修 士 論 文

形態学的計測に基づく発話機構モデルの個人化

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

西村 奈々

2012 年 3 月

修 士 論 文

形態学的計測に基づく発話機構モデルの個人化

指導教官　党建武 教授

審査委員主査　党建武 教授
審査委員　鵜木祐史 准教授
審査委員　小谷一孔 准教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

1010046 西村 奈々

提出年月: 2012 年 2 月

概 要

音声生成過程において生じる個人差の要因を明らかにするには、発話における発話器官の形態や動きを詳細に調査する必要がある。しかし、倫理の問題などにより、人間の発話器官の静的な特性と動的な特性を直接計測するには限界がある。生理学的発話機構モデルを用いることでこれらの系統的な調査が可能になることが期待されるが、話者の形態学的情報からモデルを構築するには多くの時間と労力を要するため、個人差の調査に必要となる複数話者に対応する個別的なモデルの構築が困難である。そこで本研究は、話者の特徴を反映したモデルを効率的に構築するための手法を提案する。既に構築されている生理学的発話機構モデルをプロトタイプモデルとして、任意話者の形態学的情報をプロトタイプモデルに適用することによりその話者に特化したモデルを構築することで、モデルの個人化を試みる。複数の目標話者の MR 画像を用いモデルの個人化を行った結果、提案手法によりプロトタイプモデルの構造を維持したまま、各話者の形状を反映したモデルを効率的に構築できることが確認された。また、個人化によりモデルの動作に違いが生じることが確認された。

目 次

図 目 次

表 目 次

第1章　序論

1.1　研究の背景

音声には、話者の年齢や性別などの非言語情報が含まれており、話者の特徴である個人性もそのうちの一つである。音声に個人性があることにより、我々は話者の特定などコミュニケーションにおいて必要な情報を得ることができる。個人性が生じる要因には先天的（発話器官の形状や大きさ等)、後天的（育った環境等）なものがあり、個人性が生じる音声生成過程から知覚に至るまで、様々な研究が行われている [1]。特に、音声生成過程において生じる個人差の要因を明らかにするには、発話における発話器官の形態や動きを詳細に調査する必要がある。そのため、磁気共鳴画像法 (Magnetic Resonance Imaging: MRI) を用いた観測 [2] や、双曲金属線電極法による発声時の筋電信号 [3] の測定などが行われてきた。しかし、観察の限界や倫理の問題などの原因により、人間の発話器官の静的な特性と動的な特性の全てを直接計測することは難しい。これらの問題を解決し、音声生成過程の系統的な調査を行う有効な方法のひとつとして、モデルによるシミュレーションが挙げられる。

発話機構モデルに関して、2 次元発話機構モデル [4]、2 次元生体力学モデル [5]、舌骨や喉頭部を含む 2 次元動的生体力学モデル [6]、3 次元線形モデル [7]、軟口蓋と鼻咽腔に注目した 3 次元発話機構モデル [8] など、音声生成過程の解明を目指した様々な研究が行われてきた。その中でも党ら [11, 12] によって提案された生理学的発話機構モデルは、発話目標に基づいた筋の収縮により駆動され、音声を生成することができる。このモデルは、発話生成における筋活性パターン [13] や、音声と調音の一対多の関係 [14] の調査に用いられるなど、音声生成過程の詳細を議論するのに適している。しかし、このモデルは一人の話者のデータに基づいて構築されているため、話者による違いについて検討することは困難である。もし複数話者に対応する個別的なモデルが容易に準備することができれば、音声生成の様々な過程において生じていると考えられる個人差について調査が可能になる。新たな話者に基づくモデルが構築されない理由として、このようなモデルは複雑な構造を持ち、話者の形態学的情報からモデルを構築するには多くの時間と労力を要するため、個別的なモデルの構築が難しいことが挙げられる。

これらのモデルは音声生成過程の解明を目指す他に、脳による発話機構の運動制御の研究 [9, 10] に用いられるなど、様々な応用が行われている。また、形状学以外の筋の配置などの個人差を含む個人化モデルが容易に構築できるようになると、音声生成過程における動的な個人差の調査の他に、医療現場でのモデル利用が考えられる。例えば、舌や口底を

切除すると声道形状が変わり [15]、生成音声に影響を与える [16]。切除の範囲や再建手術は音声への影響に深く関係しているため、影響が最も少ない手術プランを立てることが求められる。しかし現在の治療では医師の経験に基づいているため、実際に手術が行われた後にしか結果は分からない。解決策として、舌の切除シミュレーション研究 [17] が行われており、個人化モデルによりこれらのシミュレーションが行われるようになると、最適な手術プランを容易に、そして客観的に示すことが可能になる。他に、構音障害の判定や原因の解明に用いることが考えられる。現在は言語聴覚士による聴覚印象によって、構音障害であるか判断が行われている。そのため、科学的根拠に基づいた医療 (Evidence based medicine: EBM) を行うために、モデルによる音響的な解析が必要とされている [18]。また、障害の原因は時として MR 画像等では分からないことがある。これらの問題解決に個人化モデルを用い、生成音声と発話器官の動的関係を解析することにより、原因解明が期待される [19]。

1.2 研究の目的

本研究の目的は、話者の特徴を反映した生理学的発話機構モデルを効率的に構築するための手法を提案することである。アプローチとして、これまでに構築された生理学的発話機構モデルをプロトタイプモデルとし、任意話者の形態学的情報をプロトタイプモデルに適用することにより、その話者に特化したモデルを構築することでモデルの個人化を試みる。モデルに反映すべき話者の特徴は多岐にわたるため、個人化の対象を段階的に分けることとする。本研究を個人化のファーストステップと位置づけ、人間の発話器官の形状にある個人差に着目し、形状の個人化を通して一連の個人化プロセスの構築を目指す。なお、今回個人化の対象としない筋などの解剖学的構造や生理学的情報については、プロトタイプモデルで用いられている情報を適応する。話者の外形に関する特徴が反映されたモデルが容易に構築できるようになると、話者による静的な個人差を調査することが可能になる。

1.3 論文の構成

本論文は現在の章（序論）を含め、4 つの章から成る。第 2 章では、本研究においてプロトタイプモデルとして用いる生理学的発話機構モデルについて説明する。第 3 章で提案手法の提案と、モデルの個人化を行った結果について述べる。最後に第 4 章において、まとめと今後の課題について述べる。

第2章　生理学的発話機構モデル

本研究で用いる生理学的発話機構モデルは、日本人話者の 3 次元 MR 画像と解剖学的知見に基づき、舌、下顎、舌骨、喉頭部、声道壁が構築されている。それらの発話器官は筋と腱で繋がれ、筋を収縮させることによりモデルは駆動される。この章では、定義されている各発話器官の構造や特徴について述べる。また、モデルに筋の収縮パターンを与えた際の駆動例を示す。

2.1　モデルの構造

本研究ではプロトタイプモデルとして、2004 年に党ら [12] によって提案された 2.5 次元モデルを 3 次元に拡張したモデルを用いた。モデルの外見を図 2.1 に示す。内部の発話器官が観察しやすいよう、下顎と声道壁をそれぞれ除いた図を合わせて示す。

モデルで定義されている声道壁と舌、下顎の一部は、3 次元 MR 画像で計測された形状（形態学的計測情報）に基づいて構築されている。発話器官の形状は、等間隔に選択した複数の断面（レイヤーと呼ぶ）を基準に各レイヤーから発話器官の形状を抽出し、三次元形状を表している。表 2.1 に、形態学的情報に基づいて構築されている発話器官のレイヤー数と 1 レイヤー当たりに定義されている形状を表す点の数を示す。各発話器官は正中矢状断面を中心に左右対称に定義されており、モデルの右半分は、モデルの左半分の形状を反転して作られている。

表 2.1: 各発話器官のレイヤー数と点数

器官	レイヤー数	点数/1 レイヤー
声道壁	15	55
下顎*	15	25
舌	5	77

*解剖学的知見に基づく部分を除く

声道壁

形態学的な情報が反映された声道壁は、上歯列、軟口蓋、硬口蓋など様々な器官で構成されている（図 2.3、2.2）。各器官は点によって区切られ、モデル内では別々に扱われる。発声時に振動する軟口蓋と喉頭部分は、繰り返し発声をしながら撮影する 3 次元 MR 画像では不鮮明である場合が多い。そのため、発声毎に形状が一定でないことも考慮し、形状を取得する際には少し補正が加えられている。また、モデル構築に用いられる声道壁の形状は口蓋平面が水平になるようさせ、定義されている。本論文では形状が見やすいよう、MR 画像に声道壁形状を重ねる際は回転を加える前の形状を用いる。

下顎

下顎は、形態学的情報に基づく部分と解剖学的知見に基づく部分を組み合わせて構成されている。下顎モデル (図 2.4) のうち、横に張り出している下顎枝部は MR 画像では確認できないため（MR 画像で骨は映らない。また、撮影範囲に下顎全体が含まれていない。）解剖学的知見に基づき構築されている。その他の部分は形態学的情報に基づき (図 2.5) 構築され、別途構築された下顎枝部が滑らかに繋がるよう、接続されている。下顎領域（下歯列と同領域）は、通常の MR 画像では映らない。そのため、別途歯列領域が映るよう撮影した MR 画像から歯列を補填している（詳しい方法については、3.3.1 を参照）。

舌と舌骨

舌の外形には形態学的な情報が反映されている（図 2.6、2.7）。さらに舌内部には解剖学的知見に基づいて設計されたメッシュ構造を持ち、剛体と軟組織を統一した手法で制御されている。図 2.7 下部の赤線で表示した舌の前下部は、下顎に固定している。舌に付着している舌骨は解剖学的知見に基づき設計され、舌底面の点を舌骨にある対応点に連結するよう、各々が配置されている（図 2.8）。そのため、舌骨の頂点を基準とし、下顎結合の最下点と舌根の各々を舌骨頂点から直線で結んだ面を舌の底面と定義する。

喉頭部：喉頭蓋、甲状腺軟骨、披裂軟骨、輪状軟骨

喉頭部は、喉頭蓋、甲状腺軟骨、披裂軟骨、輪状軟骨で構成されている。各形状は解剖学的知見に基づき構築されている。図 2.9 に、各器官が筋 (腱) で繋がれて喉頭部が構築されている状態を示す。

筋と腱

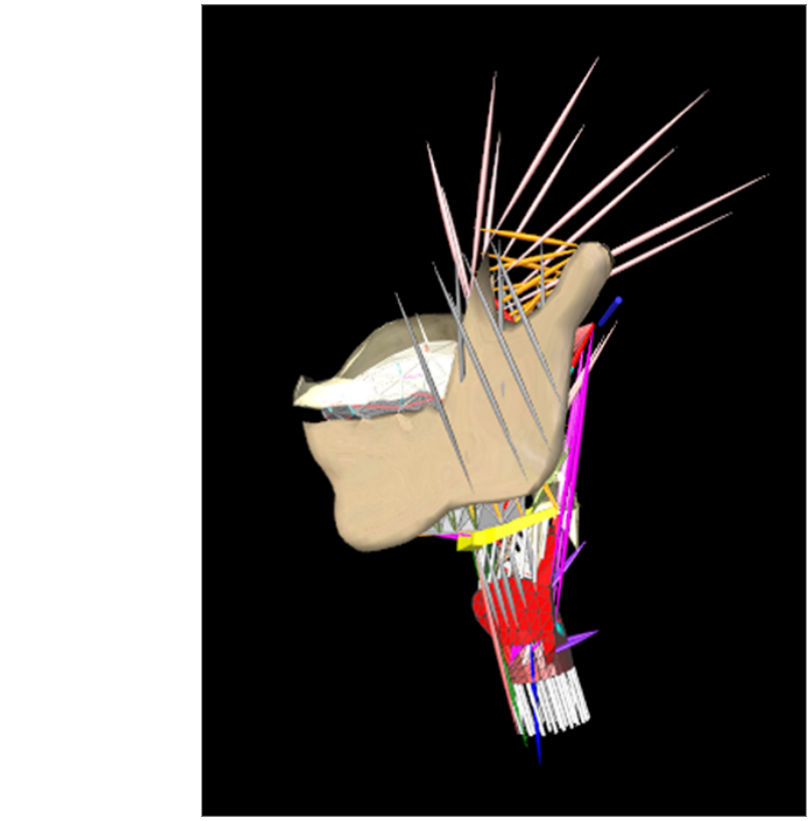
筋と腱は各発話器官を繋ぎ、モデルを駆動する際に重要な役割を担う。発話器官形状の各頂点を基準に、解剖学的知見に基づき筋 (腱) は配置され、各々に太さや長さが設定されている。

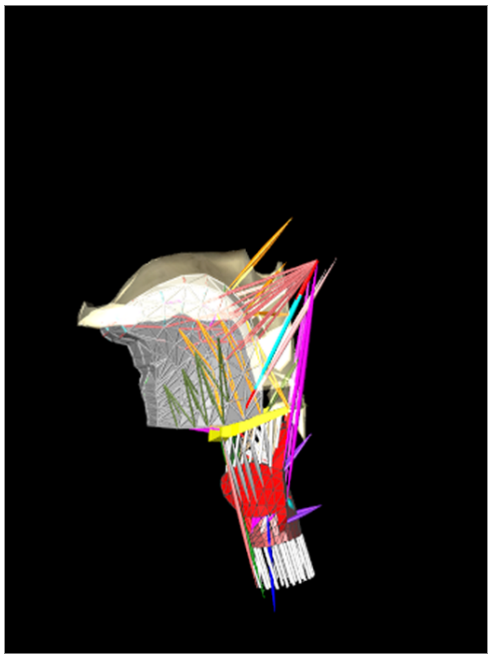
表 2.2: 日本語 5 母音を生成する筋活性パターンの例 (参照:[13])

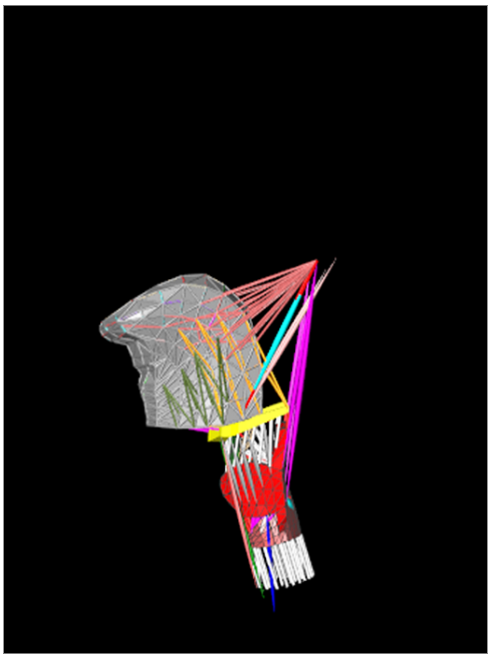
筋の名称	[a]	[i]	[u]	[e]	[o]
前部オトガイ舌筋 (GGa)	0.01				0.01
後部オトガイ舌筋 (GGp)		0.03			
舌骨舌筋 (HG)	0.01				
茎突舌筋 (SG)	0.01				0.02
(LGua)	0.05				0.02
(LGum)	0.03				0.01
上縦舌筋 (SL)	0.61		0.09		0.76
(Va)	0.17			0.01	0.15
(Tm)	0.32	0.21	0.39	0.11	0.32
(Tp)			0.05		0.09
下縦舌筋 (IL)		0.07	0.04		
顎舌骨筋 (MH)		0.02	0.02	0.09	
開顎筋群 (JawO)	0.3			0.06	0.28
閉顎筋群 (JawC)		0.005	0.005		

2.2 モデルの機能

筋に収縮（筋活性パターン）を与えることにより、モデルが駆動する。例として、プロトタイプモデルの基となっている話者の母音発声時の MR 画像と、各母音を生成した際の舌と下顎の形状を示す (図 2.10、2.11、2.12、他の母音 [e,o] は、付録 A.1 を参照)。なお、音声と調音は一対多の関係 [14] にあるため、日本語 5 母音を生成する筋収縮のパターンは一つに定まらない。今回は一例とし、Wu ら [13] がプロトタイプモデルを用いて推定した日本語 5 母音の筋活性パターン (表 2.2) を用いている。図の青は声道壁形状、黒破線は舌と下顎のオリジナル形状と位置、赤線はシミュレーション後の舌形状と位置、緑線はシミュレーション後の下顎位置を表している。党らの生理学的発話機構モデルは話者の特徴を反映しているため、適切な筋収縮を与えることで、発声時の発話機構の状態を忠実に再現することができる。

