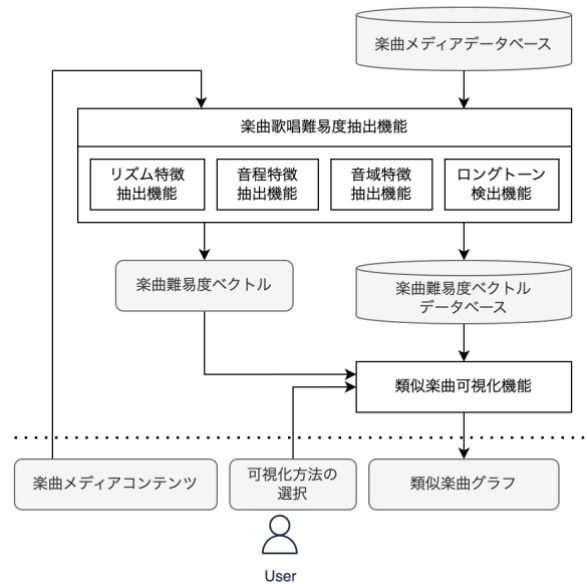


図1 提案方式の全体像

在の条件を使用し、再生される曲とアーティストを予測するシステムを提案している。

一方、田中ら[4]は、店舗周辺の地域性とリアルタイムの人口分布を用いて客層やニーズを理解することで、個人情報によらないレコメンド手法を提案している。

我々の手法は、その推薦において、歌唱演奏の相対的な難易度は、リズムの複雑度、音程の変化度、音域、およびロングトーンであると定義し、その楽曲が自身の歌唱力に合致するかを可視化する機能を持つ点で異なる。

3. 提案方式

提案方式の全体像を図1に示す。本方式は、大きく分けて楽曲歌唱難易度抽出機能と、類似楽曲可視化機能の2つの機能からなる。まず類似度を求める対象となる楽曲のデータベースを用意する。ユーザは楽曲メディアコンテンツを入力し、データベース内の楽曲との難易度の距離関係をグラフとして可視化したものを出力する。可視化の際には可視化方法を選択する。そこでユーザはどのような観点で難易度が近いのかを可視化するための評価軸を選択する。

Extraction and Visualization of Music Features Expressing User Appropriateness for Karaoke Song Selection

[†]Saya Miyaji[†], Ryotaro Okada[†], Ayako Minematsu[†], Takafumi Nakanishi[†]

[†]Musashino University, Department of Data Science

3.1 楽曲歌唱難易度抽出機能

楽曲歌唱難易度抽出機能は、入力した楽曲や楽曲メディアデータベース内の楽曲から、楽曲の難易度を表す 4 つのデータを抽出し、楽曲一つあたり一つの楽曲難易度ベクトルを出力する。

3.1.1 リズム特徴抽出機能

歌唱演奏の難易度に影響する要素として、歌唱パートのリズムの複雑度に注目し、リズム特徴抽出機能を導入する。ここでのリズムの複雑度は、メロディの音価の平均長であるとする。メロディの音価の平均帳が短いほど複雑であり歌唱が難しくなり、長いほど単純である歌唱が容易になると考える。Songle API [5]に登録されている情報を参照することで、楽曲メディアコンテンツからメロディの音の平均長を抽出し、リズムの複雑度を計算する。

3.1.2 音程特徴抽出機能

歌唱演奏の難易度に影響する要素として、歌唱パートの音程の変化の大きさに注目し、音程特徴抽出機能を導入する。音程の変化度が大きいほど歌唱演奏が難しくなり、小さいほど歌唱演奏が容易になると考える。

3.1.3 音域特徴抽出機能

歌唱演奏の難易度に影響する要素として、歌唱パートの音域に注目し、音域特徴抽出機能を導入する。音域がユーザの音域と異なるほど歌唱演奏が難しくなり、合致しているほど歌唱演奏が容易になると考える。

