

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データサイエンス学科 2年

情報処理学会 第86回全国大会
2024年3月16日(土) 5V-07



目次

1. 研究背景・研究目的
2. 提案手法
3. 実験例
4. まとめ・今後の展望



目次

1. 研究背景・研究目的
2. 提案手法
3. 実験例
4. まとめ・今後の展望



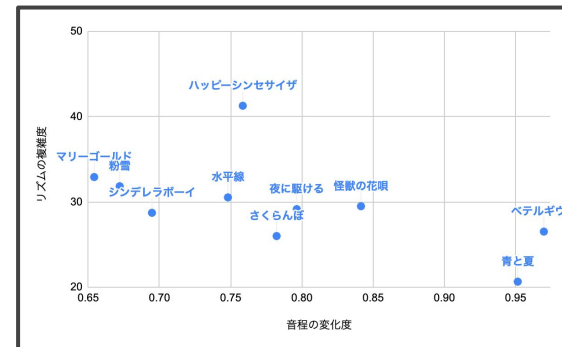
研究背景

- カラオケが世界的に広まっており、ユーザ自身が気軽に 歌唱演奏する機会が増大してきている。
- しかしながら、ユーザ自身の歌唱演奏のスキルに合致した楽曲を選定することが難しい。
- 歌唱演奏の難易度は一つの指標で表されるようなものではなく、複数の指標の組み合わせであると考えられる。

ユーザが、歌唱演奏のレパートリーを増やすためには、
楽曲の難易度を表す楽曲特徴の検討が必要であるとする

研究目的

- 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
- 歌唱演奏をする上で相対的な難易度を歌唱パートの下記の項目で決定されると仮定
 - リズムの複雑度
 - 音程の変化度
 - 音域
 - ロングトーン
- 上記の難易度に基づいてユーザが自身に合った楽曲メディアコンテンツを選択するための可視化を実現





目次

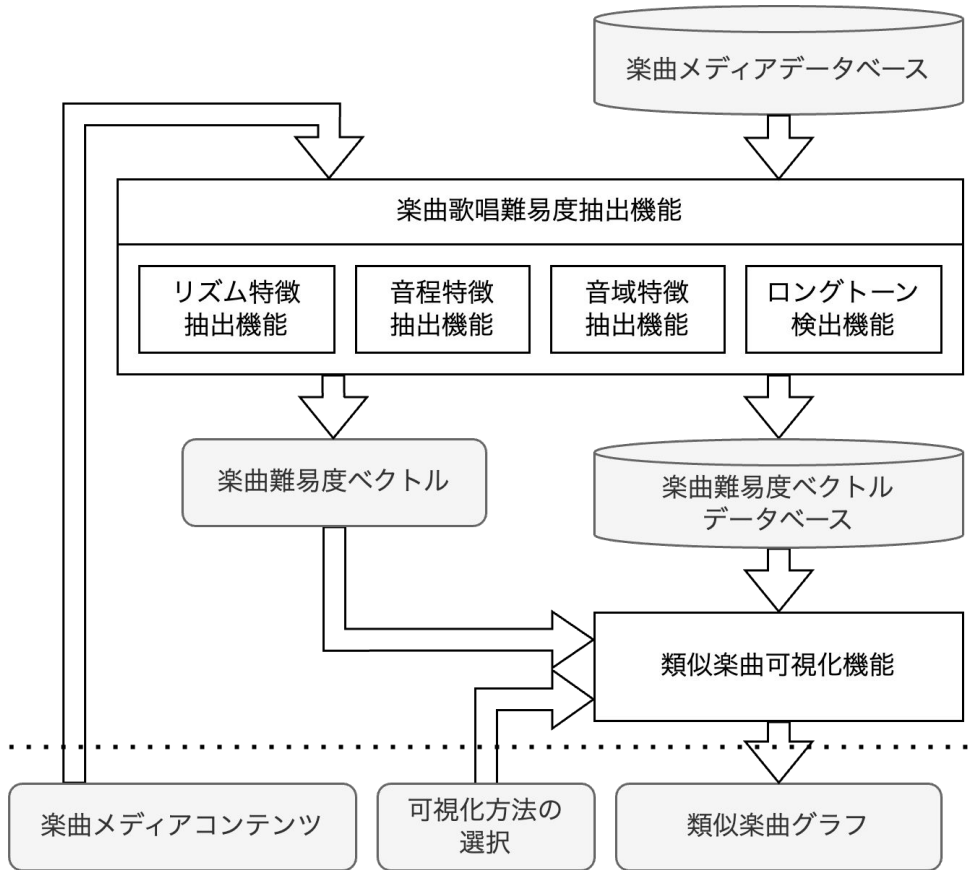
1. 研究背景・研究目的
- 2. 提案手法**
3. 実験例
4. まとめ・今後の展望



