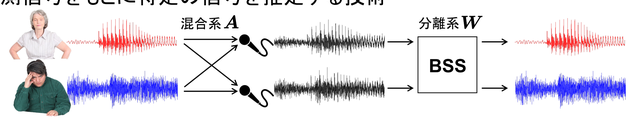
1-Q-28

深層パーミュテーション解決法の汎化性能に関する実験的評価

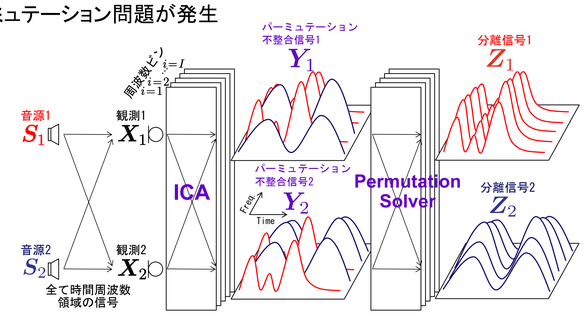
☆蓮池郁也，北村大地（香川高専），渡辺瑠伊（北陸先端大）

1. 研究背景

- **ブラインド音源分離（blind source separation: BSS）**
  - 観測信号をもとに特定の信号を推定する技術

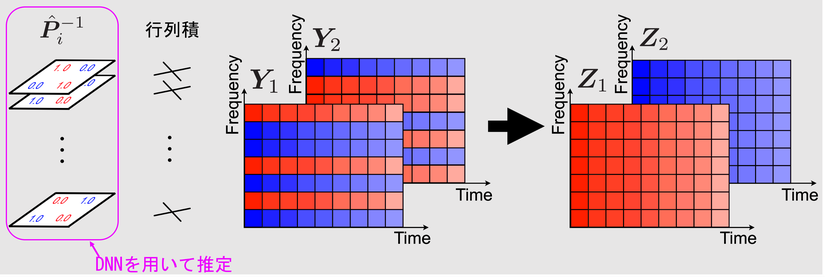


- **独立成分分析（independent component analysis: ICA）**
  - 混合行列が未知の条件で分離行列を推定
  - 分離信号の順序が不定
- **周波数領域ICA（frequency-domain ICA: FDICA）**
  - 各周波数ビンの複素時系列に対して独立なICAを適用
  - パーミュテーション問題が発生



本研究の動機及び目的

- **時間周波数領域での時不変分離行列による音源分離（線形分離）の実現**
  - 歪みが少なく音源分離後の処理に対する影響が小さい
- **汎化性能の高いパーミュテーション解決法を実現** [Hasuike+, 2022]
  - 深層学習（DNN）を用いてパーミュテーション行列を推定

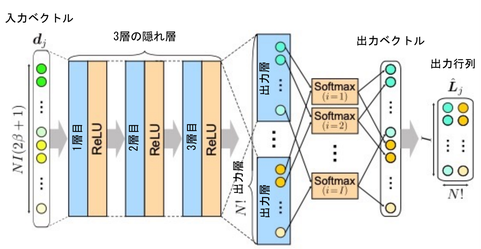


2. 深層パーミュテーション解決法

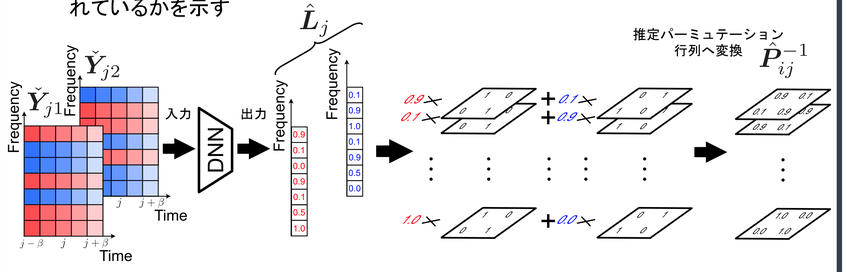
- **パーミュテーション不整合信号に対する正規化処理** [Sawada+, 2007]
  - 同一音源の成分の相関を強調かつDNNの入力の値を区間 [0, 1]に制限
  - DNNの学習が安定する

$$\overline{\boldsymbol{Y}}_{n'} = \frac{|\boldsymbol{Y}_{n'}|^{.2}}{\sum_{n'=1}^{N} |\boldsymbol{Y}_{n'}|^{.2}} \in [0,1]^{I \times J}$$

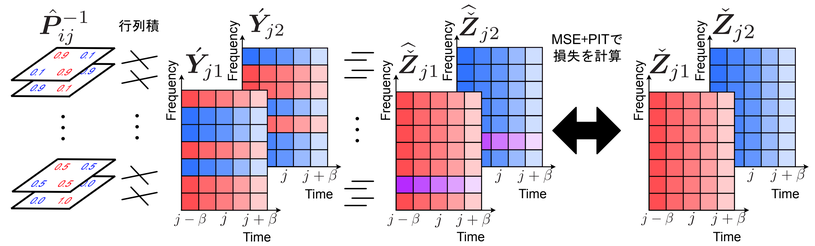
- **DNNの構造**
  - 入力層，隠れ層3層，出力層の計5層で構成
  - 出力層にSoftmax関数を用い各周波数成分の値が足して1になるよう制約
  - 活性化関数にReLU関数を使用
  - 入力層の次元は55350
  - 隠れ層の次元は全て4096
  - 出力層の次元は周波数ビン数である1025
  - 出力層は音源の階乗分を用意
  - Softmax関数の出力を周波数ごとに並べることで出力行列を作成



