

連続母音発話MRIデータに 基づく舌断面形状モデルの 評価

北村 達也 (甲南大学)

蒔苗 久則 (科学警察研究所)

伊藤 仁 (東北工業大学)

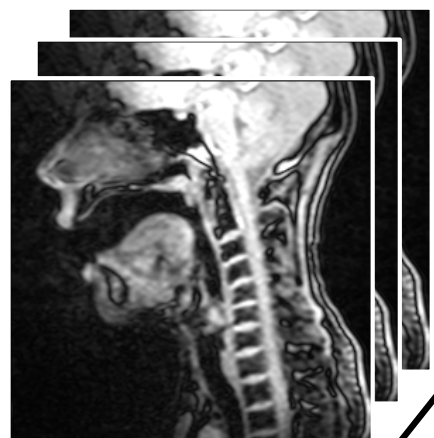
研究の背景

- 音声生成の研究にとっては調音形態/動態の3次元観測が重要
- 2次元リアルタイム観測技術は急速に進歩
 - MRIのリアルタイム撮像
 - 例: 103 fps, Fu *et al.*, (2015)
 - 磁気センサシステム, 超音波撮像の普及
- 3次元リアルタイム観測はまだ難しい

2次元調音データから3次元調音データを推定できないか？

概要

正中面



frame

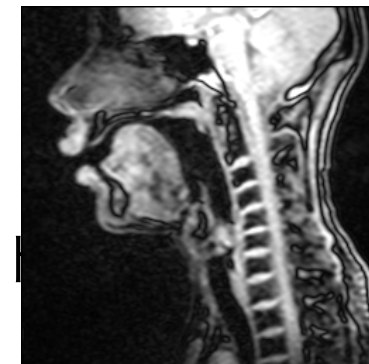
舌断面形状抽出

平滑化

主成分分析

平均形状, 固有ベクトル

