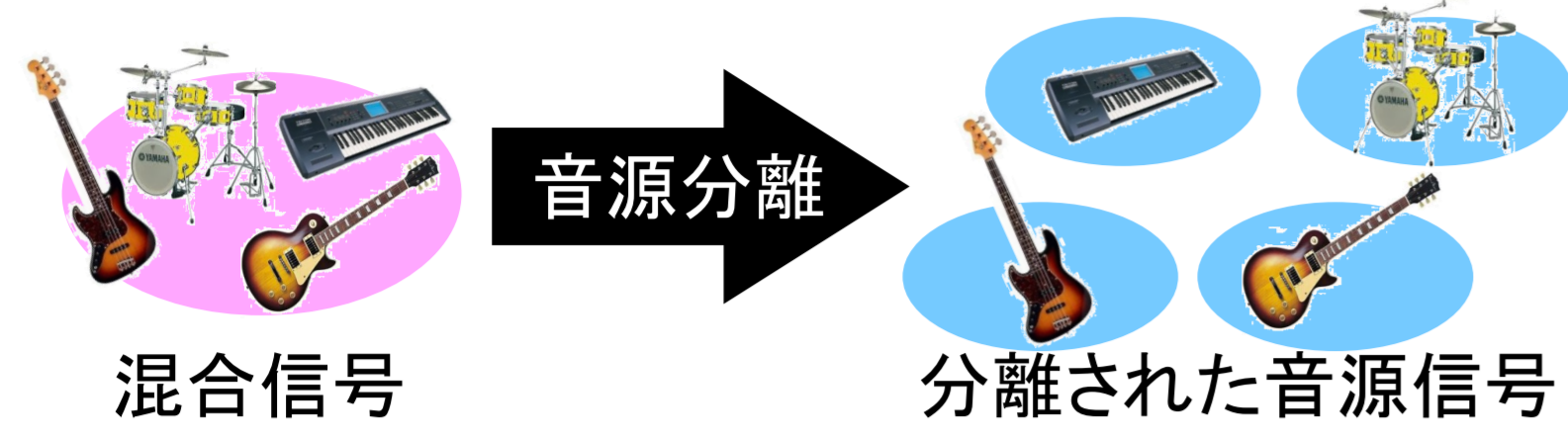


# コサイン類似度罰則条件付き非負値行列因子分解に基づく音源分離の実験的評価

☆岩瀬佑太, 北村大地 (香川高専)