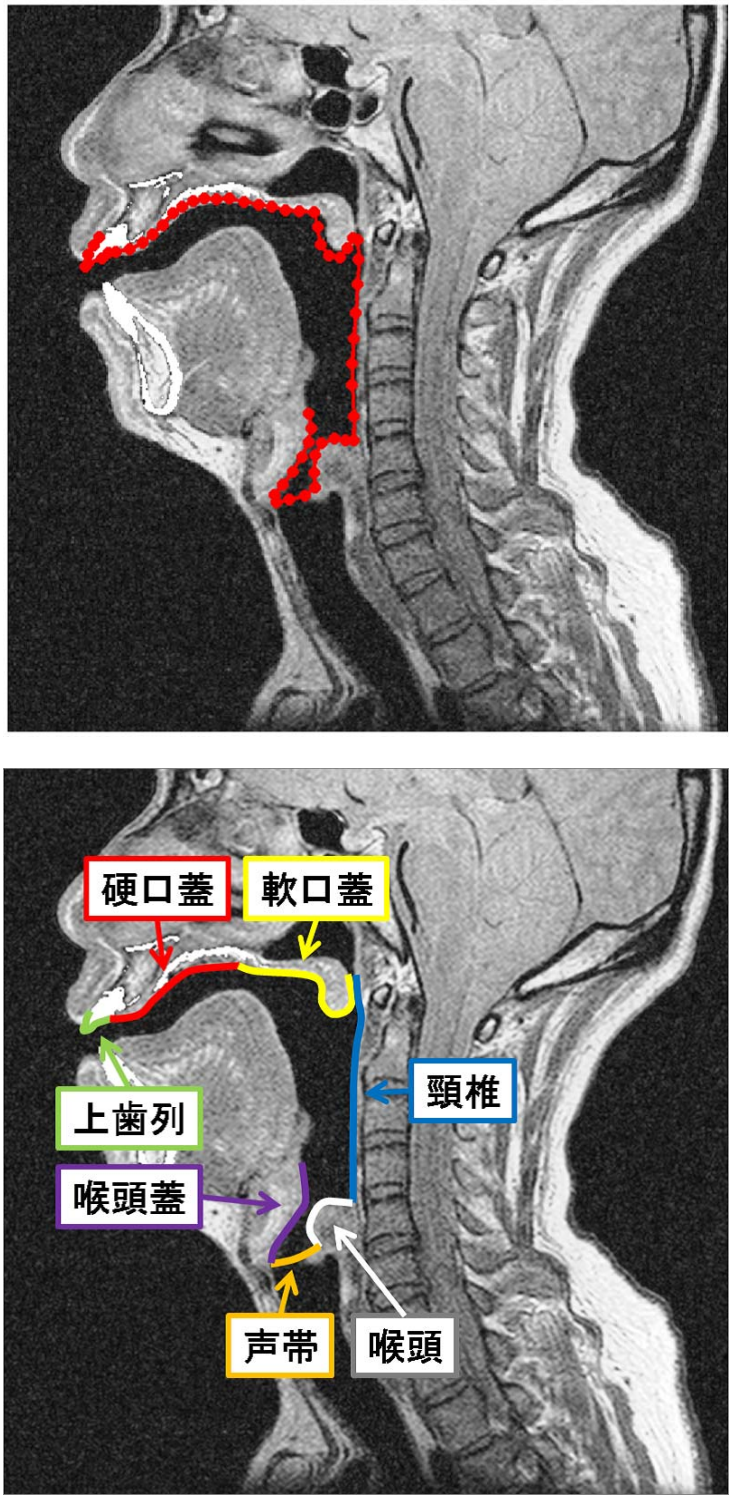






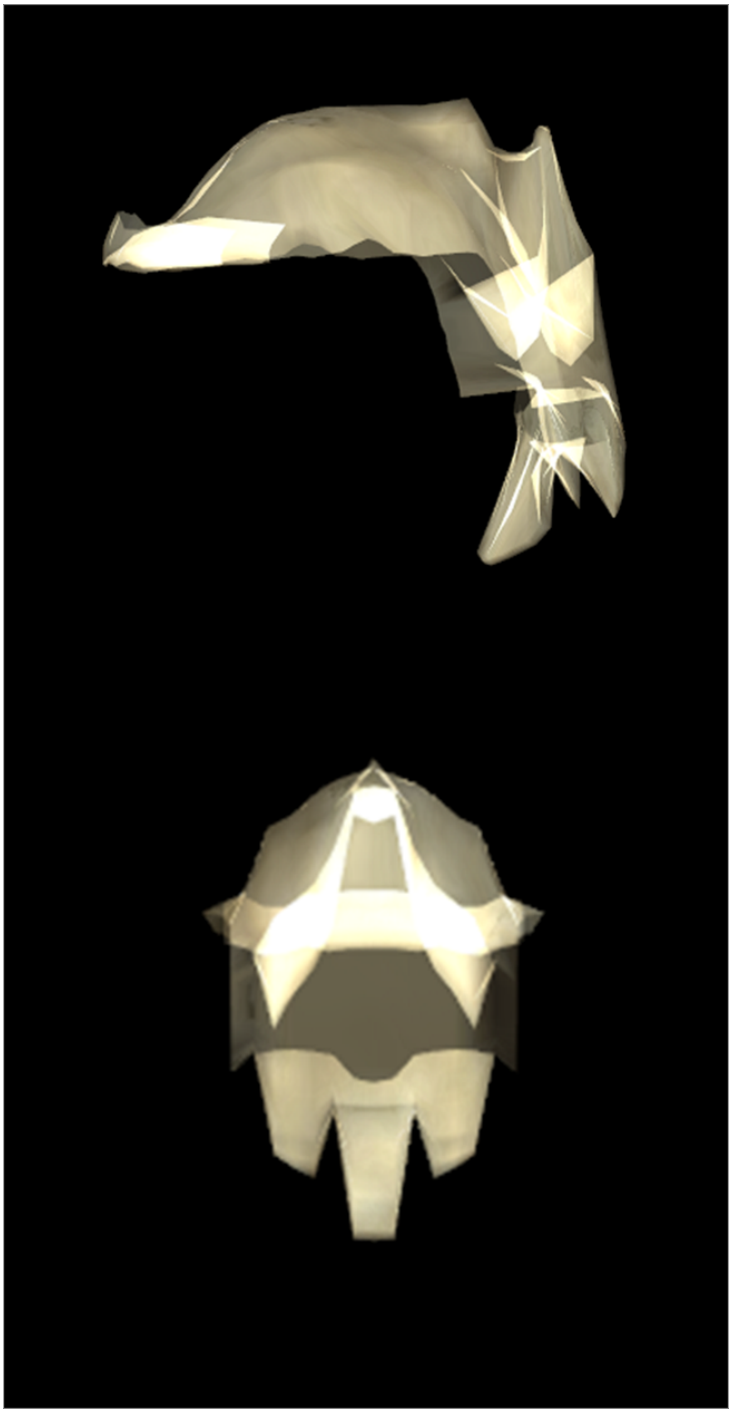
下顎なし　　下顎と声道壁なし

図 2.1: 生理学的発話機構モデル



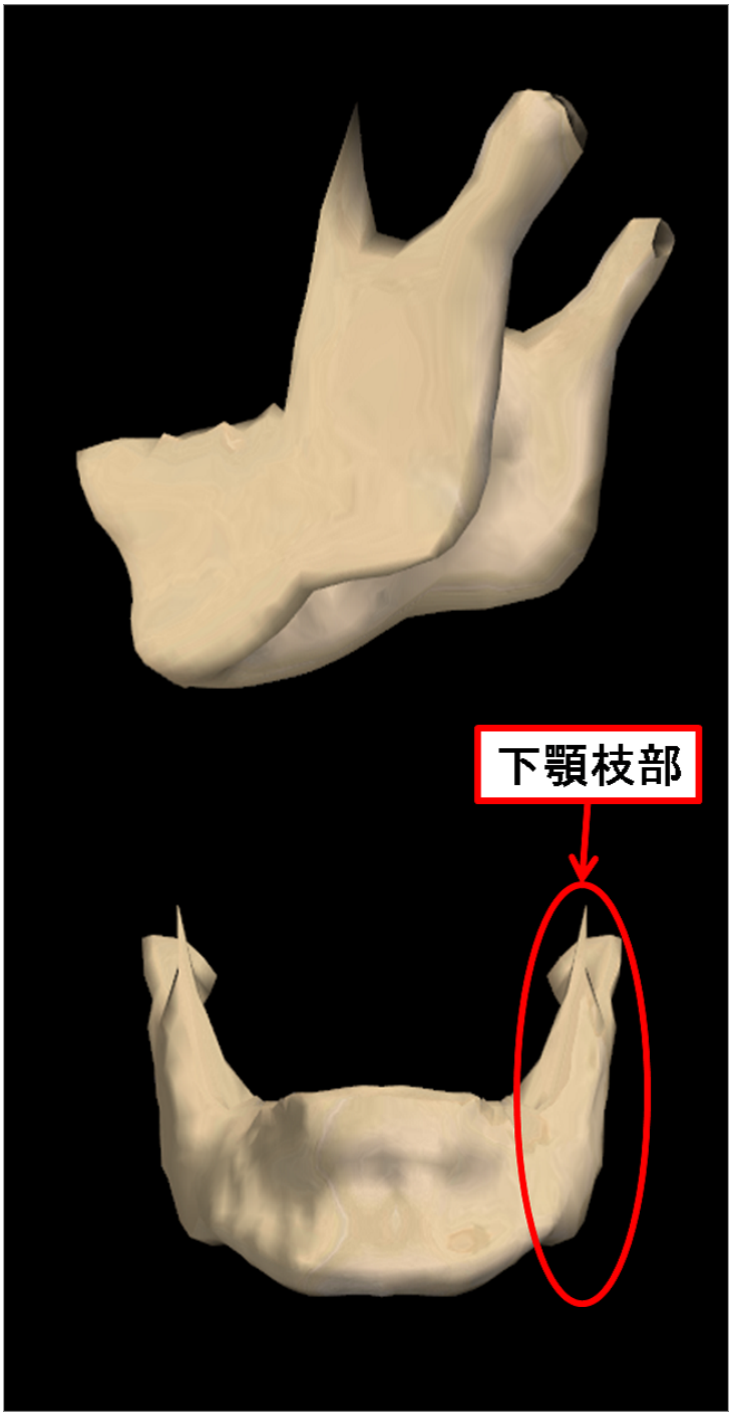
(上) 声道壁の外形形状、(下) 声道壁の領域定義

図 2.2: 声道壁の定義



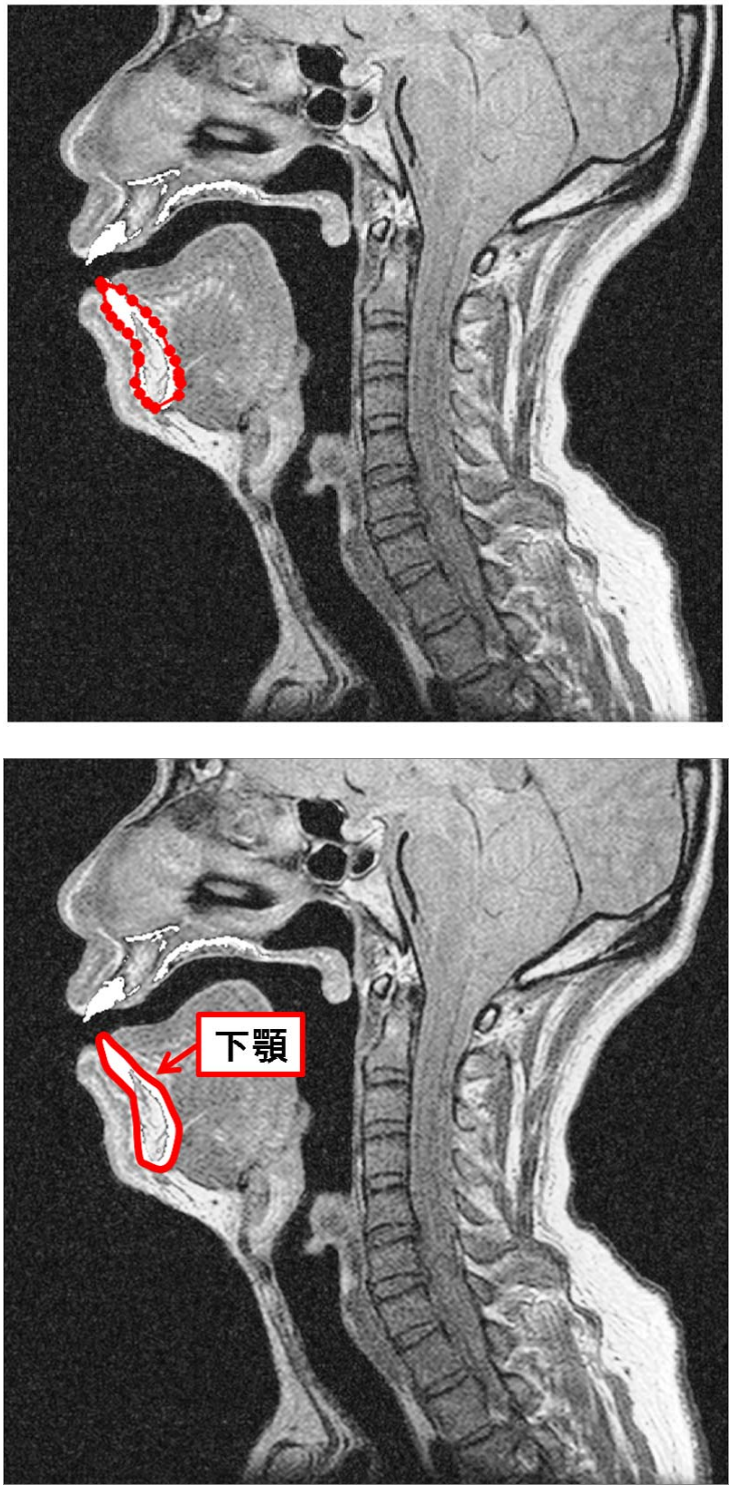
(上) 横面、(下) 正面

図 2.3: 声道壁モデル



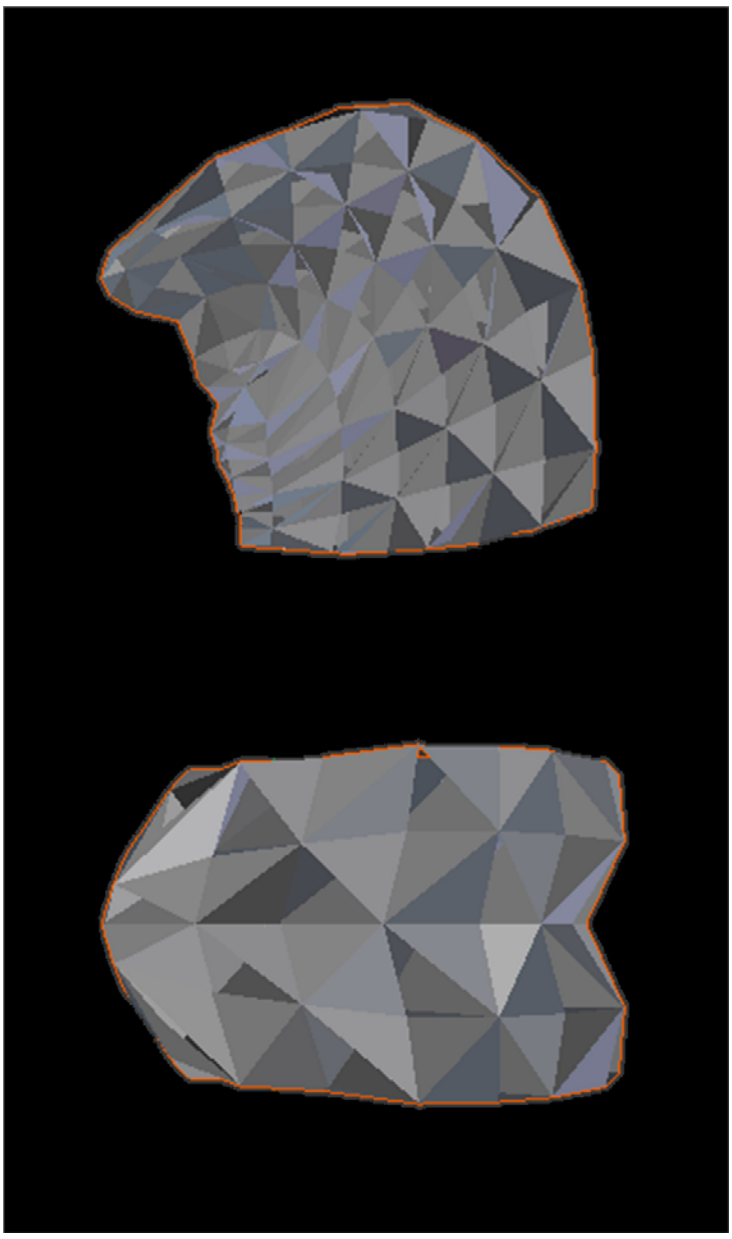
(上) 横面、(下) 正面

図 2.4: 下顎モデル



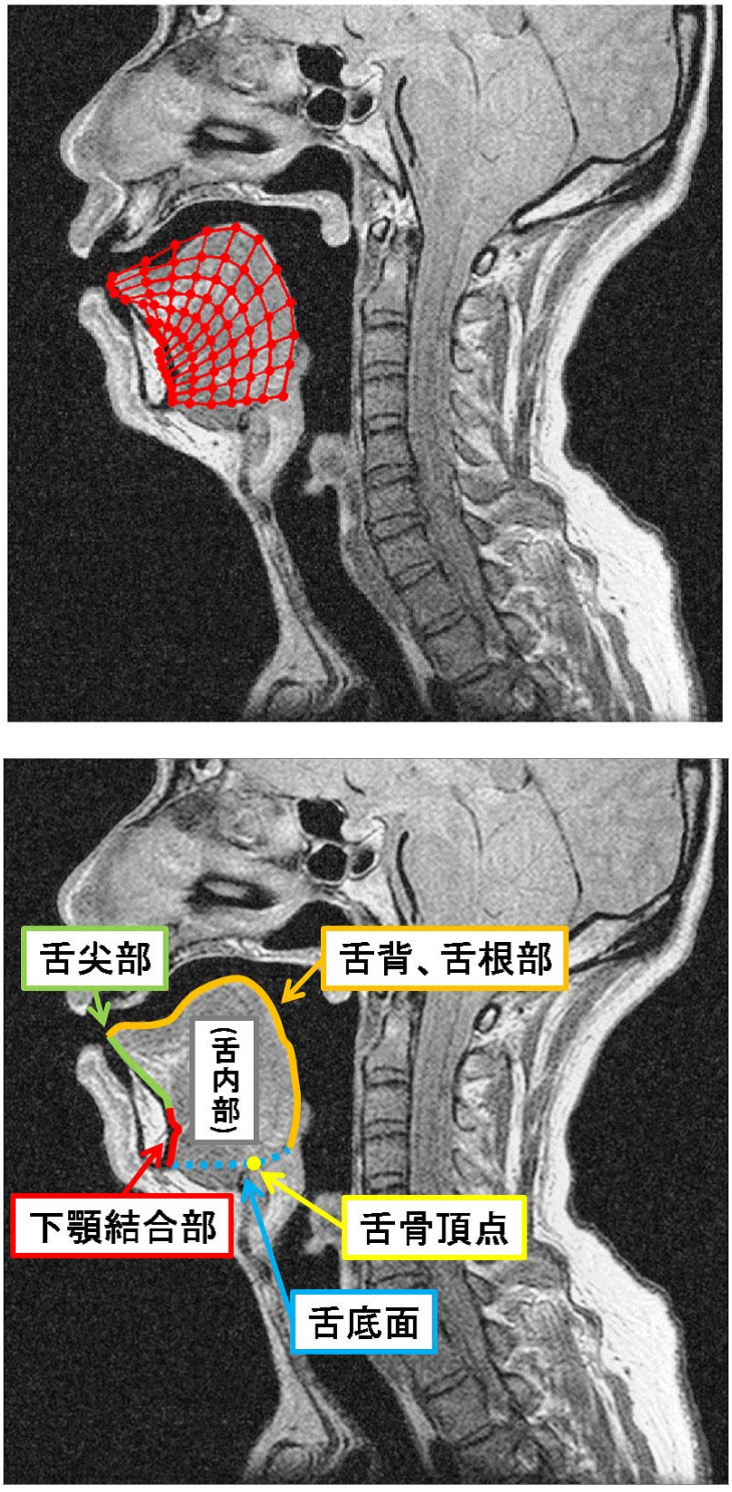
(上) 下顎の外形形状、(下) 下顎の領域定義

図 2.5: 下顎の定義



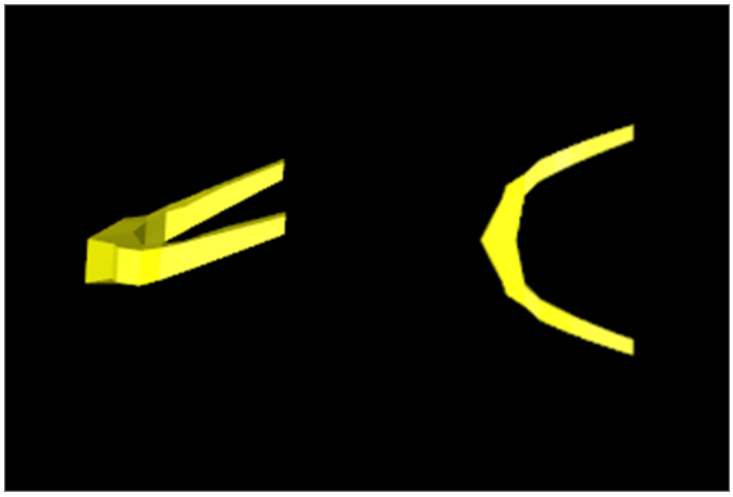
(上) 横面、(下) 上面

図 2.6: 舌モデル

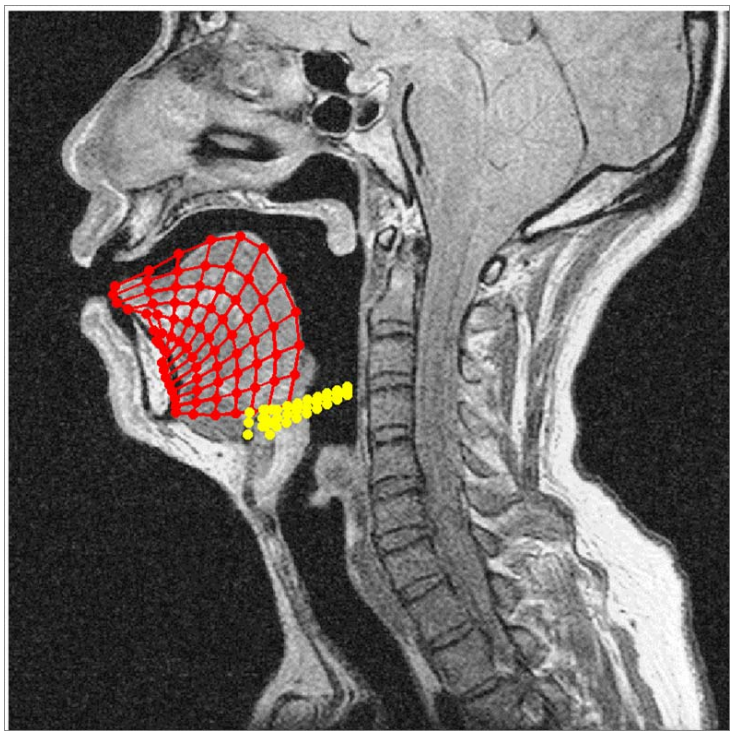


(上) 舌の形状、(下) 舌の領域定義

図 2.7: 舌の定義



(左) 横面、(右) 上面

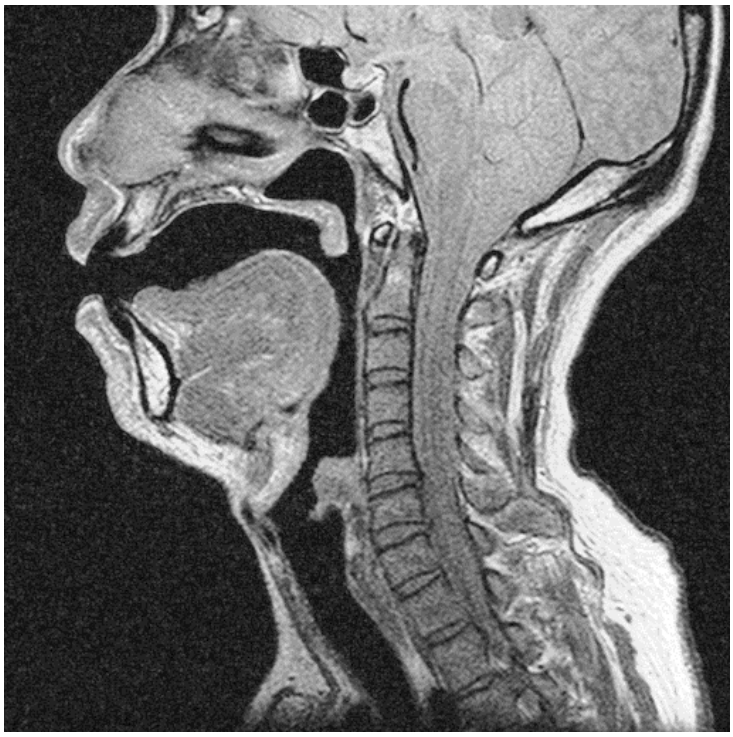


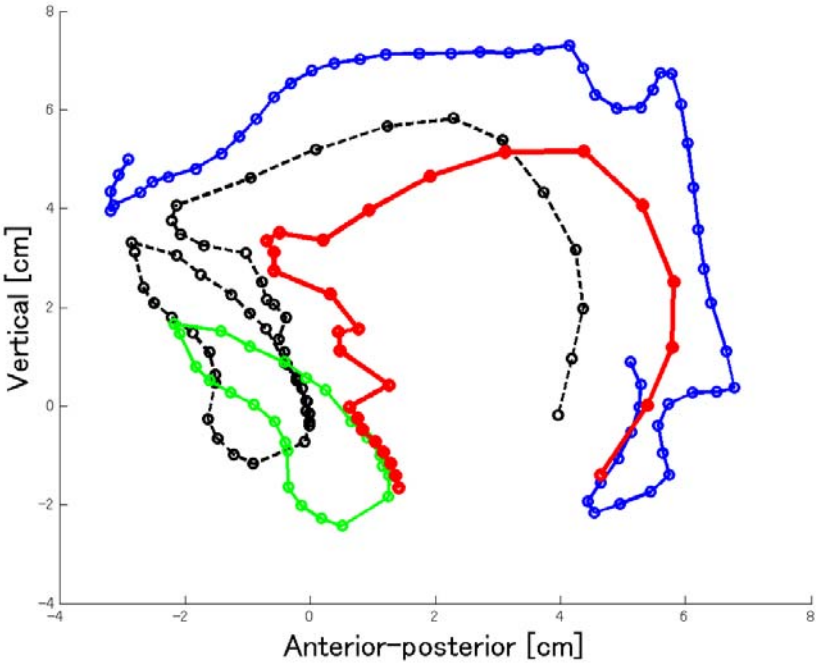
(上) 舌骨モデル、(下) 舌骨モデルの配置位置、赤：舌、黄：舌骨

図 2.8: 舌骨モデル

(上) 横面、(下) 後面

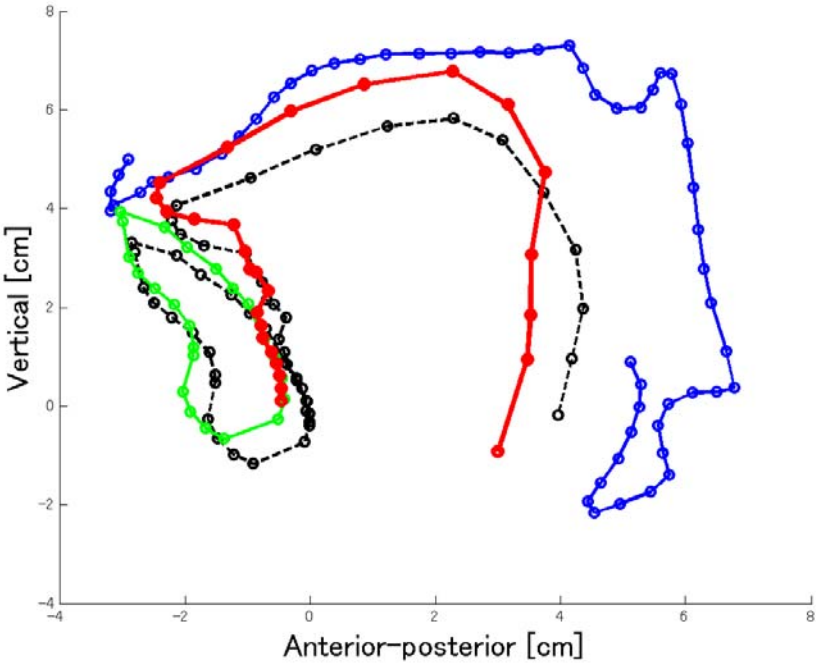
図 2.9: 喉頭部モデル





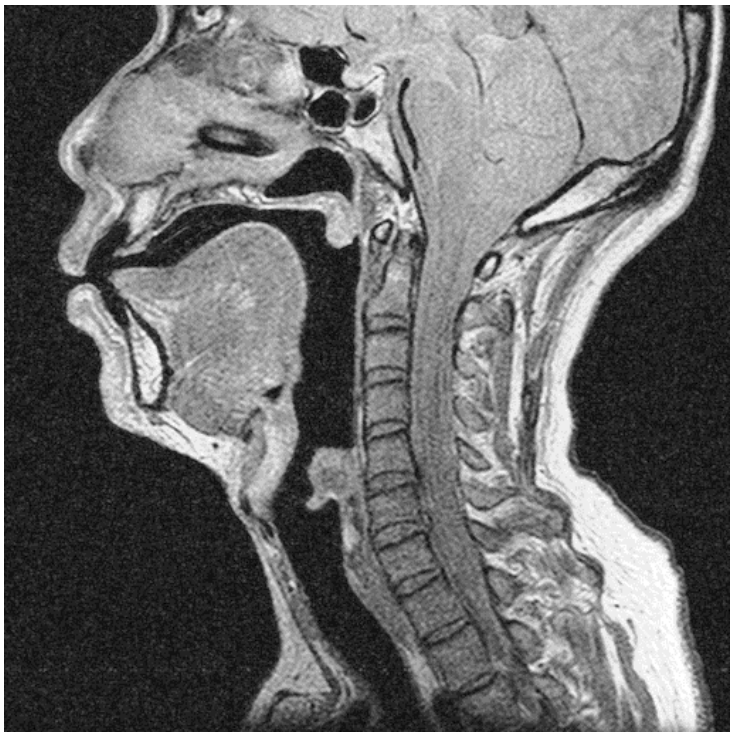
黒：初期の舌形状・下顎位置、青：声道壁、緑：移動後の下顎、赤：変形後の舌

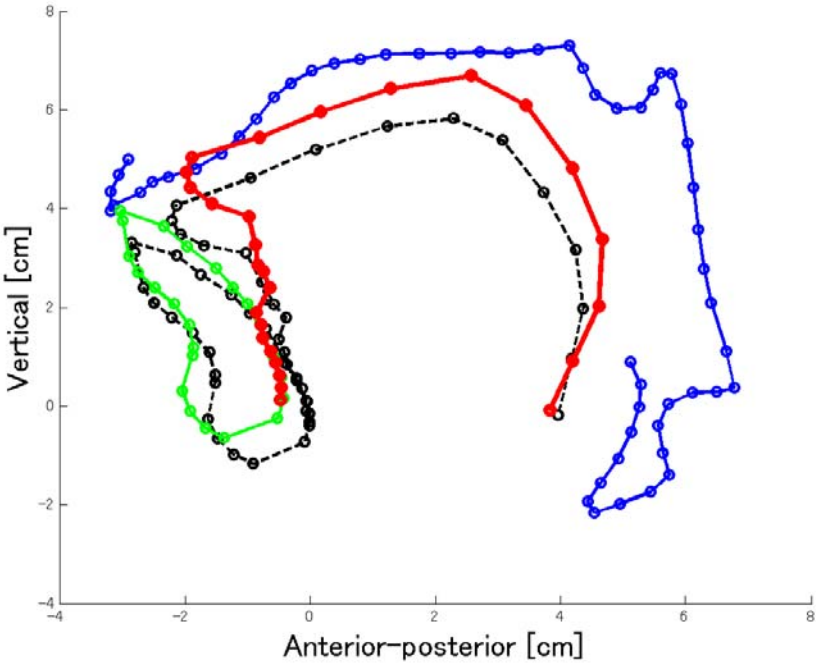
図 2.10: 母音 [a] 生成時の MR 画像とモデルの舌形状



黒：初期の舌形状・下顎位置、青：声道壁、緑：移動後の下顎、赤：変形後の舌

図 2.11: 母音 [i] 生成時の MR 画像とモデルの舌形状





黒：初期の舌形状・下顎位置、青：声道壁、緑：移動後の下顎、赤：変形後の舌

図 2.12: 母音 [u] 生成時の MR 画像とモデルの舌形状

第3章　モデルの個人化

3.1　モデル構築の問題点

生理学的発話機構モデルは、音声生成過程を再現することが可能な反面、モデルの構造が非常に複雑である。また、形態学的情報をモデルに反映させるために、形状の詳細な情報を MR 画像等から取得する必要がある。そのため、モデルを一から構築するには沢山の時間と労力が必要となり、新たなモデルの構築が容易ではない。

形状抽出について、MR 画像から対象の器官を全自動で取得することが理想ではあるが、現実的ではない。理由として、MR 画像は撮影条件や被験者によって画像が不鮮明であること（撮影には時間が掛かり、取り直すにはコストが高すぎる）、各発話器官の境界を見分けるにはある程度の解剖学的な知識を必要とすることなどが挙げられる。例えば、周波数領域を用いて声道壁領域の各器官を自動で分離する研究 [22] などが行われているが、舌と軟口蓋の領域が付着している画像等での抽出、声道壁における各器官の正確な境界位置の取得には様々な問題が残されている。以上より、形状抽出は人間が目視により作業を行い、その過程をサポートすることが望ましいと言える。

モデルの構造について、発話器官の形状や大きさに個人差はあるが、人間の発話機構の構造自体は同じと言える。そのため、新たなモデルを構築する際にはオリジナルのモデル構造をそのまま生かし、細かい部分をカスタマイズすることが考えられる。

3.2　提案手法