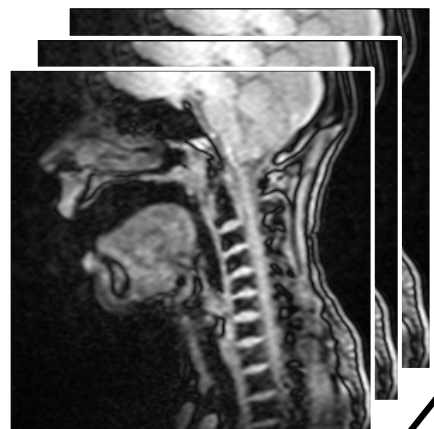
記述パラメータ



記述パラメータ

線形モデル

正中面の隣の断面



frame

舌断面形状抽出

平滑化

主成分分析

平均形状, 固有ベクトル

記述パラメータ

舌断面形状の抽出と平滑化

1. 舌断面形状を手作業でトレース
2. 舌断面形状を200点で再標本化
3. 舌断面形状上の点の座標を複素数で表現

$$\mathbf{f}_i = \mathbf{X}_i + j\mathbf{Y}_i$$

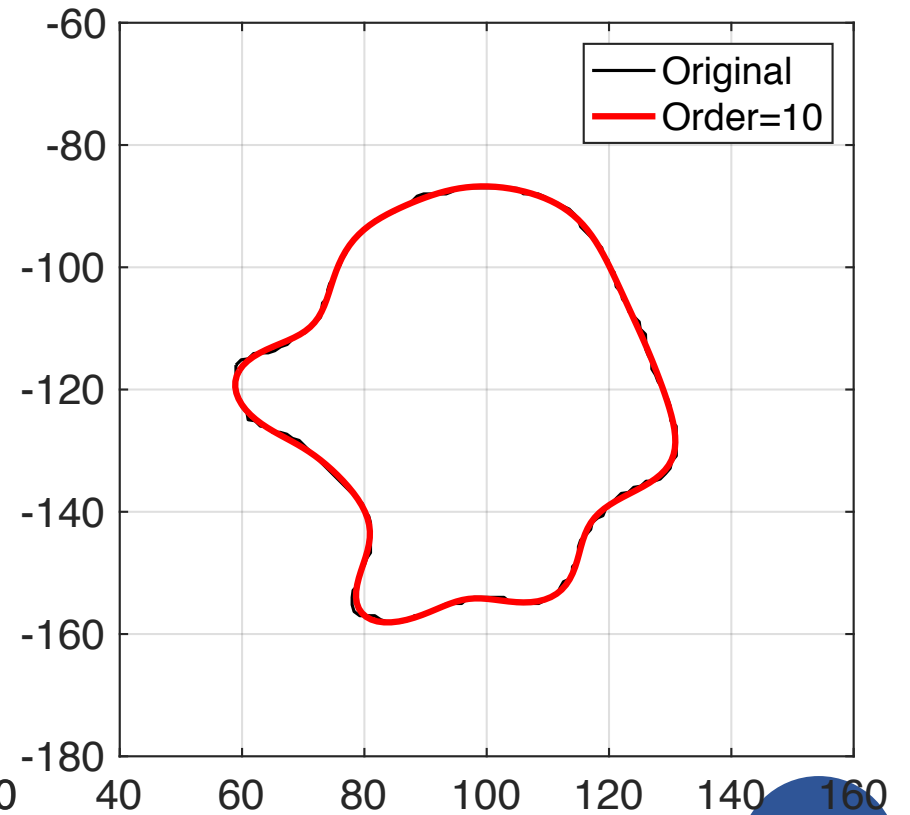
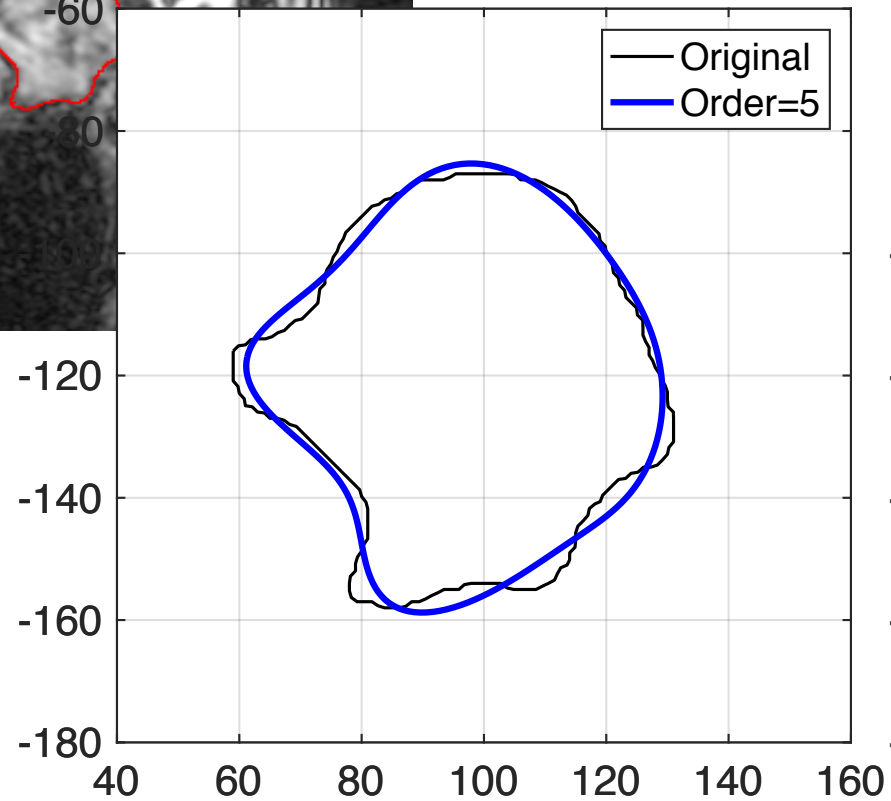
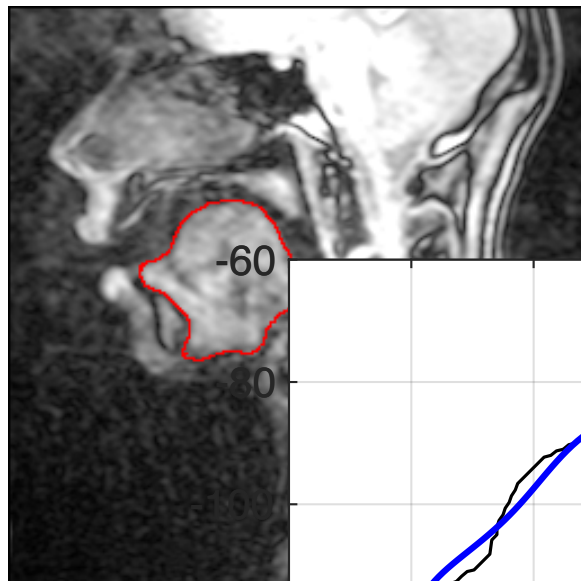
(Park & Lee, 1987)

