

Title	口腔疾患を有する複雑な声道形状と音声スペクトルの関係に関する基礎的研究
Author(s)	西本, 博則
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1454
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

修士論文

口腔疾患を有する複雑な声道形状と 音声スペクトルの関係に関する基礎的研究

指導教官 赤木 正人 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

西本 博則

2001年3月

要 旨

本研究は口腔疾患を有する患者に見られるような健常者とは異なる特徴を有する複雑な声道形状の構造的特徴と、音声スペクトルなどの音声の音響的特徴の関係を導くための基礎的調査を行うことを目的としている。本研究の特色としては、1次元近似モデルでは十分な近似ができないような口腔疾患患者の声道構造が、音声スペクトルなどの音声中の音響的特徴に及ぼす影響を調査することが挙げられ、そのために MR 画像から 3次元声道モデルを構築し有限要素法 (FEM) を用いた分析を行う。

最初に、健常例で FEM によるシミュレーション結果と実音声の分析結果を比較し、FEM によるシミュレーション法の妥当性を検証した。また、声道の一部に変形を与えることによる声道伝達特性の変化を調査した。その結果、シミュレーション結果の妥当性が示された。また、梨状窩に関する影響を考慮すればさらに実音声との一致が言えるようになる。

次に、口腔疾患を有する患者の 3次元声道モデルの FEM によるシミュレーションを行った。その結果、3次元声道モデルの口唇部に放射空間を取り付けることによってシミュレーション結果と音響分析結果との整合性がとれ、放射空間を考慮した声道モデルでのシミュレーションを行う必要性があることが分かった。また、1次元近似モデルの分析結果と比較することで、本手法の優位性が示された。口腔疾患を有する声道モデルに流路をずらすような変形を与えることによる影響が分かり、PAP 設計の一つの指針となる可能性が示された。

目次

1	緒論	1
1.1	本研究の背景	1
1.1.1	口腔疾患による音声スペクトルの歪み	1
1.1.2	個人性による音声スペクトルの違い	1
1.1.3	声道形状と音声スペクトルの関係	2
1.1.4	有限要素法による分析	2
1.1.5	核磁気共鳴撮像法	3
1.2	本研究の目的	3
1.3	本研究の特色	3
1.4	本論文の構成	4
2	研究手法	5
2.1	MR 撮像方法	6
2.1.1	被験者	7
2.1.2	撮像条件	7
2.1.3	被検音	8
2.2	有限要素法	9
2.2.1	波動方程式	9
2.2.2	連立方程式の解法	11
2.2.3	粒子速度	11
2.3	1次元近似モデル	12
3	音響分析	14
3.1	フォルマント周波数について	14
3.2	分析法	14

3.3	録音条件	15
3.4	分析結果	15
4	有限要素法の計算精度	26
4.1	はじめに	26
4.2	3次元声道形状モデルの構築	26
4.3	シミュレーション条件	26
4.4	シミュレーション結果	27
4.4.1	点間隔 1.0[mm] のモデル (Normal-A)	27
4.4.2	口腔部の点間隔 0.5[mm] のモデル (Normal-A)	31
4.4.3	放射球面付き声道モデル (Normal-A)	34
4.4.4	点間隔 0.5[mm] のモデル (Normal-B)	37
4.5	声道後室がフォルマントに与える影響	40
4.5.1	仮声帯を音源としたモデル	40
4.5.2	梨状窩を除去したモデル	43
4.6	1次元モデルとの比較	46
4.7	まとめ	47
5	舌・口底切除症例の3次元声道形状モデルによるシミュレーション	48
5.1	はじめに	48
5.2	Patient-A によるシミュレーション	48
5.2.1	PAP 非装着時のモデル	48
5.2.2	PAP 装着時のモデル	53
5.3	Patient-B のモデルによるシミュレーション	57
5.3.1	1次元近似モデルとの比較	60
5.4	口腔内の声道形状を変化させたモデルによるシミュレーション	61
5.5	まとめ	64
6	結論	65
6.1	本研究の成果	65
6.2	今後の課題	66

目 次

2.1	A flow chart of a simulation.	5
2.2	An MR image(mid-saggital image).	6
2.3	Surgical procedure of Patient-A.	7
2.4	Surgical procedure of Patient-B.	8
3.1	Power Spectrum (Normal-A).	18
3.2	Power Spectrum (Normal-B).	19
3.3	Power Spectrum (Patient-A (PAP (-))) (1).	20
3.4	Power Spectrum (Patient-A (PAP (-))) (2).	21
3.5	Power Spectrum (Patient-A (PAP (+))) (1).	22
3.6	Power Spectrum (Patient-A (PAP (+))) (2).	23
3.7	Power Spectrum (Patient-B) (1).	24
3.8	Power Spectrum (Patient-B) (2).	25
4.1	A Normal-A vocal tract model uttering Japanese vowel /i/. This model is constructed by a vocal tract data (1.0[mm] sampling).	29
4.2	A transfer function of a Normal-A vocal tract model by an 1.0[mm] sampling vocal tract data.	30
4.3	A Normal-A vocal tract model uttering Japanese vowel /i/ (0.5[mm] sampling). This model is constructed from a vocal tract data (0.5[mm] sampling).	32
4.4	A transfer function of a Normal-A vocal tract model by an 0.5[mm] sampling vocal tract data.	33
4.5	A Normal-A vocal tract model with hemispherical surface of radiation.	35
4.6	A transfer function of a Normal-A vocal tract model with hemispherical surface of radiation.	36

4.7	A Normal-B vocal tract model uttering Japanese vowel /i/. This model is constructed by a vocal tract data (1.0[mm] sampling).	38
4.8	A transfer function of a Normal-B vocal tract model by an 1.0[mm] sampling vocal tract data.	39
4.9	Glottis.	40
4.10	A Normal-A vocal tract model uttering Japanese vowel /i/. This model is cut at ventricular fold.	41
4.11	A transfer function of a Normal-A vocal tract excited at the ventricular fold.	42
4.12	A Normal-A vocal tract model removed pyriform fossas.	44
4.13	A transfer function of a Normal-A vocal tract removed pyriform fossas.	45
4.14	Transfer functions of Normal-A vocal tract estimated by FEM and 1-D model.	46
5.1	A 3-D vocal tract model of Patient-A(PAP(-)).	50
5.2	A 3-D vocal tract model of Patient-A(PAP(-)) with with hemispherical surface of radiation.	51
5.3	A transfer function of the Patient-A(PAP(-)) vocal tract model with hemispherical surface of radiation and without hemispherical surface of radiation.	52
5.4	A 3-D vocal tract model of Patient-A(PAP(+)).	54
5.5	A 3-D vocal tract model of Patient-A(PAP(+)) with with hemispherical surface of radiation.	55
5.6	A transfer function of Patient-A(PAP(+)) vocal tract model with hemispherical surface of radiation and without hemispherical surface of radiation.	56
5.7	A 3-D vocal tract model of Patient-B.	58
5.8	A transfer function of a Patient-B vocal tract model with hemispherical surface of radiation and without hemispherical surface of radiation.	59
5.9	Transfer functions of a Patient-A(PAP(-)) vocal tract model simulated by FEM and 1-D model.	60
5.10	A modified 3-D vocal tract model of Patient-A (PAP(-)).	62
5.11	A transfer function of a modified vocal tract model of Patient-A (PAP(-)).	63

表 目 次

2.1	Subjects	7
2.2	MR image acquisition	9
2.3	Variables used in 1-D model	13
3.1	Acoustic analysis.	15
3.2	Formant frequencies (Normal).	16
3.3	Formant frequencies (Patients).	17
4.1	Peak frequencies of estimated vocal tract function and formant frequencies.	28
4.2	Peak frequencies of estimated vocal tract function and formant frequencies.	31
4.3	Peak frequencies of estimated vocal tract functions.	34
4.4	Peak frequencies of estimated vocal tract function and formant frequencies.	35
4.5	Peak frequencies of estimated vocal tract function and formant frequencies.	37
4.6	Peak frequencies of estimated vocal tract function and formant frequencies.	43
5.1	Peak frequencies of Patient-A (PAP(-)) vocal tract function without hemispherical surface of radiation and formant frequencies.	49
5.2	Peak frequencies of Patient-A (PAP(-)) vocal tract function with hemispherical surface of radiation and formant frequencies.	49
5.3	Peak frequencies of Patient-A(PAP(+)) vocal tract function without hemispherical surface of radiation and formant frequencies.	53
5.4	Peak frequencies of Patient-A(PAP(+)) vocal tract function with hemispherical surface of radiation and formant frequencies.	54
5.5	Peak frequencies of Patient-B vocal tract function with hemispherical surface of radiation and formant frequencies.	57

第 1 章

緒論

1.1 本研究の背景

1.1.1 口腔疾患による音声スペクトルの歪み

口腔領域の疾患，あるいはその手術の後遺症によって，しばしば言語障害が生じる．特に口腔癌術後症例では，構音障害により音声スペクトルに独特の歪みが生じる．これは，舌・口底の欠損と運動障害により，声道形状が健常者と比較して著しく異なるためである．

口腔疾患の各症例を比較した場合には，声道形状には一定した傾向が見られず，症例毎の違いが大きい．そのため，言語障害の歪音の診断や，その治療に使われている舌接触補助床 (Palatal Augmentation Prosthesis 以下 PAP) などの人工補綴物の設計には客観的指標が無く，医療従事者の聴覚印象に頼っているのが現状である．そこで，言語障害の診断を正確に行うため，そして PAP の設計等による治療のために，口腔疾患を有する声道構造と音声スペクトルを中心とする音響的特徴の関係について導き出し，舌・口底切除症例の言語障害の原因を解明することが重要となってきた．

1.1.2 個人性による音声スペクトルの違い

同じ音声を発話する場合において，各個人により音声スペクトルはそれぞれ異なった特徴を有する．この音声スペクトルの特徴は，各個人の声道形状，構音の違いにより生じる声道の構造的特徴に起因すると考えられている．そのため，音声の個人性の研究や，自然な音声合成のために，声道形状と音声スペクトルの関係を調査することが重要になってきている．

1.1.3 声道形状と音声スペクトルの関係

声道形状は各構音器官による狭窄や閉鎖により変化することで共鳴の特性が変化し、それが音声スペクトルなどの音響的特徴となって現れる。そのため、声道形状と音声スペクトルには密接な関係がある。ところが、声道形状から音声スペクトルは一意に決定できるが、逆に音声スペクトルから声道形状を推定しようとする場合、一対多の不良設定性逆問題であり、声道形状を一意に決定できないということが知られている。口腔疾患の患者の言語障害の診断や治療のため、また、自然な音声合成のために、声道形状の構造的特徴が音声スペクトルへ与える影響を調査し、声道形状と音声スペクトルの関係を導くことが重要となっている。

1.1.4 有限要素法による分析

これまでの声道形状と音声の特徴に関する研究では、声道形状の断面積を用いて、異なる断面積の円筒音響管を接続したモデルで近似した1次元解析が行われてきた。しかし、この近似法では声道内を伝搬する音波は平面波として扱われてしまうため、それ以外の場合ではうまく近似できないという問題がある。また、実際の声道は途中でおよそ90°曲がっていたり、喉頭蓋や梨状窩などにより分岐したりと、非常に複雑な形状である。そのため、実際の声道内では複雑な伝搬をしていると考えられるが、この近似法ではこの声道形状の複雑さが考慮されていない。さらに、舌・口底切除患者の声道は、声道が中心軸に対してずれており、健常者と比較しても異常な形状をしているが、断面積関数による近似ではこの3次元形状による影響が十分に考慮されない。

これに対して、近年の計算機能力の向上により、数値解析法の一つである3次元モデルを用いた有限要素法 (Finite Element Method 以下 FEM) による解析方法が用いられることが多くなった。この解析方法は声道形状の複雑さを吸収できる解析方法であり、3次元の声道形状の特徴と、音響的な特徴の関係を調査する上では、最も適切な方法であるとされている。有限要素法を用いた声道とその音響的特徴に関する研究としては Lu ら [1]、松崎ら [2][3]、Niikawa ら [4] などがあり、3次元の形状の特徴と音響的な特徴の関係が調査されている。しかし、これらの研究で用いられている3次元声道形状モデルは健常者の声道形状モデルを対象としており、口腔疾患を有する声道形状のような、複雑な3次元声道形状モデルでの分析は行われていない。一方、声道の中心軸に対して非対称な声道管モデルによる研究としては富山の研究報告 [5] がある。ここでは、音響管の軸をずらすことによる声道伝達特性の変化が示されているが、使用した声道モデルは口腔疾患患者の声道断面積から作成された音響管を用いており、声道の複雑な構造を表現したモデルであ