本研究では人間の発話器官の形状にある個人差に着目し、筋などの解剖学的構造における差を対象にしない。そのため、プロトタイプモデルの内部構造を維持しつつ、形状のみを目標話者に適応することで、形態学的情報を反映した個人化を行うことにした。MR 画像から発話器官の形状情報の他、目標話者の発話機構の寸法情報（横幅や奥行きなど）を取得し、これらの形態学的情報を基に各発話器官の個人化を行う。変形による個人化は、骨格標準モデルを用いた各個人の骨格モデルの構築に関する研究 [23, 24, 25] などで行われている。

表 3.1: 各発話器官に用いる個人化手法

発話器官	線形変形	細部の変形
声道壁		○
下顎	△	△
舌		○
舌骨	○	
喉頭部	○	

表 3.2: 寸法情報を取得する際の基準点の名称

点	名称
A	下顎骨正中矢状断面における海綿骨最上点
B	下顎骨正中矢状断面における海綿骨最下点
C	正中矢状断面における舌骨
D	正中矢状断面における第二頸椎
E	(名称不明)
F	下顎右第二大臼歯の中心
G	下顎左第二大臼歯の中心

3.2.1 線形変形

モデル全体の寸法の適応と解剖学的知見に基づき構築されている発話器官の寸法適応には線形変換を用いる。MR 画像から寸法情報を取得しプロトタイプモデルとの比を用いることで、モデルの各発話器官の大きさと目標話者の発話器官の大きさの整合性を保つ。図 3.1 と表 3.2 に、寸法情報を取得する際の基準点を示す。これらの点の座標値から、以下の式を用いて寸法を求める。

$$\mathrm{Height} = \left| y_B - y_E \right| \tag{3.1}$$

$$\mathrm{Width} = \left| z_F - z_G \right| \tag{3.2}$$

$$\mathrm{Depth} = \left| \frac{x_A + x_B}{2} + \frac{x_F + x_G}{2} \right| \tag{3.3}$$

$$\mathrm{TongueDepth} = \left| x_C - x_D \right| \tag{3.4}$$

3.2.2 細部の形状の適応

形状の細部の変形には、RBF(Radial Basis Function) 法 [28]（式 3.5）を用いる。この手法は、表情の合成 [26] やスキンアニメーションと呼ばれる皮膚の形状変化 [27] に関する研究など、CG の分野等で広く用いられている。また、青木ら [25] は各個人の顎顔面骨格モデルを生成する際に RBF を用い、少ない特徴点から各個人の特徴を反映した骨格モデルを構築した。形状情報の内、幾つかの点を特徴点とし、プロトタイプモデルの形状の特徴点と、それに対応する目標話者の特徴点を与えることで、目標話者のモデル形状を計算により得ることができる。

$$S_{i,k} = (A(\bar{P}_i))_k + \sum_{j=1}^{N} a_{j,k} g(||\bar{P}_i - \bar{P}_j||) \tag{3.5}$$

$$g(r) = \begin{cases} r^2 \log(r) & r \neq 0 \\ 0 & r = 0 \end{cases} \tag{3.6}$$

ただし、S_i は目標話者のモデルの座標値、$\bar{P}_i$ はプロトタイプモデルの座標値、$\bar{P}_j$ はプロトタイプモデルの特徴点の座標値、$k = 1, 2, 3$ は次元である。カーネル関数 g には様々な種類があるが、本研究ではカーネル関数に関するパラメータ調整や最適化を必要としない、Thin-plane カーネル関数 (式 3.6) を用いている。

3.2.3 モデル個人化の手順

モデルの個人化は、各器官を個別に変形し統合することで実現する。しかし、舌、舌骨、下顎については、舌骨は舌骨頂点において舌と、また舌は下顎結合部分で下顎に付着しているため、個人化の順番に注意が必要である。