提案手法

大きく以下の2つの機能からなる

- 楽曲歌唱難易度抽出機能
- 類似楽曲可視化機能



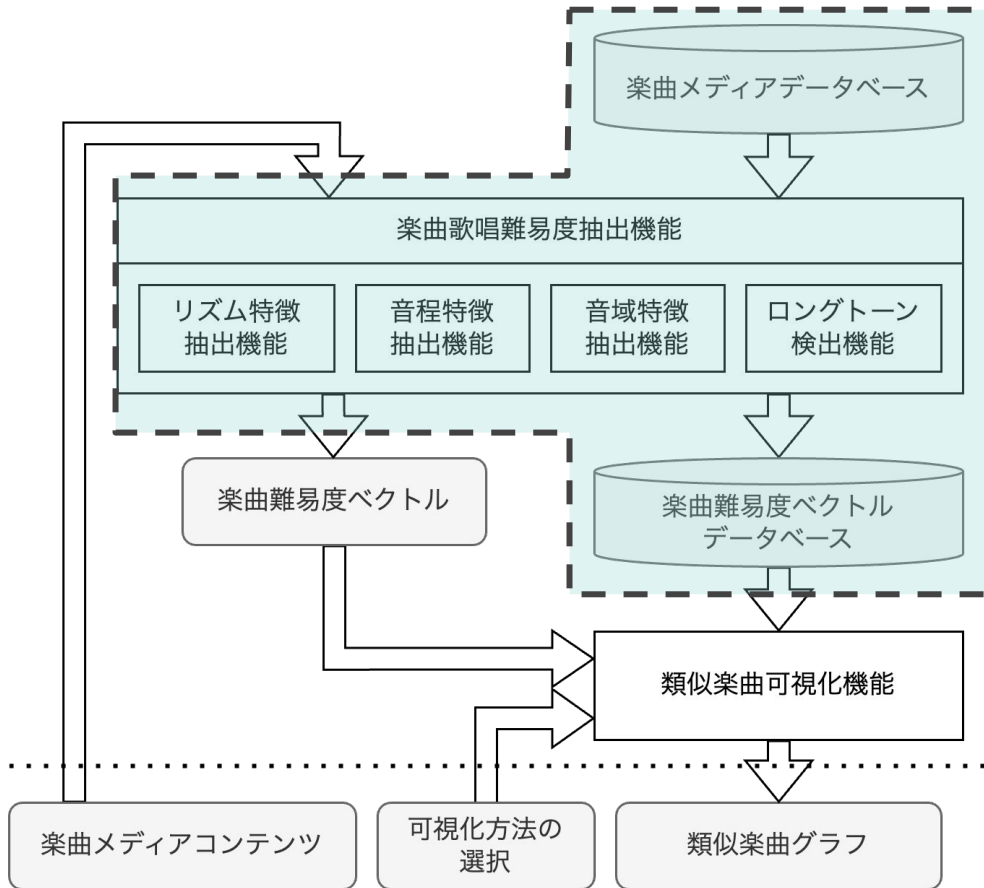
User

提案手法

事前に行う

の部分は、事前に実行しておき、楽曲の特徴を抽出して楽曲難易度ベクトルデータベースを作っておく。

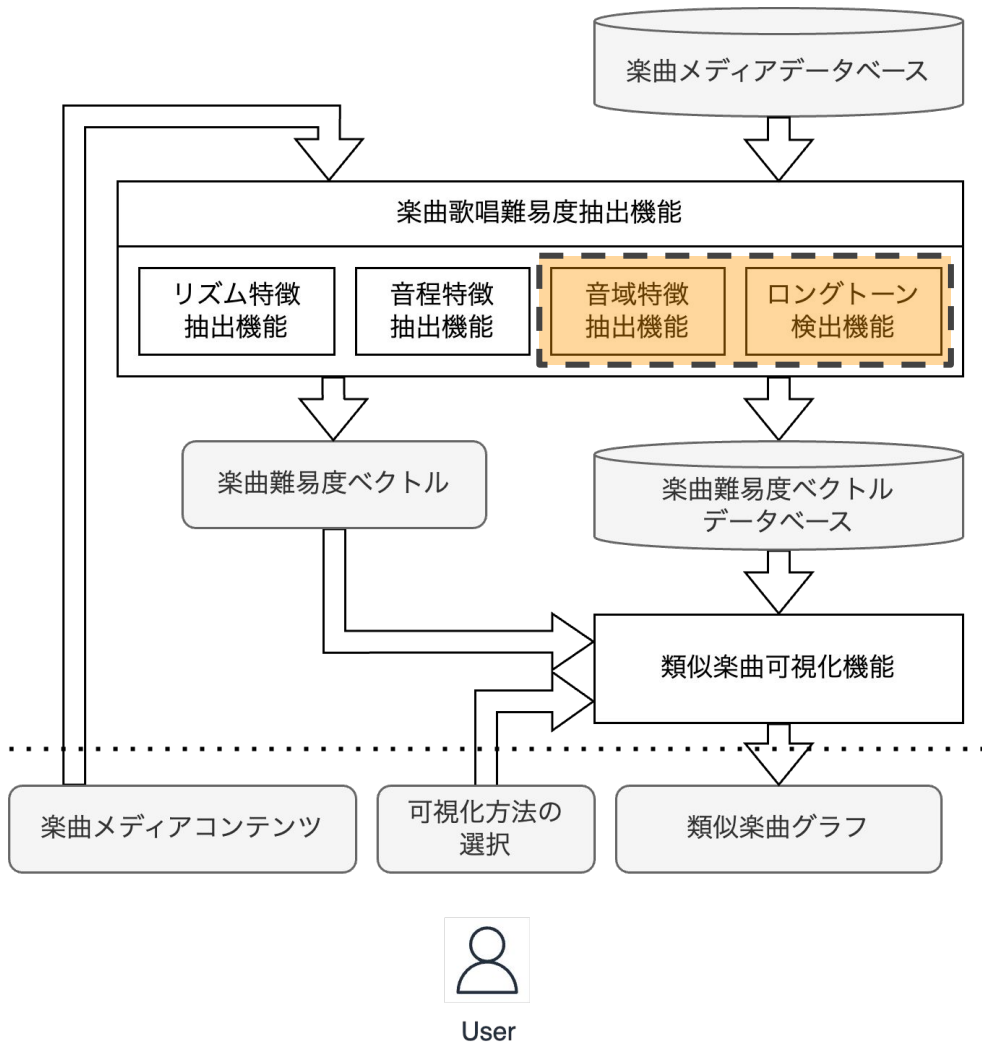
実行時には、楽曲メディアコンテンツを入力して、楽曲難易度ベクトルを抽出し、データベース内の楽曲との比較をして可視化を行う。