## 1. 研究背景

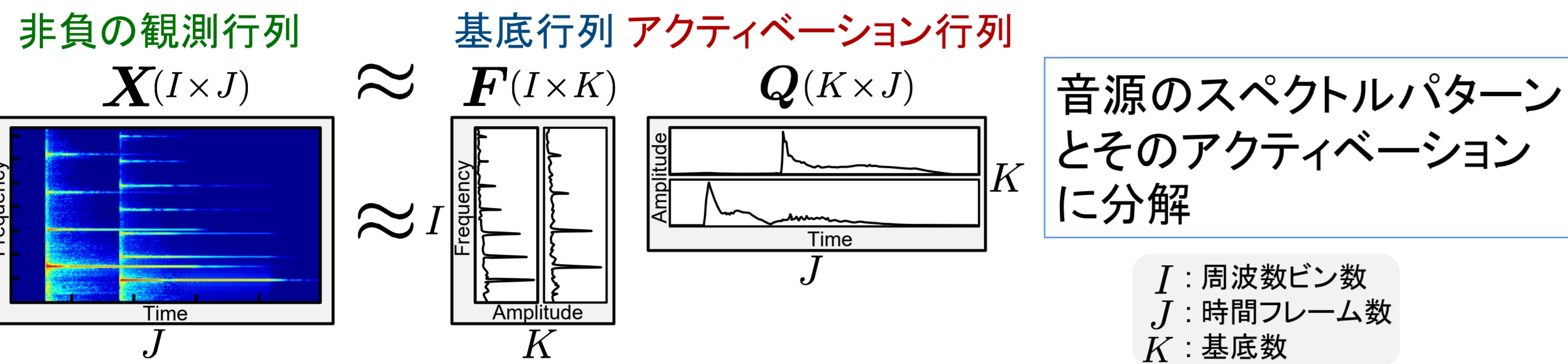
- 音源分離: 複数の音声や楽器音の混合音から個々の音源を推定する技術



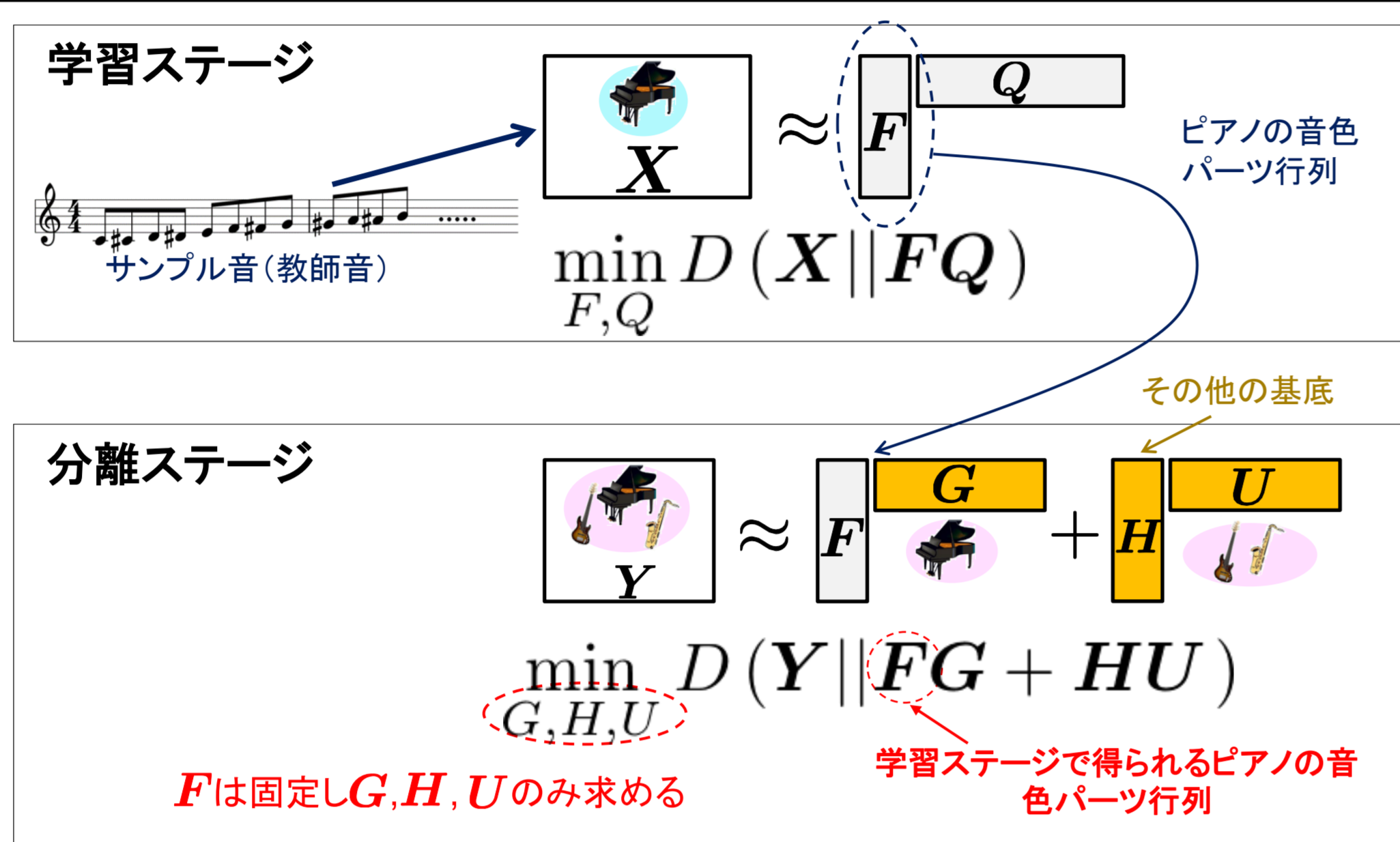
### 応用例

- 自動採譜の前段処理
- 音声認識の精度向上etc.