声道壁は、形態学的情報に基づく形状の特徴が反映されている。MR 画像から形状情報を取得し、RBF で変形することにより、個人化を行う。声道壁以外の主要な発話器官の個人化は、舌骨、舌、下顎の順番で行う。初めに舌骨に対して、一様な線形変形を行う。その後、舌骨の頂点に舌の接続点が合うよう、形状の個人化を行う。続いて、舌の下顎結合に合わせて下顎を個人化することにより、各発話器官の位置関係を保ったまま、個人化を行うことができる。下顎は舌や声道壁と比べ発話に対する影響度が低いため、計算を行う際に重要な舌と接する面を中心に、話者のおおよその形状を合わせた。正確な形状を反映させなかった理由として、MR 画像の解像度の低さが挙げられる。特に歯列補填用に撮影された MR 画像はノイズがひどく、歯列細部の形状を正確に抽出することが困難であった。また、モデルの特性から、下顎と舌の接続には多少の修正が加えられている。今回は出来る限りプロトタイプモデルの計算上の条件を優先したため、見た目の形状に多少話者との違いが生じた。しかし、これらの誤差はシミュレーションに対して大きな影響を与えない。その他の器官については、MR 画像から得た寸法情報に基づき、線形変形を行う。

その際、原点からの位置に注意する必要がある。そのまま各座標軸に対してプロトタイプモデルとの比を掛けると、原点からの距離、すなわち発話器官の位置が正しく定義されない。そのため、寸法に関する個人化を行った後に、発話器官を正しい定義位置へ移動させる必要がある。

3.3 モデル個人化実験

提案した個人化手法を用い、複数人の MR 画像から形態学的情報を取得し、モデルの個人化を行う。

3.3.1 MR 画像

本研究では目標話者を中国人話者とし、それらの 3 次元 MR 画像を用い、個人化手法の検討を行った。プロトタイプモデルでは日本語母音 [e] を発話時の声道形状から構築したが、中国語には日本語母音 [e] と同じ母音が存在しない。そのため日本語母音 [e] に最も近い中国語母音 [ɣ] の発話 MR 画像を採用した。

使用する 3 次元 MR 画像は、発話同期撮像法 [20] により撮影されている。発声と撮影を同期させることにより、呼吸等で振動する軟口蓋等を比較的鮮明に写すことが可能になる。MR 画像の撮影は、(株)ATR-Promotions 脳活動イメージングセンタ (ATR-BAIC) で行われた。島津 Marconi 社製 MAGNEX ECLIPSE 1.5T PowerDrive 250 を用い、TE=3.4[ms](エコータイム)、TR=2200[ms]（リラックスタイム）にて、1.5[mm] 厚、1.5[mm] 間隔の矢状断面スライス画像を 44〜51 枚撮影した。撮像領域は 256[mm] × 256[mm]、分解能は 512[pixels] × 512[pixels] であり、DICOM 形式で保存されている。また、全ての MR 画像に対し、使用する前に DICOM 形式から TIF 形式への変換等 [21] を行った後、ボクセルサイズを 0.5[mm] × 0.5[mm] × 0.5[mm] に正規化した。本実験で用いる目標話者の、正中矢状断面 MR 画像を図 3.2 に示す。話者により、発話器官の形状や大きさが異なることが分かる。特に硬口蓋の形状には個人差があり、形状によって舌の可動域等が変わるため、発話に大きな影響を与えていると考えられる。

歯列補填画像の作成

モデルの声道壁に含まれている歯列領域の形状を正確に抽出するため、竹本らの手法 [20] を用いて歯列補填を行った MR 画像を作成した。歯列領域の境界が分かる様、ブルーベリージュースを口に含み撮影した歯列抽出用の画像から歯列領域を抽出し、アフィン変換を用いて歯列を対象の MR 画像に補填した（図 3.3）。

正解形状の作成

変形による形状の誤差等を求めるために、予め声道壁と舌の正解形状を取得した。目視により、3 断面（矢状断面：sagittal plane、横断面：transverse plane、冠状断面：coronal plane）の全スライス面において対象器官の領域を塗りつぶし、複数断面で対象器官であると判断された領域を正解領域とし、スムージング処理と手修正を行って正解形状を作成した。図 3.4 に、矢状断面のみから得た正解領域（上、スムージング等なし）と 3 断面を用いて得た正解領域（下、スムージング等あり）を示す。1 断面から領域を得る場合、各断面から死角となる領域が存在する。例えば矢状断面から舌を抽出する場合、左右の端がどこまでなのかが分かりにくい。また、舌と軟口蓋の領域が付いている場合、矢状断面以外からその境界を推定することは難しい。そこで 3 断面からの情報を用いることにより、これらの問題を解決し、より正確な正解形状を取得することを実現した。

図 3.5 に、実際に誤差計算等で使用する際の正解形状を示した。本研究で個人化の対象としている発話器官の形状の形状が評価しやすいよう、対象領域とその他領域の境目のみを情報を持たせてある。これらの形状情報は 2 値のマスク画像データとして扱い、評価を行う際は 3 次元の正解形状として扱う。

3.3.2 実験 A：特徴点の配置位置の検討

発話器官の形状を形態学的情報を基に RBF で変形する際、与える特徴点が多いほど正しい形状へ変形することができる。しかし、必ずしも全ての点を与える必要はなく、効率的な特徴点の与え方をすることで、少ない点数から必要な形状を取得することができる。モデル構築の際に問題点の一つとして挙げられる手間を軽減するため、特徴点の与える数と得られる形状の誤差の関係について検証する。検証の手順は、以下の通りである。

1. RBF で用いる形状の特徴点を、必ず与える点と組み合わせにより与える点、特徴点として与えない点に分ける
2. 組み合わせにより与える点の全ての組み合わせを作成し、必ず与える点と共に RBF で用いる特徴点とする
3. 全ての組み合わせに対し、RBF により形状の変形を行う
4. 事前に用意した 3 次元正解形状と変形により得られた形状を比較し、各点における距離の誤差を求める
5. 全ての組み合わせを与える特徴点の数で分類し、与える特徴点の数ごとに、誤差が最小となる特徴点の与え方を求める。また、全ての組み合わせに対し目標話者間の平均と誤差が最小となる特徴点の与え方を求める
6. 以上の結果より、話者間で共通の、誤差が十分小さく与える特徴点数が少ない特徴点の与え方（効率的な特徴点の与え方）を求める

「必ず与える点」は、生理学的・解剖学的に重要な意味を持ち（例えば、各発話器官の境目を示す点、モデルを構築する際に他の発話器官と結合する点など)、MR 画像と比較した際に必ず一致する必要がある点である。「組み合わせにより与える点」は、特徴点を与える際の間引きの対象とする点である。「特徴点として与えない点」は、予備実験（2 次元 RBF を用い、与える点を数パターン用意し、変形にどの様な影響がでるかを簡単に検証）により、変形に影響を殆ど与えず、形状の中でも重要度が低いという結果が得られた点である。これらの点が変形に与える影響はとても低く、また変形により生じると予想される誤差は十分小さく、発話に殆ど影響を与えないと判断できるため、検証の対象外とした。

「全ての組み合わせの作成」では、「組み合わせにより与える点（全部で m 点とする)」から実際に与える点 (n 点とする) を 1 から順に m まで 1 点ずつ増やし、各 n 点で考えられる全ての組み合わせ ${}_mC_n$ 通りを求める。本研究では、検証の対象とする全てのレイヤーで共通の個数・位置の特徴点を与えることとする。理由は、検証の対象とする各レイヤーの形状は似ているため、その形状を取得する際に必要な特徴点の数や位置は共通であると考えられるためである。与える特徴点を最適化するには、レイヤー間で異なる特徴点を与えることを考える必要があるが、本研究での目的は、RBF で与える特徴点の数を減らしても誤差の少ない形状が得られるか、そして 1 レイヤーに対してどの程度の特徴点を与えることで、十分な形状が得られるかを検証し、モデル構築の効率の向上が可能であるか検討することにある。モデルに対し話者の特徴を何処まで反映させる必要があるのか、モデルを構築する際に基準とする誤差の値についてはシミュレーションによる音声や動作への影響を考慮する必要があるため、今後の課題である。今回は、与える点数と誤差の関係に注目し、求める精度に対してどの程度の特徴点をどの位置に与えるべきかについて検討する。

検証の対象は、形状に話者の特徴が反映されている、舌と硬口蓋の一部とする。図 3.6 に、与える特徴点の種類とその位置を示す（全レイヤー共通)。舌は、全レイヤーを対象とし、舌底面を除く形状のうち、下顎と接続している下顎結合部分の 4 点、下顎結合部と舌尖部の境界点、舌尖、舌根の必ず与える点 7 点を除く 20 点を、組み合わせにより与える点とする。必ず与える点には先に述べた 7 点の他に、舌底面にある舌骨と接続している 1 点と舌底面の 2 点を含める。与える舌底面の 2 点は、予備実験により、与えることで舌内部のメッシュが安定することが経験的に分かっている点を選択した。これらの点は、形状に対しての影響がとても小さいことも予備実験により確認済みである。声道壁の形状には、MR 画像の撮影の特徴から、目視でも正しい形状が分かりにくい器官が含まれている。また、軟口蓋など、モデルの計算特性に合わせた修正が加えられている器官も存在する。これらの理由から、特に発話時の舌の動作に影響を与え、安定した形状の評価が行える硬口蓋を形状誤差を計算する対象とした。レイヤーは、目視により十分正確な形状が得られる正中矢状断面から 3 面とする。特徴点は上歯列から軟口蓋の中から選択し、各発話器官の境界と特に特徴的な形状点（硬口蓋と軟口蓋は大きなカーブを有しているため、カーブの頂点）を必ず与える特徴点とし、その他から 19 点を組み合わせにより与える点

とした。

形状誤差は、組み合わせにより与える点とした舌の 20 点と硬口蓋の 11 点に対して事前に与える正解形状との最短なユークリッド距離を計算し、その平均（全レイヤーの平均）とする。話者間の平均値は各組み合わせ毎に話者の平均をとり、同じ与える特徴点数の組み合わせの中で誤差が最小となる組み合わせを選び、その特徴点数での話者平均としている。

3.3.3 実験 B：変形による個人化モデルの構築

実験 B の目的は、提案手法によりモデルの構造を維持したまま個人化ができるか検証することである。話者 3 人分の MR 画像（目標話者 B, C, D）を用いることで、提案手法の一般的を確認する。検証の手順は以下の通りである。

1. 目標話者の MR 画像から形態学的情報（寸法情報、形状情報）を取得する
2. 得られた情報を基にプロトタイプモデルの各発話器官の形状を変形させ、個人化モデルを構築する
3. 構築した個人化モデルの各筋を一つずつ駆動させ、モデルの構造が維持されているか確認する
4. 個人化モデルで日本語 5 母音を生成し、舌形状の変化を考察する

モデルの構造が正しく維持されていると、各筋を一つずつ駆動した際の舌や下顎の動作が、プロトタイプモデルと同様の傾向を示すことが期待される。また、個人化モデルにプロトタイプモデルを駆動させる際に用いた日本語 5 母音を生成するための筋パラメータ（表 2.2）を与えることで、形状を個人化したことによる舌の動き方の違いを検証する。

3.4 結果と考察

3.4.1 結果 A：特徴点の数と配置位置の関係

図 3.7 に、RBF により発話器官の形状を変形する際に与える特徴点の数と、その際得られる変形形状の誤差の関係についてのグラフを示す。上が硬口蓋について、下が舌の形状についての結果である。誤差はある程度特徴点を与えると収束し、その傾向は各話者で大きな違いがない。舌よりも硬口蓋で誤差が早くに収束しているが、これは形状が直線に近いため、曲面である舌よりも少ない点で十分に形状を表現することが出来るためだと考えられる。今回は全レイヤーで共通の特徴点を与えているため、さらに最適化を行うことも可能である。

続いて、各点数における特徴点の配置位置について考察する。舌について、図 3.8 に、話者平均が各特徴点数で最小となった組み合わせの、特徴点の配置位置を示す。赤線は正解形状、黄線は変形によって得られた形状、青点は必ず与える点、黄点が組み合わせて与える点である。例として、MR 画像は話者 A の正中矢状断面を用いている。結果より、特徴点の番号（インデックス）は多少前後するが、組み合わせて与える点の増え方には規則性があることが確認された。特徴点は舌背・舌根部に等間隔で与えられた後舌尖部に与えられ、形状誤差が急激に下がる。その後、特徴点の間隔が大きな場所を中心に特徴点が与えられるが、下顎結合部へ特徴点が与えられるのは、誤差が収束する段階に入ってからである。これより、下顎結合部分は形状誤差に大きな影響を与えないといえる。また、与える点には安定性があり、話者に依らず、一般的な特徴点の与え方を導くことが可能であることを示している。

3.4.2 結果 B：個人化モデルの動作確認

図 3.9 に、個人化モデルの形状を示す。内部の発話器官が観察しやすいよう、下顎は除いてある。図から、プロトタイプモデルの構造が維持されたまま、話者 A の形状に個人化されていることが確認できる。話者 A のモデルでは舌底面部にピンク色の筋（または腱）が見えているが、舌と他の発話器官との位置関係にプロトタイプモデルと多少の差があるためだと考えられる。今回は筋の配置や太さ等について個人化を行っていないため筋や腱に多少の不自然さが残るが、全体の構造や各発話器官の位置関係は維持されていることが確認できる。

次に、各筋を個別に駆動した際の結果を図 3.10〜3.13 に示す。図は、筋の中でも代表的な動作を行う GGa, GGp, SG, HG の結果である。GGa（前部オトガイ舌筋）は舌を前下方向に、GGp（後部オトガイ舌筋）は舌を前方と上方向に、SG（茎突舌筋）は舌を後方と上方向に、そして HG（舌骨舌筋）は舌を後方と下方へ移動させる働きがある。駆動させる各筋には、一律に値 0.4 を与えている。図より、舌は正しい方向へ駆動しており、モデルの計算構造が正しく維持されていることが確認できた。プロトタイプモデルと話者 A の駆動後の舌形状が異なる理由について、各々の舌形状が異なること、そして話者 B の舌モデルは内部のメッシュ構造が最適化されていない事が考えられる。同じ力で舌を引っ張っても、話者によって舌の大きさや筋の配置位置が異なるため、このような違いが生じたと推測することができる。これらの結果は、2.5 次元で定義された党らのモデルを用いて行われた筋と舌の形状を比較した研究 [30] の結果と傾向が一致している。

以上より、個人化モデルはプロトタイプモデルの計算構造を維持したまま個人化され、形状を個人化した結果、筋の働きに違いが生じていることが確認された。

3.4.3　結果 B：5 母音生成時の舌形状

個人化モデルに、プロトタイプモデルで用いられる日本語 5 母音を生成する際の筋活性パラメータを与えた結果を図 3.14〜3.16（他の母音については、A.2 を参照）に示す。図より、各母音を生成する際の舌形状の特徴が現れているが、同じ舌活性パターンでも、プロトタイプモデルと個人化モデルでは違いが生じていることが確認できる。要因として実験結果 3.4.2 と同様のことが言えるが、各筋を個別に駆動した場合と比べ、その違いは顕著ではなくなったと言える。複数の筋を同時に駆動することにより舌が複雑に制御され、目標の母音を生成するために各筋のバランスがとられているのではないかと考えられる。母音発生時の舌筋の長さは MR 画像の観察により共通性と個人差がある [31] ことから、プロトタイプモデルと個人化モデルの舌筋の長さに差が生じていることが考えられる。これは、各器官の位置関係や距離に個人差が含まれていることを示しているといえる。

以上の結果より、個人化モデルには話者の特徴が反映されており、形状の違いによる個人差が確認された。本実験から、発話機構の個人差は各発話器官の形状だけではなく他の要因にも含まれるため、各話者の音声生成過程を再現するには、形状以外の個人化が必要であることが改めて確認された。

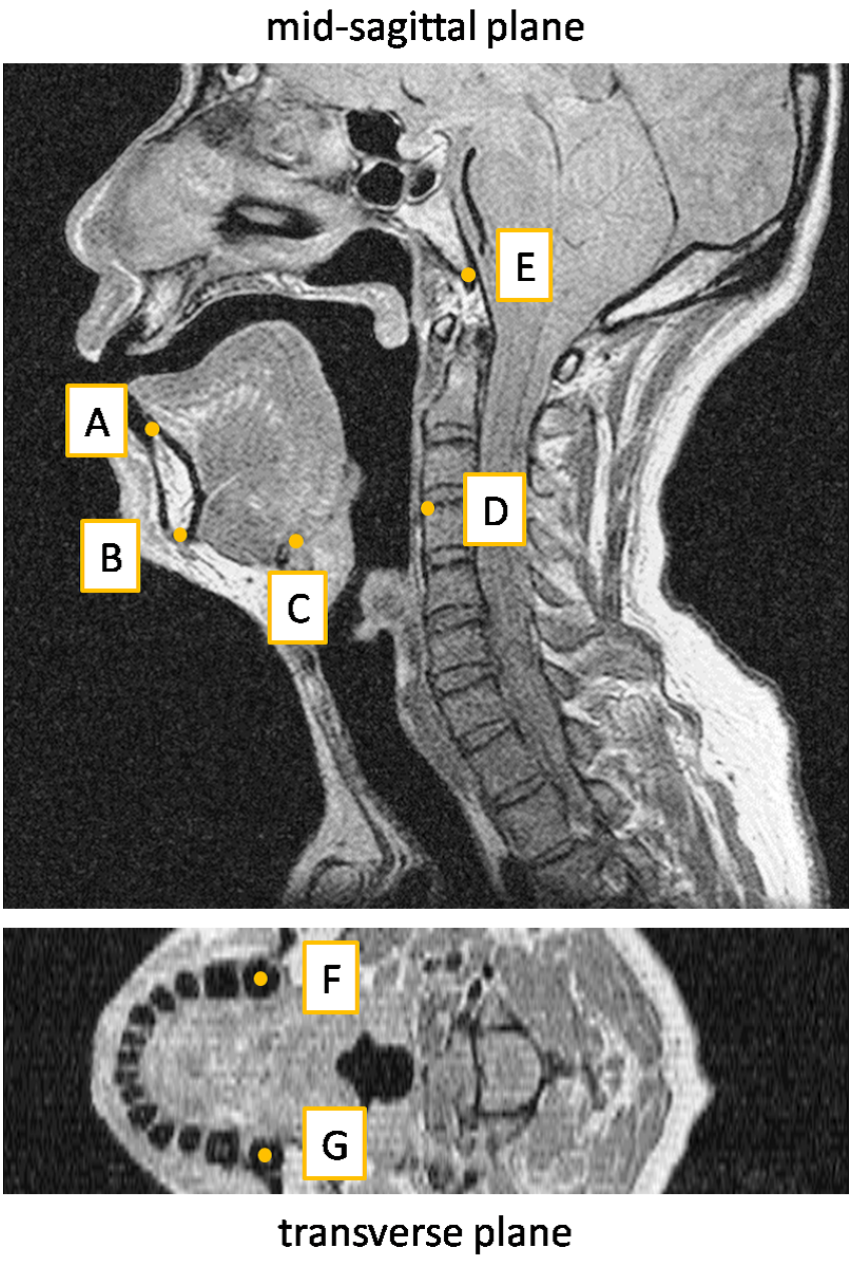
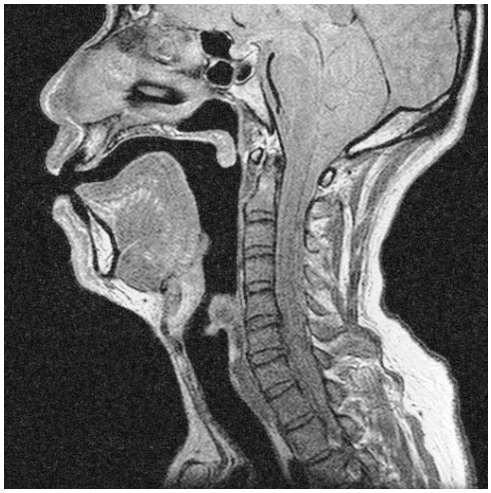
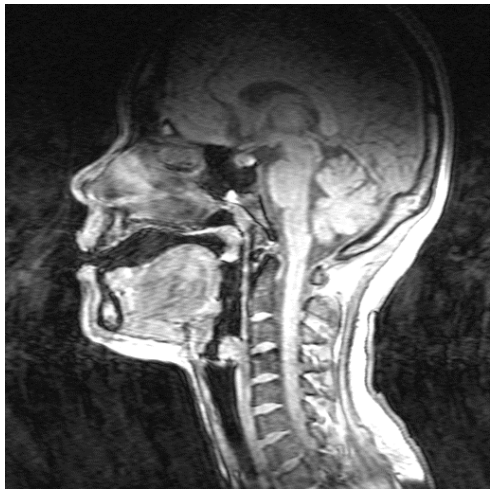


