

研究概要報告書

資料 - 1

情

( / )

研究題名	磁気共鳴映像法(MRI)による音声生成機構の解析と個人性情報を含む自然音声の合成	報告書作成者	松村雅史
研究従事者	松村雅史(大阪電気通信大学)		
研究目的	<p>本研究では、個人性情報がどのようにして音声に含まれているかを音声生成機構の解析に基づき明らかにし、個人性情報を含む自然音声を合成することを目的とする。すなわち、磁気共鳴映像法(MRI)を用いて声帯、鼻腔を含む声道の3次元形状を計測し、発声器官の機能および声道形状が個人によりどう異なるかを明らかにする。また、3次元声道形状データより正中面舌形状のモデル、声道断面形状のモデルを構築し、前舌部の2点の位置情報より声道面積関数を推定する3次元声道形状モデルを提案する。この3次元声道形状データに基づいた精密な音声生成モデルにより、個人性情報を含む自然音声の合成を行い、合成音と実音声のスペクトルを比較、評価する。</p>		
研究内容	<p>(1) 磁気共鳴映像法(MRI)による3次元声道および鼻腔形状の計測： 声道形状の計測には、超伝導MRIシステムSMT-100(島津製作所、静磁場1.0T)を用いる。正中面の撮像には、スピンエコー法のフリップ角を変化させた高速撮像法(FBSE法)を用いる。声道の3次元形状を得るために、声帯から5mm間隔で鼻腔まで24スライスの水平面MR画像を収集する。声帯や口蓋位置の校正を行うために正中面および冠状面MR画像も採取する。さらに、鼻腔の3次元形状を得るために鼻尖から環椎まで5mm間隔で20スライスの冠状面MR画像を採取する。</p> <p>成人男子10名、検査語5母音で各2サンプルの正中面、水平面、冠状面MR画像を採取した。撮像時には高周波磁場中でコイルの振動音が発生するため、音声とMR画像との同時計測は困難である。そこで、被験者には2秒程度母音発声を行った後、数秒間隔で断続的に発声を繰り返してその調音状態を維持するように指示し、その間に撮像を行った。</p> <p>MR画像から声道輪郭を抽出するアルゴリズムを図1に示す。本アルゴリズムは適応しきい値法に基づいた方法である。MR画像から声道の輪郭を抽出した結果を図2に示す。同図(a)は声門付近、(b)は声門から4.8cm上方の水平面である。図3は鼻尖より3.5cm後方の冠状断面であり、上鼻甲介、中鼻甲介の形状が観測でき、鼻腔が中鼻道、下鼻道に分岐しているようすがわかる。また、この被験者は鼻風邪をひいており、右側の鼻腔の一部が閉鎖されている様子が観測できた。この結果より声道輪郭抽出アルゴリズムの有効性が示された。</p> <p>(2) 声道および鼻腔形状の解析： 正中面の舌形状が発声母音によりどう変形するかを調べた。舌形状を特徴を抽出するために喉頭蓋基部から</p>		

3

## 説 明 書

( / )

舌先端まで舌表面の形状の曲率関数を求めた。その結果、舌形状の曲率関数は最大点が音韻により変化する上に凸の関数で表されることが示された。また、調音位置（声道が最も狭くなる位置）で舌形状の曲率が最大となることが明らかとなった。

軟口蓋の閉鎖について調べた結果、母音発声においても軟口蓋により鼻咽喉が完全に閉鎖されない母音の鼻音化が認められた。鼻腔の影響により高次ホルマント周波数が増えるため、鼻腔の影響を考慮しない従来の母音生成モデルを用いて個人性情報を含む音声を作成するには無理があることが示唆された。

主鼻道の形状を調べた結果を図4に示す。鼻尖からの距離Lが60～65mmで鼻腔が左右に分岐している様子がわかる。また、鼻風邪により右側の鼻腔が閉鎖されていることも観測できた。鼻腔は主鼻道以外に副鼻腔があるが、今回の観測では、主鼻道と副鼻腔とのつながりを十分に把握するには至らなかった。鼻腔の音響特性を解析するためには必要なデータであるので今後とも研究を継続する予定である。