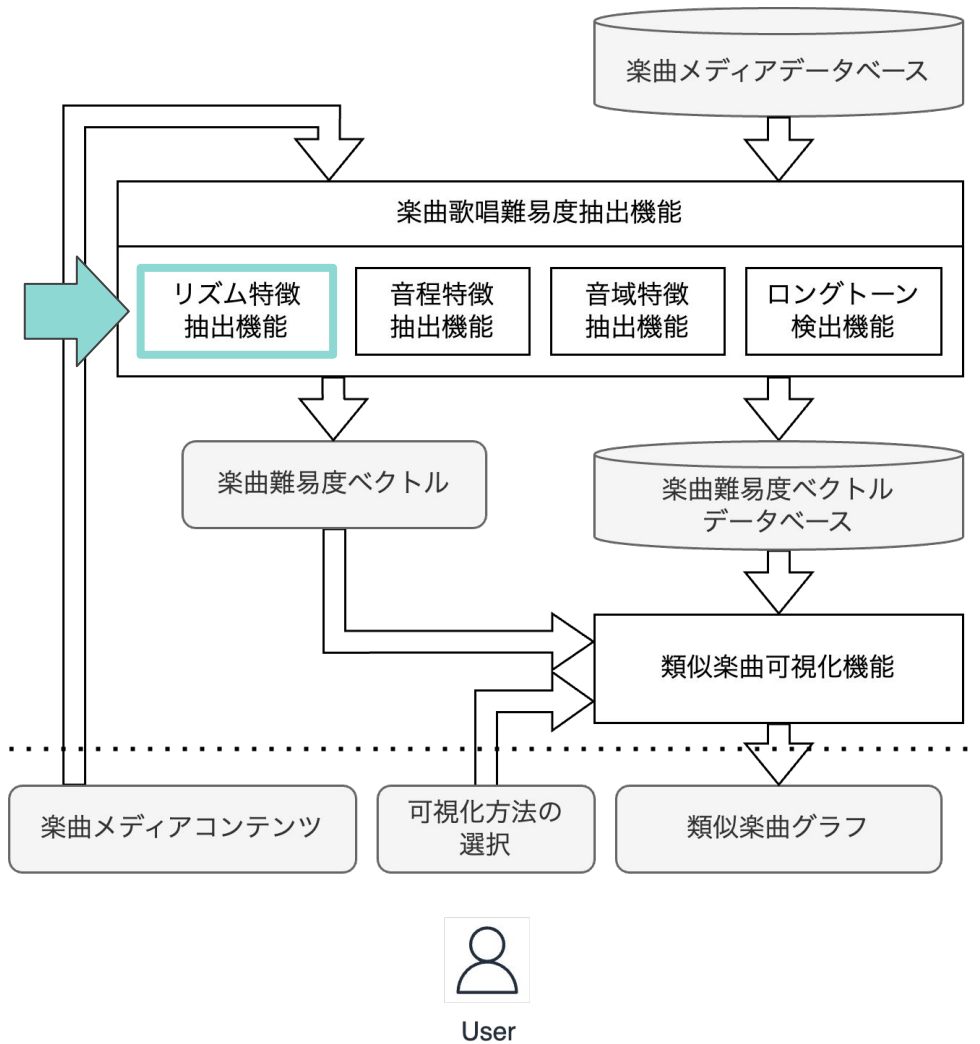
User

提案手法

未実装



提案手法





リズム特徴抽出機能

難易度の一つとして、歌唱パートの**リズムの複雑度**を抽出する。

リズムの複雑度は、メロディの音の長さの平均が短いほど複雑であり、長いほど単純であると仮定した。

また、リズムが複雑であるほど、歌唱演奏が難しく、リズムが単純であるほど、歌唱演奏が易しいと仮定した。



リズム特徴抽出機能

リズム特徴抽出機能では、

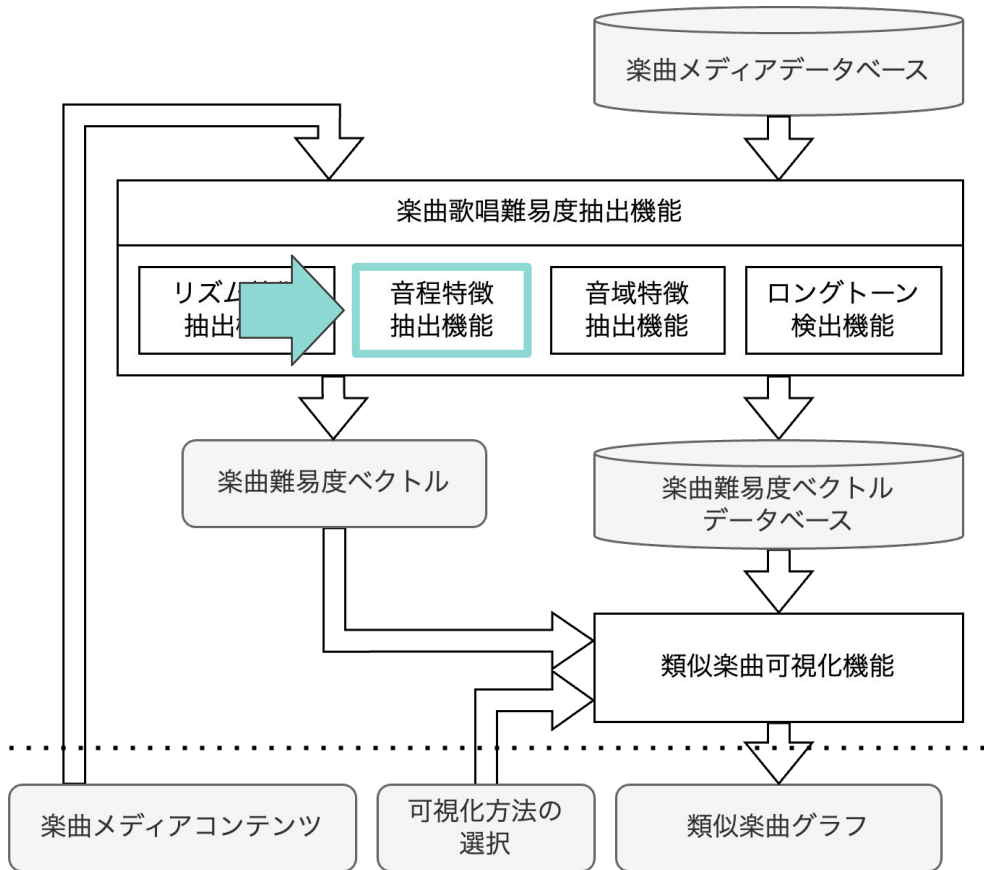
Songle API^[1]を用いて、楽曲メディアコンテンツのメロディの音の長さを抽出し、それらの平均をリズムの複雑度とする。

- ・Songle APIでは音の開始点と長さの情報が取れるためこれが可能である
- ・(音高の情報はない)

[1] Songle API <https://songle.jp/>



提案手法



User



音程変化特徴抽出機能

難易度の一つとして、歌唱パートの音程の変化度を抽出する。

音程の変化度は、前の音との差を数値化して平均を取ったものが大きいと音程の変化が大きく、小さいと音程の変化が小さいと仮定した。

また、音程の変化が大きいほど、歌唱演奏が難しく、音程の変化が小さいほど、歌唱演奏が易しいと仮定した。



音程変化特徴抽出機能

音程変化特徴抽出機能では、

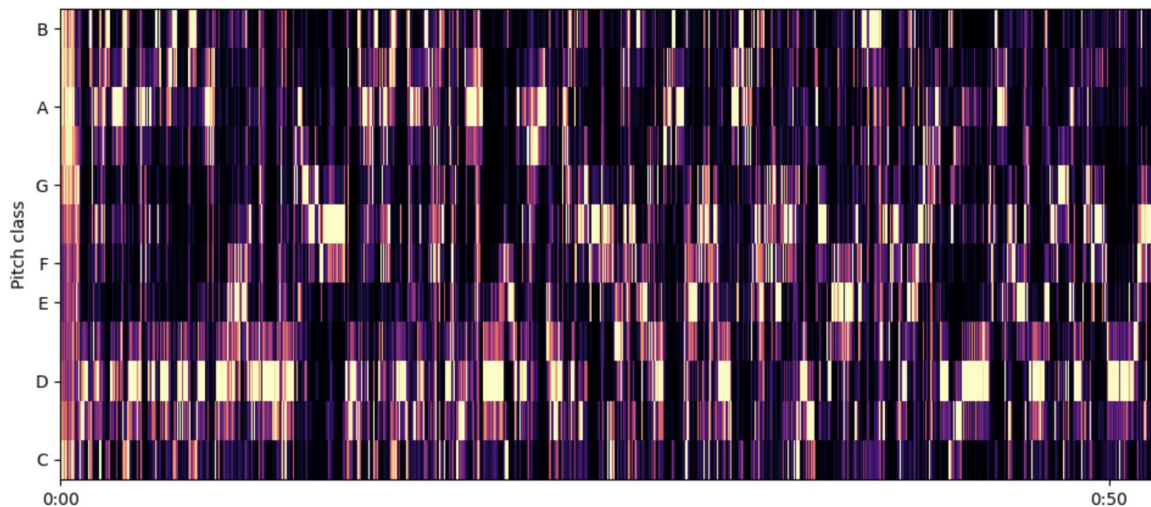
- ① 楽曲メディアコンテンツから Spleeter[2]でボーカルのみを抽出。
- ② クロマシフトを作成し、どの音が鳴っているかを抽出する。
- ③ 前の音との差の絶対値の平均をとることで、
音程の変化の特徴を抽出する。

