

# 23. 歌声声質変換における歌手の知覚年齢制御法

◎小林 和弘\*, 戸田 智基\*, 中野 倫靖\*\*, 後藤 真孝\*\*,  
ニュービッグ グラム\*, サクリアニ サクテイ\*, 中村 哲\*

\*奈良先端科学技術大学院大学 (NAIST), \*\*産業技術総合研究所 (AIST)

日本音響学会 関西支部  
第16回 若手研究者交流研究発表会  
2013/12/8(日) 前半(13:45~13:55)

## 1. 研究背景

▶ デジタル音楽コンテンツの普及



プロ・アマ問わず自身の歌声をインターネットに投稿する事が可能

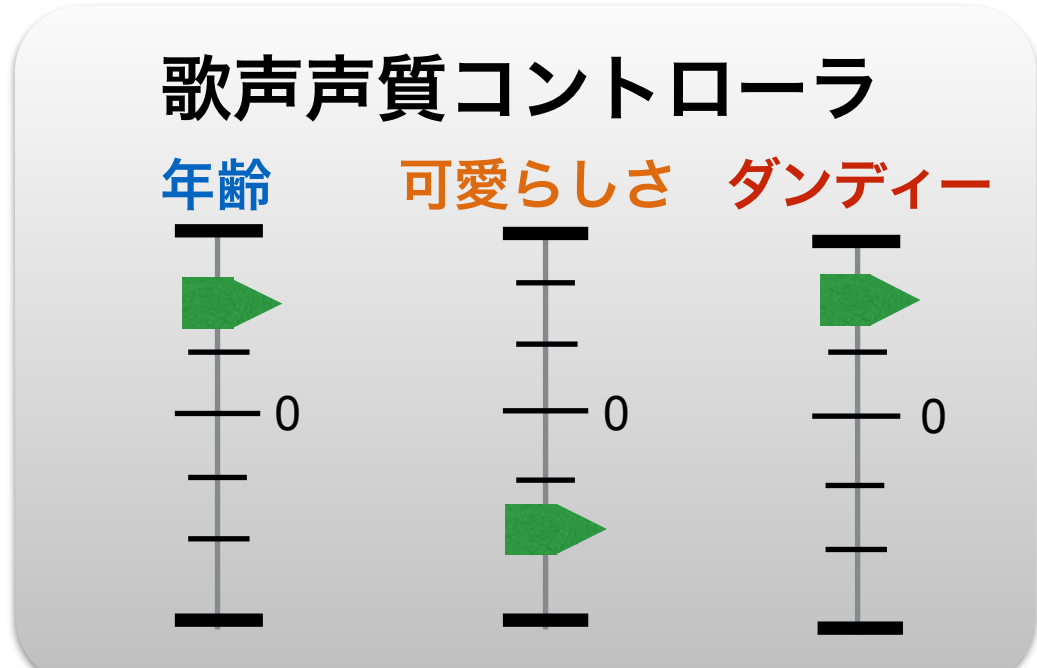
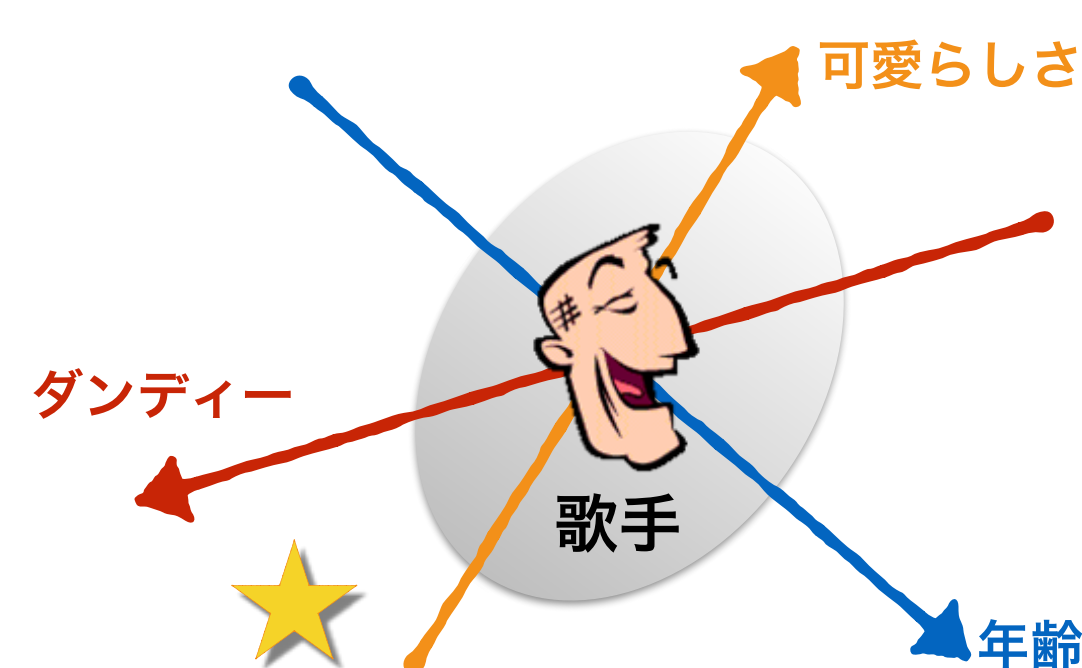
▶ 注目されるために必要なスキル

