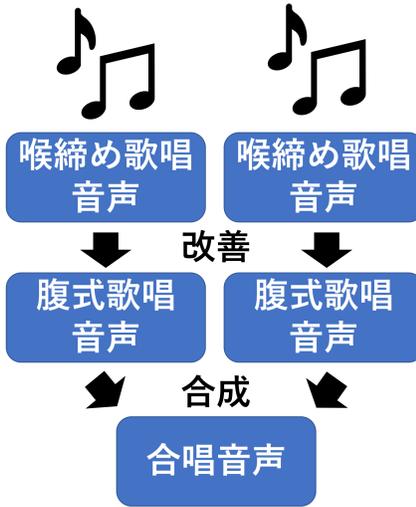


1. 研究背景・目的

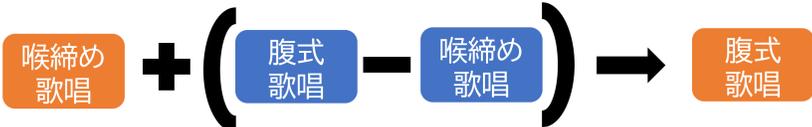
- 合唱: 複数人で行う歌唱
 - ✓ 声質をそろえることが重要
- 喉締め歌唱
 - ✓ 声帯に負担がかかる歌唱方法
 - ✓ 声質の個人差が大きい
- 腹式歌唱
 - ✓ 声帯への負担が小さい歌唱方法
 - ✓ 声質の個人差が小さい
 - ✓ 合唱に適した歌唱方法



