

音楽からボーカルの音程を測定するプログラムの作成

京都産業大学附属高等学校3年 山本 煌

概要

特定の楽曲の譜面がほしい際、基本的には自分で音声を聞き取り書き起こす必要があった。そのため、AIによる音源の分離とピッチ分析を使用し、ボーカルの音程を測定するプログラムを作成した。

1. ボーカルの分離

初めに、Spleeterという音源分離ライブラリを使用してボーカルとその他の音の分離を行う。今回はボーカルとその他の分離のみを目的としているため、2stemsモデルを使用した。また、今回は海上自衛隊東京音楽隊により公開されている行進曲「軍艦」の音源を使用させていただいた。

元の音声ファイル ボーカル抽出した音声



かすかにボーカル以外の音が混じっているものの、基本的にきれいに分離できている。

2. ピッチ推定

分離したボーカルファイルに対してCREPEを用いてピッチの推定を行う。今回は、largeモデルを使用し、viterbiによる後処理を有効化して推定を行った。これによって、以下のようなピッチの推移を得ることができる。



3. ドレミへの変換

先ほど測定したピッチをもとに、ドレミへの変換を行う。周波数と、音階は以下のように定義される。

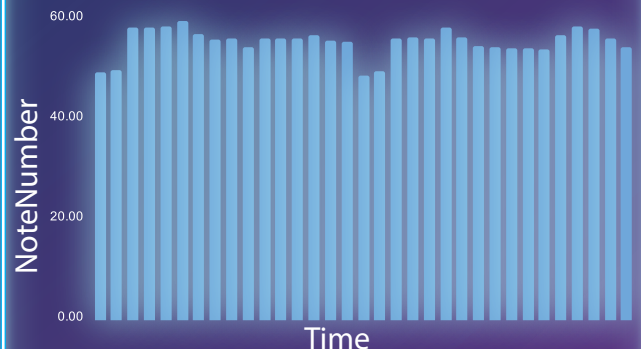
$$f = 440 \times 2^{\frac{n-69}{12}} \quad \begin{array}{l} f - \text{周波数} \\ n - \text{ノート番号} \end{array}$$

そのため、以下のような式を使用してドレミに変換する。

$$n = 12 \times \log_2 \frac{f}{440} + 69$$

4. 一定間隔ごとの音程

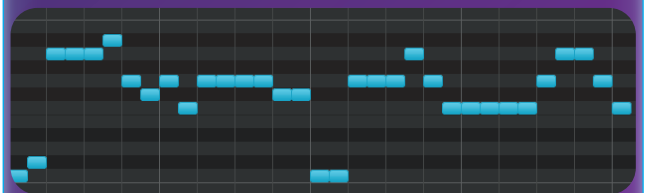
ピッチ推定で得られたデータには細かい変動が多く、そのまま使用するには使いにくい。そこで、一定間隔で先ほど変換したドレミの値から中央値をとる。また、今回は8分音符の間隔ごとに中央値を取得した。



ピッチ推定で得られたデータよりはがたつきが減っていることが分かる。

5. MIDIデータへの変換

これらの処理によって得られたデータをもとにMIDIデータを生成する。今回はpretty_midiというライブラリを使用して書き出しを行った。



上記の画像は、生成したMIDIデータをCakewalkというDAW(デジタル・オーディオ・ワークステーション)で表示した。

6. 課題と改善点

生成したMIDIデータはしっかりメロディーラインを認識することはできるもののところどころに不自然な音が入っていることもあり、改善の余地があるといえる。また、四分音符の位置で中央値をとっているため、四分音符より細かいリズムを含む曲では間隔を短く設定する必要がある。しかし、間隔を短くすると不自然な音が発生しやすくなるので調整が必要である。今後は精度をあげ、またボーカルを持たない楽曲のメロディー書き出しにも対応させたいと考えている。