# 歌唱力 + 独自性

身体的制約を超えた声質での歌唱は困難

研究目的

歌手が自身の声質を自在に制御できる技術の実現

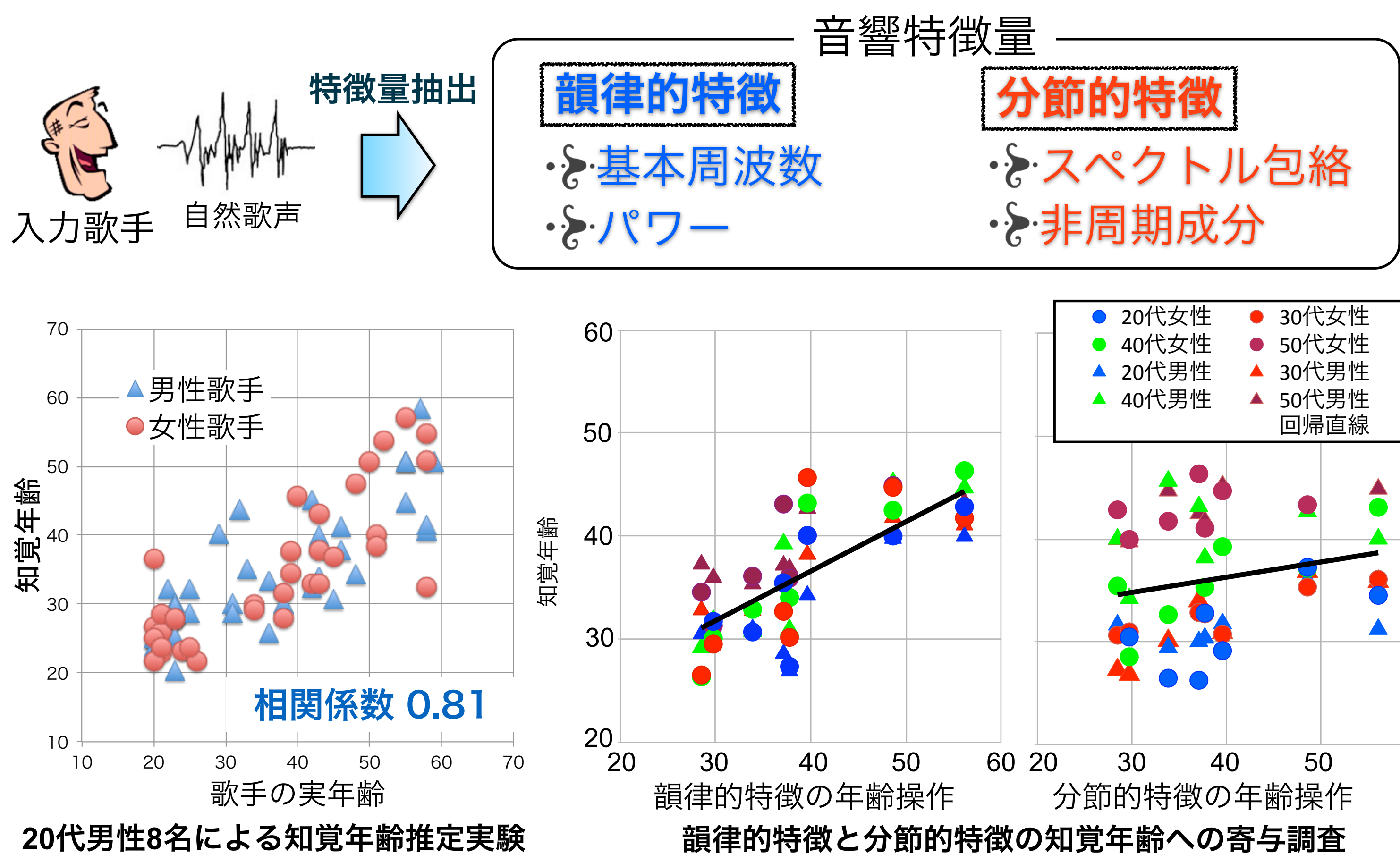


研究内容：個人性を保持した知覚年齢に基づく声質制御



## 2. 関連研究

\* 知覚年齢に寄与する音響特徴量の調査 [小林 他, 2013]



知覚年齢に関する調査結果

歌声の知覚年齢と歌手の実年齢に高い相関

音響特徴量の寄与調査

|       | 知覚年齢の変化 | 歌手の個性 |
|-------|---------|-------|
| 韻律的特徴 | 大きい     | 多い    |
| 分節的特徴 | 小さい     | 少ない   |

特徴量操作をした場合

|       | 知覚年齢の変化 | 歌手の個性 |
|-------|---------|-------|
| 韻律的特徴 | 大きい     | 失う    |
| 分節的特徴 | 小さい     | 保持される |

分節的特徴の操作による知覚年齢制御の実現

## 3. 知覚年齢に基づく声質制御

重回帰GMMに基づく歌声声質変換 (MR-GMM)

$$P(X_t, Y_t | \lambda^{(MR)}, w^{(s)}) = \sum_{m=1}^M \alpha_m \mathcal{N} \left( \begin{bmatrix} X_t \\ Y_t \end{bmatrix}; \begin{bmatrix} \mu_m^{(X)} \\ \mu_m^{(Y)}(s) \end{bmatrix}, \begin{bmatrix} \Sigma_m^{(XX)} & \Sigma_m^{(XY)} \\ \Sigma_m^{(YX)} & \Sigma_m^{(YY)} \end{bmatrix} \right)$$

MR-GMM

出力歌手の平均ベクトル:  $\mu_m^{(Y)}(s) = b_m^{(Y)} w^{(s)} + \bar{\mu}_m^{(Y)}$

代表ベクトルとバイアスベクトルでモデル化

知覚年齢スコア  $w^{(s)}$  により出力平均ベクトルが決定

問題点：特定の歌手の個人性を持つ声質を表さない

個人性を保持した重回帰GMMに基づく歌声声質変換 (Modified MR-GMM)

$$\begin{aligned} \text{MR-GMMの出力平均ベクトル } \mu_m^{(Y)}(s) &= b_m^{(Y)} w^{(s)} + \bar{\mu}_m^{(Y)} \\ &= b_m^{(Y)} (w^{(i)} + \Delta w) + \bar{\mu}_m^{(Y)} \\ &= \frac{b_m^{(Y)} w^{(i)} + \bar{\mu}_m^{(Y)}}{\Delta w} + b_m^{(Y)} \Delta w \end{aligned}$$