フレーム i における
舌断面形状の x 座標

フレーム i における
舌断面形状の y 座標

4. \mathbf{f}_i にDFTを施し, Harmonic amplitude profiles (HAP) を得る
5. HAPの10次以上を0で置換し, IDFTを施すことによって平滑化

平滑化の例



PCAによる表現(1): 分析

より少ないパラメータで形状を表す (永井ら, 1996)

- 差分形状ベクトルを求める

$$\Psi_i = \mathbf{f}_i - \boxed{\bar{\mathbf{f}}}$$

フレーム間平均

- Ψ_i の分散行列の固有ベクトル a_{ki} を求め, 固有形状ベクトルを得る

$$\mathbf{u}_k = \sum_{i=1}^N a_{ki} \Psi_i$$

ここで, $N < M$ (M はフレーム数) とすれば M 個より少ないベクトルで差分形状ベクトルを代表させられる

PCAによる表現(2): 再合成

- 記述パラメータを求める

$$r_k = \Psi_t \cdot \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$$

↑ 推定対象の舌断面形状の差分形状ベクトル

- N 個の \mathbf{u}_k と r_k を用いて \mathbf{f}_i を再合成する

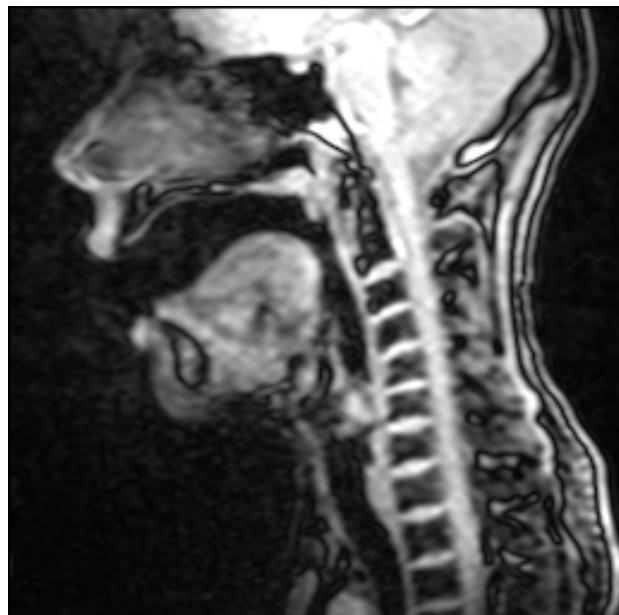
$$\mathbf{f}'_t = \sum_{k=1}^N \left(r_k \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|} \right) + \bar{\mathbf{f}}$$

