45

音響信号から音の“設計図”をひもとく
Extracting “sound map” from audio signal

SCIENCE PLAZA 2010

Motivation どんな問題に取り組むのか？

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によって推定する計算理論を研究しています．今回は**混合音の構成要素となっている音を抽出する**複素NMFと呼ぶ手法を紹介します．

Originality 得られた結果はどう新しいのか？

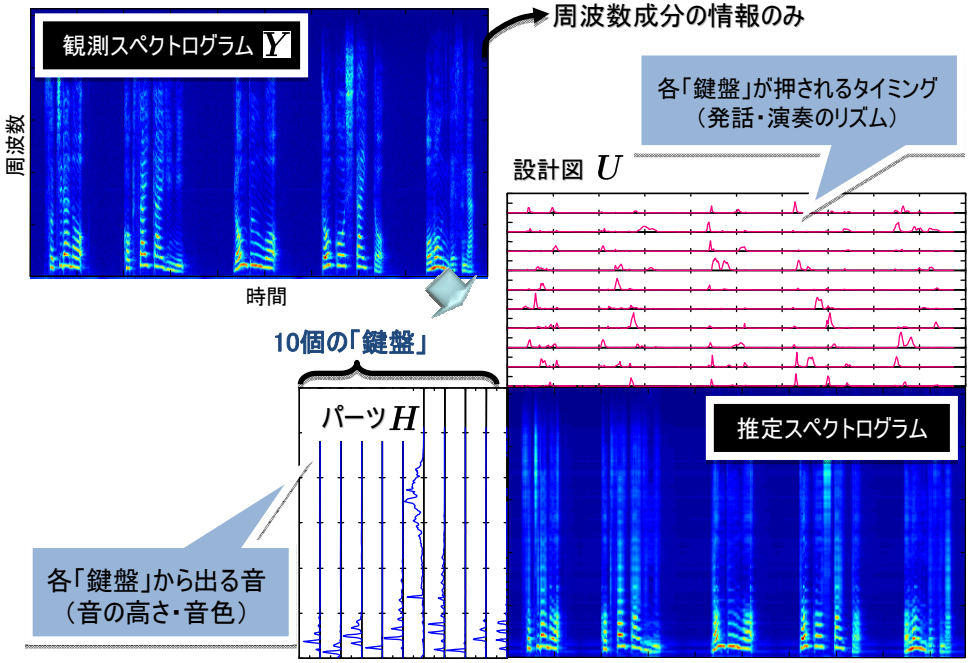
単純な周波数分析では様々な周波数成分が複雑に重畳するため個々の音の分離抽出は困難でしたが，提案法では繰り返し生起する音を重要な構成要素として学習するアイディアにより，これを可能にします．

Impact この研究が成功した場合のインパクトは？

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，押された時どんな音を出すか，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。



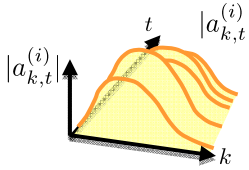
音響信号モデル

(仮定1) 音響信号はI 種類の要素信号だけから構成:

$$F_{k,t} = \sum_{i=1}^{I} a_{k,t}^{(i)}$$

(周波数, 時間)

(仮定2) 各要素信号の周波数成分比は時不変:



$$|a_{k,t}^{(i)}| = H_k^{(i)} U_t^{(i)} \quad (H_k^{(i)} \geq 0, U_t^{(i)} \geq 0),$$

$U_t^{(i)}$: 時刻tでのゲイン

$H_k^{(i)}$: 周波数成分比 $\sum_k H_k^{(i)} = 1$

位相スペクトルは時変: $\arg(a_{k,t}^{(i)}) = \phi_{k,t}^{(i)}$.

$$Y_{k,t} \simeq F_{k,t} = \sum_i H_k^{(i)} U_t^{(i)} e^{j\phi_{k,t}^{(i)}}$$

(観測時間周波数成分)

$$\theta := \{H_k^{(i)}, U_t^{(i)}, \phi_{k,t}^{(i)}\}$$

最適化アルゴリズム

$$\text{minimize} \quad f(\theta) := \sum_{k,t} |Y_{k,t} - F_{k,t}|^2 + 2\lambda \sum_{i,t} |U_{i,t}|^p$$

$$\text{subject to} \quad \sum_k H_k^{(i)} = 1 \quad (i = 1, \cdots, I)$$

$$f^+(\theta, \bar{\theta}) := \sum_{i,k,t} \frac{1}{\beta_{k,t}^{(i)}} \left| \bar{Y}_{k,t}^{(i)} - H_k^{(i)} U_t^{(i)} e^{j\phi_{k,t}^{(i)}} \right|^2 + \lambda \sum_{i,t} \left\{ p|\bar{U}_t^{(i)}|^{p-2} U_t^{(i)2} + (2-p)|\bar{U}_t^{(i)}|^p \right\}$$

$$\bar{\theta} := \{\bar{Y}_{k,t}^{(i)}, \bar{U}_t^{(i)}\} \quad \text{ただし,} \sum_i \bar{Y}_{k,t}^{(i)} = Y_{k,t}$$

任意の定数 $\beta_{k,t}^{(i)}$, p :

$$\sum_i \beta_{k,t}^{(i)} = 1, \beta_{k,t}^{(i)} \in (0,1), p \in (0,2]$$

$$f(\theta) \leq f^+(\theta, \bar{\theta})$$

局所最適解への収束が保証

Step 1) $\bar{\theta} \leftarrow \underset{\bar{\theta}}{\operatorname{argmin}} f^+(\theta, \bar{\theta})$

Step 2) $\theta \leftarrow \underset{\theta}{\operatorname{argmin}} f^+(\theta, \bar{\theta})$

いずれも閉形式で求まる！

NTT

NTTコミュニケーション科学基礎研究所
NTT Communication Science Laboratories

連絡先
亀岡弘和(Hirokazu Kameoka) kameoka@cs.brl.ntt.co.jp
ルルージョナトン(Jonathan Le Roux) leroux@cs.brl.ntt.co.jp
大石康智(Yasunori Ohishi) ohishi@cs.brl.ntt.co.jp