るとは言えない．そのため，口腔疾患患者の声道形状と音響的特徴の関係を調べる上で，口腔疾患患者の複雑な声道形状を再現したモデルによる解析が必要である．

1.1.5 核磁気共鳴撮像法

FEM による分析を行う際に，3次元声道モデルの構築が重要となってくる．3次元声道モデルの構築法の一つに，核磁気共鳴撮像法 (Magnetic Resonance Imaging 以下 MR 撮像法) による MR 画像から声道形状情報の抽出を行い，それをもとに3次元声道モデルを計算機上に構築するという手法がある．MR 撮像法は従来の X 線を応用した撮像法とは異なり，人体に対して非侵襲性であり，得られる MR 画像はスライス画像のデータであるという特徴がある．MRI の計測精度は齋藤ら [6] で検討されており，MR 観測結果から3次元声道モデルの構築が可能であり，十分に声道形状の複雑さを再現できる．

1.2 本研究の目的

本研究では，口腔疾患患者に見られるような健常者とは異なる特徴を有する複雑な声道形状の構造的特徴と，音声スペクトルなどの音声の音響的特徴の関係を導くための基礎的調査を行うことを目的とする．そのために，まず MR 画像より得られる精度の高い声道形状情報を利用して，計算機上に3次元モデルを構築し，FEM によるシミュレーションを行う．そのシミュレーション結果と実音声の分析結果を比較し，シミュレーションの妥当性の検討を行う．そして，さらに多くの症例の3次元声道モデルのシミュレーションや3次元声道形状モデルに任意の変形を与えたモデルのシミュレーションを行うことで，声道形状の構造的特徴と音声の音響的特徴の関係について導く．声道形状から音声スペクトルを推定する順問題を多数解くことにより，逆問題を解く指針を得ることで，音声スペクトルと声道形状の関係を導く．

1.3 本研究の特色

これまでの声道形状と音声の特徴に関する研究として，3次元声道モデルを用いた FEM による解析が行われてきた．しかし，過去の FEM による解析で用いられている3次元声道モデルの多くは，声道断面を円や楕円等で近似した，声道の軸に対して左右対称な断面形状をもつ声道モデルや，健常者を対象にした声道モデルが多く，本研究で目的としている口腔疾患患者のような異常な形状を有する声道モデルによる解析はほとんど行われて

いない。

本研究の特色としては、第一に、1次元近似モデルでは十分な近似ができないような口腔疾患患者の声道構造が、音声スペクトルなどの音声中の音響的特徴に及ぼす影響を調査することが挙げられる。そのために口腔疾患患者の3次元声道形状モデルでのFEMを用いた分析を行う。第二に、3次元声道モデルに対して口腔疾患を有する声道や、PAPを装着した声道のような形状になるように任意に変形を与えることで、言語障害の診断やPAPの設計への応用に貢献できるということが挙げられる。第三に、以上の手法により、声道形状から音声スペクトルを推定する順問題を統計的に解くことにより、逆問題を解く指針を得ることで、音声スペクトルと声道形状の関係を導くことが挙げられる。

1.4 本論文の構成

第1章では本研究の背景、目的、特色について述べる。第2章では本研究で用いるシミュレーション法について述べ、その妥当性や有用性を検討するための各種分析法について述べる。第3章では音響分析法について解説し、口腔疾患患者の音声の特徴について述べる。第4章では実際にシミュレーションを行い、その結果を示し、実音声の分析結果との比較や従来手法による分析結果との比較を行うことで本シミュレーション法による分析結果の妥当性、有用性について議論する。第5章では口腔疾患を有する患者の3次元声道形状モデルに変形を与えてシミュレーションを行い、PAPの有無による影響や声道形状の複雑さが音声スペクトルに与える影響について議論する。第6章では第5章までのまとめをし、得られた知見、成果について述べる。

第 2 章

研究手法

本研究では 3 次元声道モデルの FEM による声道伝達特性の計算を行うことで、口腔疾患を有する声道形状とその音響的特徴の関係を導くことを目的としている。本研究の手法のフローチャートを図 2.1 に示す。まず、昭和大学病院で撮像された MR 画像から、静岡大学の北村 [7][8] により声道形状情報が抽出される。その声道形状情報をもとにして、3 次元声道モデルを計算機上に構築し、FEM で分析するというシミュレーション法を用いる。以上のシミュレーション法で声道形状から音響的特徴を求める。このシミュレーション法により声道形状から音響的特徴を計算するという順問題を多数扱うことで、不良設定の逆問題である音響的特徴から声道形状を導くための指針を得ることを目標としている。本章では本研究でのシミュレーションに用いる各手法について述べる。また、本シミュレーション法の妥当性を検討するために用いる方法について述べる。

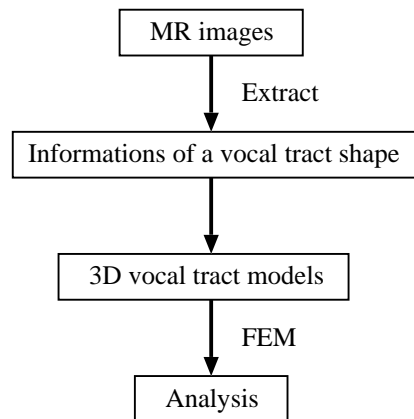


図 2.1: A flow chart of a simulation.

2.1 MR 撮像方法

構音器官の観測，計測の種々の方法のうち，発音時の声道形状を測定する代表的な方法として，X 線を用いた方法がある．この方法は 2 次元的な観察にとどまっており，3 次元的な声道形状，とくに本研究で目的としているような複雑な声道形状を獲得することはできない．一方で 3 次元的に声道形状を測定する方法としては，X 線による撮像法を応用した 3 次元 CT による撮像法，超音波断層法，そして MR 撮像法がある．3 次元 CT による撮像法は解像度は良好であるが，X 線を応用した撮像法であるため，人体に有害であり，多数の画像を得るための長時間の X 線の照射はできない．超音波断層法では，人体に対して非侵襲性があるが，得られる画像の解像度が低く，本研究には向いていない．MR 撮像法は撮像に長時間を要し，大きな騒音を発するため撮像時に音声の同時採取はできないが，非侵襲的であり，画像の解像度が良好で，計測誤差も改善されてきている．また，声道観測データはスライス画像として得られるため，本研究で用いる声道形状情報の抽出，それにより構築される 3 次元声道モデルの精度という点から有効な撮像法である．

健常者の MR 画像の例を図 2.2 に示す．



図 2.2: An MR image(mid-sagittal image).

2.1.1 被験者

撮像の対象となる被験者は次の通りであり，健常者 2 名，患者 2 名の計 4 名である．なお，健常者は矯正治療の既往，歯の欠損及び機能的な構音障害がない，正常咬合であるという基準で 2 名選択した．各被験者の症例等の詳細を表 2.1 に示す．患者の症例は Patient-A は舌・口底切除例，Patient-B は口底切除例である．なお，Patient-A は術後の機能障害に対して PAP が適用されている．

表 2.1: Subjects

Subject	Age	Sex	Case
Normal-A	26	male	Normal
Normal-B	24	male	Normal
Patient-A	37	male	Tongue and mouth floor resection
Patient-B	72	male	Mouth floor resection

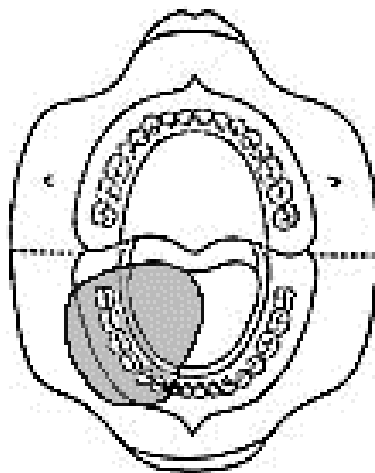


図 2.3: Surgical procedure of Patient-A.

2.1.2 撮像条件

MR 撮像は昭和大学病院のシーメンス社ビジョン Version 31B により行った．この MRI 装置は 1 回約 21[sec] の撮像で 250[mm] × 250[mm] × 80[mm] の領域のデータを得ること

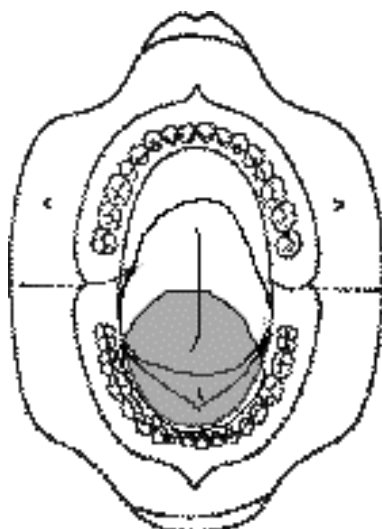


図 2.4: Surgical procedure of Patient-B.

ができる．撮像条件を表 2.2 に示す．

撮像の際，被験者を仰臥位にさせ，CP-NECK-Array コイル（シーメンス社製）を被験者の頭部に装着，固定した．その後，撮像の位置を投光器により定め，位置決めのための仮のスキャンを行った．それにより得られる MR 画像から正中面の位置，及び撮像範囲を指定し，本撮像を行った．この一連の過程は診療放射線技師によって行われた．

歯や PAP は水の陽子が非常に少ないために MR 画像からの形状の認識が困難である．そのため MR 撮像時には和久本らの方法 [9] に従って作成された MRI 用造影剤（マグネヒスト）を封入した厚さ約 $1.0[mm] - 1.5[mm]$ の歯冠プレートを口腔内に装着することで，MR 画像に歯の形状を描出できるようにしている．また，Patient-A は術後の機能障害に対して PAP が適用されており，PAP 装着，非装着の場合でそれぞれ撮像を行った．PAP 装着時の撮像には PAP の外形抽出した PAP 歯冠造影プレートを用い，MR 画像に PAP の形状が写り込むようにした．

2.1.3 被検音

被検音は日本語母音 /i/ とした．日本語母音 /i/ は発話時に狭めを作るための舌の持ち上げが必要である．そのため，舌・口底切除患者の場合は構音が困難であり，音声の歪みが著しくなる音である．実際の撮像においては，舌の初期位置を一定にするために，まず先行母音 /a/ を短く発声し，続けて被検音である後続母音 /i/ を被験者に発声させる．

表 2.2: MR image acquisition

MR device	SIEMENS MAGNETOM VISION Version 31B
Sequence	Fisp3D TR 5msec, TE 2msec FA 15° FOV 250mm
Image size	250mm × 250mm × 80mm (32split)
Image matrix	256 × 256 × 128 pixels (1 pixel = 1.95mm × 1.95mm × 0.98mm)
Required for one slice image	21sec
Number of excitations (NEX)	1 time
Image treatment	SIEMENS MPR(Multipranar Reconstruction)

MR 撮像には 21[s] がかかるため，後続母音 /i/ の発声はできるだけ長くするようにさせ，息が切れた後も発話時と同じ構音を保つようにさせ，息継ぎを行わないようにさせた．

2.2 有限要素法

有限要素法 (FEM) とは，解析の対象となる問題形式や構造物を，領域を要素 (element) と呼ばれる有限の大きさを持つ図形の集まりのモデルで置き換えて微分方程式の近似解を得る，数値計算法の一種である．このモデルの表現として，節点，要素，境界条件を使用する．要素は数値的には節点の各自由度間での相互作用の関係を行列で表現したものであり，形状的には面 (2 次元) や立体 (3 次元) で表現される．

2.2.1 波動方程式

音響問題の速度ポテンシャル ϕ の定常状態における波動方程式は以下に示される．

$$\nabla^2 \phi = -k^2 \phi \quad (2.1)$$

ここで， $k(= \omega/c, \omega: \text{角周波数 } c: \text{音速})$ は波長定数である．速度ポテンシャル ϕ から計算される音圧 p および粒子速度 u は，それぞれ

$$p = j\omega\rho\phi \quad (2.2)$$

$$u = -\nabla\phi \quad (2.3)$$

となる．ここで ρ は大気密度である．以上の式を FEM シミュレーションに適用する．

声道内の音圧分布を求めるためには，この方程式と境界条件からガラーキンの法を用いて，この境界値問題を解く．

$$\int_{V \text{ 内}} (-\omega_i)(\nabla^2 p + k^2 p) dv = 0, \quad i = 1 - n \quad (2.4)$$

ここで n は総節点数である．ガラーキンの法では，各々の要素に関しての区分多項式を

$$p = \sum_i N_i o_i \quad (2.5)$$

とすると同時に，重み関数についても同じ関数を採用する．

$$w = \sum_i N_i p_i \quad (2.6)$$

従って，領域全体については以下のようになる．

$$\sum_{e_i=1}^n \int_{V_{e_i}} p(\nabla^2 p + k^2 p) dV = 0 \quad (2.7)$$

式 (2.7) 中の i は 1 から n まで要素番号に対応している．

しかし，未知数 p は内挿関数 N_i の 1 次式で表されているので，2 階微分すると不定になる．これを避けるためにグリーン・ガウスの定理で部分積分する．

$$\sum_{e_i=1}^n \left(\int_{V_{e_i}} (\nabla p \nabla p + k^2 p p) dv - \sum_{S_j=1}^4 \int_{S_j} p \nabla p \cdot n ds \right) = 0 \quad (2.8)$$