(3) 精密な音声生成モデルの構築と個人性情報を含む自然音声の合成： 計測結果に基づき、舌形状、鼻腔形状および声道断面形状のモデル化を行い、精密な音声生成モデルを構築する。このモデルパラメータを協調問題解決により制御し、個人性情報を含む自然音声の合成を行う。MRIによる声道形状データに基づいて合成した音声のスペクトルと実音声のスペクトルを比較した結果を図5に示す。第3ホルマント周波数までよく一致していることがわかる。また、モデルパラメータを協調問題解決により制御し、連続音声を作成した。その結果、被験者が発声した連続音声のホルマントパターンとよく一致し、3次元声道モデルに基づく音声合成法の有効性を確認した。

### (4) まとめ

現在の音声研究では言語情報を対象とした研究に比べ、音声の個人性情報に関する研究は遅れている。この原因としては個人性情報がどのようにして音声中に含まれているかが解明されていないこと、音声分析の基礎となる音声生成モデルが個人性情報を付加するにはあまりにも単純すぎたことなどが考えられる。また、様々な調音機構の観測方法が提案されているが、声帯、鼻腔を含む声道の3次元形状を計測した例は少ない。本研究では、声帯、鼻腔を含む声道の3次元形状の計測に基づき、新しい音声生成モデルを提案した。このモデルパラメータを協調問題解決により制御し、個人性情報を含む自然音声の合成を行う手法を開発した。鼻腔形状のより正確なデータとそのモデル化、鼻腔における共鳴のメカニズムの解析、調音方法における個人差など個人性情報には様々な要因があり今後とも生成メカニズムの解析に基づいた研究を推進するつもりである。

(注： フローチャート図、ブロック図、構成図、写真、データ表、グラフ等 研究内容の補足説明に御使用下さい)