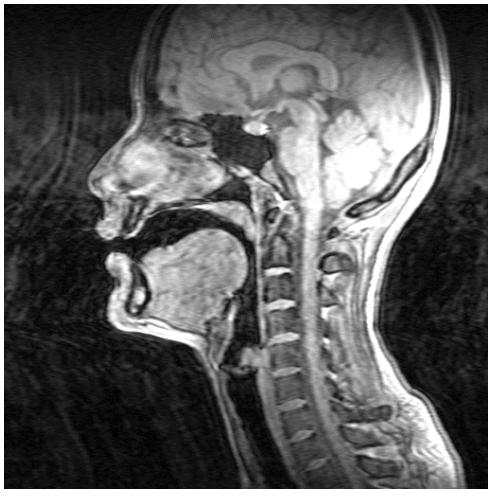
図 3.1: 寸法情報を取得する際の基準点



プロトタイプモデルの話者



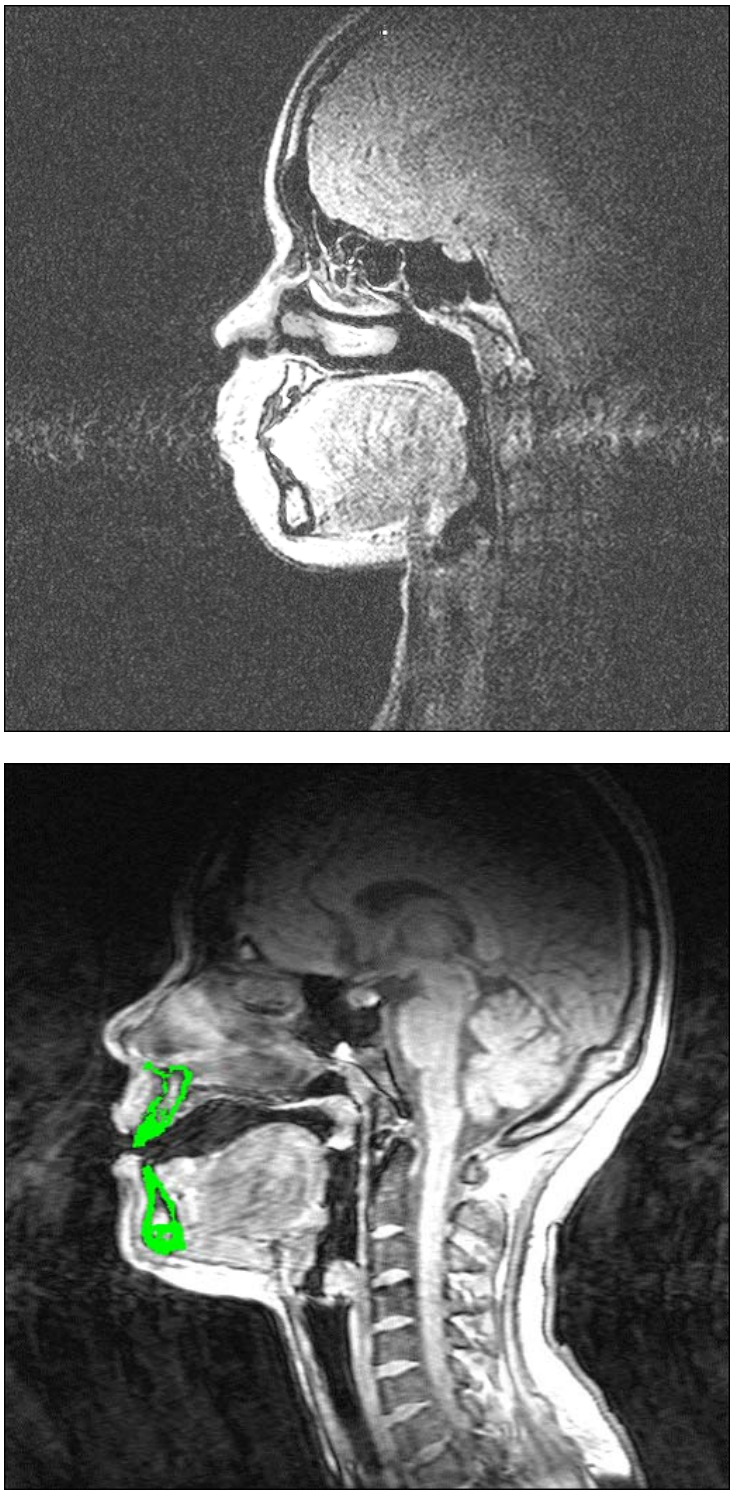
話者 A



話者 B

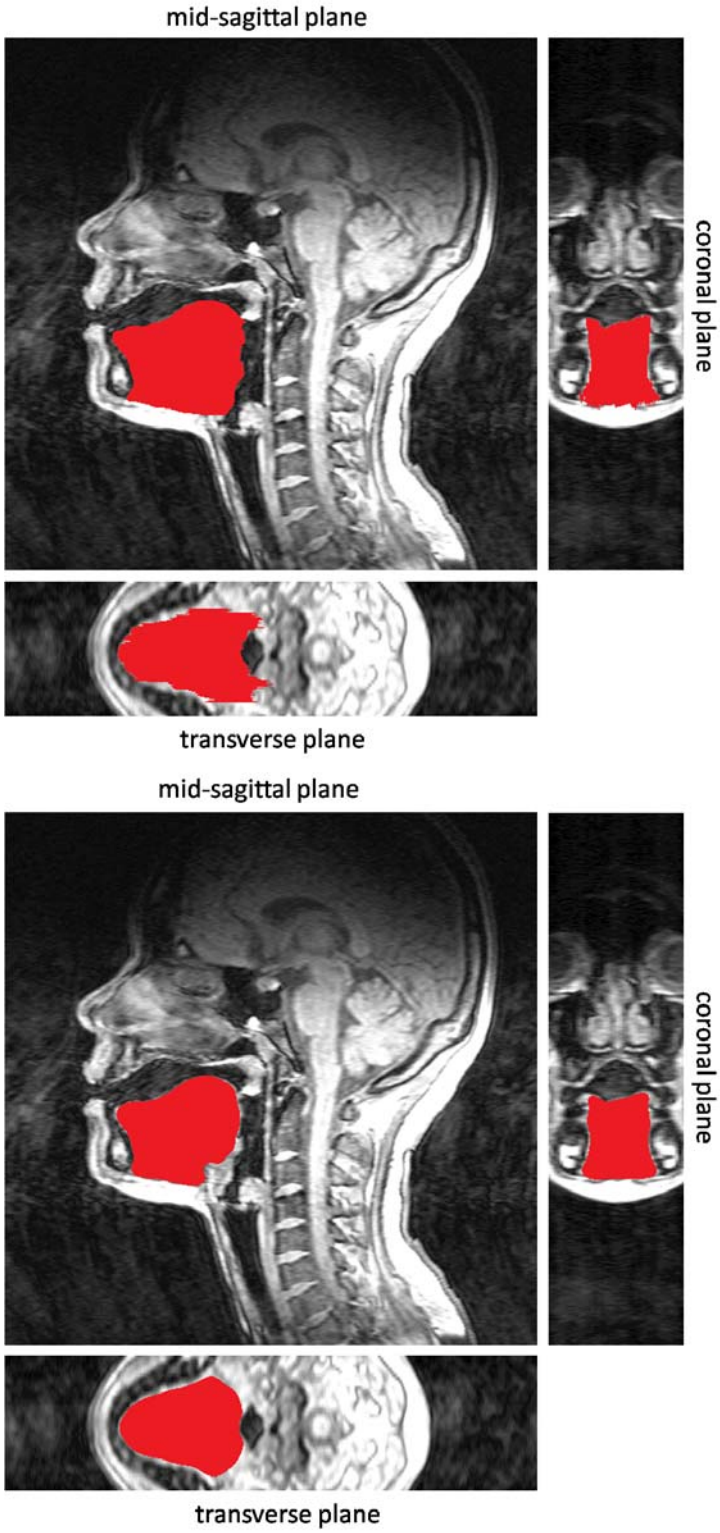
話者 C

図 3.2: 各話者の正中矢状断面 MR 画像



(上) 歯列抽出用 MR 画像、(下) 歯列補填 MR 画像

図 3.3: 歯列補填の例



(上)1 断面使用の例、(下)3 断面使用の例

図 3.4: 使用断面数による取得形状の違い

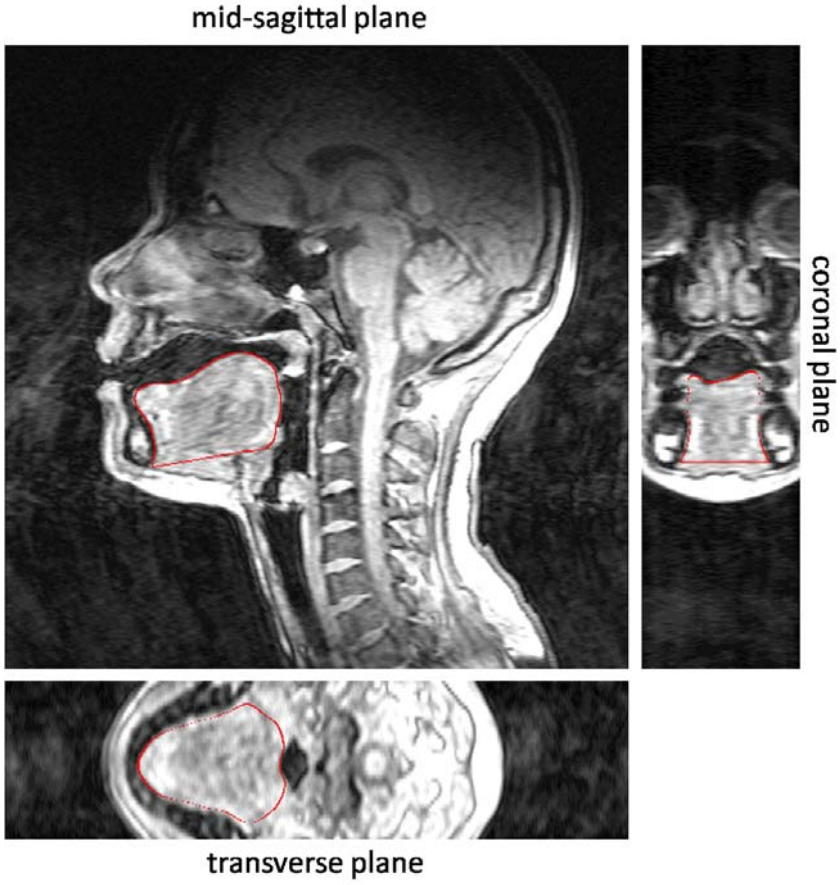
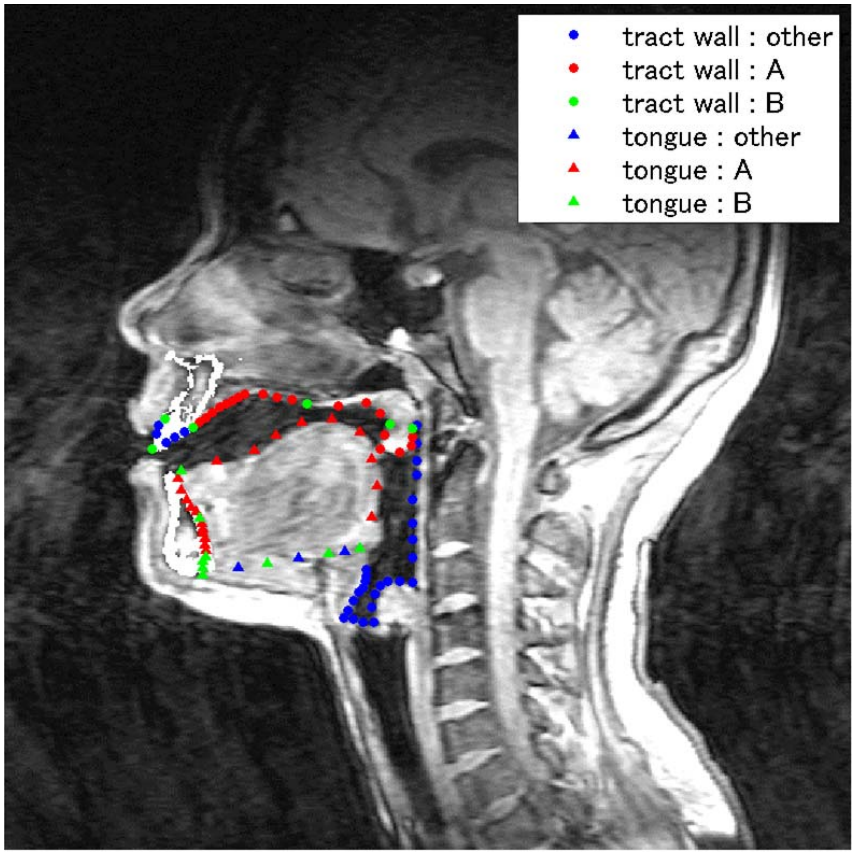
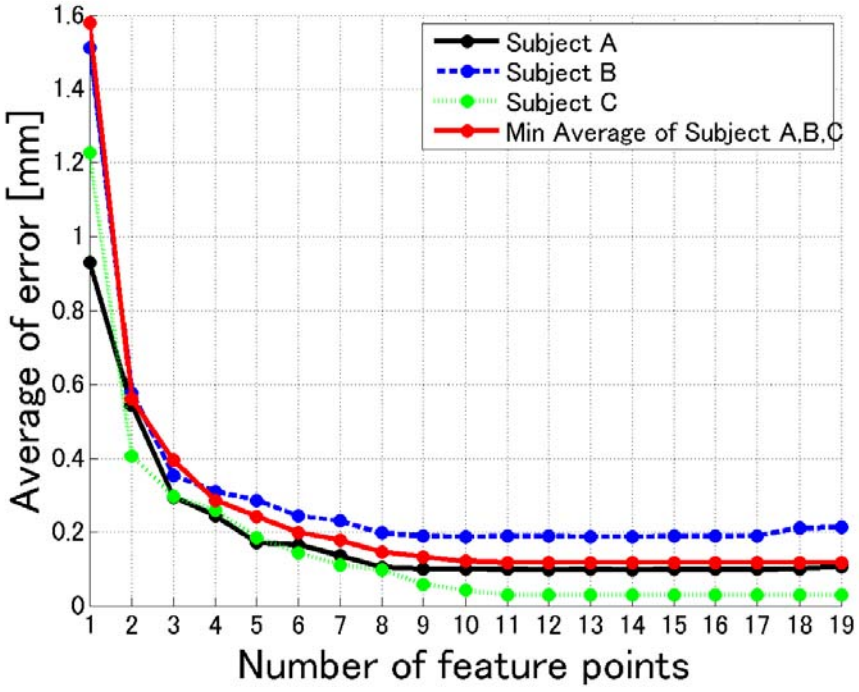