ただし式 (2.8) 中の n は法線ベクトルである．この方程式を離散化し，まとめると次式になる．

$$\sum_{e_i}^n \left([S] - k^2 [T] + jk \frac{\rho c}{z_R} [G] \right) [p] = \sum_{e_i}^n (-jk\rho c u) [F] \quad (2.9)$$

ここで $[S], [T], [G]$ は 4×4 の行列， $[p], [F]$ は 4 行のベクトルである． $\sum_{e_i}^n$ をとるとき，異なる要素でも同じ節点について足し合わせれば合計の行列のサイズを小さくできる．節点の数が n 個の場合，行列 $[S], [T], [G]$ は $n \times n$ で，ベクトル $[p], [F]$ は n 行になる．

$$\left([S_{all}] - k^2 [T_{all}] + jk \frac{\rho c}{z_R} [G_{all}] \right) [p_{all}] = -jk\rho c u [F_{all}] \quad (2.10)$$

この連立方程式を解くと，各節点の音圧が求まる．

2.2.2 連立方程式の解法

以上によりできた連立方程式は，Gauss の消去法によって解くことが可能である．この計算に最も大きな容量を占めるのが $n \times n$ の行列である．この行列の要素の特徴としては $k_{ij} = k_{ji}$ であることと，要素中に 0 が数多く含まれていることが挙げられる．よって計算時に使用するメモリ，及び計算回数節約のために圧縮した形で保存する．例えば，次のような行列があるとする．

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} & 0 & 0 & 0 \\ k_{21} & k_{22} & k_{23} & k_{24} & 0 & 0 \\ k_{31} & k_{32} & k_{33} & k_{34} & k_{35} & 0 \\ 0 & k_{42} & k_{43} & k_{44} & k_{45} & k_{46} \\ 0 & 0 & k_{53} & k_{54} & k_{55} & k_{56} \\ 0 & 0 & 0 & k_{64} & k_{65} & k_{66} \end{pmatrix} \quad (2.11)$$

これを以下の形に圧縮して置き換える．

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{22} & k_{23} & k_{24} \\ k_{33} & k_{34} & k_{35} \\ k_{44} & k_{45} & k_{46} \\ k_{55} & k_{56} & 0 \\ k_{66} & 0 & 0 \end{pmatrix} \quad (2.12)$$

この形の行列をバンドマトリックス，バンドマトリックスの列数をバンド幅と呼び，列の数は要素中の最大節点番号と最小節点番号の差まで減らすことができるので，列の数とバンド幅は一致する．従って，バンド幅を 1 減らすと ($1 \times$ 総節点数) だけ行列を小さくすることができる．バンド幅は節点番号の付け方によって変化するため，バンド幅が小さくなるように節点番号の最適化を行うことがバンドマトリックスの構成には有効である．また，バンドマトリックスの行番号は節点番号と一致しており，実際の計算は一つの要素を構成する節点同士の連立方程式となる．その際には Gauss の消去法を用い，各節点の音圧を求めることができる．

2.2.3 粒子速度

次に音圧と粒子速度の関係について述べる．前述のように各節点の音圧が求まると，この音圧の値を用いて，要素の中心の粒子速度を求めることができる．粒子速度が角周波数

ω で正弦変化するとして時間項を分離する .

$$U(x, y, z, t) = ue^{j\omega t} \quad (2.13)$$

となる . ここで

$$\nabla p = -\rho \frac{du}{dt} \quad (2.14)$$

より

$$u = -\frac{\nabla p}{jk\rho c} = -\frac{1}{jk\rho c} \left(\frac{\partial p}{\partial x}, \frac{\partial p}{\partial y}, \frac{\partial p}{\partial z} \right) \quad (2.15)$$

が得られ , 音圧から粒子速度が計算できる .

2.3 1次元近似モデル

有限要素法との比較のために , 1次元近似モデルを用いる . ここでは , 声道の音響管近似モデルとして , 声道を断面積が連続的に変化する音響管とみなした Sondhi のモデル [13][14][15] を用いる . そのモデルは式で表現することができる .

$$\begin{pmatrix} P_{out} \\ U_{out} \end{pmatrix} = \prod_i \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix} \times \begin{pmatrix} P_{in} \\ U_{in} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \times \begin{pmatrix} P_{in} \\ U_{in} \end{pmatrix} \quad (2.16)$$

ただし , 式 (2.16) で , 入力を声門での音圧 P_{in} , 体積速度 U_{in} とし , 出力を口唇での音圧 P_{out} , 体積速度 U_{out} とする .

マトリックスの各構成要素 A, B, C, D は次のように表すことができる .

$$\begin{aligned} A_i &= \cosh(\sigma \Delta l / c) \\ B_i &= -\frac{\rho c}{area(i)} \gamma \sinh(\sigma \Delta l / c) \\ C_i &= -\frac{area(i) \sinh(\sigma \Delta l / c)}{\rho c \gamma} \\ D_i &= \cosh(\sigma \Delta l / c) \end{aligned} \quad (2.17)$$

なお , 第 i セクションの断面積 $area(i)[cm^2]$, セクション長 $\Delta l[cm]$ とする . γ と σ は次の式で定義される .

$$\gamma = \sqrt{\frac{\alpha + j\omega}{\beta + j\omega}} \quad (2.18)$$

$$\sigma = \gamma(\beta + j\omega) \quad (2.19)$$

ただし ,

$$\alpha = \sqrt{j\omega c_1} \quad (2.20)$$

$$\beta = \frac{j\omega\omega_0^2}{(j\omega + a)j\omega + b} + \alpha \quad (2.21)$$

である . シミュレーションに用いる各パラメータの意味とその値を表 2.3 に示す . [13]

表 2.3: Variables used in 1-D model

Variable	Meaning	Value	Unit
c	wave velocity	3.5×10^4	cm/s
ρ	density of air	1.14×10^{-3}	g/cm^3
Δl	length of elementary section	0.5	cm
a	ratio of wall resistance to mass	130π	rad/s
b	squared angular frequency of mechanical resonance	$(30\pi)^2$	$(rad/s)^2$
c_l	correction for thermal conductivity and viscosity	4	rad/s
ω_0^2	lowest squ. ang freq. of acoustic resonance	$(406\pi)^2$	$(rad/s)^2$

声道伝達関数 $H(\omega)$ は

$$H(\omega) = \frac{1}{A - CZ_L} \quad (2.22)$$

で得られる . ただし , Z_L は口唇での放射インピーダンスであり , 口唇部の半径を $r[cm]$ とすると次の式で表現できる .

$$Z_L = \frac{(\omega r)^2}{2c^2} + \frac{j\omega(8r)}{3\pi c} \quad (2.23)$$

第 3 章

音響分析

本研究で用いるシミュレーション法の妥当性を検討するために、シミュレーション結果と実音声の分析結果との比較を行う。ここでは実音声の分析に用いる音響分析法、分析条件について述べる。

3.1 フォルマント周波数について

フォルマント周波数とは、音声に含まれる成分のうち、声門から唇や鼻腔までの声道管の共鳴によって強められた部分であり、声道形状によって決まる。フォルマントは音声中に多数存在し、低い方から順に第一フォルマント (F_1)、第二フォルマント (F_2)、第三フォルマント (F_3)、... と名付ける。母音では F_1 , F_2 , F_3 によって母音固有の音色が決まる。[16]

口腔疾患を有する患者の音声は、構音器官の欠損や運動障害により声道形状が健常者のものとは異なり、複雑になる。そのためにフォルマントの位置が移動し、音声の歪みの原因になる。

本研究で用いる FEM による 3 次元声道モデルの声道伝達特性推定の結果とは、音響管の共振、反共振周波数である。これは音声のフォルマント周波数と対応させることができ、シミュレーション法の妥当性を検討するためのパラメータとして用いることができる。

3.2 分析法

音響分析の方法としては不偏推定法によるケプストラム平滑化対数スペクトル [10][11] を用いる。これは、音声波形のスペクトル包絡を全極形有利スペクトルと仮定して、 N

個の波形標本値 $x = (x_1, x_2, \dots, x_n)$ の対数尤度 $L(x|\omega)$ を最大にするようなパラメータ $\omega(\sigma a_1, a_2, \dots, a_p)$ を求めることによってスペクトル包絡を推定する方法である．不偏推定法では対数スペクトルの推定値にバイアスが生じないように非線形平滑化を行うことにより，微細構造を表現するスペクトル包絡を得ることができる．

3.3 録音条件

音声の録音は MR 撮像後に防音室内にて行った．被験者の姿勢は Normal は仰臥位，Patient は立位とした．MR 撮像時と同一条件にするため，被検音は日本語母音 /i/ とし，MR 撮像時に歯冠プレートや PAP を使用していた場合は同様に装着した状態で録音した．その音声を DAT (SONY TCD-D 10 PRO II, SONY 社製) に録音した．発話時間は Normal は MR 撮像時の条件と同一になるように，息継ぎをすることなく息が続く限り発話させた．Patient の場合はおよそ 1[sec] の発話時間とした．

3.4 分析結果

録音された音声は DAT-LINK+ (Townshend Computer Tools, Inc) を介して WS 上に音声ファイルとして取り込み，視察により /ai/ 発音時の /i/ の部分を切り出し，不偏推定法によるケプストラム平滑化対数スペクトルによりパワースペクトル形状の抽出を表 3.1 の設定で行った．切り出した音声 /i/ の長さは Normal で 10[sec]，Patient で 200[ms] である．

表 3.1: Acoustic analysis.

Sampling frequency	20[kHz]
FFT size	512[point]
Frame length	25.6[ms]
Frame period	2[ms]
Cepstrum order	40
Window type	Hanning

得られたパワースペクトルから各被験者の日本語母音 /i/ 発話時のフォルマント周波数の平均と標準偏差を求めた．不偏推定法によるケプストラム平滑化対数スペクトルによる

分析の結果得られたパワースペクトルを示す．そして，そのパワースペクトルから得られるフォルマント周波数の平均値を示す．

まず，Normal-A のパワースペクトルを図 3.1 に，Normal-B のパワースペクトルを図 3.2 に示す．これは分析のために切り出した音声 (10[sec]) の各時間におけるスペクトル包絡を示している．また，各時間におけるパワースペクトルから得られる各フォルマント周波数の平均値と標準偏差を表 3.2 に示す．

表 3.2: Formant frequencies (Normal).

Subjects	Formant frequencies (Average \pm S.D. [Hz])	
	Normal-A	Normal-B
F_1	260 \pm 20	320 \pm 25
F_2	2320 \pm 33	2320 \pm 29
F_3	2880 \pm 79	3320 \pm 39
F_4	3780 \pm 112	3980 \pm 104

Normal の場合は発話時間が長いためにフォルマントの位置にずれが生じていることが分かる．なお，各フォルマントの最大値，最小値の差は平均値に対して約 ± 10 [%] ある．

次に Patient-A の PAP 非装着時 (以下 PAP(-) とする) の 1 回目の発話時のパワースペクトルを図 3.3 に，2 回目の発話時のパワースペクトルを図 3.4 に示す．また，Patient-A の PAP 装着時 (以下 PAP(+) とする) の 1 回目の発話時のパワースペクトルを図 3.5 に，2 回目の発話時のパワースペクトルを図 3.6 に示す．Patient-B の 1 回目の発話時のパワースペクトルを図 3.7 に，2 回目の発話時のパワースペクトルを図 3.8 に示す．各時間におけるパワースペクトルから得られるフォルマント周波数の平均値と標準偏差を表 3.3 に示す．