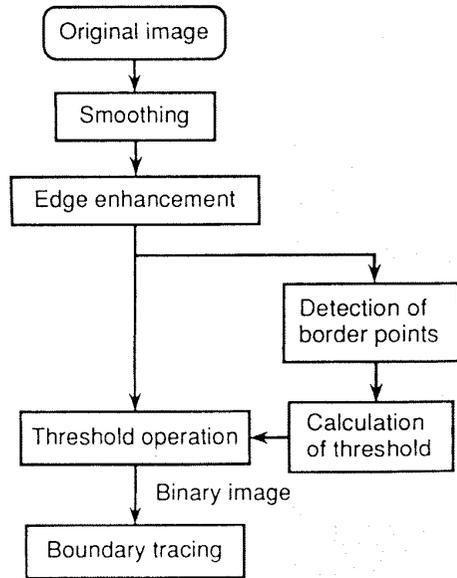


図 1 声道輪郭抽出の処理手順

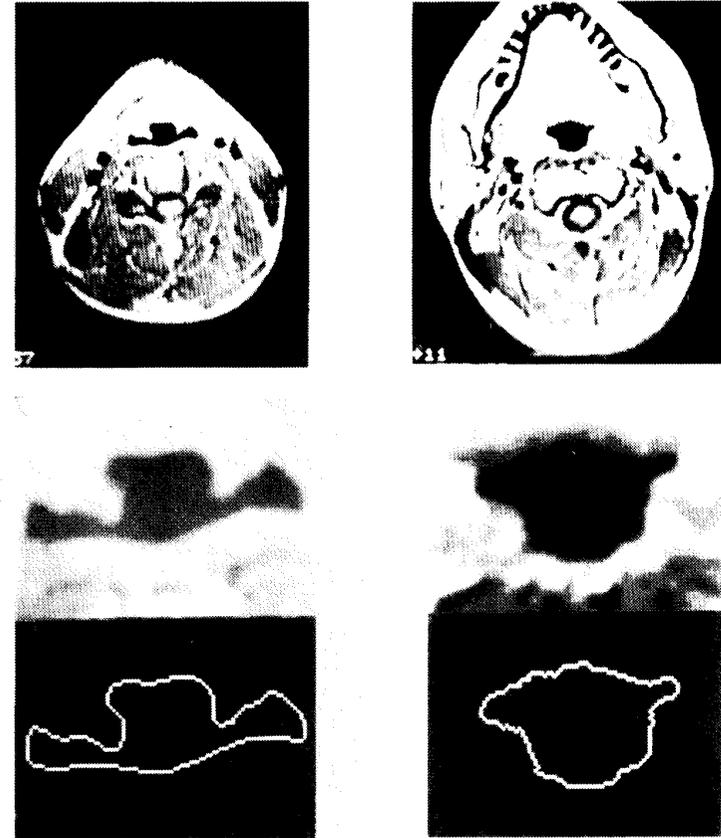
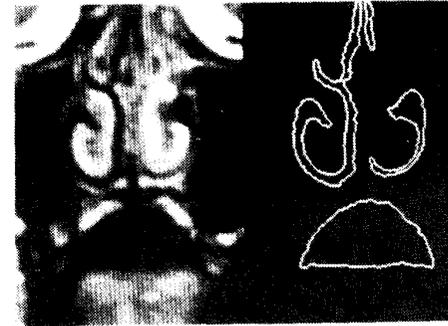
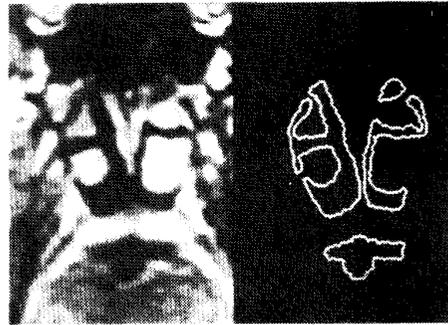
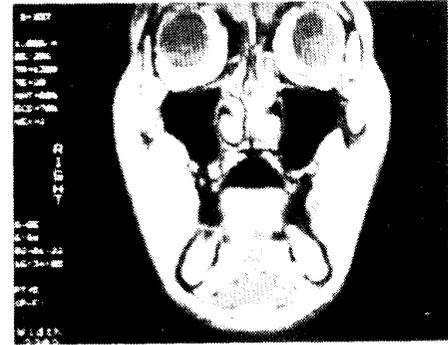
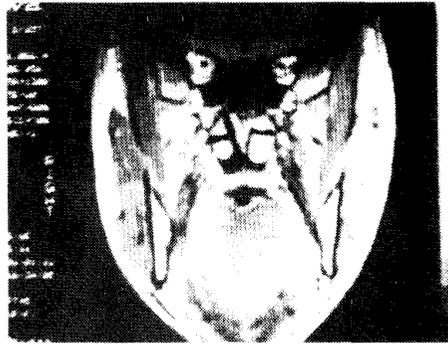


図 2 水平面MR画像と声道輪郭抽出結果、  
(a) 声門付近、(b) 声門より4.8cm上方

(注： フローチャート図，ブロック図，構成図，写真，データ表，グラフ等 研究内容の補足説明に御使用下さい)



(a)

(b)

図3 冠状面MR画像と声道および鼻腔の輪郭抽出結果  
(a) 鼻尖から6.0cm、(b) 鼻尖から3.5cm

(注： フローチャート図，ブロック図，構成図，写真，データ表，グラフ等 研究内容の補足説明に御使用下さい)

