

審査の結果の要旨

氏 名 中村 友彦

近年、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音楽アプリケーションを意識した音響信号加工技術の研究開発が望まれている。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、加工前段で高精度な分離を行うことが重要であり、そ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。その一つとして、調波性などの音源分離に有用な手がかりを与える時間周波数表現（スペクトログラム）に基づくアプローチは有効と考えられる。

本論文は、このような問題認識のもとで、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したもので、「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おけるモノラル音楽音響信号加工の研究）」と題し、全9章からなる。

第1章「Introduction」では、本研究の背景を説明した後、本研究の方針が以下の3点にあり、これらを同時に考慮した手法を提案することが本論文の目的であると述べている。(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の利用、(ii) 楽音の生成過程モデルの活用、(iii) スペクトル漏れの考慮。

第2章「Audio Source Separation Approach with CWT Representations」では、上記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ことを述べ、具体的研究課題を説明している。

第3章「Shifted NMF with Source-Filter Model」では方針(i)と方針(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について検討している。まず、対数周波数領域では基本周波数と高調波周波数の間隔が基本周波数によらず一定である性質に着目し、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ている。ついで、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し、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ている。

第4章「Harmonic Temporal Factor Decomposition」では方針(i)と方針(iii)を同時に考慮するために、CWT領域でのスペクトル漏れについて検討している。まず、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することにより、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ている。ついで、補助関数法を用いて、調波時間因子分解 (harmonic-temporal factor decomposition, HTFD) と呼ぶ提案モデルのパラメータ推定アルゴリズムを導出している。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それぞれ確認している。

第5章「HTFD with Source-Filter Model」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している。まず、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している。ついで、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し、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ている。

第6章「Fast Signal Reconstruction from Magnitude CWT Spectrogram」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している。まず、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している。補助関数法に基づいて、収束性の保証された最適化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ている。

第7章「 L_p -Norm NMF for Singing Voice Enhancement」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している。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ている。

第8章「Timbre Replacement of Drum Components in Music Audio Signals」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し、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ている。

第9章「Conclusion」では、本論文の結果をまとめ、今後の課題を述べている。

以上を要するに、本論文は、音楽アプリケーションを意識した音響信号加工技術で重要となる加工前段での高精度な音源分離を実現する手法ならびにそれに基づく合成手法を提案し、その有用性を検証したもので、工学上貢献するところ大である。よって本論文は、博士(情報理工学)の学位請求論文として合格と認められる。