図 3.5: 正解形状の例

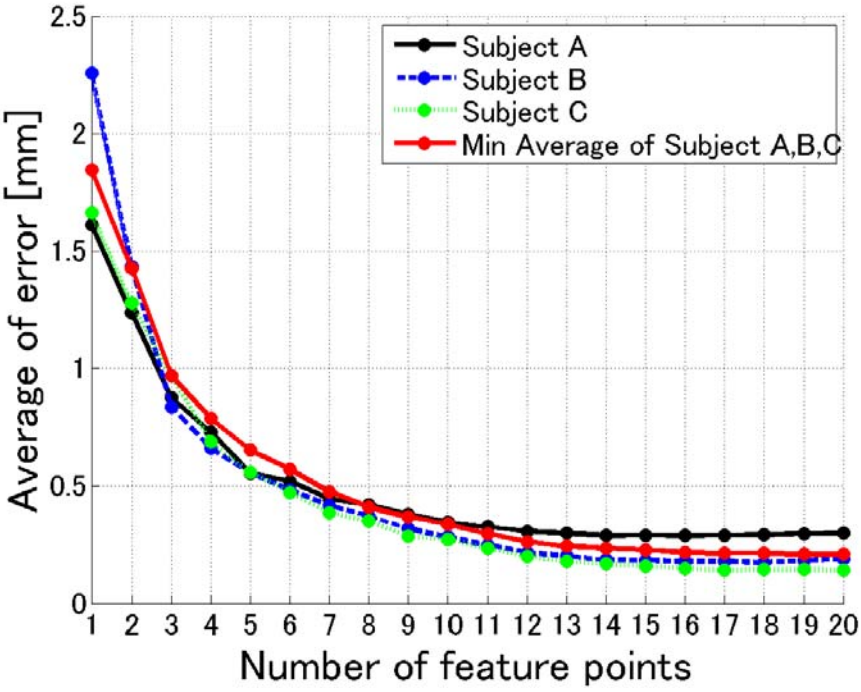


A:組み合わせにより与える特徴点、B:必ず与える特徴点、other：特徴点として与えない形状点

図 3.6: 与える特徴点の種類と位置

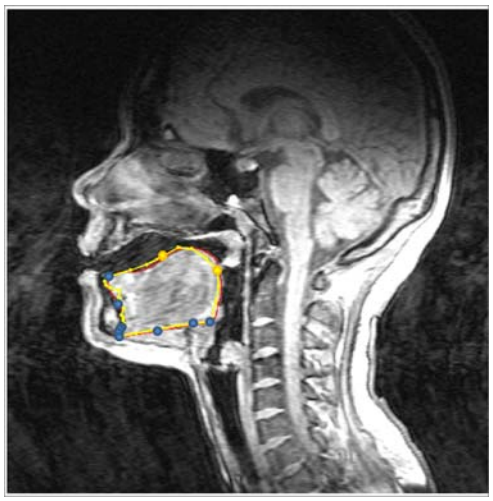


硬口蓋における特徴点の数と最小平均誤差の関係

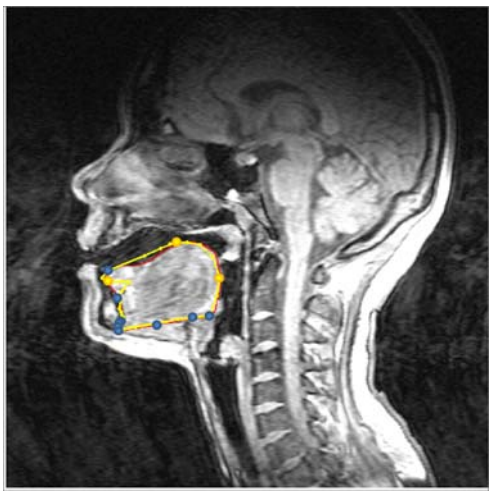


舌における特徴点の数と最小平均誤差の関係

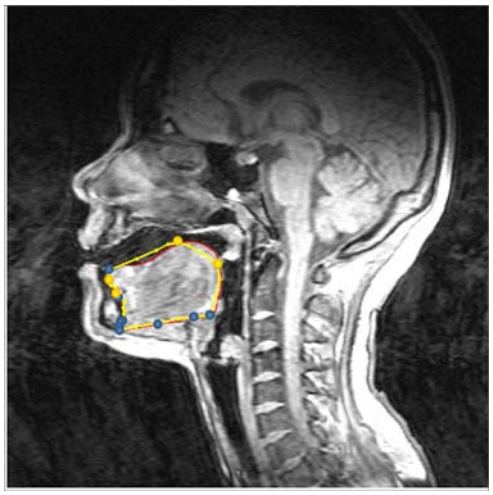
図 3.7: 特徴点の数と形状誤差の関係



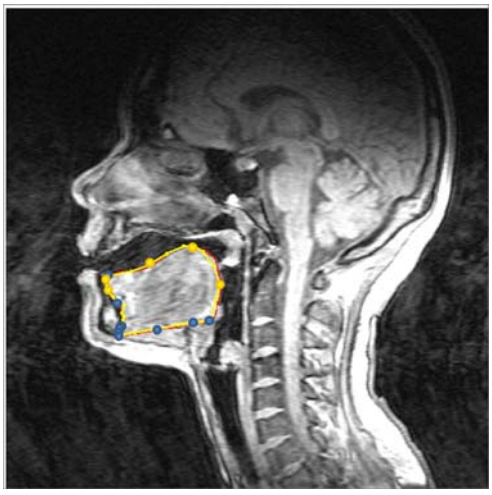
+2 点



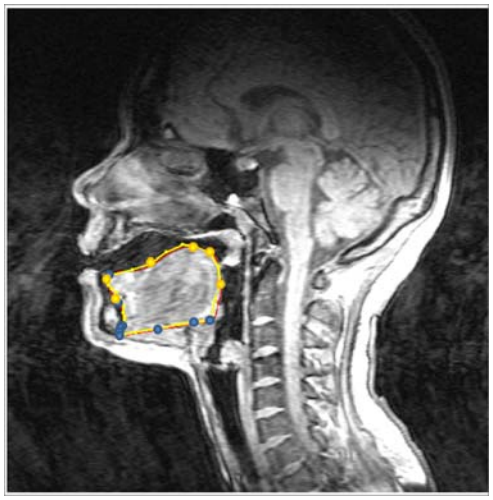
+3 点



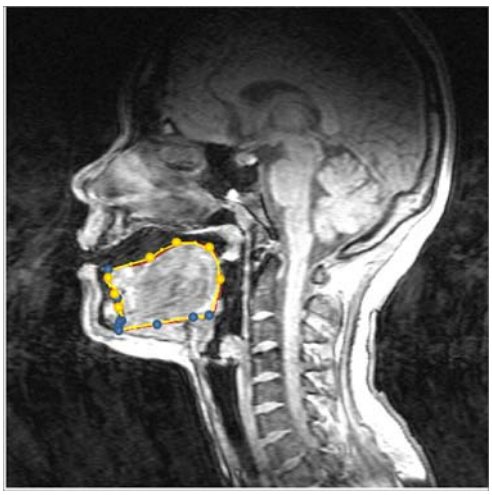
+4 点



+5 点

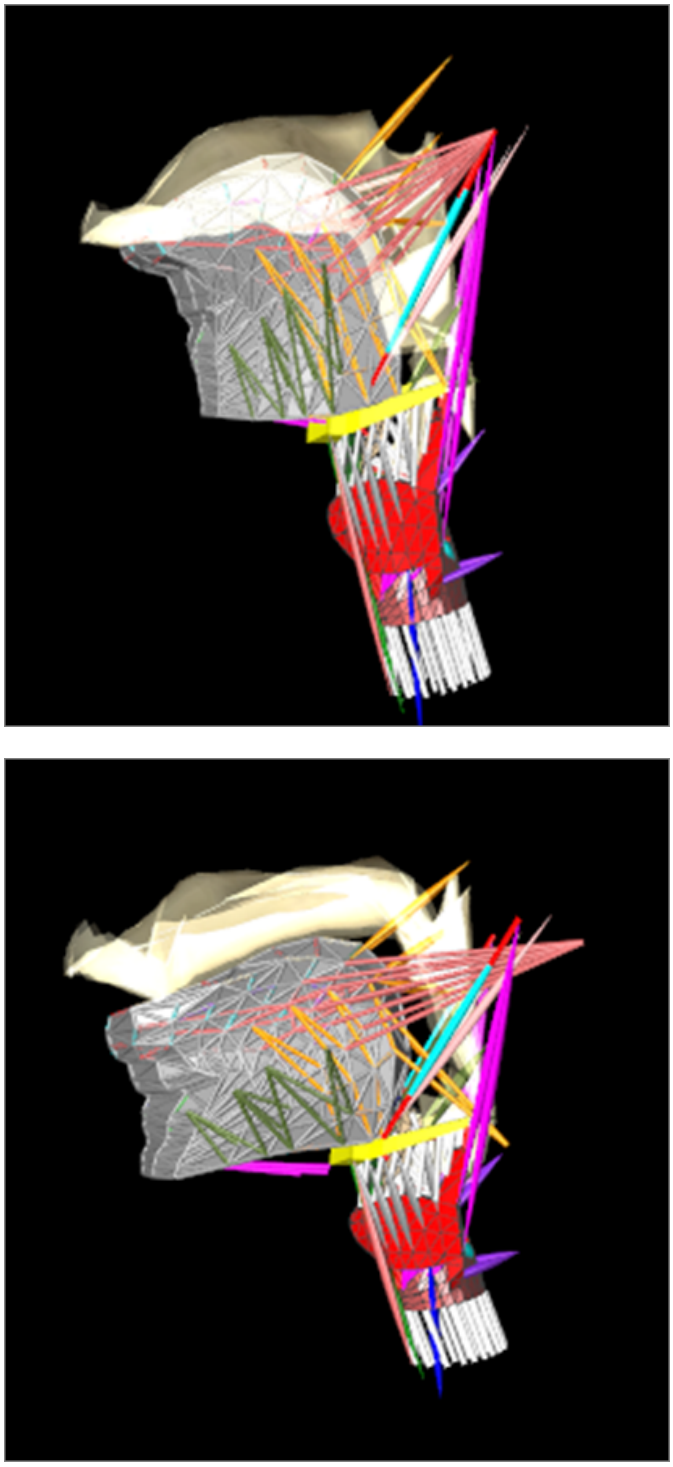


+6 点



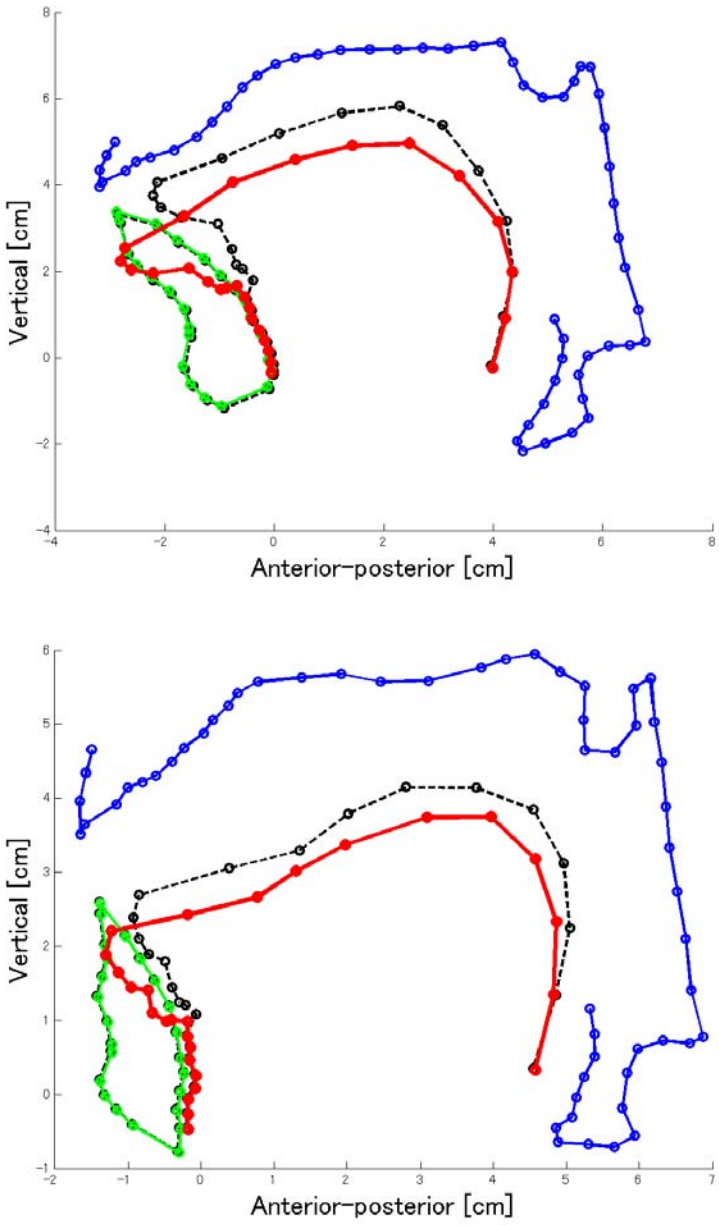
+7 点

図 3.8: 誤差が最小となる特徴点の配置位置



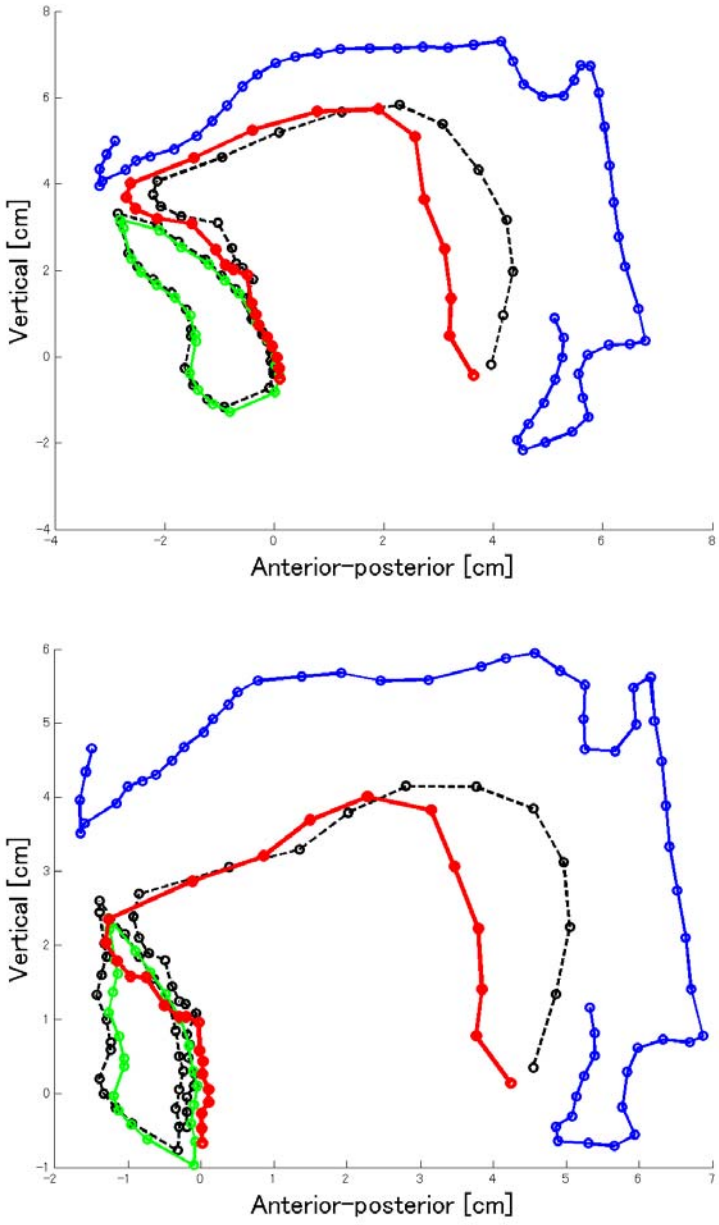
(上) プロトタイプモデル、（下）話者 A

図 3.9: 個人化モデルの形状（下顎を除く）



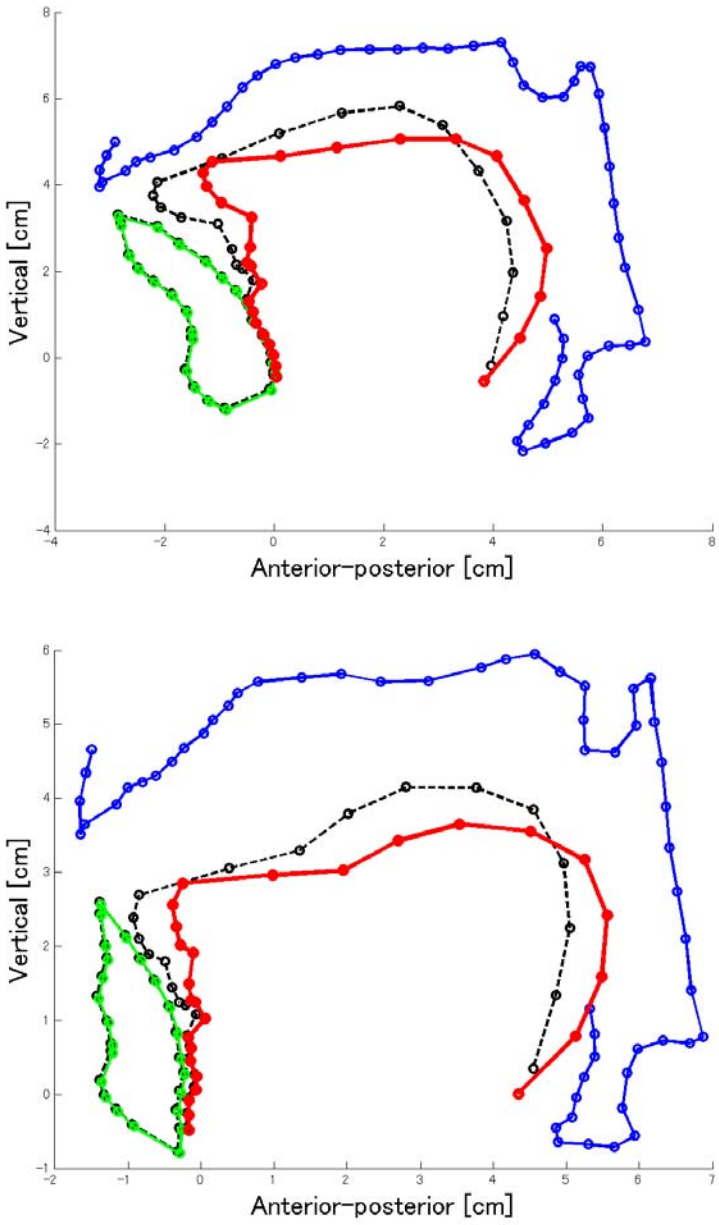
(上) プロトタイプモデル、(下) 話者 A

図 3.10: 筋 GGa を駆動させた際の舌形状



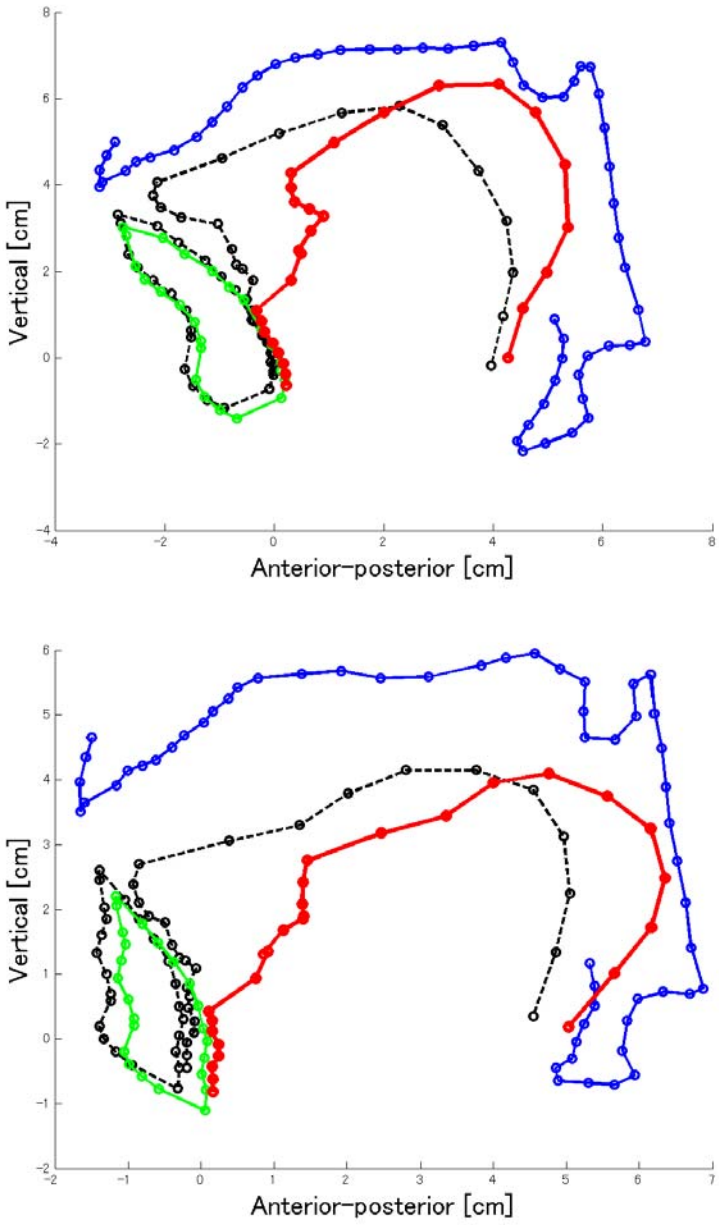
(上) プロトタイプモデル、（下）話者 A

図 3.11: 筋 GGp を駆動させた際の舌形状



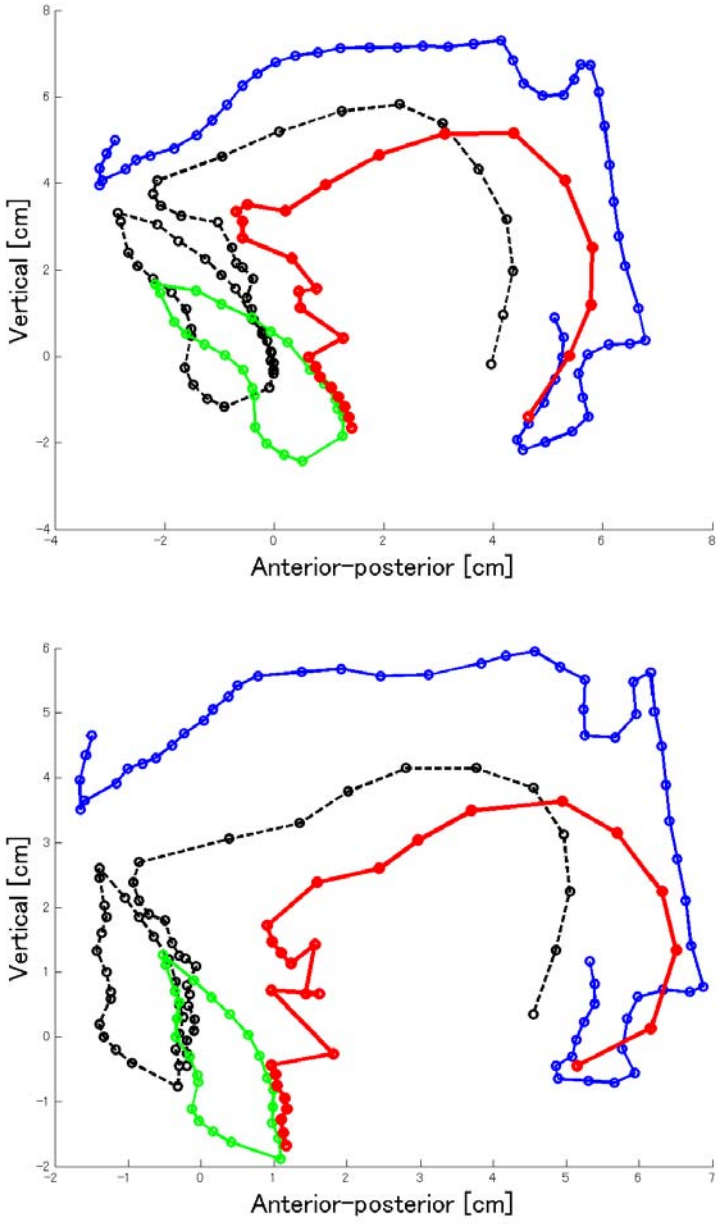
(上) プロトタイプモデル、（下）話者 A

図 3.12: 筋 HG を駆動させた際の舌形状



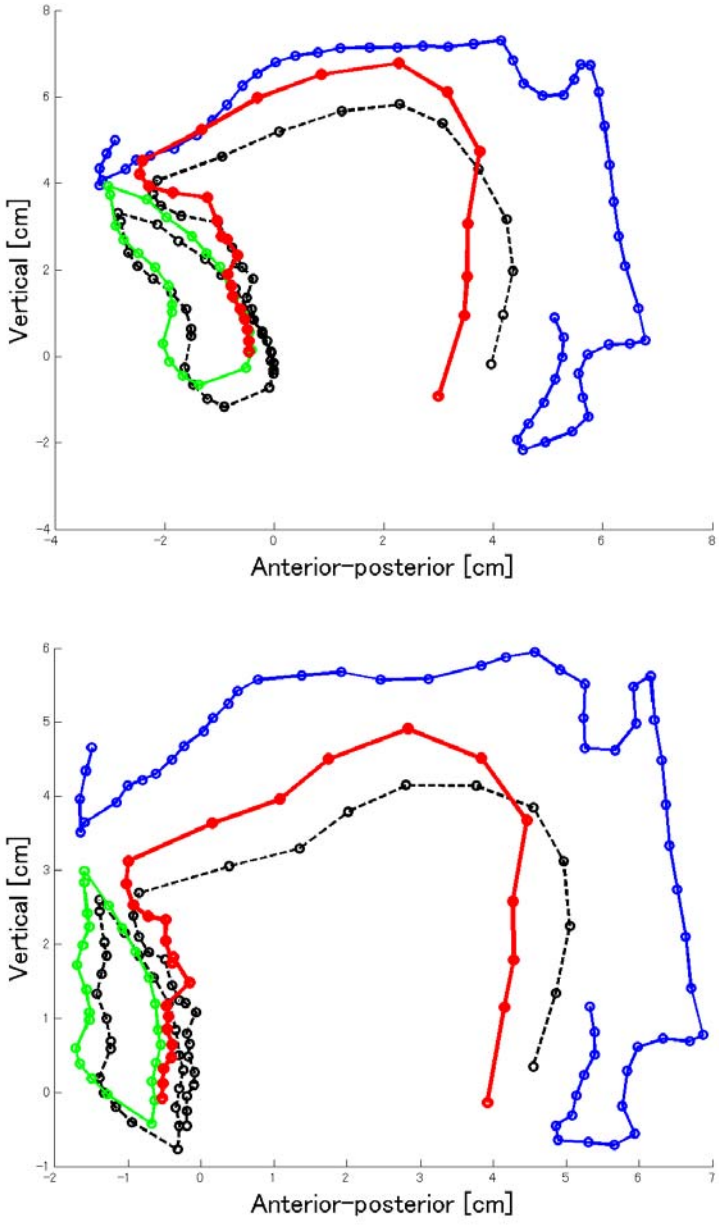
(上) プロトタイプモデル、（下）話者 A

図 3.13: 筋 SG を駆動させた際の舌形状



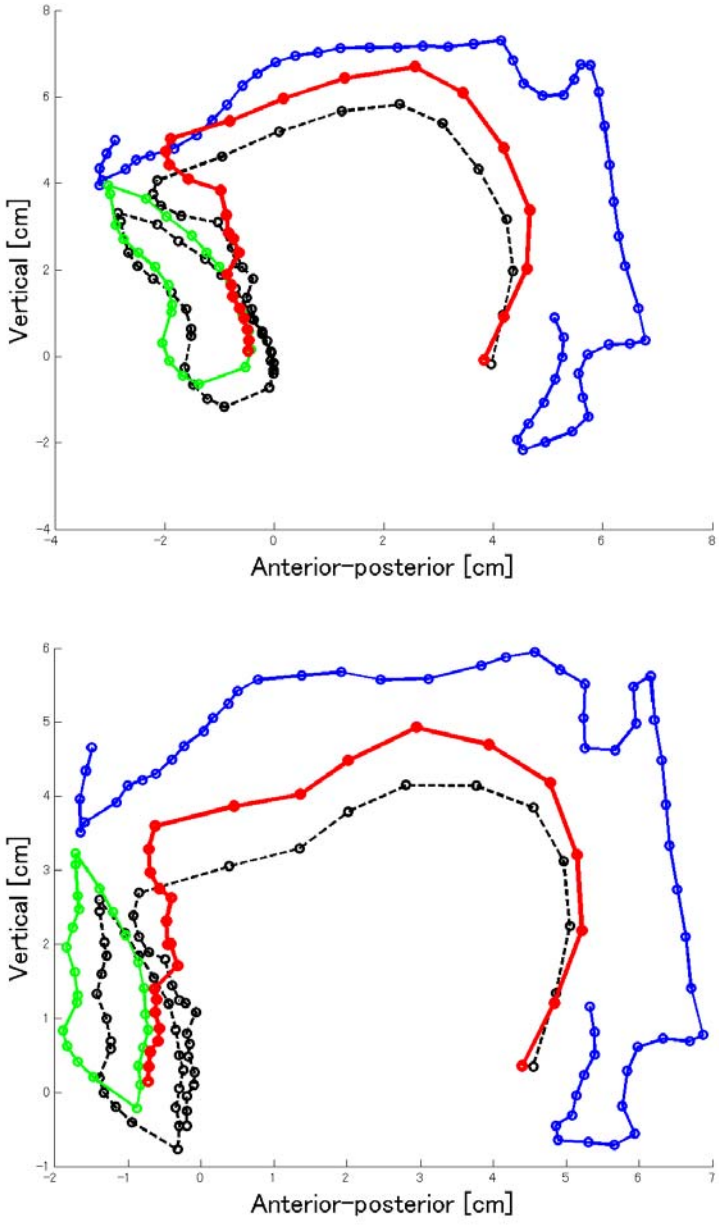
(上) プロトタイプモデル、(下) 話者 A

図 3.14: 個人化モデルによる母音 [a] 生成時の舌形状



(上) プロトタイプモデル、(下) 話者 A

図 3.15: 個人化モデルによる母音 [i] 生成時の舌形状



(上) プロトタイプモデル、(下) 話者 A

図 3.16: 個人化モデルによる母音 [u] 生成時の舌形状

第4章　結論

4.1　まとめ

