

歌唱テクニックの自動検出に向けた 技術基盤の構築

Establishing foundations for automatic singing technique detection

学籍番号：201921652

氏名：山本 雄也

Yamamoto Yuya

歌唱は人類の音楽文化の根源である。我々が最も気軽に行うことができる表現の一つでもあり、歌を歌う・聴くことは我々の生活を豊かなものにしてきた。しばしば歌手は、歌唱テクニックを用いて歌い方を変化させ、個性豊かな音楽表現を行う。歌唱テクニックとは歌手が瞬時的に発音・声質・音高等を変化させる音楽表現技法のことをいう。歌唱テクニックの頻度・出現箇所・種類等は歌手の声質以上に個性が発揮される部分であり、それらを調べることは歌手の歌い方の理解につながる。本研究では、歌唱テクニックの歌声からの自動検出を目指す。歌唱テクニックの種類と出現箇所を歌声から自動検出することができれば、歌手の歌い方の観察・分析を大規模に行うことが可能になり、楽曲検索への応用が期待できる。

しかし、歌唱テクニック検出は以下に示す3つの障壁が存在するために挑戦的な課題である。1) 検出対象についてこれまで議論が不十分であり、ポピュラー音楽の既存曲のどの部分に何のテクニックが現れているのか実際に調べられていなかったこと。2) 多様な歌唱テクニックのモデリングが難しく、性能が頭打ちになりやすいこと。3) 歌唱テクニックの出現時刻付きラベルデータが不足しており、新たな作成にも多大な手間・負担を要するために困難であること。

この課題に対し、それぞれ以下のような基礎検討を行い、歌唱テクニックを歌声から検出するための技術基盤を構築する。1) 歌唱テクニックについて網羅的な文献調査を行い、検出する候補を列挙したうえで、実際にポピュラー音楽のデータセットから歌唱テクニックを発見し集計する。2) 深層学習により歌唱テクニックのモデリング（主に特徴抽出）を自動化し、従来の専門知識に基づくモデリングと比較する。3) 大規模な歌唱テクニックの出現時刻なしラベルデータセットを用い、弱教師あり学習の枠組みで検出モデルを構築する。その結果、1) について、歌唱テクニックの定義を明確にした上で歌唱テクニックのラベルが含まれたデータセットを構築した。2) について、分類実験において深層学習を用いた自動モデリングが正解率を2.6%上回った。特にモデリングの難しい、通常の歌唱と乖離した極端な歌唱テクニックにおいて識別性能を上回ったことから、歌唱テクニックのモデリングの困難を解消するために深層学習が有用である可能性を示した。3) について、提案手法の検出性能はF-microが0.215と低かった。問題の枠組みによって検出できない歌唱テクニックもあり、いまだデータ・ラベルの不足が検出のための障壁であることを示した。本論文の終結部では主に3)に対する解決策となり得る手法の列挙も含め、残存する課題を取りまとめ、歌唱テクニックの自動検出にむけた出発点を示した。

研究指導教員：平賀 譲

副研究指導教員：寺澤 洋子