MRIデータ

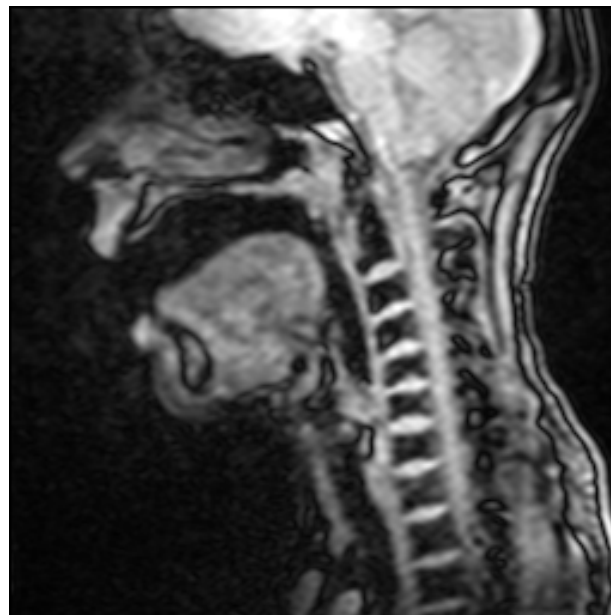
- 実験協力者: 20代男性, 日本語母語話者1名
- 発話資料: /aiueo/, /oeuia/
- MRI装置: Siemens社製MAGNETOM Verio
- 断面: 正中面を含む5断面
- 撮像パラメータ

時間分解能:	60 fps	Repetition time:	857.3 ms
フレーム数:	113フレーム	Number of averages:	1
スライス厚:	5 mm	Flip angle:	15°
スライス間隔:	1 mm	Field of view:	256 mm x 256 mm
Echo time:	1.28 ms	Pixel size:	1 mm x 1 mm

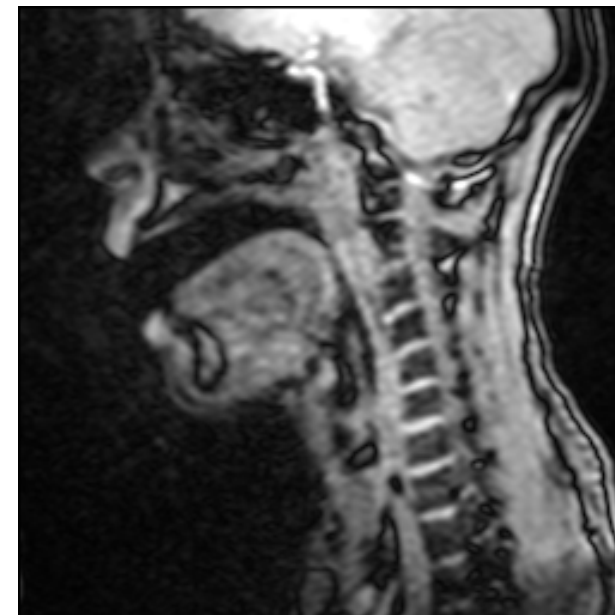
MRIデータの例



Plane S_m (正中面)



Plane S_{m+1}



Plane S_{m+2}

/aiueo/ の /a/ に対応するフレーム (第35フレーム)

評価実験

- 舌断面形状セット (3断面/フレーム)
 - /aiueo/ 系列: D_t
 - /oeuia/ 系列: D_e
- D_t を用いて線形モデルを学習
- クローズドテスト: D_t を推定
- オープンテスト: D_e を推定
- 面積一致率で評価

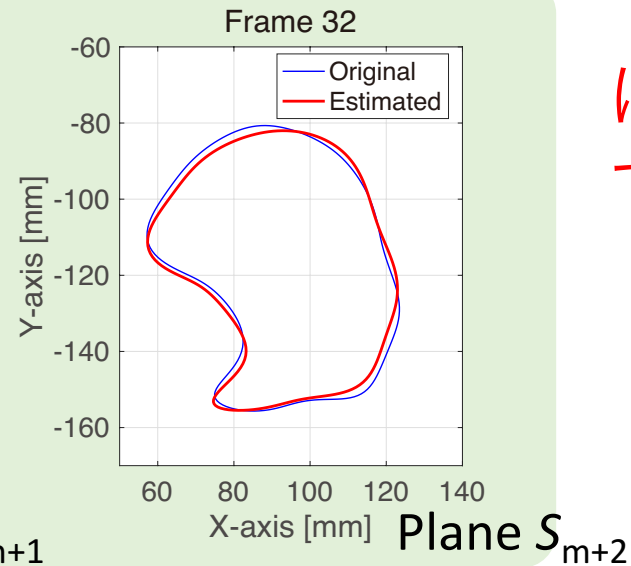
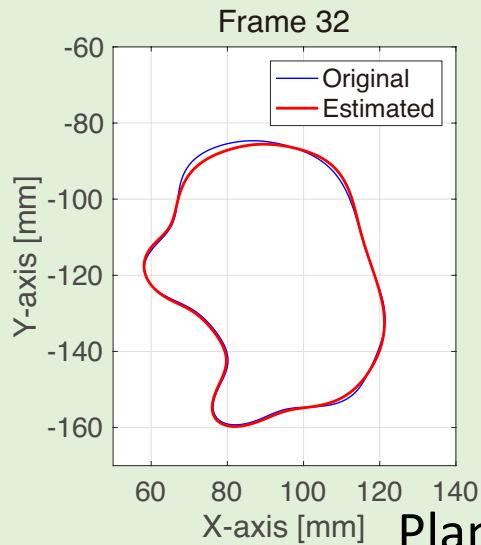
$$r_a = \frac{A_u}{A_t}$$