入力歌手の特定モデルへと変更

入力歌手の知覚年齢に基づく平均的な声質

$\Delta w$  歳の年齢変動

$$\mu_m^{(Y)}(s) = \hat{\mu}_m^{(Y)} + b_m^{(Y)} \Delta w$$

修正した出力平均ベクトル

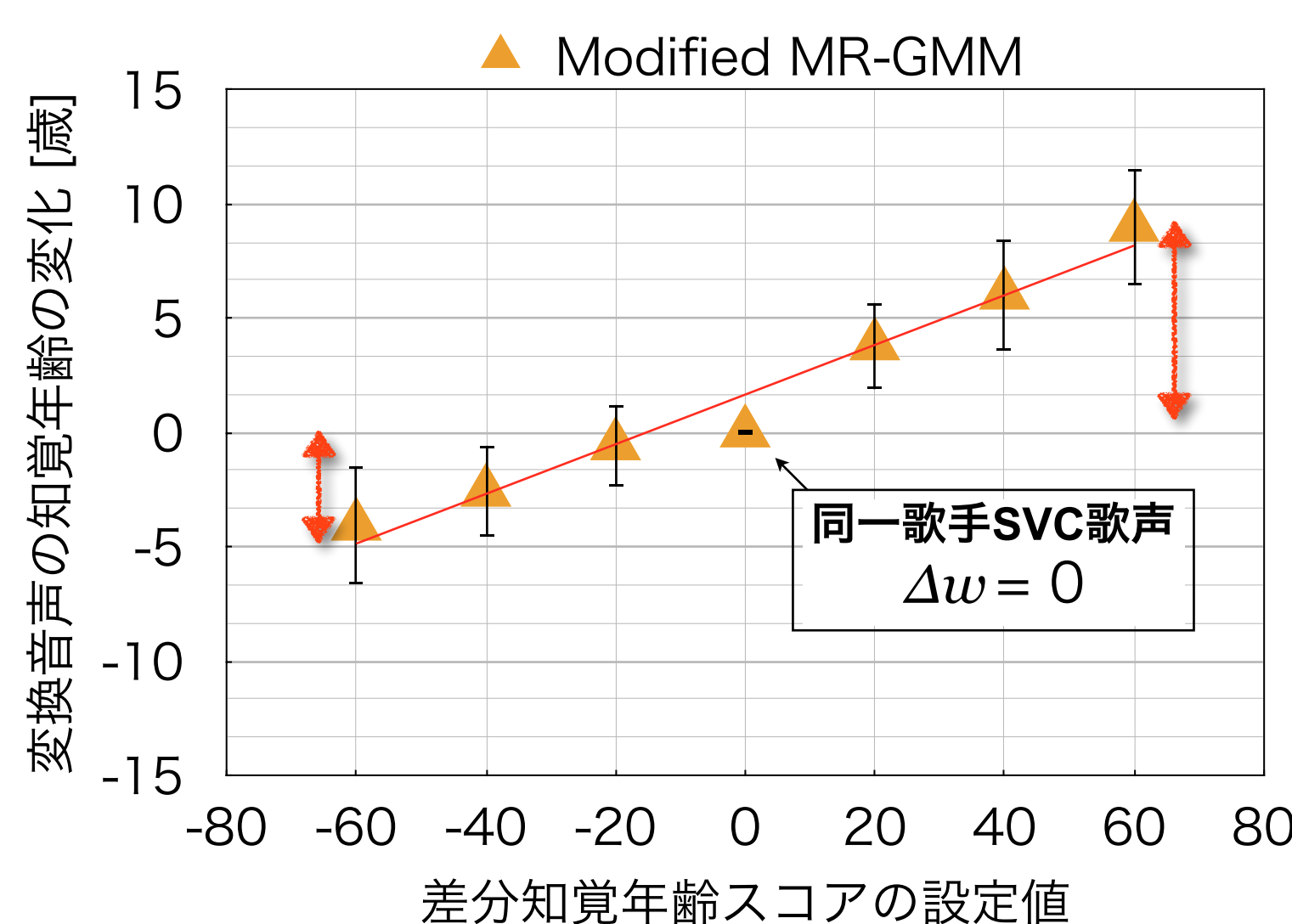
入力歌手の特定モデル 年齢変動

$\Delta w = 0$   $\Delta w = 40$

歌手の知覚年齢を基準に差分知覚年齢スコアに基いて年齢操作

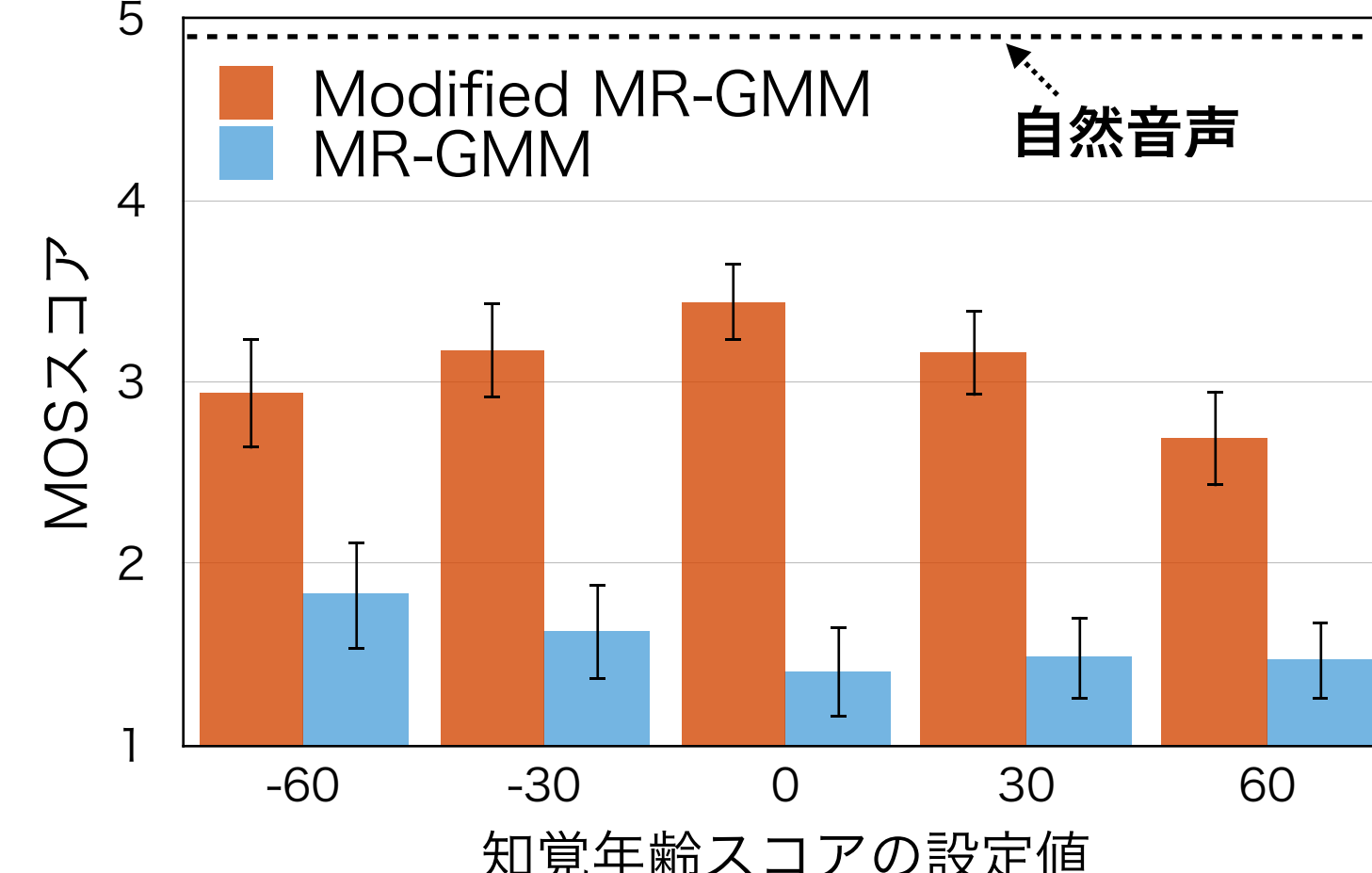