3.1.4 ロングトーン検出機能

歌唱演奏の難易度に影響する要素として、歌唱パートのロングトーンに注目し、ロングトーン検出機能を導入する。ロングトーンが長いほど歌唱演奏が難しくなり、短いほど歌唱演奏が容易になると考える。この機能では、楽曲メディアコンテンツからメロディの最長音を抽出する。

3.2 類似楽曲可視化機能

ユーザが入力した楽曲とデータベース内の楽曲の難易度の類似する楽曲をグラフにして可視化する。本機能は 2 つのモードを持っており、一つは、楽曲難易度ベクトルのすべての特徴量を用いて、正規化したうえで主成分分析を行って、2 次元のグラフ上にプロットするモードである。この場合は総合的な難易度の距離がグラフで表現される。もう一つは、楽曲難易度ベクトルから 2 つの特徴量を選んで、それを縦横の軸にしてプロットするモードである。この場合は、着目する要素についての距離がグラフで表現される。ユーザはモードを選択し、後者のモード

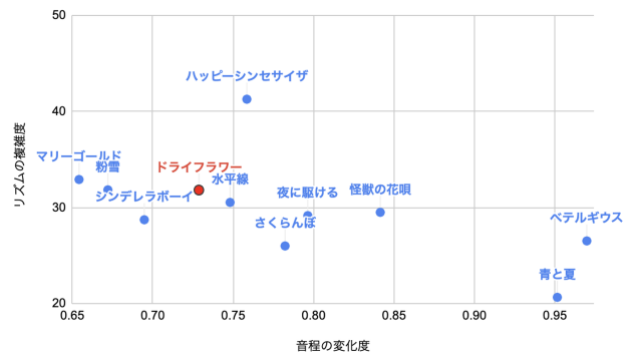


図 2 リズムの複雑度と音程の変化度を軸とした場合の「ドライフラワー」の類似楽曲グラフ

の場合はさらに自分が着目する要素を軸として 2 つ選ぶことで、グラフが生成される。

4. 実験例

実験による出力の例を図 2 に示す。楽曲メディアデータベースに J-POP の楽曲を 10 曲用意した。優里による「ドライフラワー」を入力として、グラフの軸には「リズムの複雑度」と「音程の変化度」を選択してグラフを出力した。図 2 より、軸として選択した観点においては「ドライフラワー」に最も難易度が近い楽曲は back number の「水平線」であることが分かる。

5. おわりに

本研究では、ユーザが自身に適した難易度の歌唱曲を容易に選定できるような楽曲メディアコンテンツの可視化方式を提案した。

今後の課題としては、難易度を表現する特徴量の網羅性を高めることが挙げられる。

参考文献

- [1] M. He, H. Guo, G. Lv et al., "Leveraging proficiency and preference for online Karaoke recommendation," *Frontiers of Computer Science*, vol. 14, pp. 273-290, 2020. [Online]. Available: <https://doi.org/10.1007/s11704-018-7072-6>
- [2] Y. Wang, S. Tanaka, K. Yokoyama, H.-T. Wu, Y. Fang, "Karaoke Key Recommendation Via Personalized Competence-Based Rating Prediction," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, pp. 286-290, 2021, doi: 10.1109/ICASSP39728.2021.9414524.
- [3] S. Tanaka, R. Wakamoto, Y. Fukazawa, "Real-time Karaoke Recommendations: Session-based Multi-Task Recommendations with Multivariate RNNs," in *2020 IEEE International Conference on Big Data (Big Data)*, Atlanta, GA, USA, pp. 1402-1409, 2020, doi: 10.1109/BigData50022.2020.9378439.
- [4] 田中茂樹, 深澤佑介, "地域性とリアルタイムの周辺状況に応じたカラオケ楽曲推薦," *情報処理学会論文誌*, vol. 62, no. 5, pp. 1184-1192, 2021.
- [5] Songle API <https://songle.jp/>