推定対象と推定結果の
共通領域の面積

推定対象の面積

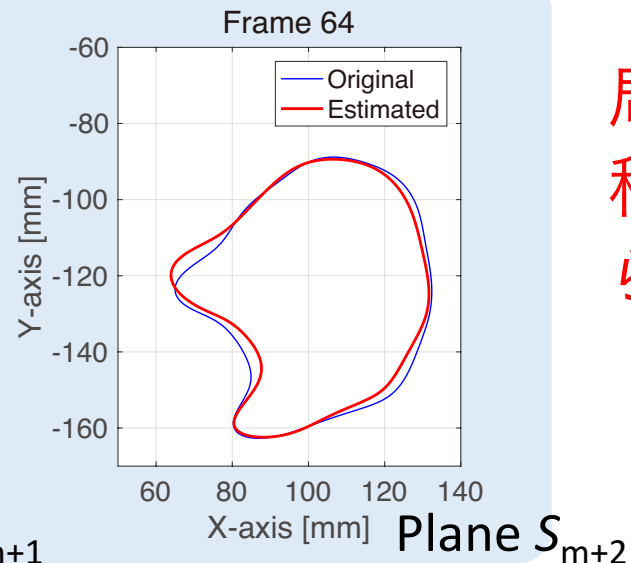
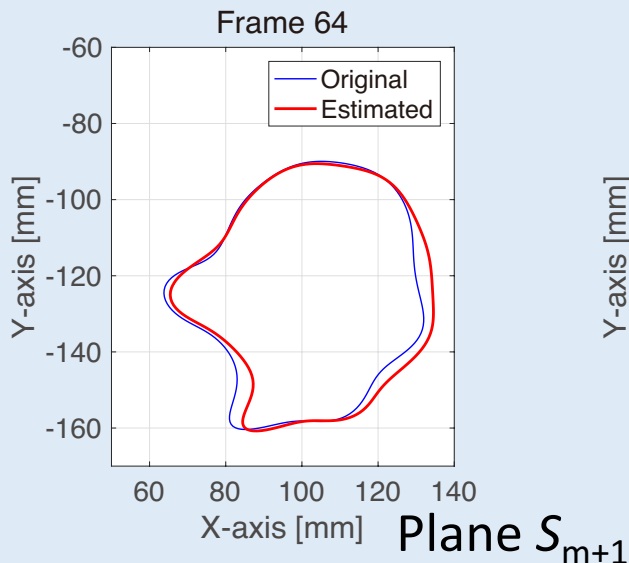
結果(1): 推定された舌断面形状の例