## 4. 実験的評価

知覚年齢変換精度の評価



知覚年齢制御の実現

5段階MOSによる自然性の評価



自然性の高い変換音声

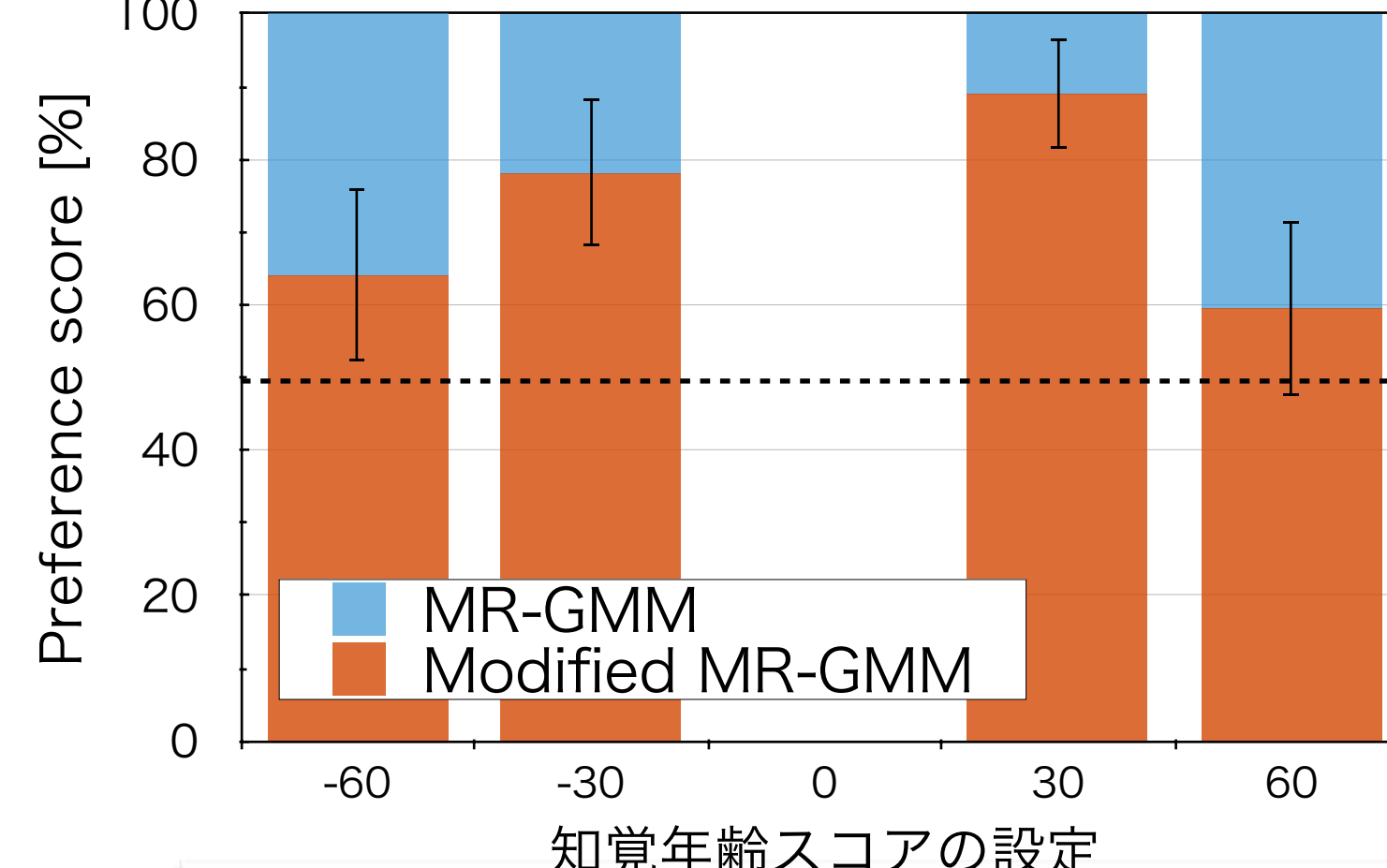
今後の課題

統計的歌声声質変換における自然性の向上

実験環境

| 歌声データベース        | AISTハミングデータベース         |
|-----------------|------------------------|
| メルケプストラム, 非周期成分 | 24次元 (1st-25th), 5周波数帯 |
| 事前収録歌手          | 54名 (20代~50代)          |
| 評価歌手            | 各年代の男女16名 (オープンテスト)    |
| 学習データ           | 25曲/人                  |
| 被験者             | 20代男性8名                |
| 評価歌手の組合せ        | 各年代, 性別が含まれる2セットに分割    |

対比較実験による個人性の評価



個人性保持の実現