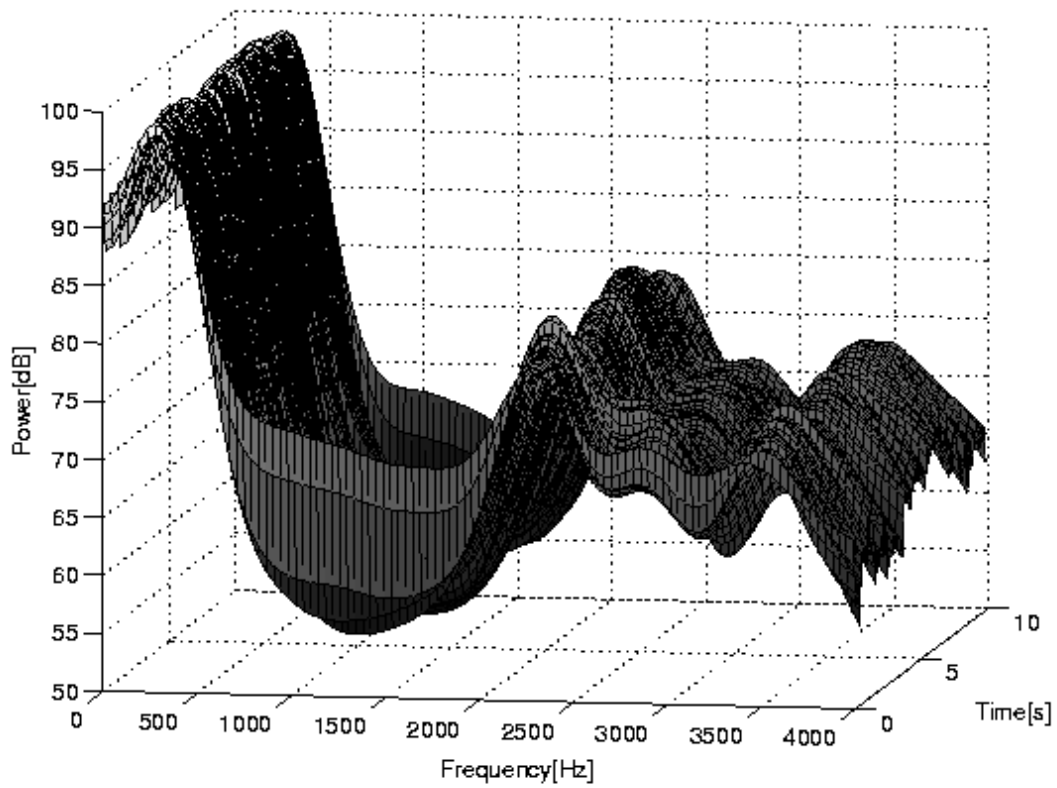
フォルマントの標準偏差は Patient の PAP(-) の各症例において，大きな幅があることが分かる．特に，Patient-A の音声スペクトルである図 3.3 と図 3.4 の違いからも分かる通り，発話毎に，そして発話時間中にスペクトルが一定していないことが分かる．これは構音器官の欠損や運動障害により，構音が一定しないために起きると考えられる [19]．その場合の各フォルマントの最大値，最小値の差は平均値に対して約 20[%] となっている．Patient A(PAP(+)) の場合は，PAP を装着することにより構音が安定し，標準偏差，平均値と最大値，最小値の差ともに PAP 装着前より改善されている．

症例 Patient A (PAP(-)) の音声は 1500[Hz] 付近にフォルマントが存在していること

表 3.3: Formant frequencies (Patients).

Subjects	Formant frequencies (Average \pm S.D. [Hz])		
	Patient-A (PAP(-))	Patient-A(PAP(+))	Patient-B
F_1	340 \pm 20	310 \pm 12	290 \pm 25
F_2	1450 \pm 65	2100 \pm 35	1860 \pm 134
F_3	2300 \pm 65	3420 \pm 55	3070 \pm 69
F_4	3920 \pm 621	—	3380 \pm 40

が分かる。これは、日本語母音 /i/ 発話時に必要な、舌を持ち上げることによりできる狭めが舌・口底を切除したために十分にできないために起こっていると考えられる。これが、音声の歪みの主要な原因であると考えられる。これに対して、術後の機能障害に対する治療として用いられている PAP を装着した場合は、1500[Hz] 付近にあったピークが消えていることがわかる。スペクトル形状としては、健常者の例である Normal-A, Normal-B のスペクトル形状に似ていることが分かる。また、4000[Hz] 付近にあったフォルマントが低い周波数方向にシフトしていることも分かる。



☒ 3.1: Power Spectrum (Normal-A).

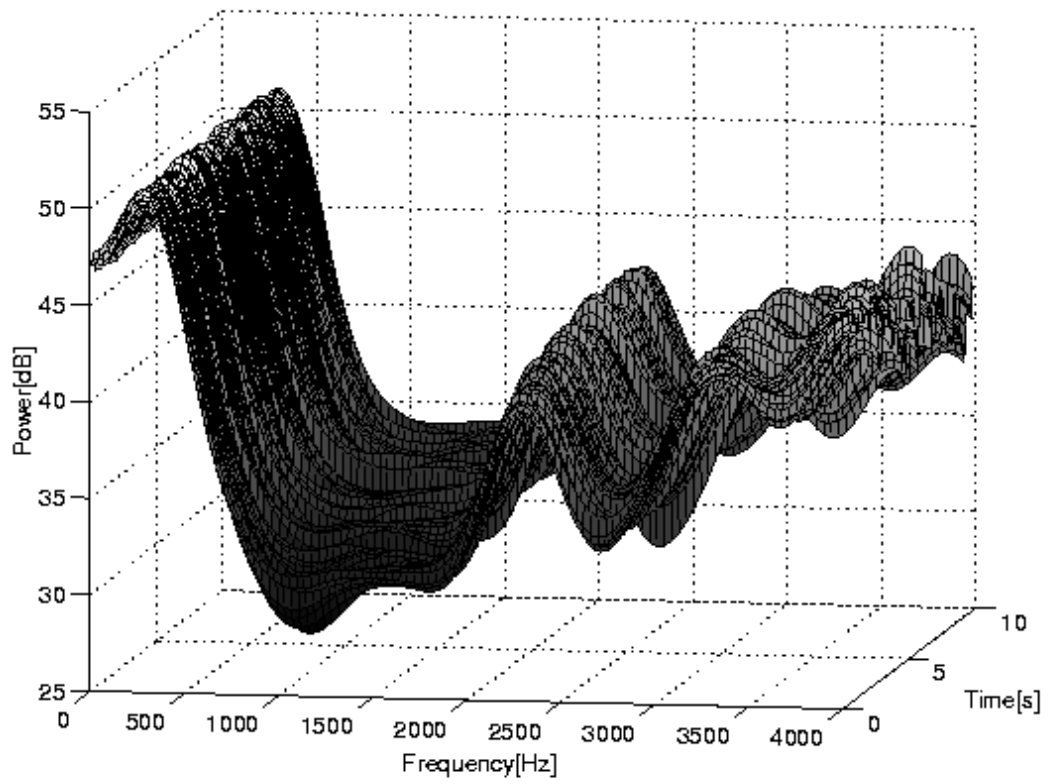


图 3.2: Power Spectrum (Normal-B).

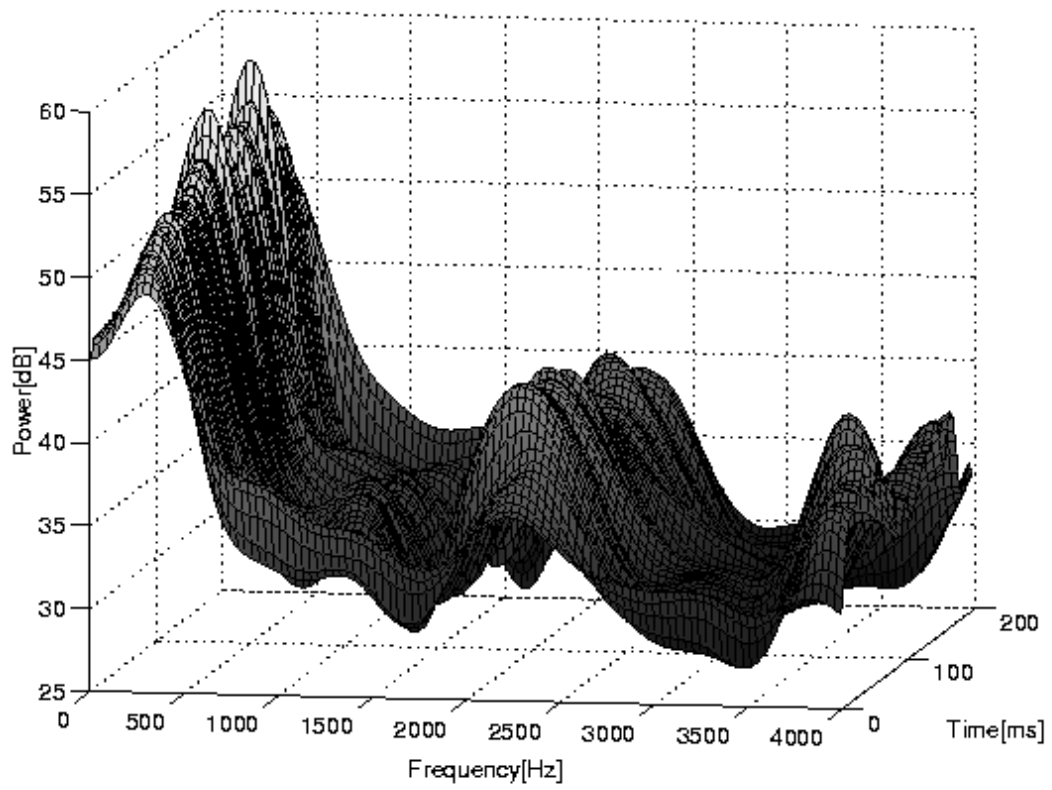


图 3.3: Power Spectrum (Patient-A (PAP (-))) (1).

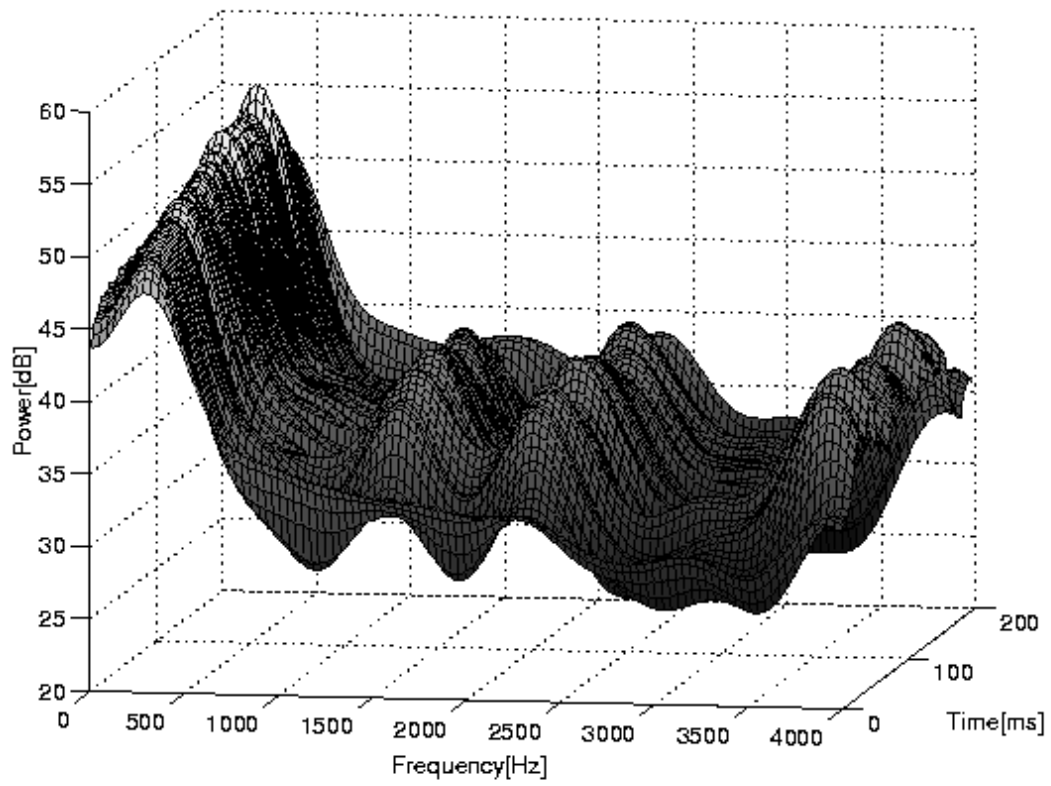
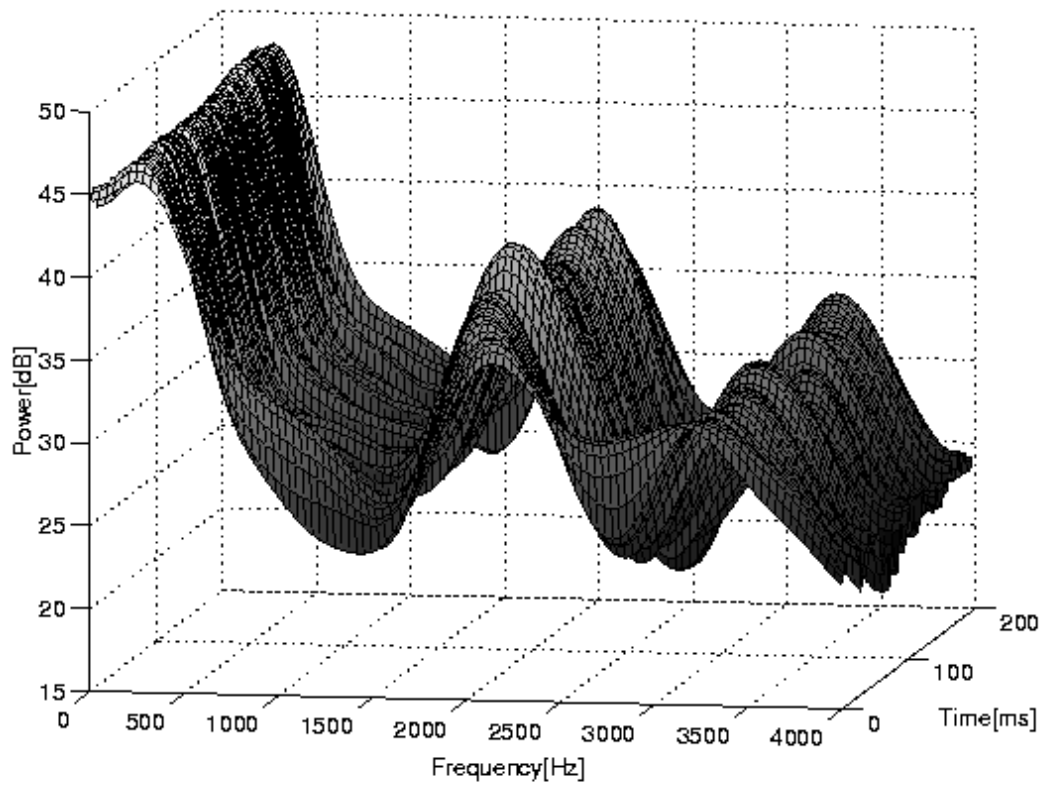
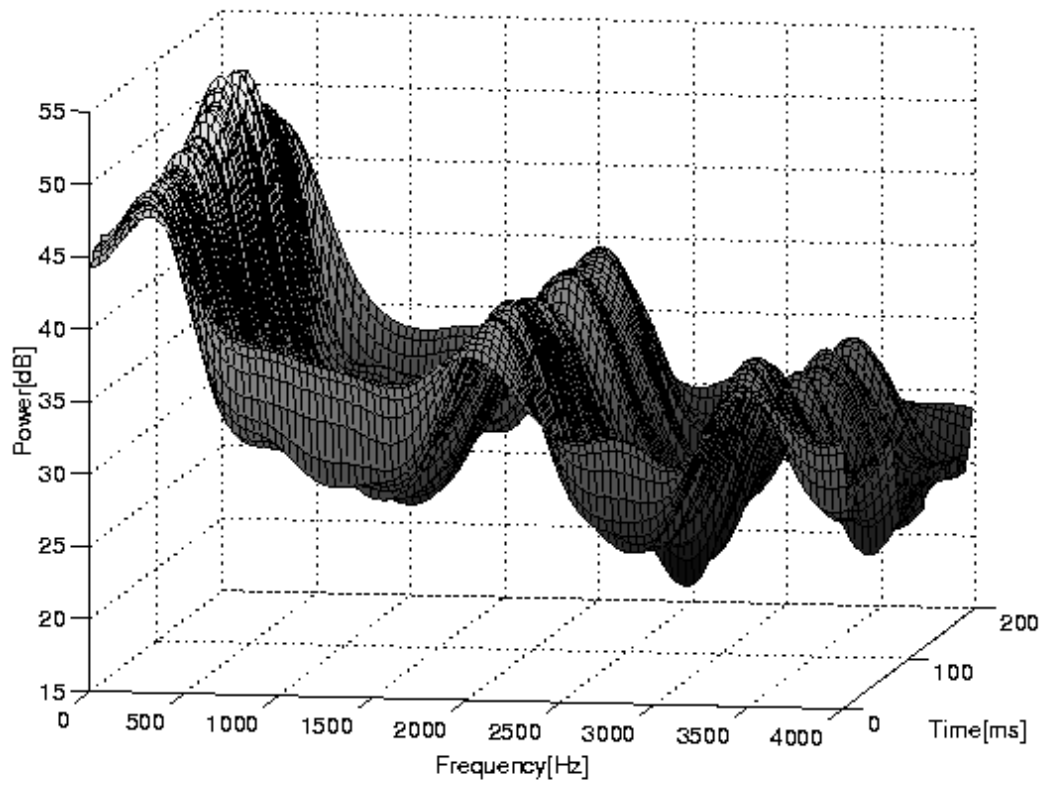