Closed test



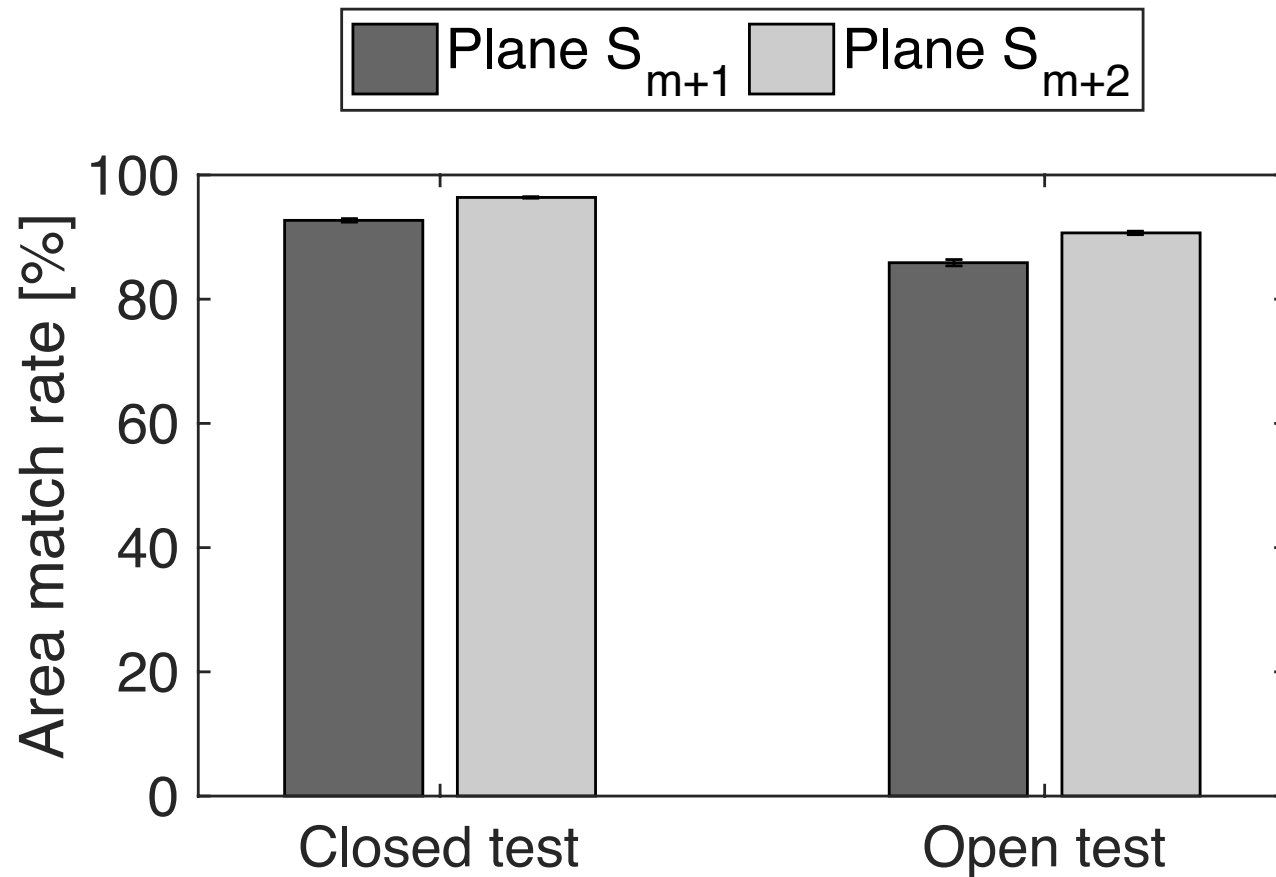
ほぼ正確に推定
できている

Open test



局所的に 3 mm
程度の誤差が見
られる

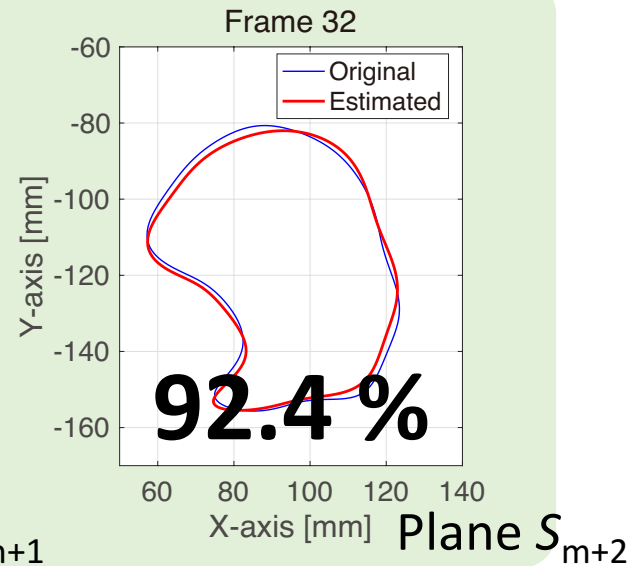
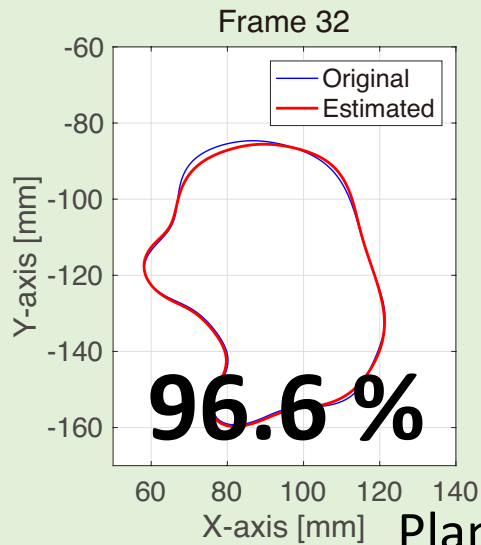
結果(2): 平均面積一致率



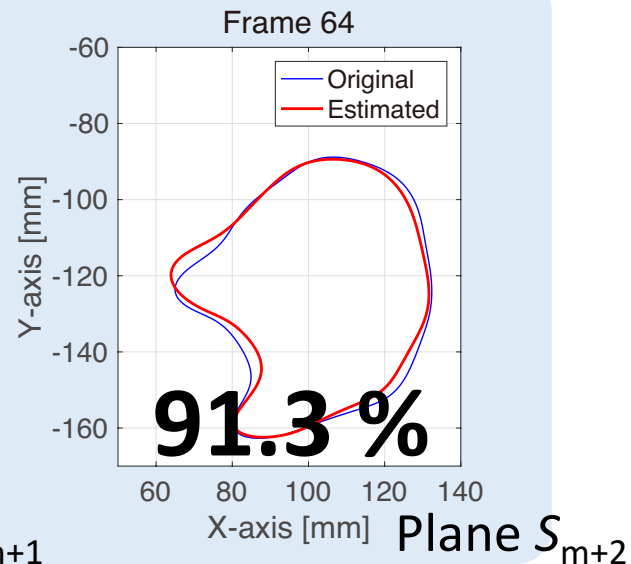
