

人間が歌った音声で音楽を作る

理工学部数理情報学科

T130083 三橋 宏哉

指導教員 佐野 彰

概要

音楽を作るには楽譜の読み書きができることが望ましい。本研究では楽譜の読み書きができない人でも音声によって楽譜を自動生成し、音楽制作ができるのシステムの開発を開始した。

本研究でのシステム制作は Python を使って制作している。

音階は周波数(Hz)の大きさに決まる。その音程を調べるため本研究では人の発生から離散フーリエ変換にかけることで周波数成分を取り出し、音階を求めていく。まず、音声をマイクから入力し、wave ファイルとして保存する。wave ファイル形式の音声データは圧縮されておらず、そのまま周波数解析の処理が可能となる。wave ファイルには 1 秒間で何回かのデータを取るかを示すサンプリング周波数があり、サンプリング定理によりサンプリング周波数の半分まで周波数を表現することが可能であるが、標準のサンプリング周波数(44,100Hz)では本実験において不要な周波数まで扱ってしまう。その為、サンプリング周波数を変更可能な録音システムを製作した。

wave ファイルに保存された音声を、短時間(0.025ms)毎に離散フーリエ変換することで短時間ごとのフーリエスペクトルが求まる。ここから短時間毎の音階の候補を得て、これらの統計処理することで発声ごとの音階を決定する。

本研究では、短時間ごとの周波数スペクトルのデータから、人間の音声の音階を正確に決定する為の方法を決定した。人の音声のような音から得られる周波数スペクトルから音階を決定するためにはおおくの問題点がある。ノイズが多く入ったり、個人差による変化が問題となってくる。これらの問題点は、周波数スペクトルの一番大きな点を取るのではなく、複数選択することで誤処理を減らす。そして選択した周波数スペクトルを平均値で処理するのではなく重み付け平均値で処理したり、選択する周波数スペクトルの個数を変更し、ノイズや個人差による変化を無視できるように実験を行い、処理方法を変更することで改善していく。

本研究では一つの音階を決定することは可能になった。しかし、発声の変わり回などで複数の音階が重なるような時間で音階を決定することは困難であった。たとえば人が「ミ」と「ラ」を続けて発声した場合、「ミ」から「ラ」に変わる時の音階の決定は不正合となってしまった。

今後の課題はこのように連続した発声から楽譜として適した音長を決定することである。本研究を進めていく上で連続した音声の間には綺麗に決定できる音階の時とは違った周波数スペクトルの特徴が存在するように思えた。その特徴を見つけることにより、発声の変わり目を適切に処理を可能にすることで楽譜を自動生成するシステムを次の段階に進めていきたい。