图 3.4: Power Spectrum (Patient-A (PAP (-))) (2).



⊗ 3.5: Power Spectrum (Patient-A (PAP (+))) (1).



⊗ 3.6: Power Spectrum (Patient-A (PAP (+))) (2).

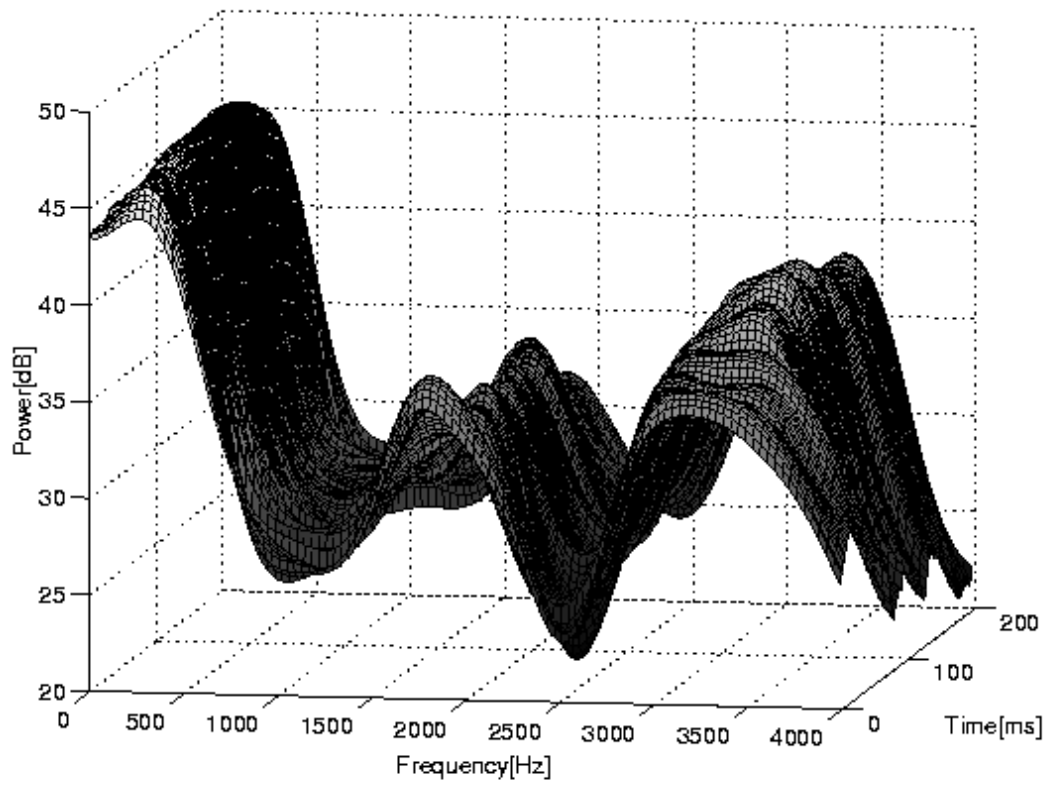


图 3.7: Power Spectrum (Patient-B) (1).

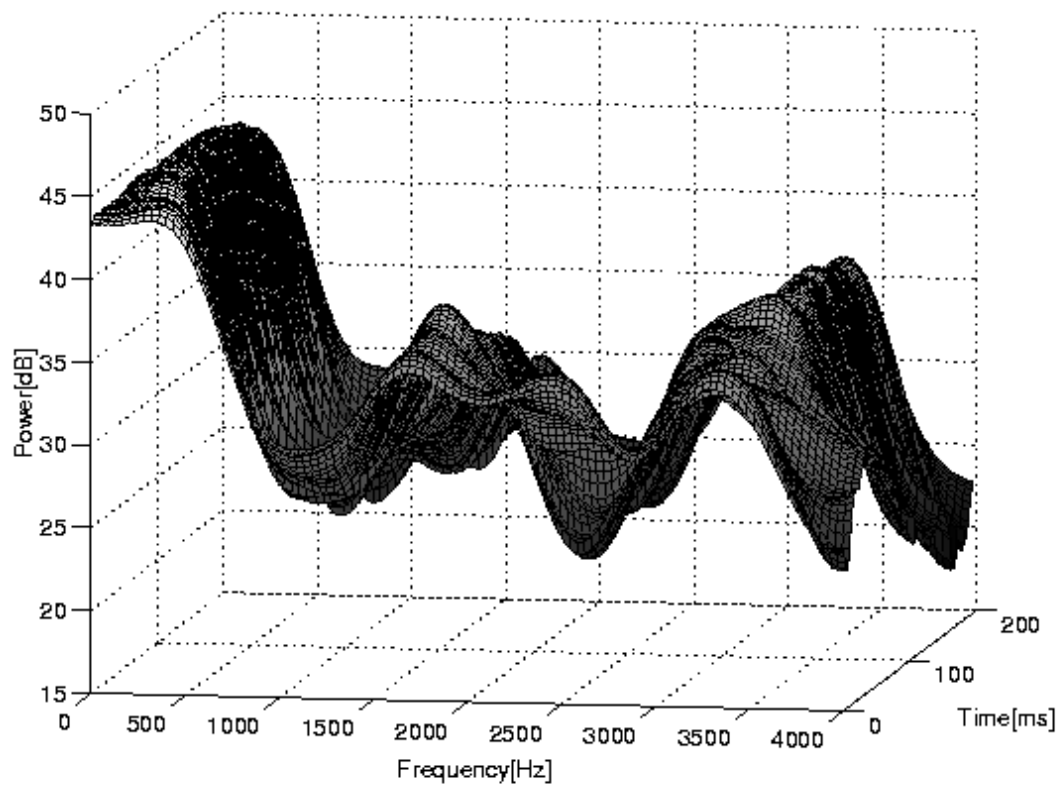


图 3.8: Power Spectrum (Patient-B) (2).

第 4 章

有限要素法の計算精度

4.1 はじめに

本章では，MR 画像より抽出された声道形状情報より 3 次元声道形状モデルを構築する手法を解説し，その 3 次元声道モデルを FEM によりシミュレーションを行った結果を示す．シミュレーション結果と第 3 章で示した音響分析結果とを比較し，本シミュレーション法の妥当性を検討する．なお，FEM による分析結果の妥当性を検証するためにここで対象とする 3 次元声道モデルは，構音が一定しているため，音声のフォルマントの揺れが少ない健常者 (Normal) のモデルを使用する．さらに，Normal-A の 3 次元声道モデルの形状に変形を与え，それが声道伝達特性に与える影響について示す．

4.2 3 次元声道形状モデルの構築

静岡大学の北村により MR 画像から抽出された声道形状情報 [7][8] をもとに，3 次元の声道形状モデル (サーフェスモデル) を作成した．それには Geomagic 社のサーフェスモデル作成ソフト Wrap を使用した．さらに，このサーフェスモデルから Altair 社の汎用プリポストシステム HyperMesh により，ソリッドモデルである 3 次元有限要素モデルを作成した．

4.3 シミュレーション条件

本研究では NIT(Numerical Intergration Technologies) 社の音響解析プログラム SYS-NOISE の FEM による解析を使用した．入力端では正弦波振幅が 1.0 の粒子速度で励振

させた．また出力となる開放端での放射インピーダンスは 0.001^1 を与えた．それら以外の境界面は剛壁 (無損失) とするため，インピーダンスは ∞ を与えた．管内では、音速 $c = 340[m/s]$ ，気体密度 $\rho = 1.225[kg/m^3]$ で一定とした．それによる声道伝達特性 H は次の式で求めた．

$$H = 20 \log_{10} \frac{U_{out}}{U_{in}} \quad (4.1)$$

ここで， U_{in} は入力面 (声門部) の中心付近の粒子速度であり， U_{out} は出力面の中心付近の粒子速度である．分析範囲は F_1 から F_3 までのピークを含むようにするため，特に注意がない限り $10 - 4000[Hz]$ で $10[Hz]$ 刻みとした．また，シミュレーション結果と実音声の比較に用いる誤差 $Error$ の計算式は次の式で表される．なお，シミュレーションの結果得られる声道伝達関数のピーク周波数を F_s ，音響分析結果のフォルマント周波数を F_a とする．

$$Error = \frac{|F_a - F_s|}{F_a} \times 100 \quad (4.2)$$

4.4 シミュレーション結果

4.4.1 点間隔 $1.0[mm]$ のモデル (Normal-A)

ここでは，Normal-A の MR 画像から声道形状情報を $1.0[mm]$ 間隔で抽出したデータから作成された 3 次元声道形状モデルを用いて行ったシミュレーション結果を示す．なお，このモデルの口唇部は口角までであり，歯の形状は考慮されていない．この 3 次元声道モデルを図 4.1 に示す．有限要素モデルの要素数は 81865，節点数は 19888 である．入力面は声帯部であり，出力面は口唇部としており，その中心付近，右側，左側の各点を出力としている．FEM によるシミュレーション結果を図 4.2 に示す．また，シミュレーションの結果得られる声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 4.1 に示す．