- 非負値行列因子分解 (nonnegative matrix factorization: NMF): 低ランク行列分解



- 半教師ありNMF (semi-supervised NMF: SNMF)  
事前学習した教師基底行列を用いて混合信号の振幅スペクトログラムを分解し目的音源成分  $FG$  と非目的音源成分  $HU$  を推定



### 解決すべき問題

混合信号中の目的音源と非目的音源が類似した音色を含む場合、2つの基底行列  $F$  と  $H$  に類似したスペクトルパターンが現れ、分離精度は劣化

### 本研究の目的

非目的音源の基底行列  $H$  に、教師基底行列  $F$  と類似したスペクトルパターンが現れないように最適化する新しい罰則条件付きSNMF (penalized SNMF: PSNMF) を提案し、性能を調査

## 2. 従来手法

- 内積型PSNMF [Kitamura et al., 2014]

- 教師基底行列  $F$  と非目的音源の基底行列  $H$  ができるだけ直交するように、内積に基づく罰則条件を付与したSNMF

$$\min_{G,H,U} D(Y || FG + HU) + \mu \|F^T H\|_{Fr}^2$$

内積罰則項

- 小さな  $\alpha$  を用いて  $H \leftarrow \alpha H$  及び  $U \leftarrow (1/\alpha)U$  とすれば、 $FG + HU$  の値を変えずに内積罰則項  $\|F^T H\|_{Fr}^2$  の値を最小化可能

➡ 本質的には、内積罰則項は最適化に影響を与えていない

- Log-cos型PSNMF [岩瀬ら, 2020]

- 内積ではなく、大きさに非依存なコサイン類似度の対数を罰則条件に用いる
- 内積型PSNMFの内積罰則項が最適化に影響しない問題を解決

$$\min_{G,H,U} D(Y || FG + HU) + \mu \sum_{k,l} \log \frac{\sum_i f_{ik} h_{il}}{\left(\sum_i f_{ik}^2\right)^{\frac{1}{2}} \left(\sum_i h_{il}^2\right)^{\frac{1}{2}}}$$

対数コサイン類似度罰則項

- コサイン類似度の「対数」を用いた理由は、補助関数法に基づく最適化の反復更新式導出を容易にするため(分数関数を対数で引き算に分解すると簡単)

➡ 対数が原因で罰則項の下限値は  $-\infty$  となり、最適化が不安定

## 3. 提案手法

- Cos型PSNMF

- 対数を用いない純粋なコサイン類似度を罰則条件に用いる
- 上記の内積型PSNMF及びlog-cos型PSNMFの問題を同時に回避可能
- 補助関数法に基づく反復更新式の導出は多少複雑になる(後述)
- 最適化問題は次式

$$\min_{G,H,U} D(Y || FG + HU) + \mu \sum_{k,l} \frac{\sum_i f_{ik} h_{il}}{\left(\sum_i f_{ik}^2\right)^{\frac{1}{2}} \left(\sum_i h_{il}^2\right)^{\frac{1}{2}}}$$

コサイン類似度罰則項

最適化の目的関数を  $\mathcal{J}$  とし、 $\mathcal{J}^+ \geq \mathcal{J}$  なる上限関数  $\mathcal{J}^+$  を設計する  $\mathcal{J}$  中の凸関数にJensenの不等式を用いると、次式の上限関数がえられる

$$\mathcal{J}^+ = \sum_{i,j} \left[ y_{ij} \log y_{ij} - y_{ij} \sum_k \alpha_{ijk} \log \frac{f_{ik} g_{kj}}{\alpha_{ijk}} - y_{ij} \sum_l \beta_{ijl} \log \frac{h_{il} u_{lj}}{\beta_{ijl}} - y_{ij} + \sum_k f_{ik} g_{kj} + \sum_l h_{il} u_{lj} \right] + \mu \sum_{k,l} \left[ \left( \sum_i f_{ik}^2 \right)^{-\frac{1}{2}} \left( \sum_{i,i'} f_{ik} \gamma_{i'l}^{\frac{3}{2}} \frac{h_{il}}{h_{i'l}} \right) \right]$$

