

## 論文の内容の要旨

論文題目    **Source-Filter Representation and Phase Estimation in Continuous Wavelet Transform Domain for Monaural Music Audio Editing**  
(連続ウェーブレット変換領域におけるソースフィルタ表現と位相推定によるモノラル音楽音響信号加工の研究)

氏 名    中村 友彦

本研究では、モノラル音楽音響信号を音高や楽器などの単位に分解し、分解成分を個別に加工することを可能にする音楽信号分離および合成手法を提案する。これは、ユーザによる音楽制作や既存楽曲の加工支援システム、楽音それぞれ加工可能な音楽プレイヤー、計算機による自動編曲システムなどの音楽アプリケーションに応用できる。

音響信号加工の性能は音源分離の精度に大きく依存するため、加工の前段で高精度な分離を行うことが重要である。一般に高精度な分離のためには解を適切に限定するための手がかりが必要となる。時間周波数表現（スペクトログラム）領域では調波性などの音源分離に有用な手がかりを利用できるが、適切な周波数解像度のスペクトログラムを選択するためには対象がどのような音響信号であるかを考慮することが重要である。また、そのスペクトログラム上で対象の音響信号をどのように表現できるかという点も考慮する必要がある。したがって、時間周波数表現を意識したアプローチをとるべきである。

そこで、本研究では以下の3つの方針を考えこれらを同時に考慮した手法を提案する。まず第1の方針として (i) 対数周波数解像度を与える連続ウェーブレット変換 (continuous wavelet transform, CWT) によるスペクトログラムを利用する。これは音楽における各音高の基本周波数 ( $F_0$ ) は対数尺度で均等に並ぶ性質があるからである。次に、第2の方針として (ii) 楽音の生成過程モデルを活用する。ソースフィルタ理論によると楽音の生成過程は楽器の振動体と共鳴体に分離して考えることができ、楽音のスペクトルに関する仮定が見通しよく立てられるためである。第3の方針として (iii) スペクトル漏れを考慮する。実際に観測されるスペクトルが取ることを許される形状には制約があり、もしスペクトル漏れの具体的な形状や関数が分かっているならば、近接した異なる音源の  $F_0$  成分や高調波成分を分離する手がかりとなるからである。これ

ら3つの方針を同時に考慮したアプローチを実現するためには、スペクトログラム領域とソースフィルタモデルなどの時間信号領域のモデルとの対応関係を得る必要がある。しかし、CWTの基底波形は直交しないためどのように対応関係を得るべきかは必ずしも知られていない。

第3章では方針 (i) と (ii) を同時に考慮するために、CWT領域へのソースフィルタモデルの導入に取り組む。対数周波数領域では $F_0$ と高調波周波数の間隔が $F_0$ によらず一定である性質を活かしつつ、ソースフィルタモデルを導入したスペクトログラムモデルを提案し、補助関数法と呼ばれる最適化原理を用いて提案モデルのパラメータを反復的に推定する収束性の保証されたアルゴリズムを導出する。ソースフィルタモデル導入による音源分離性能の向上を実験により確認した。

第4章では方針 (i) と (iii) を同時に考慮するために、CWT領域でのスペクトル漏れの記述に取り組む。音源分離に有用な手がかりである時間周波数表現の局所的な構造と大局的な構造を同時に利用しつつ、スペクトル漏れを考慮したスペクトログラムモデルを導出し、前章と同様に補助関数法を用いて提案モデルのパラメータ推定アルゴリズムを導出する。この手法を調波時間因子分解 (harmonic-temporal factor decomposition, HTFD) と呼ぶ。音源分離性能の評価実験によりスペクトル漏れの考慮とCWT領域での分離の有効性を確認した。

第5章では全方針を考慮するため、第4章で提案したHTFDを拡張しCWT領域でスペクトル漏れとソースフィルタモデルを同時に考慮したスペクトログラムモデルを導出する。HTFDの解析的な時間信号モデルを介し、離散時間信号領域で定義されるソースフィルタモデルとCWT領域で定義されたHTFDのスペクトログラムモデルのパラメータの対応関係が得られることを示し、新たなスペクトログラムモデルを提案する。前章と同様に、補助関数法を用いて閉形式の更新則からなるパラメータ推定アルゴリズムを導出する。スペクトル漏れの考慮に加えソースフィルタモデルを導入することの有効性を実験により確認した。

第6章では、分離後に加工された振幅スペクトログラムを時間信号に変換するために、振幅スペクトログラムからの高速位相推定法を提案する。時間周波数表現を意識したアプローチをとることで、時間領域信号に対応する複素スペクトログラムが満たす条件を導出し、その条件を元に位相推定問題が最適化問題として定式化できることを示す。前章までと同様に補助関数法に基づき収束性の保証されたアルゴリズムを導出し、提案法の有効性を実験により確認した。

第7章では、歌声の振幅スペクトログラムがスパース行列、伴奏の振幅スペクトログラムが低ランク行列とみなせることを利用し、 $L_p$ ノルム規準の非負値行列因子分解による歌声分離手法を提案する。歌声の振幅スペクトログラムのスパース性を表すパラメータを適切に定めることで、分離性能が向上することを確認した。

最後に、第8章では2つの音楽音響信号間で調波楽器音の周波数特性の置換および打楽

器音の音色置換が可能なシステムを提案する．主観評価実験により，調波楽器音の周波数特性と打楽器音の音色どちらについても提案法の有効性を確認した．