図 4.2 から出力点を変えても推定される声道伝達関数に影響がないことが分かる．これは，健常者の声道形状が声道の軸に対して左右対称であり，出力面での音圧に差がないためである．

推定された声道伝達関数のピークと実音声のフォルマントの数が異なっている．実音声のフォルマントを分析する時にスペクトル包絡から推定しているが，包絡を取っているた

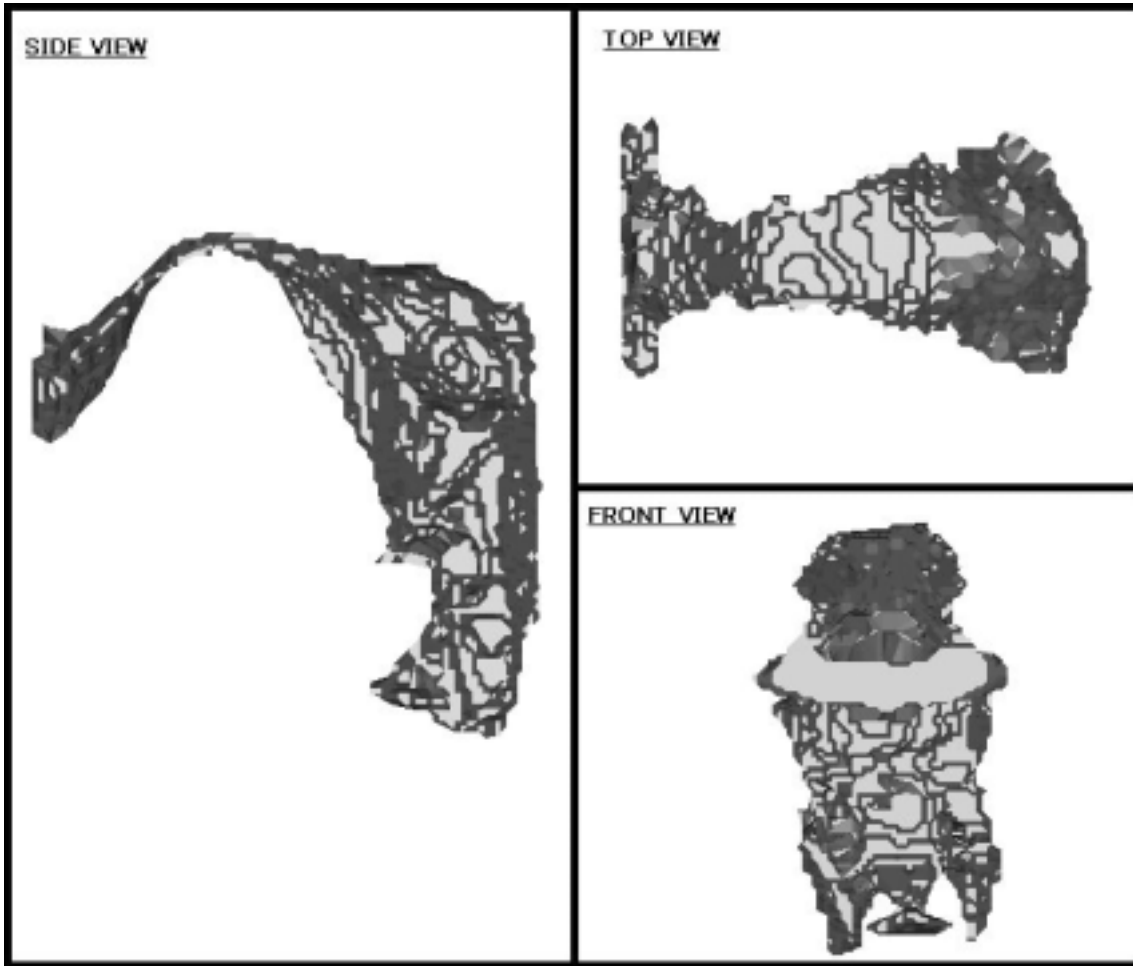
¹計算機上では 0 を入力できないため，この値を用いた

表 4.1: Peak frequencies of estimated vocal tract function and formant frequencies.

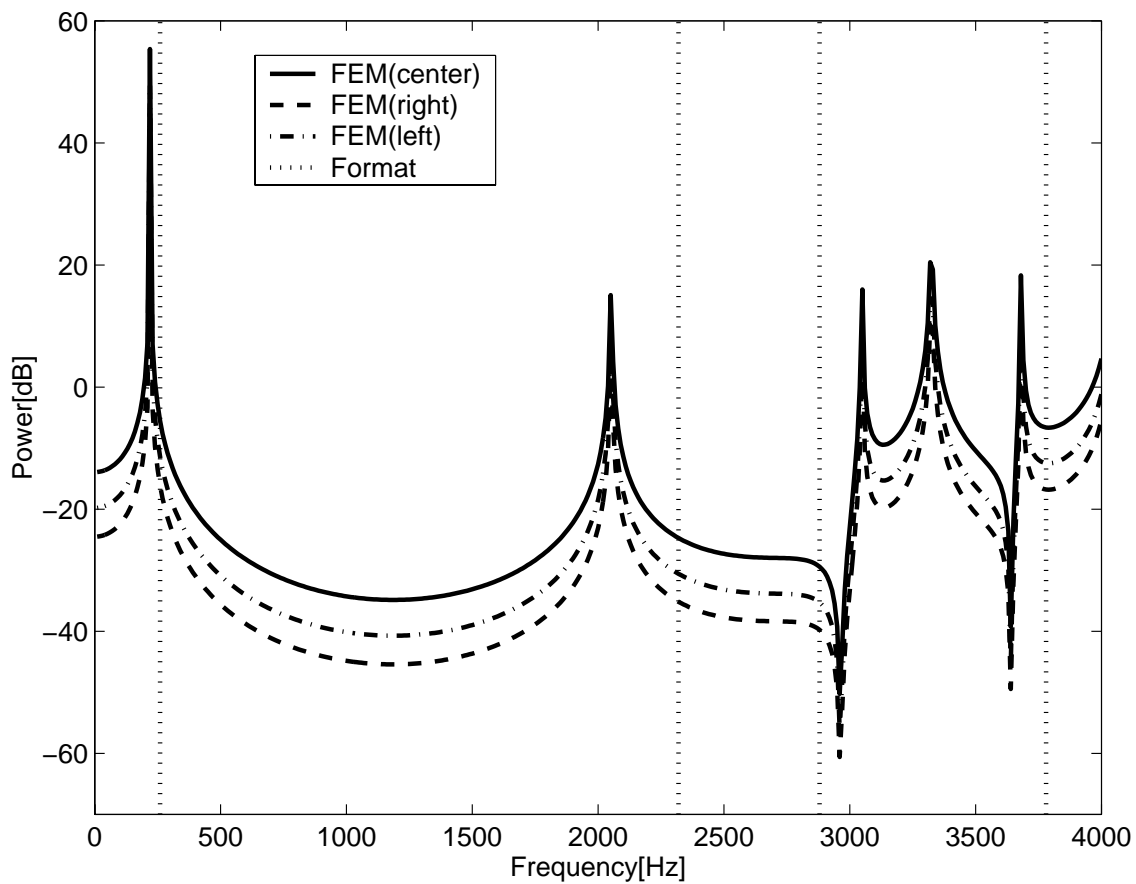
	Molel[Hz]	Formant[Hz]	Error[%]
F_1	220	260	19.2
F_2	2050	2320	11.6
F_3	3050	2880	8.9
—	3320	—	—
F_4	3670	3780	3.2

めにピークが近接している場合は実音声の分析結果がピークとピークを結ぶようななだらかな曲線となり、1つのピークとして現れることが考えられる。一方で、本シミュレーション法では声道壁を剛壁とし、空気に粘性がないと仮定しているためピークが鋭く立つ。そのため、シミュレーション結果と実音声と比較する場合は、実音声のフォルマントと対応づけるシミュレーション結果のピークに注意する必要がある。本稿では音響分析結果のフォルマント周波数に最も近い声道伝達特性のピーク周波数を比較する。

音響分析結果より、実音声は F_3, F_4 において $\pm 10[\%]$ の幅で変動しているということと、MR 画像より得られる測定値と実測値との間の誤差は $5[\%]$ であるという報告 [6] を考慮すれば、表 4.1 より F_3, F_4 のずれは誤差の範囲内であるといえる。ただし、 F_1, F_2 には大きな誤差が存在する。



⊗ 4.1: A Normal-A vocal tract model uttering Japanese vowel /i/. This model is constructed by a vocal tract data (1.0[mm] sampling).



⊠ 4.2: A transfer function of a Normal-A vocal tract model by an 1.0[mm] sampling vocal tract data.

4.4.2 口腔部の点間隔 $0.5[mm]$ のモデル (Normal-A)

4.4.1 節でのシミュレーションにより、点間隔 $1.0[mm]$ の Normal-A の声道形状情報から構築された声道モデルでは、 F_1 、 F_2 に大きな誤差が出ていることがわかった。健常者の場合、日本語母音 /i/ における F_1 は、主に舌による狭めの影響を受けているといわれている。しかしながら、4.4.1 節で用いたモデルはサンプリング精度が粗く、舌による狭めの幅が精度良く再現できていない可能性がある。そこで、ここでは口唇部から狭めの位置にかけてを $0.5[mm]$ 間隔の点で細かく取り直した声道情報をもとに、3次元声道モデルを構築し、それを FEM でシミュレーションを行った。

ここで用いるモデルの口唇部は口角までであり、歯の形状は考慮されていない。この3次元声道モデルを図 4.3 に示す。有限要素モデルの要素数は 138803、節点数は 32987 である。入力面は声帯部であり、出力面は口唇部の中心付近、口唇部の右側、および口唇部の左側の3点としている。FEM によるシミュレーション結果を図 4.4 に示す。また、シミュレーションの結果得られる声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 4.2 に示す。

表 4.2: Peak frequencies of estimated vocal tract function and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	280	260	7.6
F_2	2070	2320	10.8
—	2920	—	—
F_3	3100	2880	7.6
—	3620	—	—
F_4	3710	3780	1.8

表 4.2 から F_1 、 F_3 の音響分析結果とシミュレーション結果の誤差が改善されていることが分かる。しかし、依然として F_2 に大きな誤差が残っている。日本語母音 /i/ の場合、 F_2 の位置は舌の狭めの位置から声帯側にかけての後室に依存する [20] といわれている。よって、MR 画像からモデルを作成するときに後室を精度良く取る必要がある。ただし、発声時には声帯が振動し、梨状窩が膨張、収縮をするために MR 画像ではその周辺がぼやけてしまう。さらに、撮像時に音声が続かなくなった場合、声道内の圧力が下がることにより軟組織が変形を起こし、声道形状が変わっている可能性がある。そのため、後室側の器官の形状を3次元声道モデルに精度良く再現することが困難となっている。MR 撮像