本研究では、プロトタイプモデルの構造を維持したまま発話器官の外形を変形することによる個人化モデルの構築手法を提案した。MR 画像から得られる形態学的情報を基に変形を行うことで、モデルに目標話者の特徴を容易に反映できることが確認された。提案手法を用いてモデルの個人化を行った結果、個人化したモデルにはプロトタイプモデルの計算構造が維持されていること、そして変形を行う際は全ての形状点を抽出する必要が無く、構築を行う際の手間を、従来法で一から構築する際と比較し、大幅に減らすことが出来ることを確認した。その際、形状誤差が最小となる最適な特徴点の配置位置には規則性があり、必要な形状精度に合わせた段階的な特徴点の配置位置を一般化出来ることが示された。以上より、提案手法を用いることで話者の特徴が反映された個人化モデルを、効率良く構築できることが示された。また、個人化モデルは、プロトタイプモデルと同じ筋活性パラメータを与えても舌の動き等に違いが生じ、発話器官の形状以外にも個人差が含まれているため、各話者の音声生成過程を忠実に再現するには、形状以外の個人化も必要であることが再確認された。

4.2　今後の課題と展望

被験者を増やした、個人化手法の検討

本研究では 3 人の被験者により、提案した個人化手法を検討した。結果として、3 人共にそれぞれの特徴が反映された個人化モデルを構築することができた。次のステップとして、提案した個人化手法をさらに効率化・最適化するには、被験者数を増やした検討が必要である。

原点の統一化

本研究で用いている生理学的発話機構モデルは、[11]、[12]、そして 3 次元モデルと、段階的に提案された。そのため、各器官を定義する際の原点が、統一されていない。本研究では、形態学的情報を MR 画像から取得する際には統一の原点（画像の左上端とした）とし、形状の変形等を行う際に座標変換を行った。特に、舌や下顎、軟骨類で用いられて

いる原点は下顎結合部分にあり、MR 画像から明確な基準で抽出することが難しい。今後は、明確な基準で定義することが出来る一点に原点を定め、統一することが求められる。そのためには、モデル内部での計算過程を見直し、プログラムを修正する必要がある。

舌内部のメッシュ構造

舌内部のメッシュは、解剖学的知見に基づいて構築されている。メッシュの構造は、舌の制御に深く関係することから、今後個人化することが必要である。具体的な方法として、Tagged-MRI を用いた舌内部のモデリング [32] などが考えられる。

生成音声の評価

本研究では、モデルで音声を生成することを行っていない。理由として、プログラム内で音声を生成する際に声道断面積関数を計算する過程で、プロトタイプモデルの基となっている話者に特化した記述があることが挙げられる。また、音声の個人差に深い関係がある声帯部分が個人化されていないため、生成音声を用いて話者の個人性を議論することが難しい。以上の点に加え筋の配置や制御も合わせて個人化することで、生成音声を用いた個人差の評価が可能になると考えられる。

軟口蓋形状、梨状窩形状のモデリング

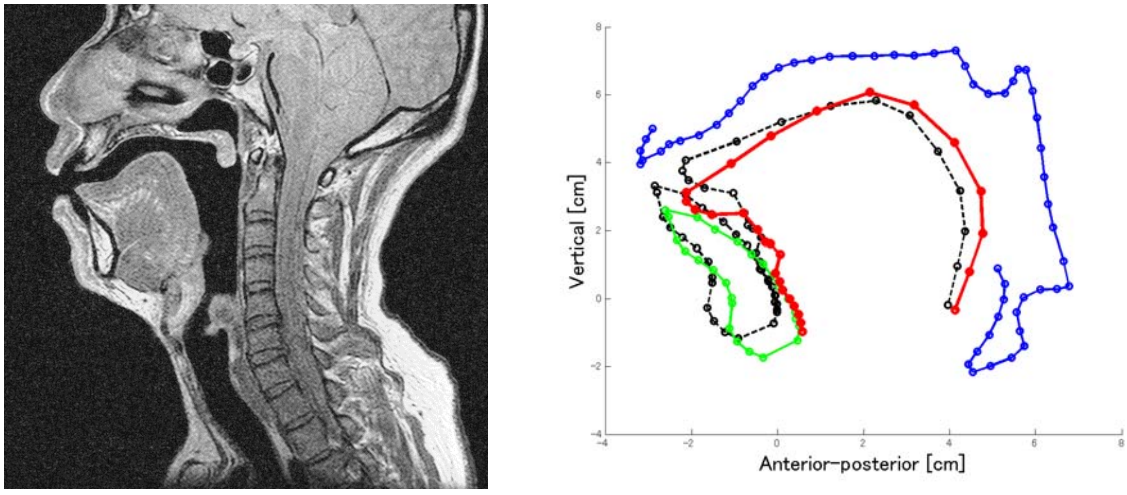
プロトタイプモデルでは、声道壁形状に軟口蓋と梨状窩が含まれている。軟口蓋の振動 [33] と梨状窩 [34] が発話に影響を与えていることが指摘されているが、これらの部分について、モデルを構築する際の明確な基準が存在しない。プロトタイプモデルではこれらの形状に、修正（スムージング等）が加えられているため、どの様な修正がなのか、正確な形状を抽出する方法の検討を含めて考える必要がある。

外形以外の個人化

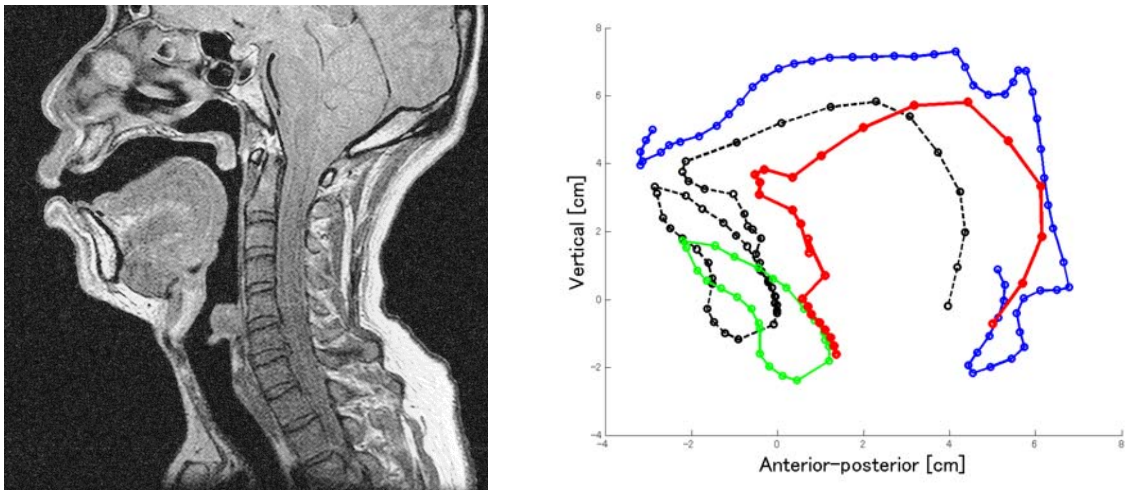
本研究では、個人化のファーストステップとして、発話器官の外形に注目した。筋の配置や太さ、舌内部にも個人差が含まれるため、これらの部分について順次個人化する必要がある。今回提案したプロセスを基盤にすることで、新たな個所の個人化をスムーズに行えるようになると考えられる。