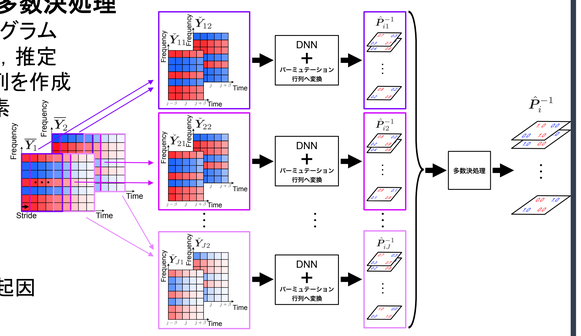
- **DNNの入出力**
  - DNNの入力に複数の（パーミュテーション問題の残る）推定信号の局所時間をとってきてベクトル化し，結合したものを使用
  - DNNの出力は各周波数における確率値
- **推定パーミュテーション行列の導出**
  - DNNの出力値（確率値）をパーミュテーション行列の係数として利用
  - 確率値は観測信号の各周波数において，自身の音源成分がどの程度の割合で含まれているかを示す



- **推定分離信号の導出**
  - 推定パーミュテーション行列とパーミュテーション不整合信号との間で行列積



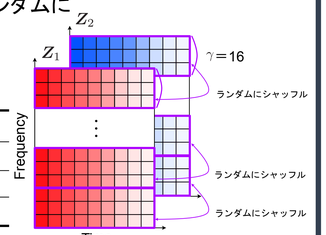
- **テストデータに対する多数決処理**
  - 各局所時間スペクトログラムに対してDNNで予測し，推定パーミュテーション行列を作成
  - 予測した行列の各要素に対して多数決処理を施す
  - 観測信号の各周波数成分は混在せず必ずどちらかの推定分離信号となる
  - DNNの精度の向上に起因



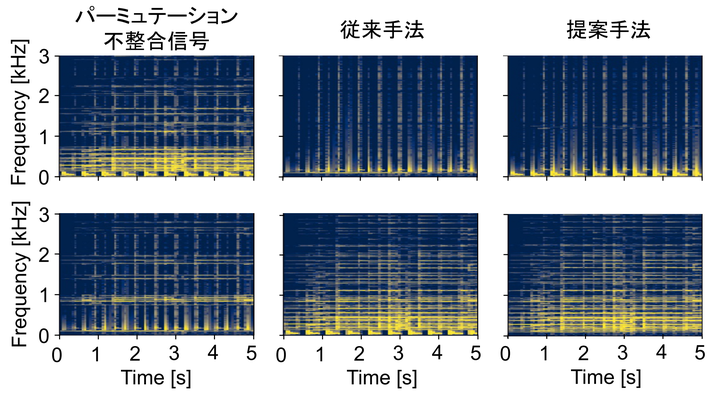
3. 実験

- **音声信号と音楽信号の2つのDNNモデルを用意し，インドメインとアウトドメインに対するsource-to-distortion-ratio (SDR)を比較**
  - 従来の深層パーミュテーション解決法 [Yamaji+, 2020] と性能を比較
- **実験条件**
  - SiSEC2011男女の音声信号及びピアノとベースの音楽信号を使用
  - 学習データは，時間周波数信号を16行1セットでランダムにシャッフルさせた300パターンの信号を使用
  - テストデータには学習データにはないパターンでシャッフルさせた10パターンの信号を使用

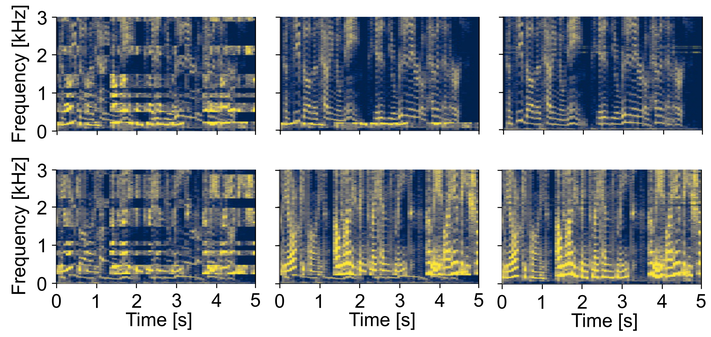
FFT長	128 ms （ハミング窓）
シフト長	64 ms
スペクトログラムのビン数	1025
エポック数	1000



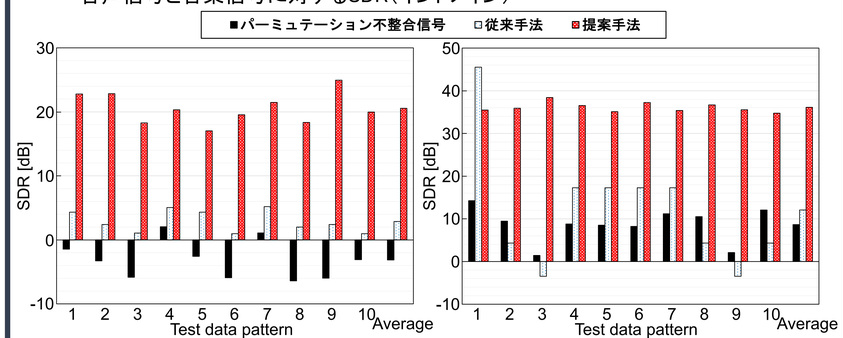
- **実験結果**
  - ベースとピアノのスペクトログラム（インドメイン）



  - 男女の音声信号のスペクトログラム（インドメイン）



  - 音声信号と音楽信号に対するSDR（インドメイン）



  - 音声信号と音楽信号に対するSDR（アウトドメイン）