時間を短縮したり，MR 撮像時に声帯部分も明瞭に写るような工夫をすることで，シミュレーション結果と実音声との整合性を取ることができるといえる．

出力端を口唇部の中心付近，口唇部の右側，および口唇部の左側の 3 点でとって見たが，4.4.1 節と同様に声道伝達特性のピーク周波数に変化はなかった．

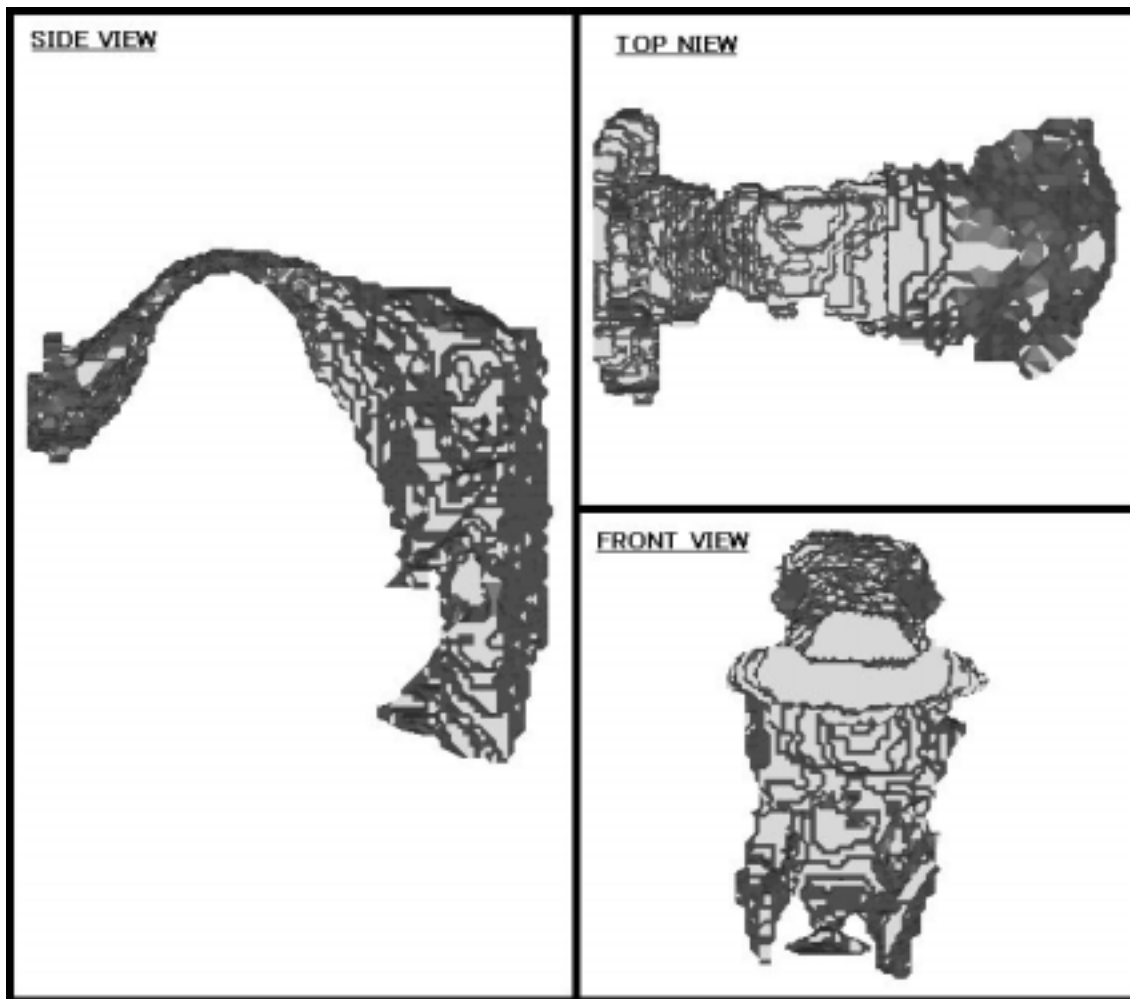
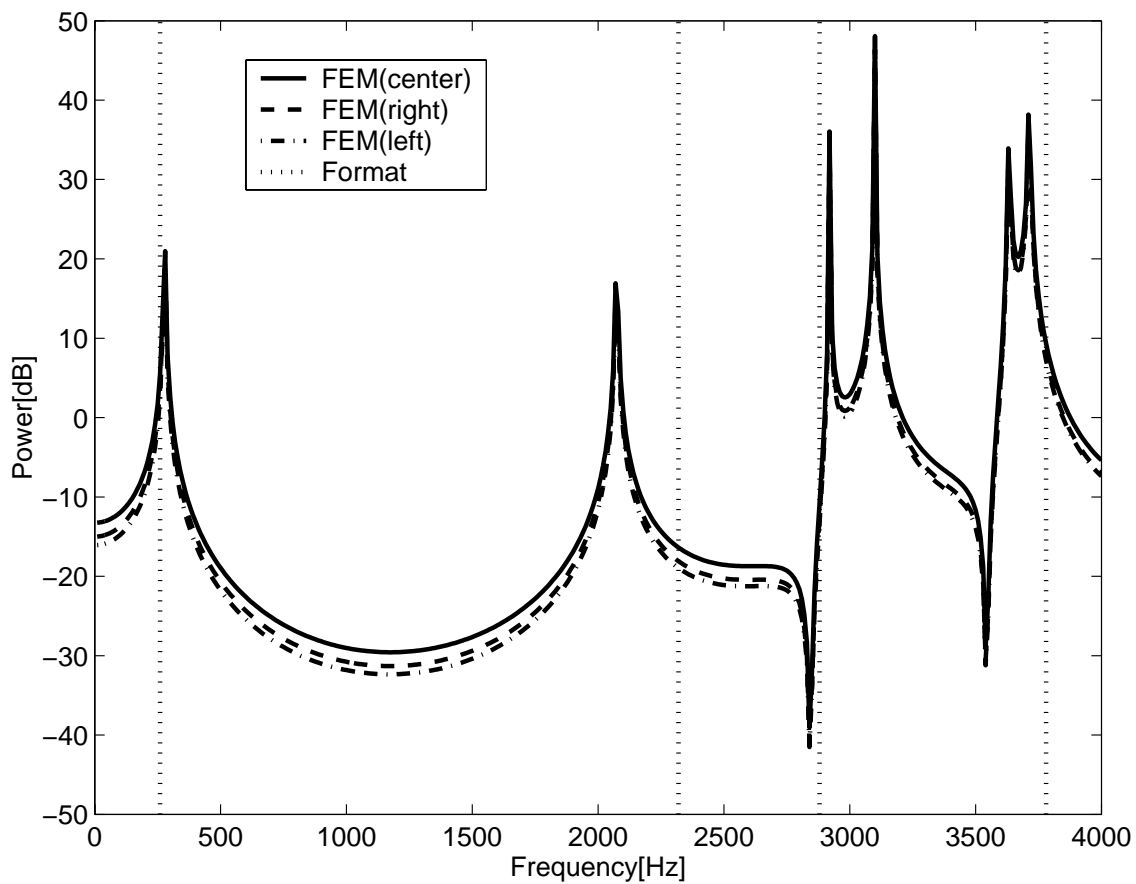


図 4.3: A Normal-A vocal tract model uttering Japanese vowel /i/ (0.5[mm] sampling). This model is constructed from a vocal tract data (0.5[mm] sampling).



⊠ 4.4: A transfer function of a Normal-A vocal tract model by an $0.5[mm]$ sampling vocal tract data.

4.4.3 放射球面付き声道モデル (Normal-A)

実際の音声は口唇部から 3 次元的に空間に放射している．そのため 3 次元声道モデルも放射を考慮したものでなくてはならない [17][18]．特に口腔疾患の場合，呼気が口腔正中から出ないために 3 次元の放射を考慮することが必要となってくる．そこで，ここでは 4.4.2 節で用いた 3 次元声道モデルの口唇部に半球状の放射球面を取り付け，3 次元的な放射を近似的に表現したモデルを作成した．これを 図 4.5 に示す．有限要素モデルの要素数は 98454，節点数は 23751 である．入力面は声帯部であり，出力面は放射球面の先端とした．また，放射球面の平面は剛壁 (インピーダンスは ∞) としている．このモデルの FEM によるシミュレーション結果を 図 4.5 に示す．また，シミュレーションの結果得られる放射球面ありのモデルの声道伝達特性と放射球面無しのなしのモデルの声道伝達特性を表 4.3 に，声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 4.4 に示す．

表 4.3: Peak frequencies of estimated vocal tract functions.

	Original model [Hz]	Model with radiation [Hz]	Difference [Hz]
F_1	280	280	0
F_2	2070	2080	10
—	2920	2770	150
F_3	3100	2940	160
—	3620	3310	310
F_4	3710	3500	210

表 4.3 のように放射球面を取り付けたことによる声道伝達関数の F_2 以下のピーク位置への影響はみられない．しかし， F_3 以上の声道伝達関数の形状は変わっており，放射球面を付けることによる影響が現れている．その変化による実音声との差は F_2 以外では誤差の範囲内であるとみなせる．

口腔疾患を有する声道形状の場合は健常者の声道形状とは異なり，口腔正中より呼気が出るとは限らず，そのことによる音響的な歪みが発生する．それを本シミュレーションで表現するためには放射球面の取り付けが重要となる．そこで，以後のモデルには口唇部に放射球面を取り付けたモデルでシミュレーションを行う．

表 4.4: Peak frequencies of estimated vocal tract function and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	280	260	7.6
F_2	2080	2320	10.3
—	2770	—	—
F_3	2940	2880	2.1
—	3310	—	—
F_4	3500	3780	7.4

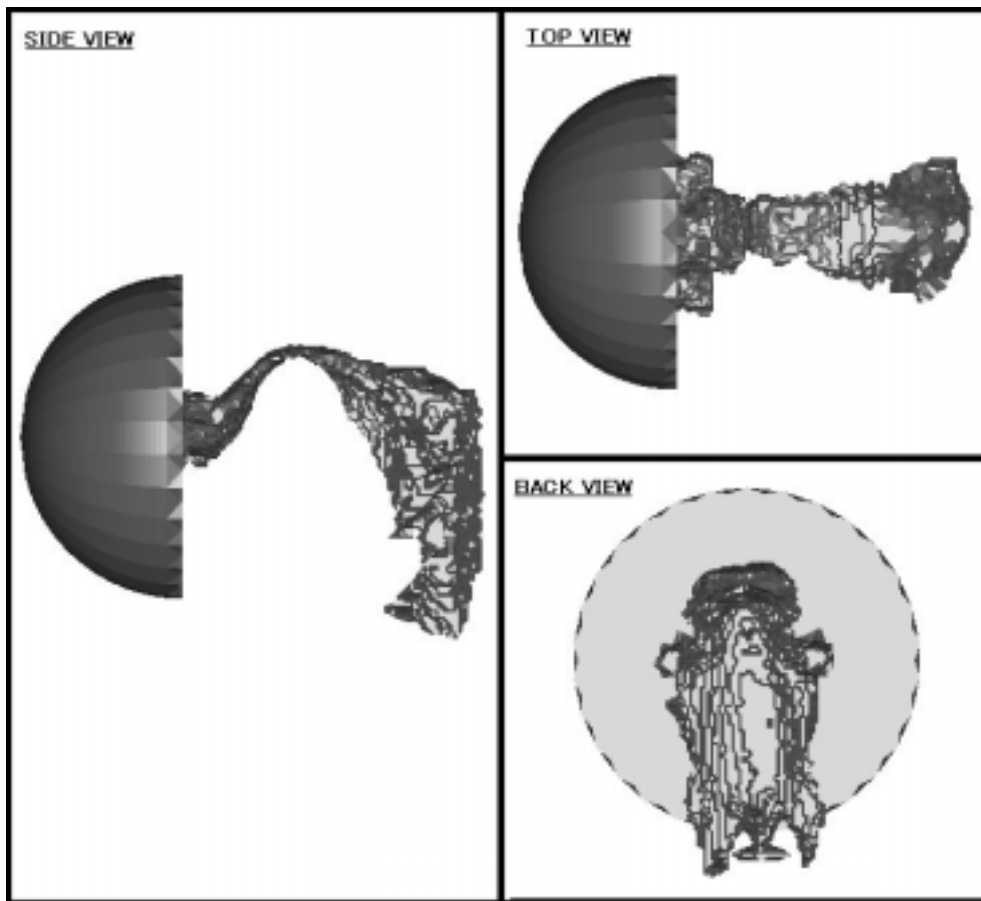
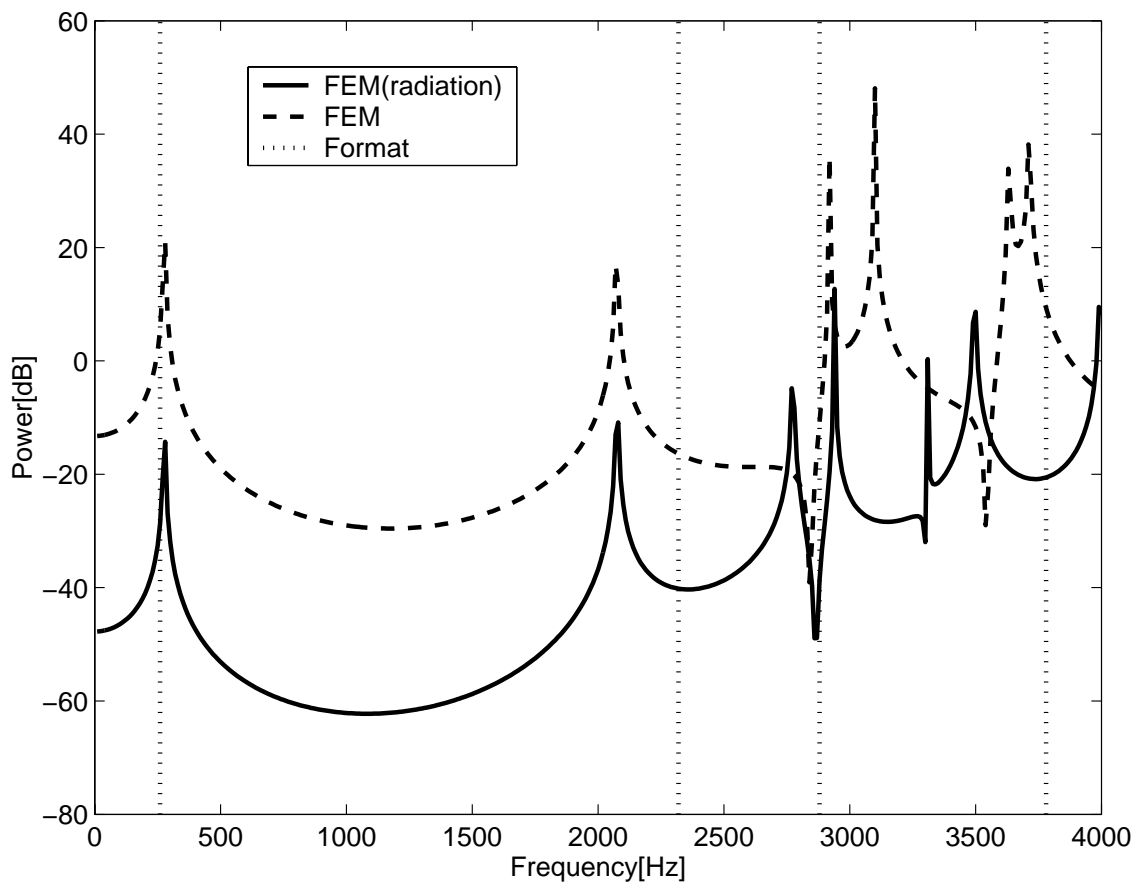


图 4.5: A Normal-A vocal tract model with hemispherical surface of radiation.