[2] <https://github.com/deezer/spleeter>



クロマシフト

クロマシフト: 音楽のオクターブの12個の異なる半音(またはクロマ)を表す12個のビンにスペクトル全体を投影したもの



- ・時間単位ごとに最大値の箇所を取り出している図中の明るい色の部分
- ・現状ではメロディーの音程そのものではなく、オクターブを無視した音高の音程を用いている



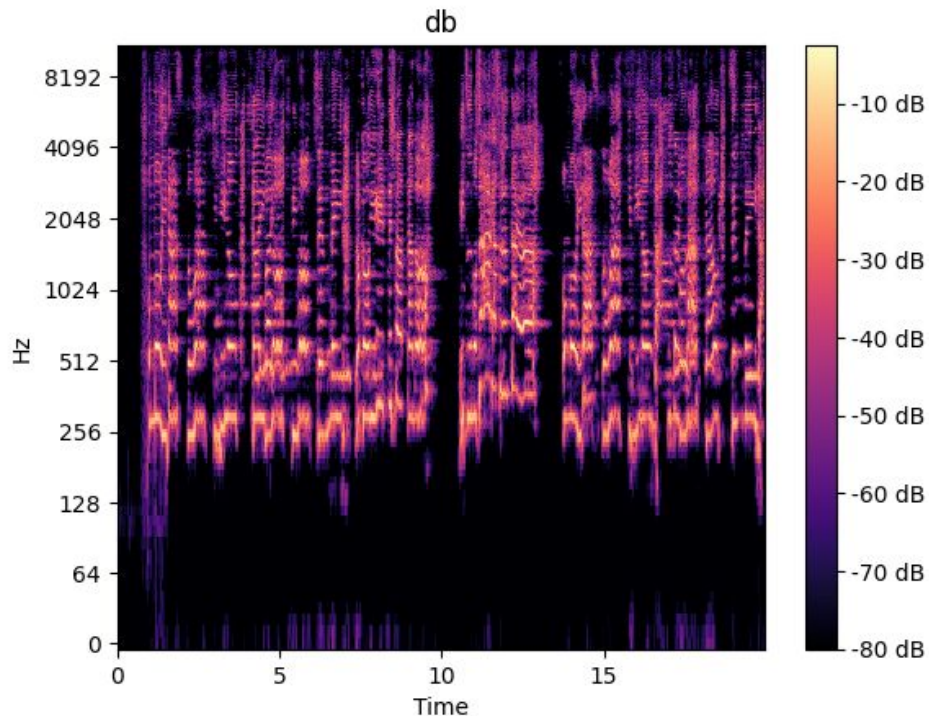
(参考)スペクトログラム

スペクトログラム: 音楽を周波数ごとに分けて時系列で可視化したもの

ここからメロディーの抽出して「音程変化」「音域」「ロングトーン」を抽出しようとしたが、高い精度でのメロディーの抽出を実現できなかった。

- オクターブ違いの音が最大音量になることがある
- ハモリパートなども抽出されている
- 伸ばしている音の中でも周波数が変化している

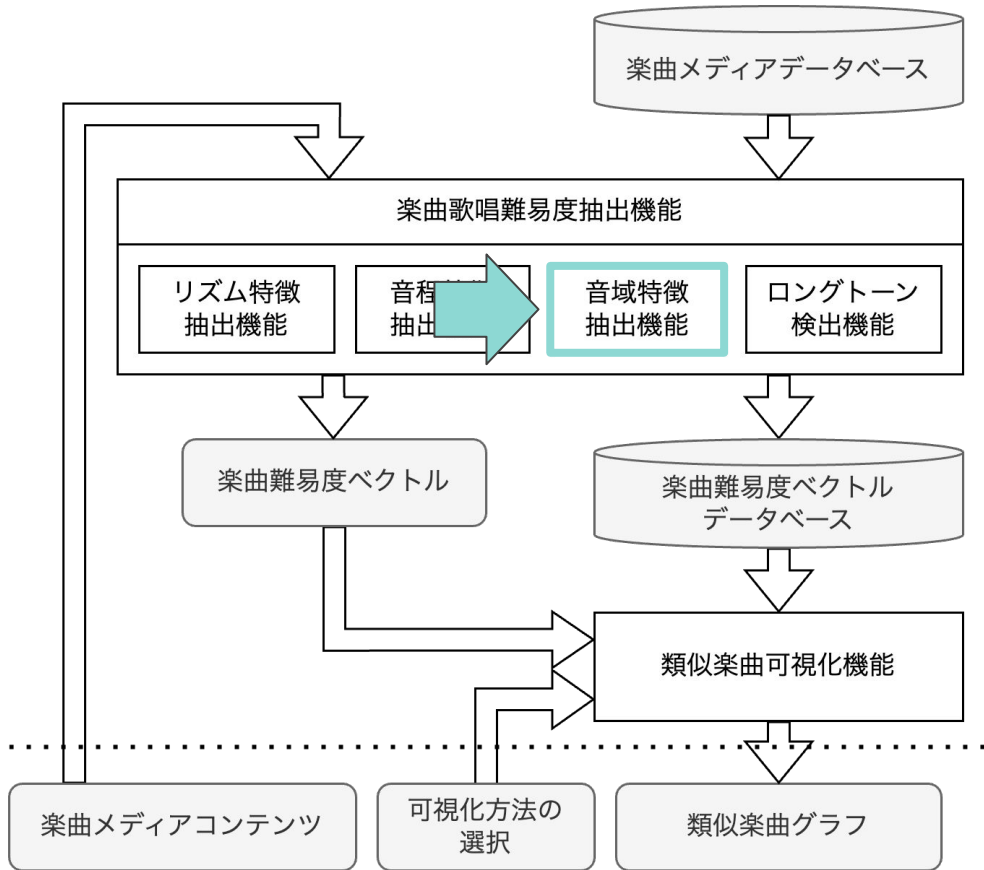
音程変化の抽出には簡易的な方法としてクロマシフトを用いた。



怪獣の花唄(Vaundy)の冒頭20秒から、Spleeterでメロディーを抽出し、スペクトログラムに変換したもの



提案手法



User



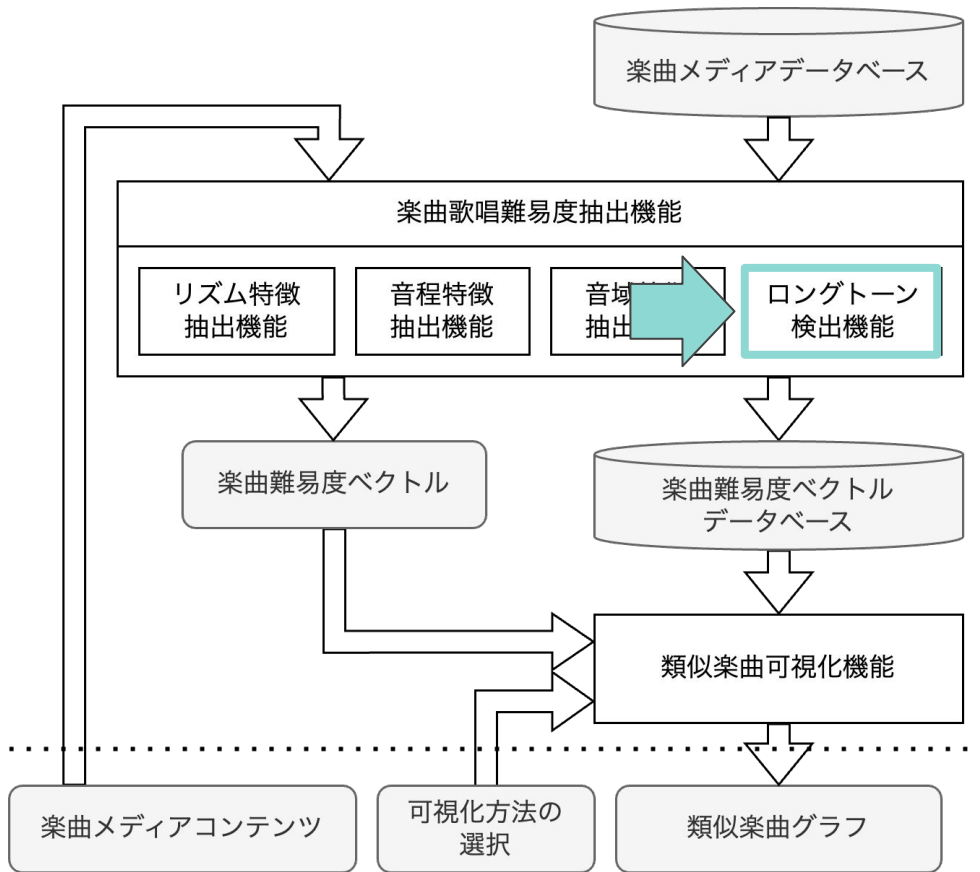
音域特徴抽出機能

難易度の一つとして、歌唱パートの音域を抽出する。

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。

楽曲の最高音と最低音と最頻値の音を抽出する。難易度は、比較のときにそれらの差分として表現する。

提案手法



User

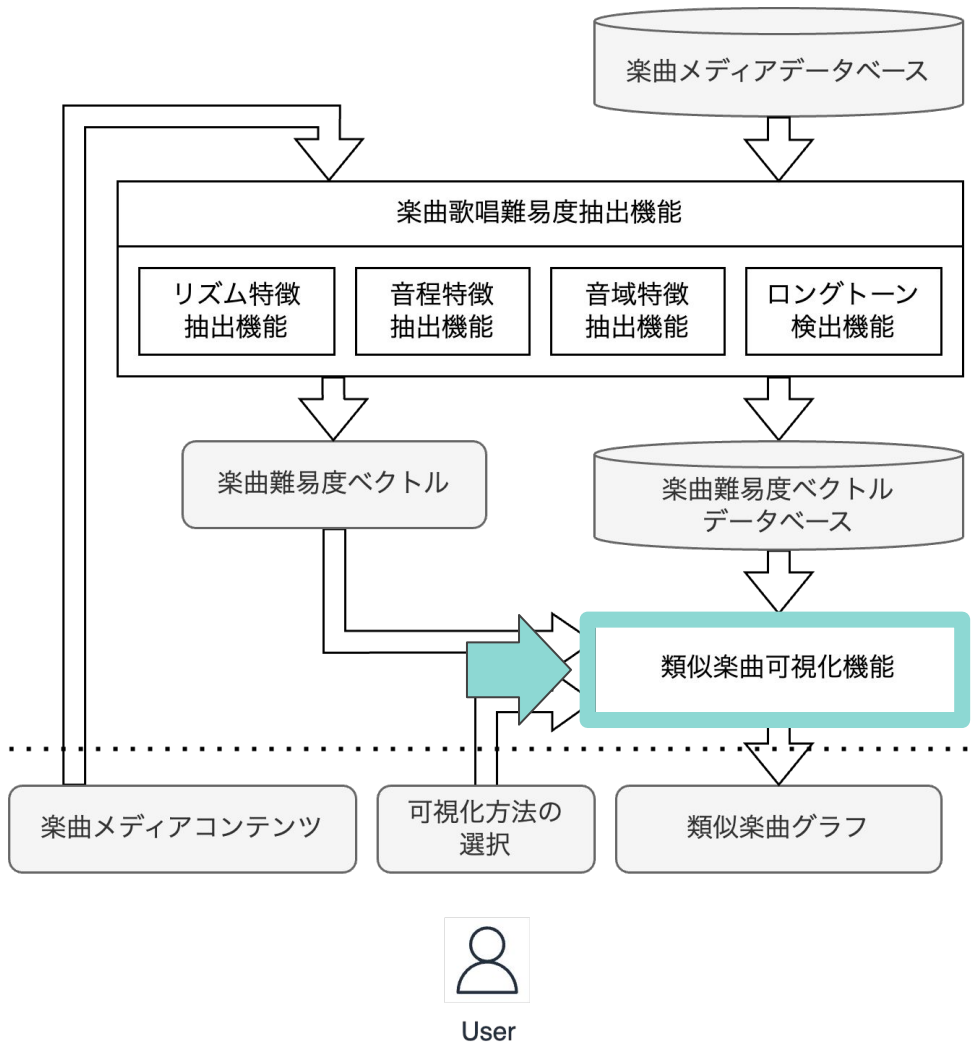


ロングトーン検出機能

難易度の一つとして、歌唱パートのロングトーンの長さを抽出する。