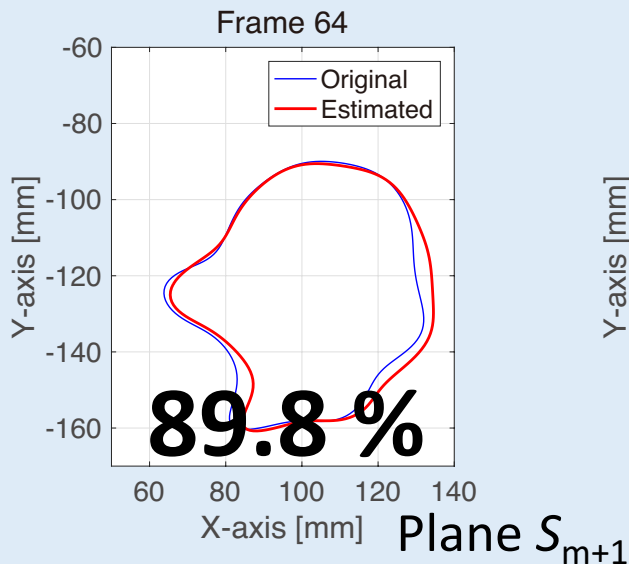
Open test でも
約 90 % の精度

結果(1): 推定された舌断面形状の例

Closed test



Open test



考察

- オープンテストでも比較的高精度の推定が可能
- 面積一致率による評価でよいのか？
- 声道の変形がその音響特性に及ぼす影響は場所によって異なる
- 学習用データの追加や統計モデル等の導入による精度向上が期待できる
 - 大量にMRIデータがあればモデル化さえ不要？
- 次の課題：他の話者のデータを用いた評価