

学位論文題目 Exploiting Regularities in Natural Acoustical Scenes for Monaural Audio Signal Estimation, Decomposition, Restoration and Modification (邦訳: 音環境に内在する規則性に基づくモノラル音響信号の推定・分解・復元・加工に関する研究)

取得年月 2009年3月 **学位種別** 博士(情報理工学) **大学** 東京大学・パリ第6大学
ルルー・ジョナトン(日本電信電話(株)NTTコミュニケーション科学基礎研究所リサーチアソシエイト)
音楽情報科学

氏名

推薦研究会

推薦文

Jonathan Le Roux 君は、フランス最高峰の大学を卒業し数学の修士学位を2つ得た後、東大の博士課程で音楽音響音声信号処理を多方面から精力的に研究した俊英で、学位論文は信号に内在する構造から音楽信号分離やスペクトルからの位相復元や断続信号の補間などを扱っていて量・質とも高く評価されました。

音声アプリケーションや通信システムなどの音を媒体としたあらゆるアプリケーションを実現する上で、非定常な背景雑音、複数音源信号の重畳、パケットロスによる信号の欠落などの問題に対処できるシステムを構築することは重要な研究課題である。我々人間は周囲の複雑な音響情景を、信号源の知識や所与の観測信号の統計的規則性に基づいて頑健に認識し理解することができる。本論文の目的は、音環境に内在する規則性のモデルに基づき、かくのごとく複雑な観測信号から音響情景を解析するための数理的アプローチを確立することである。本アプローチでは主として、(1)ある音環境モデルのもとで、何らかの原因で歪み、複数の音が混在し、あるいは部分的に欠損した観測信号データに対し、どのような音響的事象が生じているかについての最も自然な解釈を与える問題、(2)上記モデルに基づき欠損した情報を復元する問題、(3)上記モデルを観測データから自律的に獲得する問題を扱う。これらすべての問題に対し、数理的定式化を通して目的関数を立て、これによりさまざまな問題を拘束つき最適化問題として見通しよく解決することを可能にした。

まず、人間の聴覚機能の1つとして知られる音脈分凝プロセスにおいて関与しているとされる音声の規則性を拘束条件として付与した時間周波数パワー領域の統計モデルを導入し、単一音声・雑音重畳音声・複数話者による同時発話音声の基本周波数推定・音声強調・音源分離などの音響情景解析のさまざまな課題を扱えることを示した。

次は、不完全データから情景解析を行うために上記モデルを拡張する方法、欠損データを復元する方法を補助関数法と呼ぶ最適化の原理をもとに考案した。

以上では、時間周波数パワー領域において位相を考慮しないことによる各種課題が残されていた。以下では、これらの課題解決のために、複素時間周波数領域および時間領域での信号処理法を確立した。

まず、時間周波数パワー成分から信号を復元するための位相付加アルゴリズムを提案した。短時間フーリエ変換(STFT)は時間信号の冗長な表現となるため、任意に選んだ複素数の時

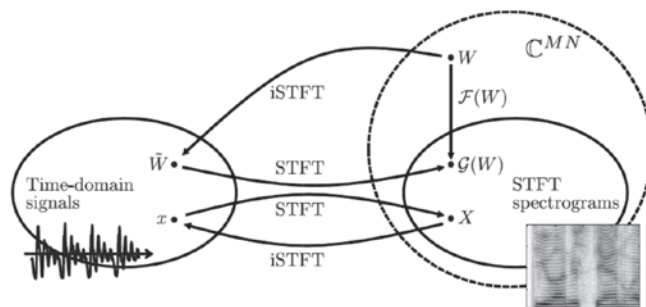


図-1

間周波数配列がある波形信号のSTFT表現であるとは限らない(図-1参照)。波形信号に対応するSTFT表現としてどれだけの矛盾のない複素数配列であるかを表す無矛盾性規準を導いた。さらに、これを局所的に計算できる近似的規準を導入することにより、高速な位相付加アルゴリズムを実現した。

時間領域の手法として、観測信号を基底波形のスパースな非負結合でモデル化した適応的テンプレートマッチング手法を提案し、音楽や生体信号を対象とした動作実験により、信号の中に繰り返し生起する基底波形とともに、それらの生起時刻と結合係数も推定できることを確認した。

また、以上ではSTFT等の古典的なフィルタバンクから得られる時間周波数表現を扱ってきたが、観測データから信号分析に適したフィルタバンクを自律的に学習する方法を検討した。人間の聴覚において重要な役割を果たしていると考えられる信号の変調構造をうまく獲得できるよう聴覚末梢系を適応させる人間の機能について仮説を立て、その学習機能を数理的に定式化した。

本論文の主要な成果は、広範囲な用途がありながら解決が難しいとされてきたモノラル入力からの音響情景解析の問題の本質を明らかにし、信号の推定・分解・復元・加工の問題を統一的な観点より解決した点にある。

(平成22年3月30日受付)