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、楽曲からメロディの最長音の長さを抽出する。

提案手法



類似楽曲可視化機能

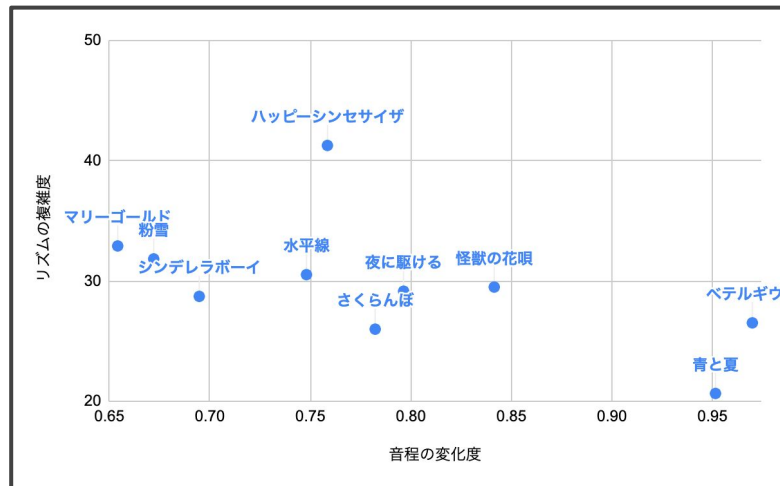
ユーザが入力した楽曲と難易度が類似している、データベース内の楽曲をグラフにして可視化する。ユーザは可視化のモードを選んで利用する。

モード1: 2つの特徴量を選んで可視化する
→ なんの特徴に差あるのかが分かりやすい

モード2: 全部の特徴量を使って可視化する(未実装)
→ 総合的に似ている曲を探することができる

モード1では、2つの特徴量を縦軸と横軸に割り当てて2次元グラフとして表示する

モード2では、特徴量をそれぞれ標準化し、主成分分析を行って第主成分までの2次元グラフとして表示する



モード1による可視化の例

楽曲の難易度の位置関係の表現

「リズムの複雑度」、「音程の変化度」、「ロングトーンの長さ」の3つの特徴量については、一つの値で表現されているため、そのままプロットすることで位置関係が表現できる。

残る「音域」については、以下の方式を定める。

ユーザが得意な曲として選んだ曲をa、比較対象の曲をbとする

曲aの最高音をa1、最低音をa2、最頻値をa3

曲bの最高音をb1、最低音をb2、最頻値をb3とする

最終的に、音域の違いによる相対的な位置を以下の式で表す。

$$k1(b1 - a1) + k2(b2 - a2) + k3(b3 - a3)$$

k1,k2,k3は係数とする。

最高音の差 + 最低音の差 + 最頻値の差

値が大きいほど、bの方が高いことを意味する。



目次

1. 研究背景・研究目的
2. 提案手法
- 3. 実験例**
4. まとめ・今後の展望



実験例