付 録A　結果：追加資料

A.1　母音生成時のMR画像とモデルの舌形状



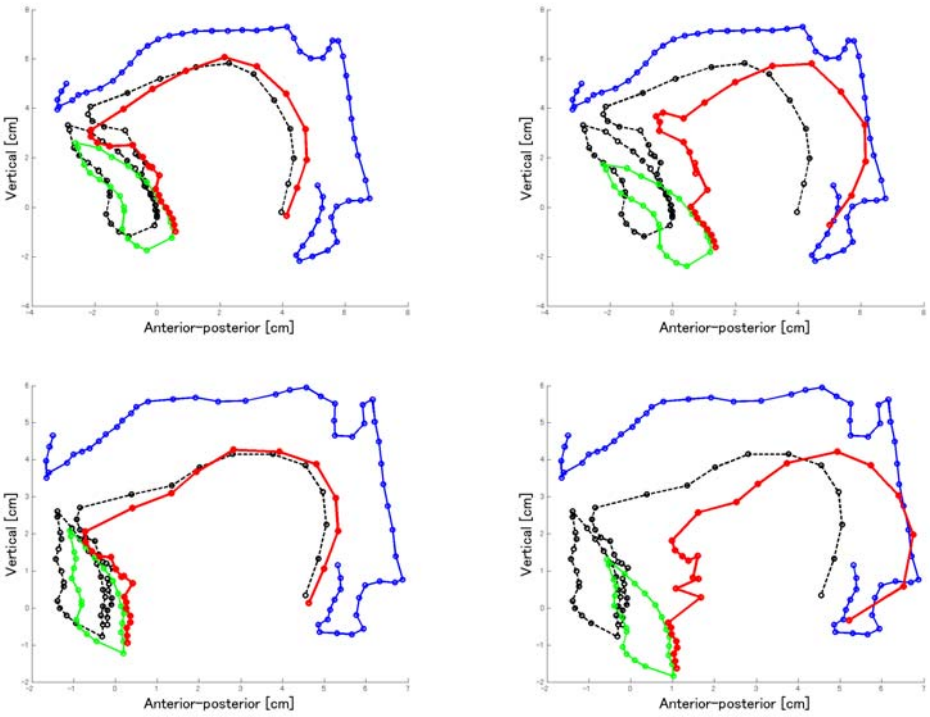
母音 [e] 生成時の MR 画像とモデルの舌形状



母音 [o] 生成時の MR 画像とモデルの舌形状

A.2　個人化モデルによる母音生成時の舌形状

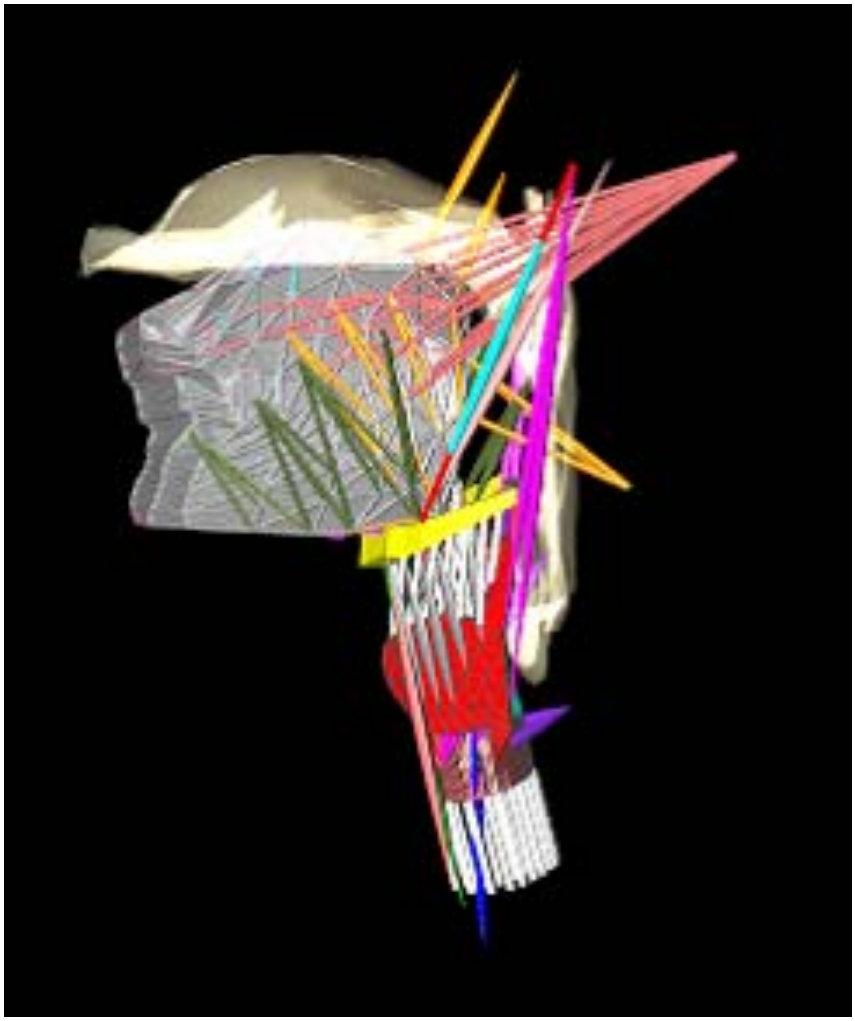
母音 [e]　　母音 [o]



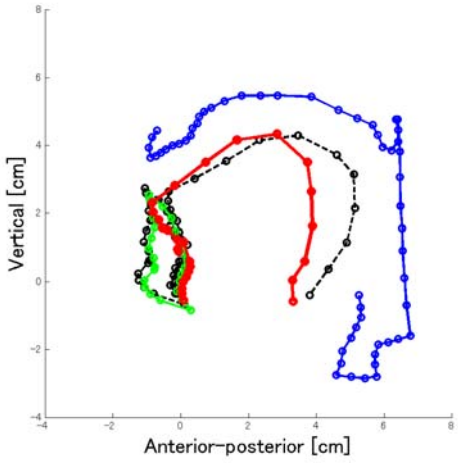
(上) プロトタイプモデル、(下) 話者 A

A.3　個人化の結果：話者 B

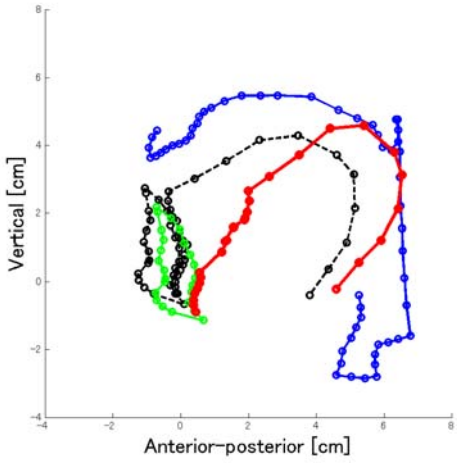
A.3.1　個人化モデル



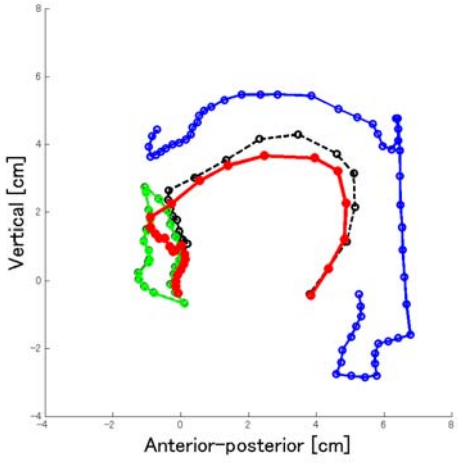
A.3.2　各筋を駆動させた際の舌形状



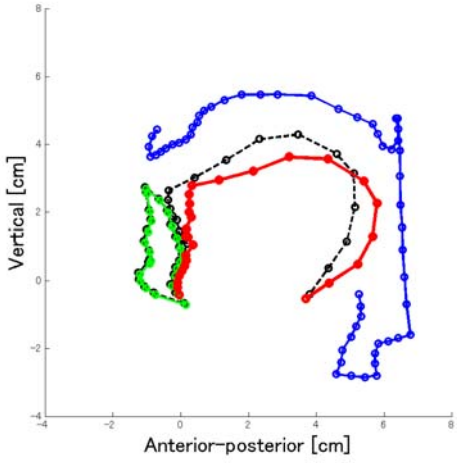
GGp



SG

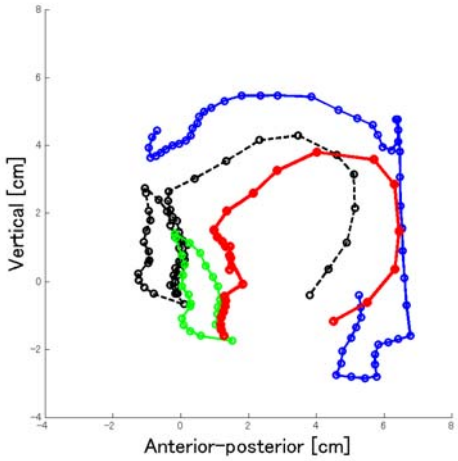


GGa

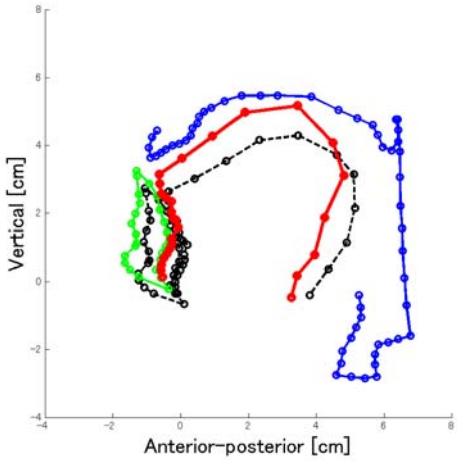


HG

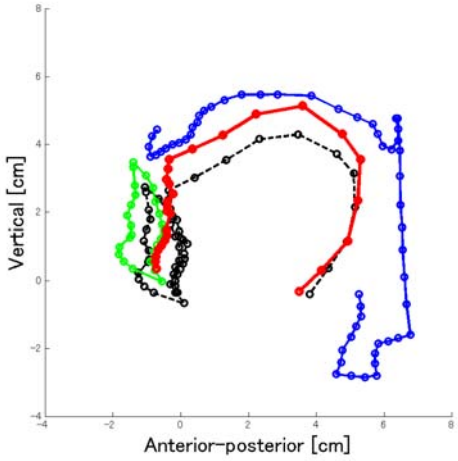
A.3.3　個人化モデルによる母音生成時の舌形状



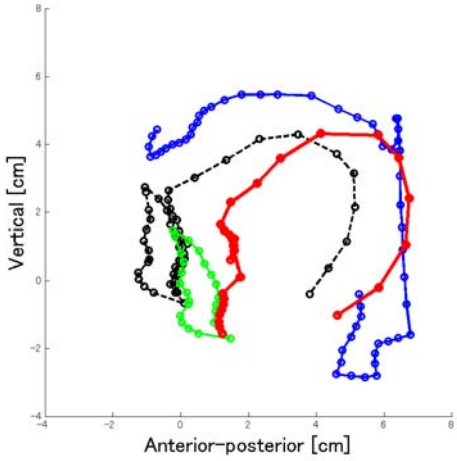
母音 [a]



母音 [i]



母音 [u]



母音 [o]

参考文献

[1] 北村達也, “音声における個人性の知覚と生成について,” Mem. Konan Univ. Intelli. and Ser., Vol.1 No.2, pp. 141–155, 2008.

[2] 北村達也, 正木信夫, “MRI 観測を基礎にした音声生成系研究の進展,” 日本音響学会誌, 62 巻 5 号, pp. 385–390, 2006.

[3] 党建武, 本多清志, “舌の三次元的変形と筋電図との関連について,” 日本音響学会講演論文集, pp. 241–242, 1997.

[4] P. Mermelstein, “Articulatory model for the study of speech production,” J. Acoust. Soc. Am. 53, pp. 1070–1082, 1973.

[5] Y. Payan and P. Perrier, “Synthesis of V-V sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis,” Speech Commun. 22(2-3), pp. 185–205, 1997.

[6] V. Sanguineti, R. Laboissiere, D. Ostry, “A dynamic biomechanical model for neural control of speech production,” J. Acoust. Soc. Am. 103(3), pp. 1615–1627, 1998.

[7] P. Badin, A. Serrurier, “Three-dimensional linear modeling of tongue: Articulatory data and models,” 7th literational Seminar on Speech Production, pp. 395–402, 2006.

[8] A. Serrurier, P. Badin, “A three-dimensional articulatory model of the velum and nasopharyngeal wall based on MRI and CT data,” J. Acoust. Soc. Am. 123(4), pp. 2335–2355, 2008.

[9] F. Guenther, S. Ghosh, J. Tourville, “Neural modeling and imaging of the cortical interactions underlying syllable production,” Brain and Language 96, pp. 280–301, 2006.

[10] B. Kroger, J. Kannampuzha, C. Neuschaefer-Rube, “Towards a neurocomputational model of speech production and perception,” Speech Communication 51, pp. 792–809, 2009.

[11] J. Dang, K. Honda, “A physiological articulatory model for simulating speech production process,” Acoust. Sci. and Tech. 22, 6, pp. 415–425, 2001.

[12] J. Dang and K. Honda, “Construction and control of a physiological articulatory model,” J. Acoust. Soc. Am. 115(2), pp. 853–870, 2004.

[13] X. Wu, Q. Fang, J. dang, “Investigation of Muscle Activation in Speech Production Based on an Articulatory Model,” Proc. ISCSLP. pp. 330–334, 2010.

[14] 錦戸信和, 党建武, “発話機構モデルに基づく音声と調音状態との一対多の関係に関する考察,” 日本音響学会誌 67(1), pp. 3–14, 2011.

[15] 北村達也, 鈴木規子, 齋藤浩人, 道健一, 高橋俊行, 赤木正人, 和久本雅彦, “MRI による舌・口底切除患者の 3 次元声道形状の分析,” 電子情報通信学会技術研究報告, SP22000-153, pp. 1–6, 2001.

[16] 齋藤浩人, 鈴木規子, 北村達也, 赤木正人, 道健一, “舌・口底切における異常構音の音響的特徴-スペクトルのピーク分析の試み-,” 電子情報通信学会技術研究報告, SP98-149, pp. 25–31, 1999.

[17] S. Fujita, J. Dang, N. Suzuki, K. Honda, “A Computational Tongue Model and its Clinical Application,” Oral Science Internationa, Vol.4 No.2, pp. 97–109, 2007.

[18] 加藤理恵, 田中誠也, 高見観, 杉山裕美, 北村洋子, 南克浩, 古川博雄, 辰巳寛, 山本正彦, “構音障害に対する治療効果の音響学的考察,” 心理科学, 第 2 巻第 1 号, pp. 25–36, 2010.

[19] 本多清志, “音声生成の生理機構と発話障害,” 認知神経心理学研究会, S-3, 2005.

[20] H. Takemoto, T. Kitamura, H. Nishimoto, K. Honda, “A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions,” Acoust. Sci. and Tech. 25(6), pp. 468–474, 2004.

[21] G. Wang, T. Kitamura, X. Lu, J. Dang, J. Kong, “MRI-based Study of Morphological and Acoustical Properties of Mandarin Sustained Steady Vowel,” J. Signal Processing, Vol.12 No.4, pp. 311–314, 2008.

[22] E. Bresch, S. Narayanan, “Region Segmentation in the Frequency Domain Applied to Upper Airway Real-Time Magnetic Resonance Images,” IEEE Transactions on Medical Imaging, Vol.28 No.3, pp. 323–338, 2009.

[23] 伊能教夫, 小林弘樹, 槇宏太郎, “X 線 CT データに基づく下顎骨の個体別モデリング手法,” 日本機械学会論文集 (C 編), 60 巻 574 号, pp. 188–193, 1994.

[24] 伊能教夫, 鈴木知, 槇宏太郎, 宇治橋貞幸, “X 線 CT データに基づく骨体の自動モデリング手法 (デラウニー分割を利用した有限要素モデルの生成),” 日本機械学会論文集 (C 編), 68 巻 669 号, pp. 139–144, 2002.

[25] 青木義満, 宮田宗樹, 寺嶋雅彦, 中島昭彦, “頭部 X 線規格画像を用いた顎顔面骨格 3 次元形状可視化システムの構築,” Medical Imaging Technology, vol.22 No.5, pp. 250–258, 2004.

[26] 祖川慎治, 四倉達夫, 森島繁生, “影響力マップ用いた任意表情モデル上での表情合成,” 電子情報通信学会技術研究報告, 102 巻 598 号, pp. 19–24, 2002.

[27] 山中健太郎, 中村槙介, 小林昭太, 森島繁生, “MRI を用いた前腕皮膚形状変化モデルの構築と運動生成,” 電子情報通信学会技術研究報告, 109 巻 29 号, pp. 85–90, 2009.

[28] N. Arad and D. Reisfeld, “Image Warping Using few Anchor Points and Radial Functions,” Computer Graphics Forum. 14(1), pp. 35–56, 1995.

[29] H. Takemoto, K. Honda, S. Masaki, Y. Shimada, I. Fujimoto, “Measurement of temporal changes in vocal tract area function from 3D cine-MRI data,” Jurnal of the Acoustical Society of America 119(2), pp. 1037–1049, 2006.

[30] Q. Fang, S. Fujita, X. Lu, J. Dang, “A model-based investigation of activations of the tongue muscles in vowel production,” Acoust. Sci. and Tech. 30(4), pp. 277–287, 2009.

[31] 高野佐代子, 本多清志, “磁気共鳴画像法に基づく母音発生時の舌筋の筋長計測,” 音楽情報科学, 39-21, pp. 145–152, 2001.

[32] M. Stone, E. Davis, A. Douglas, M. NessaAiver, R. Gullapalli, W. Levine, A. Lundberg, “Modeling the motion of the internal tongue from tagged cine-MRI images,” J. Acoust. Soc. Am. 109(6), pp. 2974–2982, 2001.

[33] J. Dang, K. Honda, “INVESTIGATION OF THE ACOUSTIC CHARACTERISTICS OF THE VELUM FOR VOWELS,” Proc. ICSLP, pp. 1–4, 1994.

[34] 党建武, 本多清志, ‘母音発声時の音声スペクトルに対する梨状窩の影響,” 電子情報通信学会技術研究報告, SP95-10, pp. 1–6, 1995.

[35] 鏑木時彦, 正木信夫, 元木邦俊, 松崎博季, 北村達也, “音声生成の計算モデルと可視化,” コロナ社, 2010.

謝辞

本研究を進めるにあたり、指導して頂いた党建武教授、全面的にサポートして頂いた川本真一助教、ならびに有意義な討論・助言をして頂いた党研究室、赤木・鵜木研究室をはじめとする皆様に心から感謝致します。

学会発表リスト

1. N. Nishimura, S. Kawamoto, J. Dang, "Morphological personalization according to human mechanism using MR images," 2012 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP'12), pp. 497–498, 2012.3.
2. 西村奈々, 川本真一, 党建武, "形態学的情報に基づく個人化発話機構モデルの構築," 日本音響学会 2012 年春季研究発表会, pp. 451–454, 2012.3.