ここで、 $\alpha_{ijk}$ 、 $\beta_{ijl}$ 、及び  $\gamma_{il}$  は補助変数であり、 $\mathcal{J}^+ = \mathcal{J}$  となる条件は次式

$$\alpha_{ijk} = \frac{f_{ik} g_{kj}}{\sum_{k'} f_{ik} g_{k'j} + \sum_{l'} h_{il} u_{l'j}}, \quad \beta_{ijl} = \frac{h_{il} u_{lj}}{\sum_{k'} f_{ik} g_{k'j} + \sum_{l'} h_{il} u_{l'j}}, \quad \gamma_{il} = \frac{h_{il}^2}{\sum_{i'} h_{i'l}^2}$$

反復最適化更新式は、各変数で  $\mathcal{J}^+$  を偏微分し、0とおくことで導出可能  $h_{il}$  の反復更新式の導出のみ下記に示す

$$\frac{\partial \mathcal{J}^+}{\partial h_{il}} = \sum_j \left[ \frac{-y_{ij} \beta_{ijl}}{h_{il}} + u_{lj} \right] + \mu \sum_k \left[ \left( \sum_i f_{ik}^2 \right)^{-\frac{1}{2}} \left( -\frac{1}{h_{il}^2} \sum_{i' \neq i} f_{ik} \gamma_{i'l}^{\frac{3}{2}} h_{i'l} + \sum_{i' \neq i} f_{ik} \gamma_{i'l}^{\frac{3}{2}} \frac{1}{h_{i'l}} \right) \right] = 0$$

$h_{il}$  の-1次項       $h_{il}$  の-2次項       $h_{il}$  の0次項

上式は  $h_{il}$  の二次方程式であるため、解の公式を用いて解析的に解ける

$$h_{il} = \frac{-b_{il} \pm \sqrt{b_{il}^2 - 4a_{il}c_{il}}}{2a_{il}}$$

ここで、各次数の係数  $a_{il}$ 、 $b_{il}$ 、及び  $c_{il}$  は次式となる

$$a_{il} = \sum_i u_{lj} + \mu \sum_k f_{ik} \left( \frac{\sum_{i' \neq i} \gamma_{i'l}^{\frac{3}{2}} \frac{1}{h_{i'l}}}{\sqrt{\sum_i f_{ik}^2}} \right), \quad b_{il} = -\sum_j y_{ij} \beta_{ijl}, \quad c_{il} = -\mu \gamma_{il}^{\frac{3}{2}} \sum_k \left[ \left( \sum_i f_{ik}^2 \right)^{-\frac{1}{2}} \left( \sum_{i' \neq i} f_{ik} h_{i'l} \right) \right]$$

## 4. 比較実験

- 実験条件

- 罰則条件のないSNMF, 内積型PSNMF, log-cos型PSNMF, 及び提案手法のcos型PSNMFの4手法でsource-to-distortion ratio (SDR)を比較
- 手法毎に最適な重み係数  $\mu$  を開発データから求め、テストデータで使用

窓関数	ハミング窓
窓長	92.9 ms
シフト長	46.4 ms
基底数	目的音源用: 27 非目的音源用: 50
初期値	乱数行列
反復回数	100回
使用する楽器音信号	公開データセット songKitamura より2楽器音の混合信号を90曲作成
最適な重み係数 $\mu$ を探索するための開発データ	90曲の中からランダムに45曲選出
最適な $\mu$ を使う音源分離テストデータ	学習データ以外の残りの45曲

- 更新毎の基底正規化 (basis normalization: BN) の有無の影響についても評価

### BNの処理

$H$  を更新する度に、 $H$  の全基底ベクトル(列ベクトル)の長さを1に正規化  
このとき、基底ベクトル毎の正規化係数  $\delta_l$  の逆数を  $U$  の行ベクトルに乗じて  $HU$  を不変とする

- 実験結果