今回実装できた「リズム特徴抽出機能」と「音程変化特徴抽出機能」を用いて、楽曲の難易度を抽出し、楽曲の相対的な難易度の関係を可視化する。

楽曲メディアデータベースとして用いる楽曲は以下の10曲とし、入力楽曲としては、「ドライフラワー(優里)」を用いた。

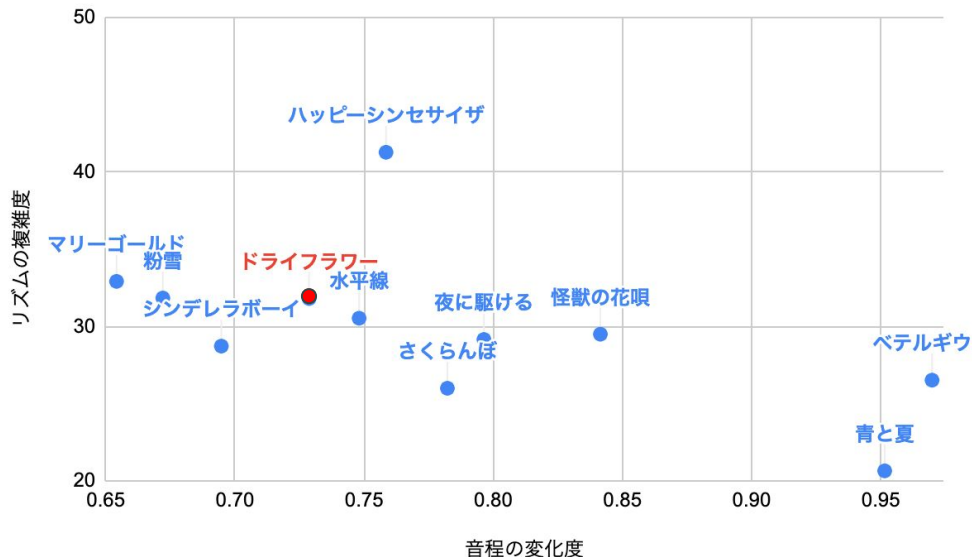
楽曲名	アーティスト名	楽曲名	アーティスト名
マリーゴールド	あいみょん	夜に駆ける	YOASOBI
粉雪	レミオロメン	さくらんぼ	大塚愛
シンデレラボーイ	Saucy Dog	怪獣の花唄	Vaundy
水平線	back number	ベテルギウス	優里
ハッピーシンセサイザ	EasyPop	青と夏	Mrs. GREEN APPLE

実験結果

例) ユーザ自身の歌唱演奏のスキルに合致した楽曲が「ドライフラワー(優里)」である場合

図の左上: 易しい
(リズムが複雑でなく、音程の変化が小さい)
図の右下: 難しい
(リズムが複雑で、音程の変化が大きい)

「ドライフラワー」と難易度の近い楽曲は「水平線」であると分かる





歌唱演奏スキルに合致した楽曲の抽出方法

ユーザ自身の歌唱演奏のスキルに合致した楽曲を抽出する方法

- ①ユーザ自身の歌いやすい曲から(リズムの複雑度、音程の変化度)を抽出し、散布図で表す。
- ②距離の近い楽曲を抽出する。



歌唱演奏のスキルに合致した楽曲を抽出できる



目次

1. 研究背景・研究目的
2. 提案手法
3. 実験例
4. **まとめ・今後の展望**



まとめ

- 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を行った。
- 歌唱演奏をする上で相対的な難易度を表現する特徴量を提案し、その中で「リズム特徴」「音程特徴」の実装を行った。
- 上記の難易度に基づいてユーザが自身に合った楽曲メディアコンテンツを選択するための可視化を実現した。