様式-10

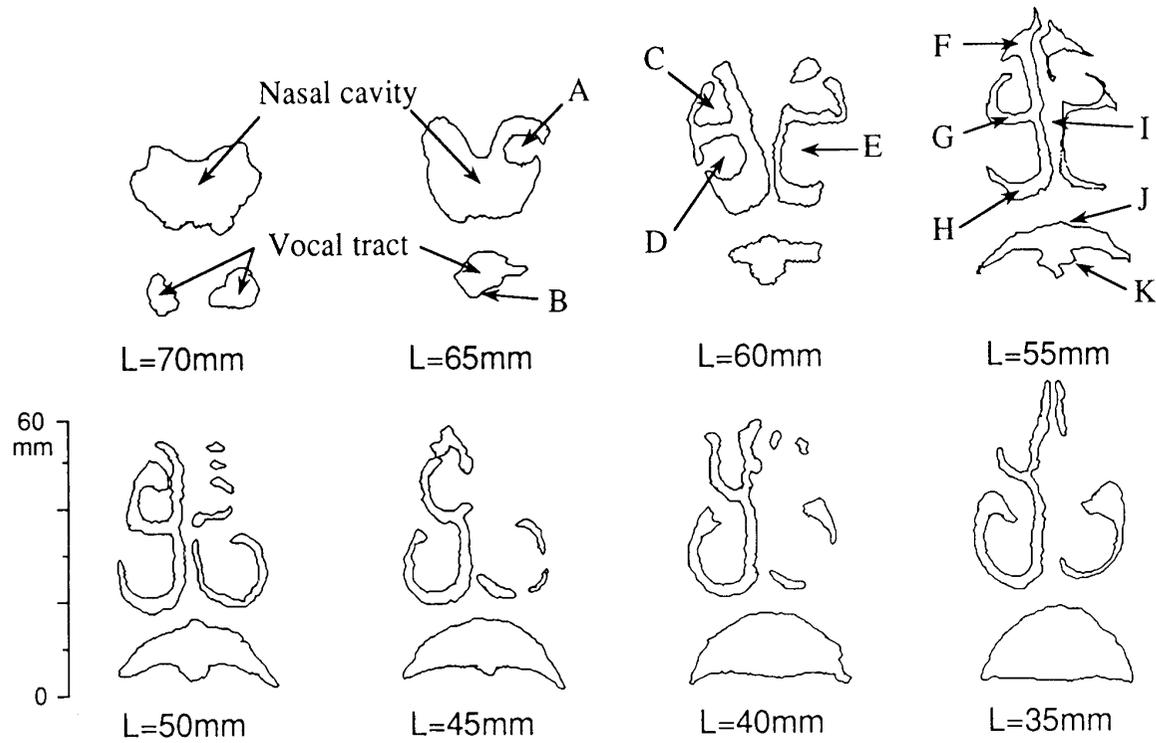
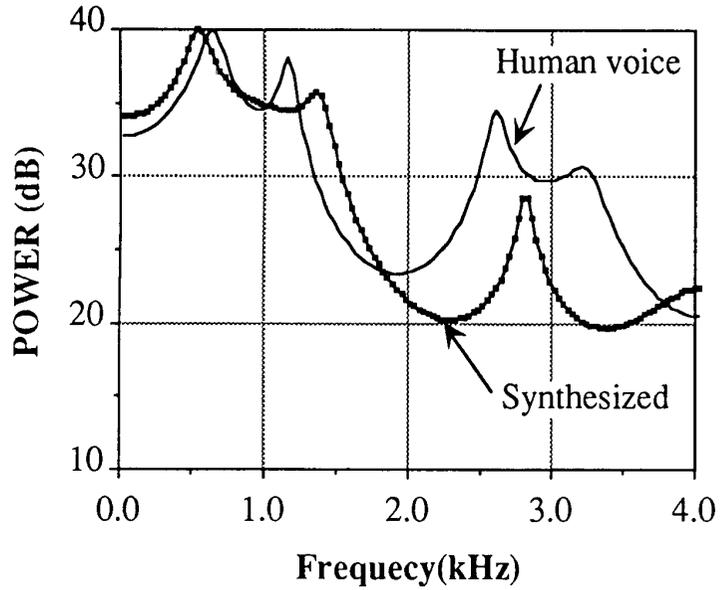


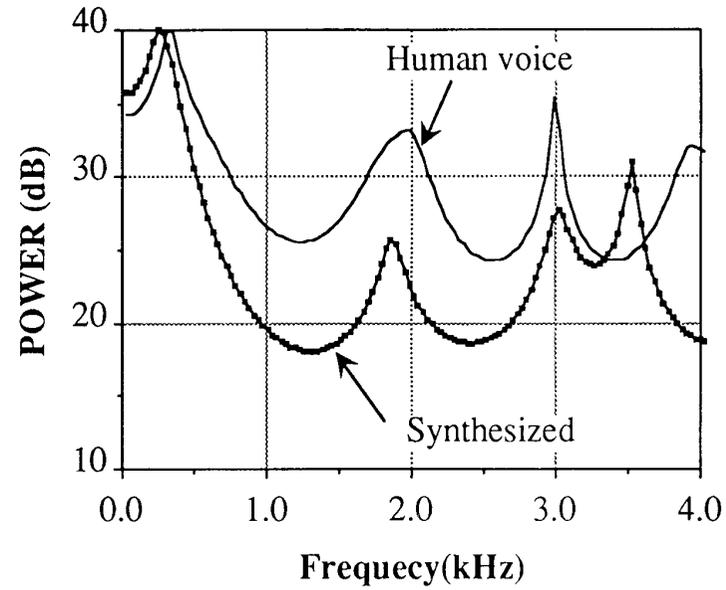
図4 鼻腔の輪郭抽出結果、L：鼻尖からの距離

(注： フローチャート図，ブロック図，構成図，写真，データ表，グラフ等 研究内容の補足説明に御使用下さい)

様式-10



(a) Power spectrum for vowel /a/



(b) Power spectrum for vowel /i/

図5 実音声と合成音のスペクトルの比較

(注： フローチャート図，ブロック図，構成図，写真，データ表，グラフ等 研究内容の補足説明に御使用下さい)