個人で収録した音声をもとに、本人の腹式歌唱に改善し、合成することで、**上質な合唱音声**を生成

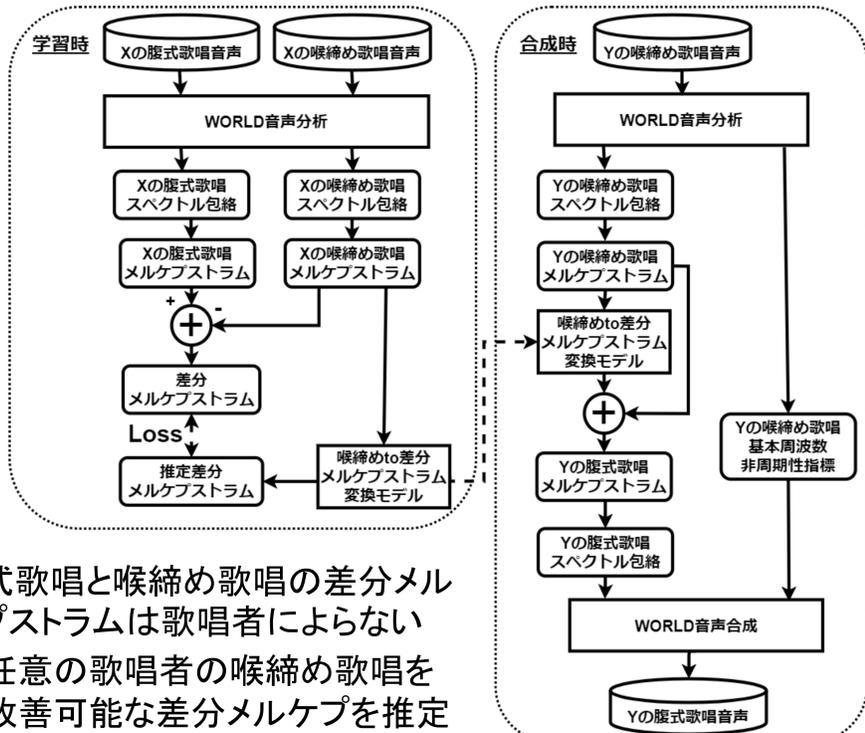
変換方式

- 差分メルケプストラム
 - ✓ 差分スペクトル補正による声質変換(Kobayashi+ 2014)の応用



2. 提案方式

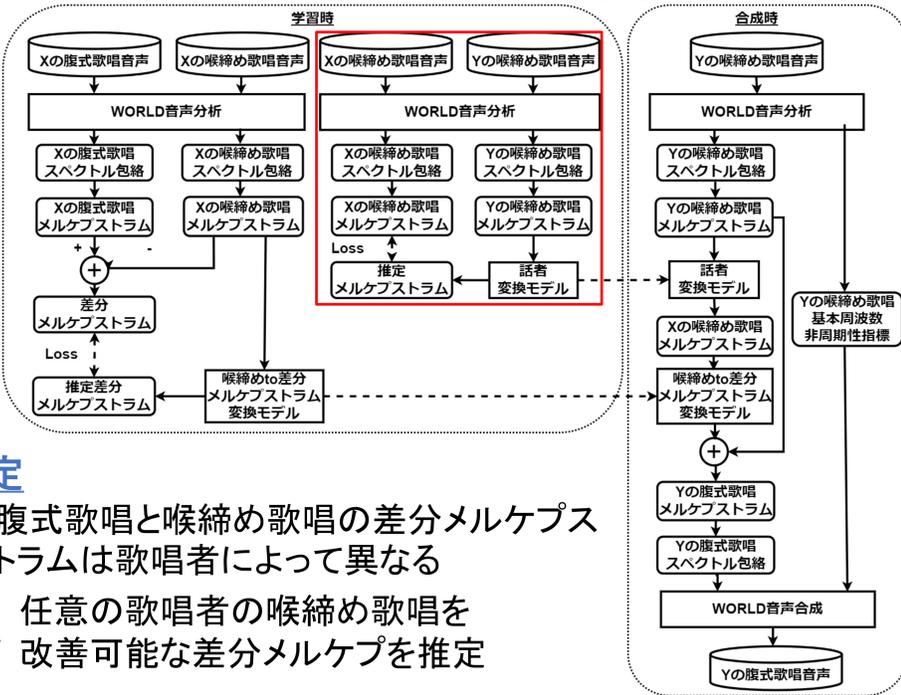
方式1: 歌唱方法の話者性を考慮しない変換方式



仮定

- 腹式歌唱と喉締め歌唱の差分メルケプストラムは歌唱者によらない
- 任意の歌唱者の喉締め歌唱を改善可能な差分メルケプを推定

方式2: 歌唱方法の話者性を考慮した変換方式



仮定

- 腹式歌唱と喉締め歌唱の差分メルケプストラムは歌唱者によって異なる
- 任意の歌唱者の喉締め歌唱を改善可能な差分メルケプを推定

3. 評価実験

実験条件

喉締めto差分メルケプストラム変換モデルの学習に用いた歌唱者をA, 用いていない歌唱者をBとする

	喉締めto差分メルケプストラム変換モデル	話者変換モデル
学習データ	Aの喉締め歌唱 125曲 Aの腹式歌唱 125曲	Aの喉締め歌唱 25曲 Bの喉締め歌唱 25曲
検証データ	Aの喉締め歌唱 10曲 Aの腹式歌唱 10曲	Aの喉締め歌唱 5曲 Bの喉締め歌唱 5曲
評価データ		Bの喉締め歌唱 25曲 Bの腹式歌唱 25曲

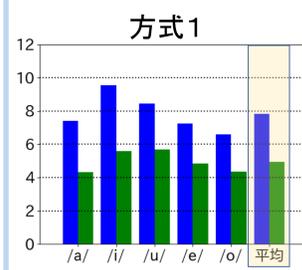
LSTMモデルパラメータ			
入力特徴量	49次元メルケプストラム	中間層	2
損失関数	平均二乗誤差	バッチサイズ	5
最適化手法	Adam	学習率	0.001

客観評価実験

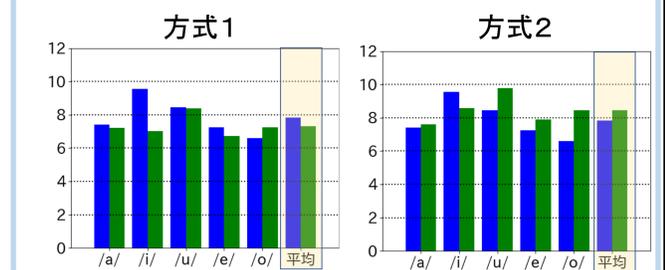
評価音声と腹式歌唱音声とのメルケプストラム歪みを評価

$$\frac{10}{\ln 10} \sqrt{2 \sum_{i=1}^P (c_i - \hat{c}_i)^2}$$

Aの歌唱変換結果



Bの歌唱変換結果



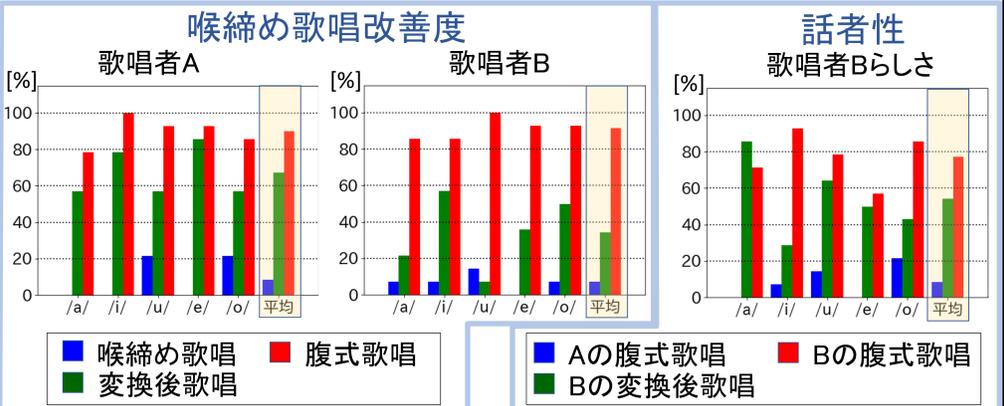
■ 喉締め歌唱と腹式歌唱とのメルケプストラム歪み [dB]
■ 変換後歌唱と腹式歌唱とのメルケプストラム歪み [dB]

✓ 方式1で母音平均のメルケプストラム歪みが**減少**

主観評価実験

3つ目の音声(X)に対し、1つ目の音声(A)と2つ目の音声(B)の近いと感じる音声を選択するABXテストで評価

- 喉締め歌唱改善度: 腹式歌唱と選択された割合
- 話者性: 歌唱者Bの腹式歌唱と選択された割合



- ✓ Bの喉締め歌唱は変換前に比べ**改善**
- ✓ 目標の腹式歌唱と比較するとまだ**改善が必要**
- ✓ 話者性は**維持されている**

4. まとめと今後の課題

まとめ

- 差分メルケプストラムを用いた喉締め歌唱改善方式を検討
- 客観評価実験より、方式1の変換後歌唱のメルケプストラム歪みが減少
- 主観評価実験より、方式1の問題点を確認

今後の課題

- 方式2の推定精度向上
- ✓ LSTM以外を用いた学習について検討