ユーザ自身が、容易に、歌唱演奏のスキルに合致した楽曲のレパートリーを増やすことができる



今後の展望

- 未実装部分の実装を行う
 - 楽曲からのメロディ推定
 - 音域特徴抽出機能
 - ロングトーン抽出機能
 - 可視化機能モード2(全ての特徴を用いた可視化)
- リコメンドシステムとしての機能の充実
 - ランキング形式(特徴別、総合)
- 実用的なアプリとしての実装



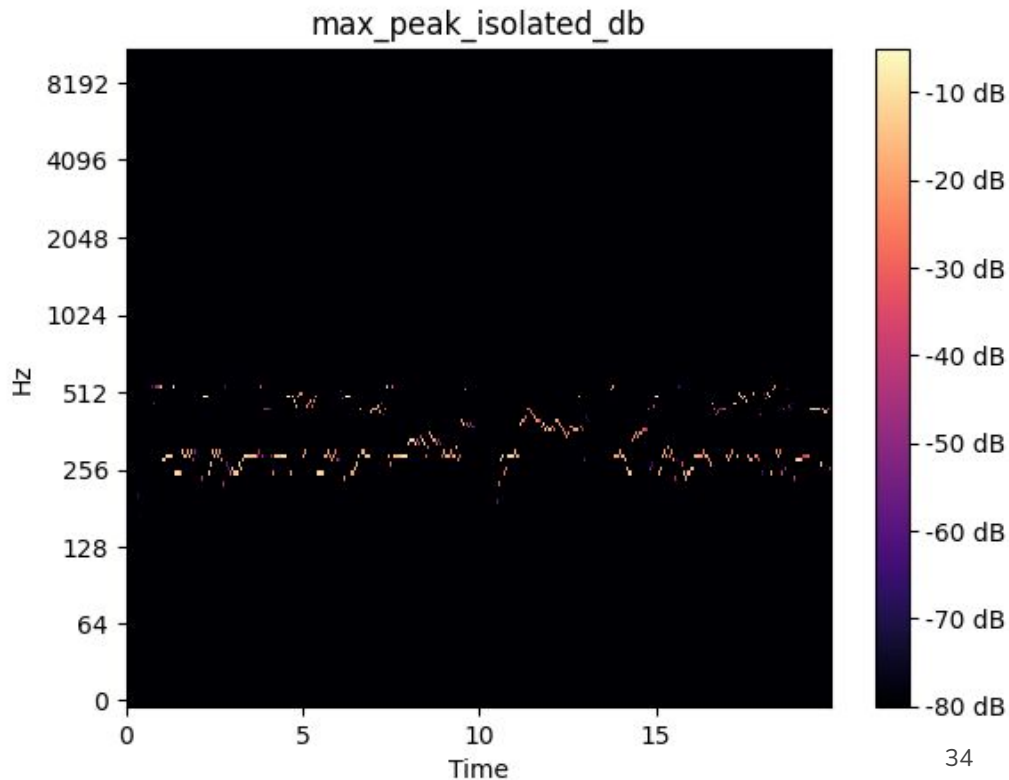
ご清聴ありがとうございました。



メロディ抽出の現状 (1/2)

スペクトログラムの中で時間ごとに音量が最大な部分を抽出した

高い音が抽出されてしまっているのが分かる



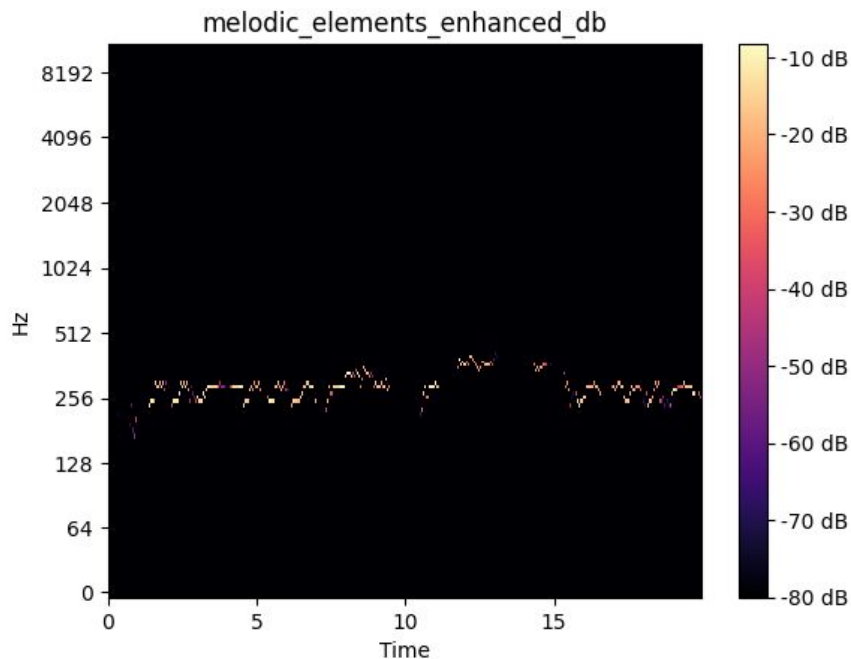
メロディ抽出の現状 (2/2)

