

音響信号から音の“設計図”をひもとく

- 複素NMFと複合自己回帰系による音響信号のスパース表現 -

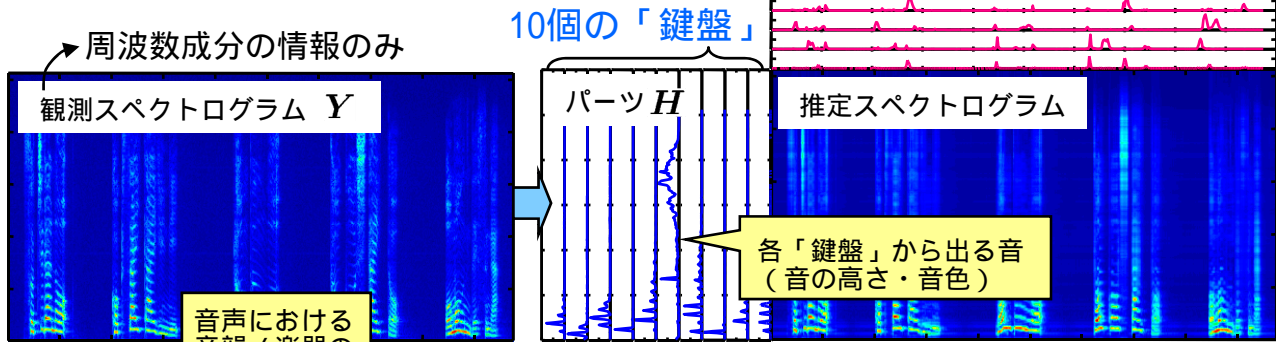
どんな研究？

音響信号がどのような音響的“パーツ”によってどのように構成されているか（“設計図”）を情報論的アプローチで推定する計算理論を研究しています。今回はこれを実現する2つの手法（複素NMFと複合自己回帰系）を紹介します。

もたらされる変革

将来、推定した“パーツ”を選択・置換したり“設計図”を書き換えて音を再構成することで、混合音から目的音を選択的に抽出したり、音の情報の一部、たとえばある音の音色だけを変えたりすることが可能になります。

複素NMF たとえるなら、実世界音響信号を一種の「鍵盤楽器」から生成されたものと捉え、そのもとで、各「鍵盤」が、押されたときどんな音を出すか、いつからいつまで押されたか、を観測信号から推定します。



周波数成分の情報のみ
観測スペクトログラム Y

10個の「鍵盤」

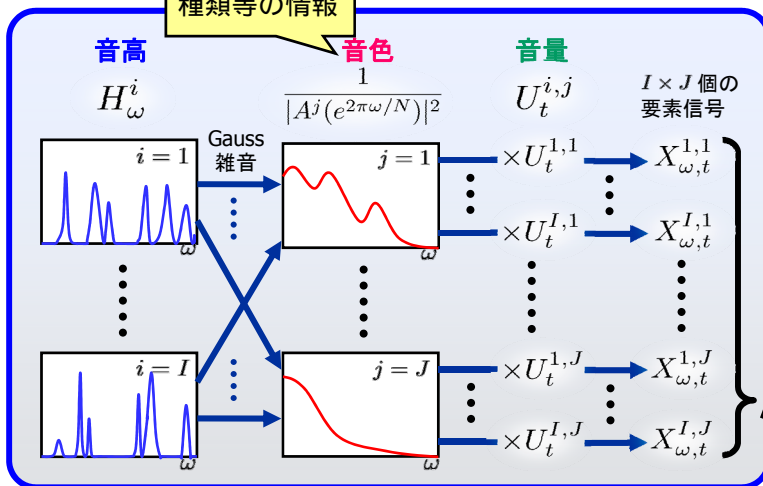
パーツ H

推定スペクトログラム

音声における音韻/楽器の種類等の情報

各「鍵盤」から出る音（音の高さ・音色）

各「鍵盤」が弾かれるタイムテーブル（発話・演奏のリズム/抑揚）



複合自己回帰系

概念的には、複素NMFで考えたモデルを高機能化したものです。たとえるなら、実世界音響信号を、複数台の「鍵盤楽器」の合奏から生成されたものと捉えます。重要なのは、各「鍵盤楽器」は一つの音色の音しか出せないという拘束があることです。例えば、音韻ごとの「鍵盤楽器」によるアンサンブルは音声に他なりません。

実世界音響信号 $S_{\omega,t}$



関連文献

H. Kameoka, N. Ono, K. Kashino, S. Sagayama, "Complex NMF: A New Sparse Representation for Acoustic Signals," In Proc. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2009), Apr. 2009.
H. Kameoka, K. Kashino, "Composite Autoregressive System for Sparse Source-Filter Representation of Speech," In Proc. 2009 IEEE International Symposium on Circuits and Systems (ISCAS2009), May 2009.

連絡先: 亀岡弘和 (Hirokazu Kameoka)

メディア情報研究部 メディア認識研究グループ