音域を指定することで正確な抽出に近づいた

しきい値の設定により歌っていない部分を消すことができた

しかしまだ汎用性に欠ける

音域を抽出するのが目的なので、音域を指定する現状の手法は不十分



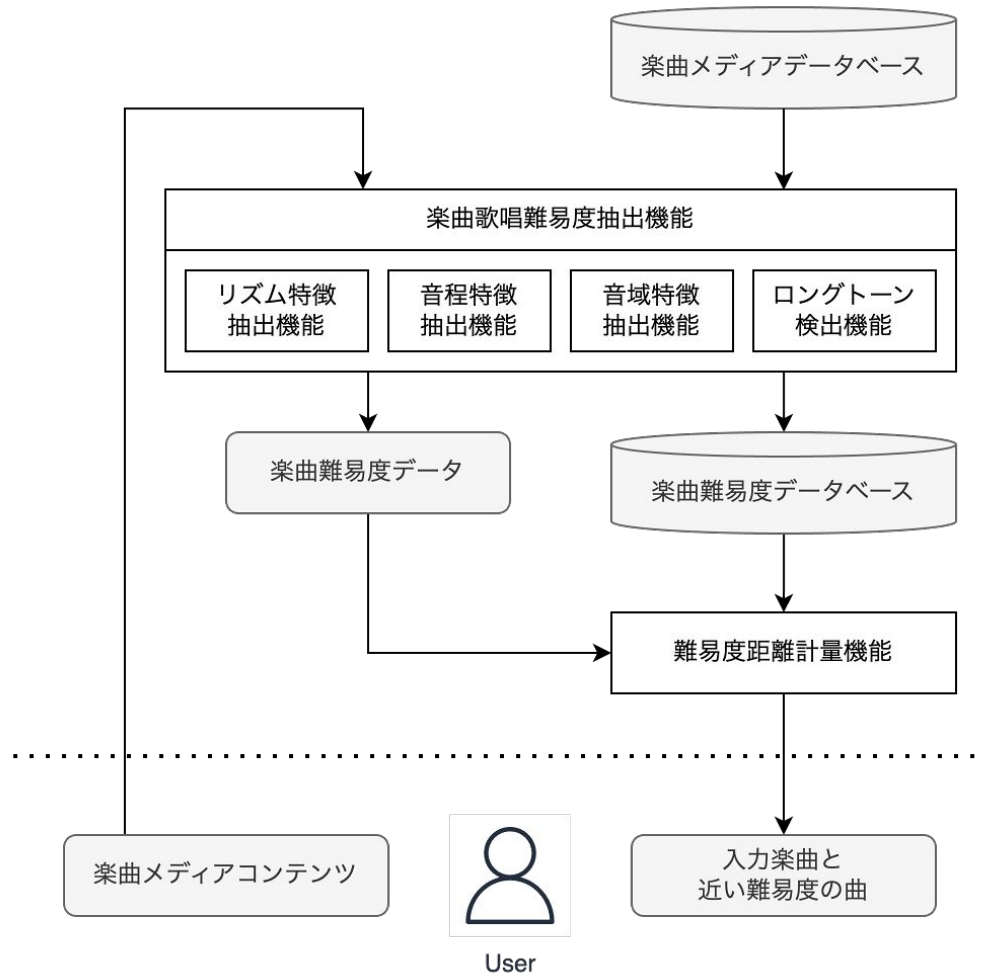


今後やること

音域とロングトーンによる難易度導出

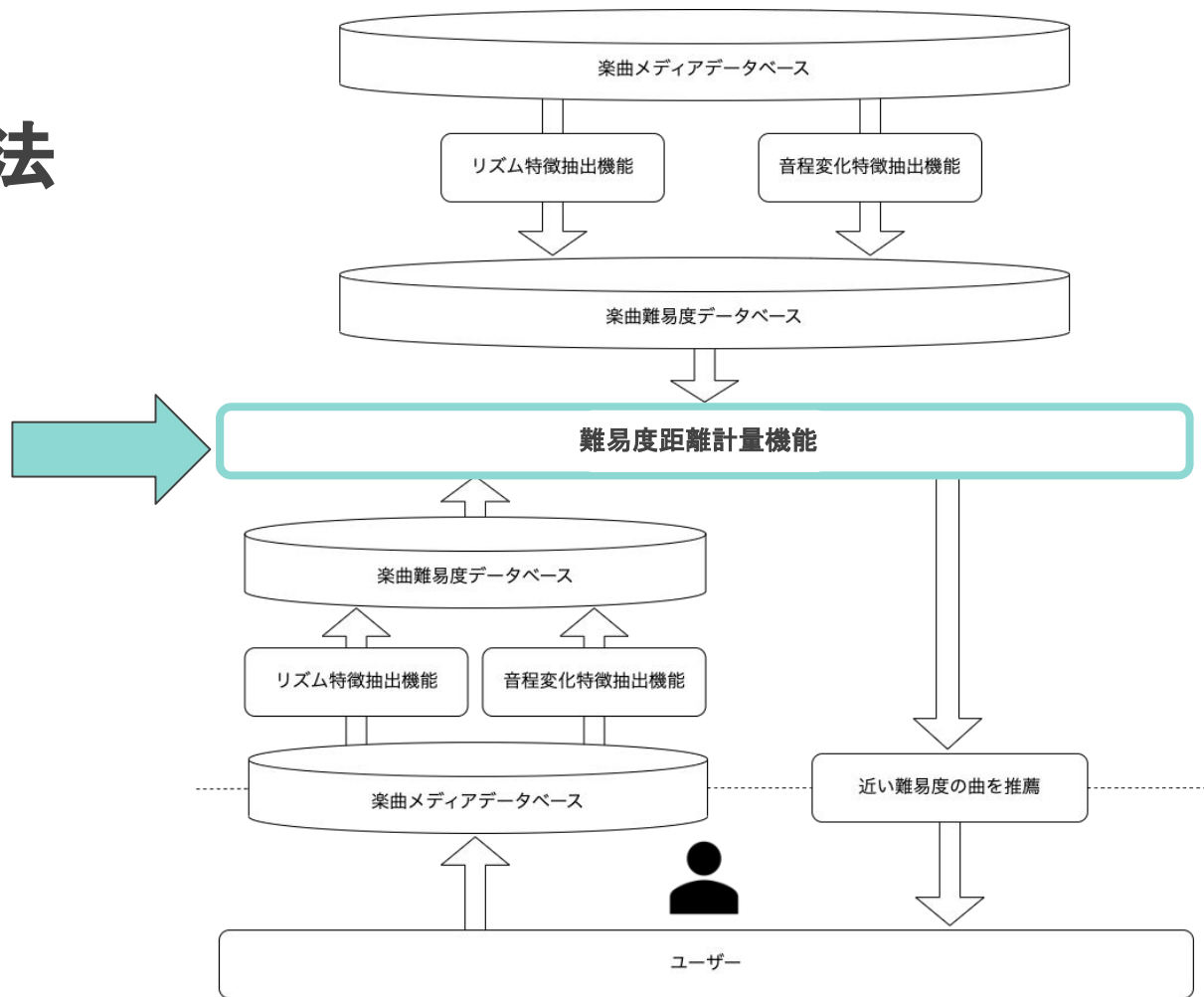
ビジュアルライゼーションを考える

リコメンドシステムの実現





提案手法





難易度距離計量機能

ユーザが選択した楽曲と難易度として近しい楽曲を導出するために、難易度距離計量機能を用いる。

難易度距離計量機能では、

楽曲メディアデータベースの全ての楽曲とのふたつの値(リズムの複雑度、音程の変化度)からなる難易度の距離を計量する。

実験結果 (データベース内の楽曲の位置関係の表示)

10曲の楽曲メディアコンテンツから(リズムの複雑度、音程の変化度)を抽出し、散布図で表した。

図の左上(リズムが複雑でなく、音程の変化が小さい)の曲が歌唱演奏をする上で易しい曲であり、図の右下(リズムが複雑であり、音程の変化が大きい)の曲歌唱演奏をする上で難しい曲である。