⊠ 4.6: A transfer function of a Normal-A vocal tract model with hemispherical surface of radiation.

4.4.4 点間隔 $0.5[mm]$ のモデル (Normal-B)

ここでは，Normal-B の MR 画像から声道形状情報を $0.5[mm]$ 間隔で抽出したデータから作成した 3 次元声道形状モデルを用いて行ったシミュレーション結果を示す．なお，このモデルの口唇部は前歯の付け根 ($/i/$ の場合は舌の先端のほぼ同じ意味) までであり，歯の形状は考慮されていない．また，放射球面は取り付け済みである．この 3 次元声道モデルを図 4.7 に示す．有限要素モデルの要素数は 114464，節点数は 27329 である．入力面は声帯部であり，出力面は口唇部としている．FEM によるシミュレーション結果を図 4.8 に示す．なお，Normal-B の F_4 は $4000[Hz]$ 付近にあるため，シミュレーション範囲は $10 - 5000[Hz]$ とした．また，シミュレーションの結果得られる声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 4.5 に示す．

表 4.5: Peak frequencies of estimated vocal tract function and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	180	320	43.8
F_2	2010	2320	13.4
F_3	3200	3320	3.6
—	3390	—	—
—	3580	—	—
F_4	4180	3980	5.0

Normal-B でのシミュレーションの場合，表 4.5 から F_1 に大きな誤差が見られる．Normal-B の音声の $/i/$ 発話時の基本周波数を計測した結果， $160[Hz]$ であった．そのため，基本周波数の倍ピッチ成分が音響分析の結果出ていることが考えられ，そのため大きな誤差が現れると考えられる．それ以外の比較では MR の計測誤差や音声の揺れを考慮すると F_2 以外は誤差の範囲内であると考えられる．

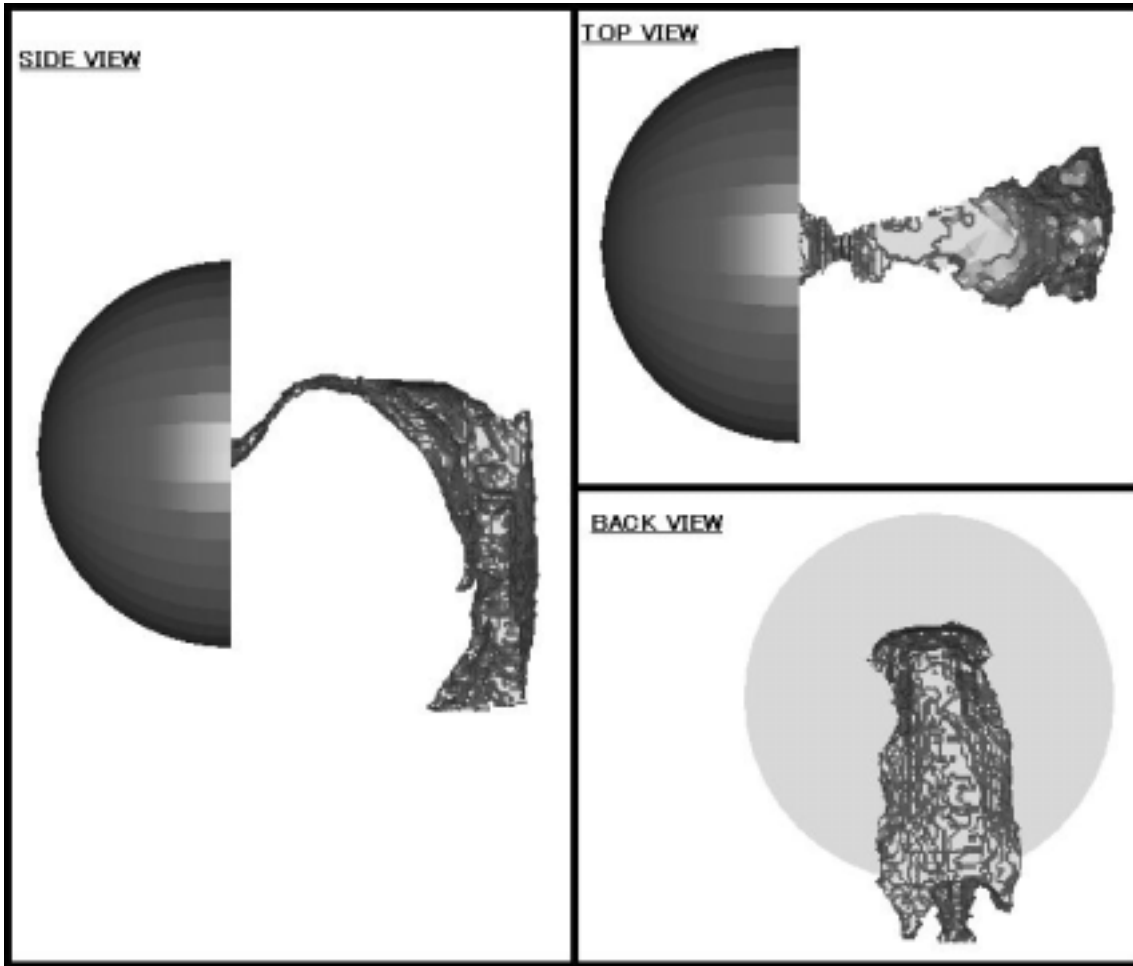
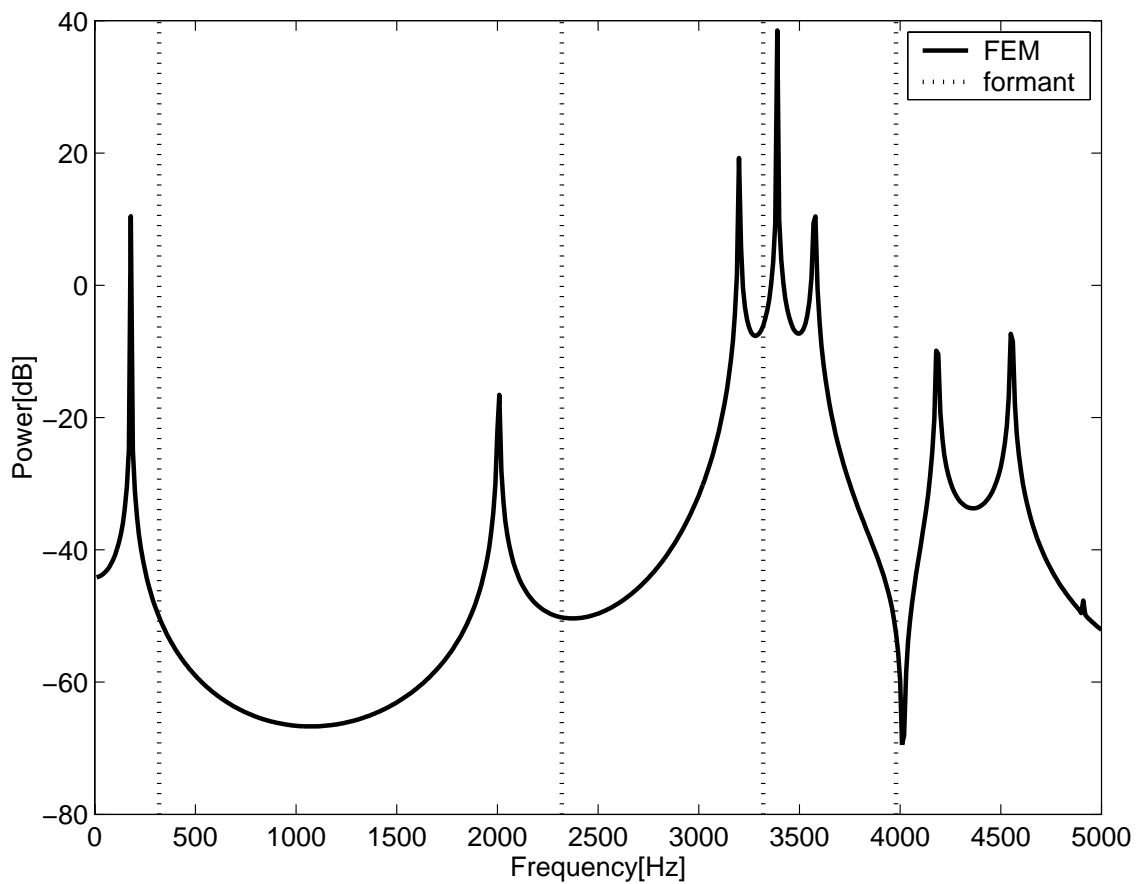


図 4.7: A Normal-B vocal tract model uttering Japanese vowel /i/. This model is constructed by a vocal tract data (1.0[mm] sampling).



⊠ 4.8: A transfer function of a Normal-B vocal tract model by an $1.0[mm]$ sampling vocal tract data.

4.5 声道後室がフォルマントに与える影響

ここまでは、健常者の声道モデルによるシミュレーション結果を示してきたが、 F_2 に誤差が残ることが分かる。 F_2 は声道の後室側の影響を受ける [20] と言われている。本節では声道後室に着目し、声道モデルの後室に変形を与えることでそれが声道伝達関数に及ぼす影響を調べ、シミュレーション結果と実音声の整合性を向上させることを目的とする。

4.5.1 仮声帯を音源としたモデル

音声の音源は普通声帯である。そのため、これまでのシミュレーションでは声帯部を音源に指定した。しかし、仮声帯近傍などの狭めがある部分でも乱流雑音源が存在すると考えられ、2 次的なエネルギー供給源になっていると考えられる [21] [22] 。

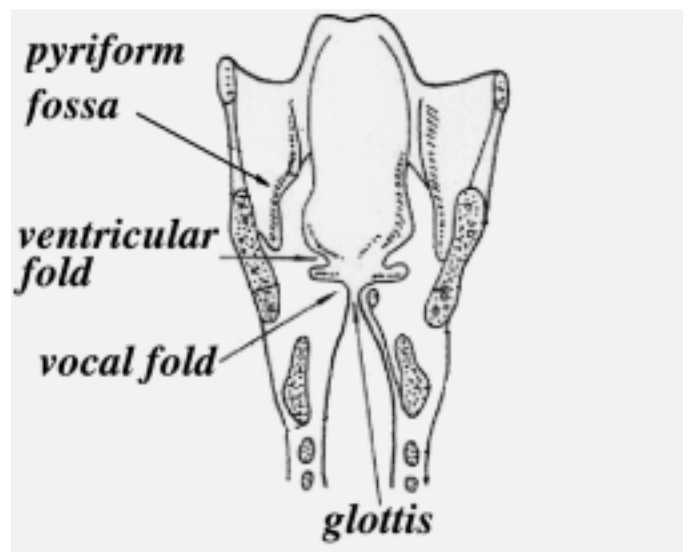


図 4.9: Glottis.

そこで、今回は 4.4.3 節で示したモデルの声門部を切断し、仮声帯を励振源としたモデルを構築した。これを 図 4.10 に示す。有限要素モデルの要素数は 94381、節点数は 22406 である。入力面は仮声帯部であり、出力面は放射球面の先端とした。このモデルの FEM によるシミュレーション結果とオリジナルのモデルでのシミュレーション結果を 図 4.11 に示す。

仮声帯を駆動面に指定した 3 次元声道モデルでは声道伝達特性の F_3 以上の周波数帯ではピーク周波数の位置に変化が見られた。しかし、目的としていた F_2 に対する変化は見

られなかった。

ここで用いている 3 次元声道モデルの声道壁は剛壁であり，空気は粘性を持っていない．そのため，仮声帯のような狭窄部が原因で生じる乱流雑音源のシミュレートができない．ささやき声では仮声帯が乱流雑音源になるが [21] [22]，通常の発話においても仮声帯が二次的なエネルギー供給源になるなら仮声帯での励振をさせる必要がある．

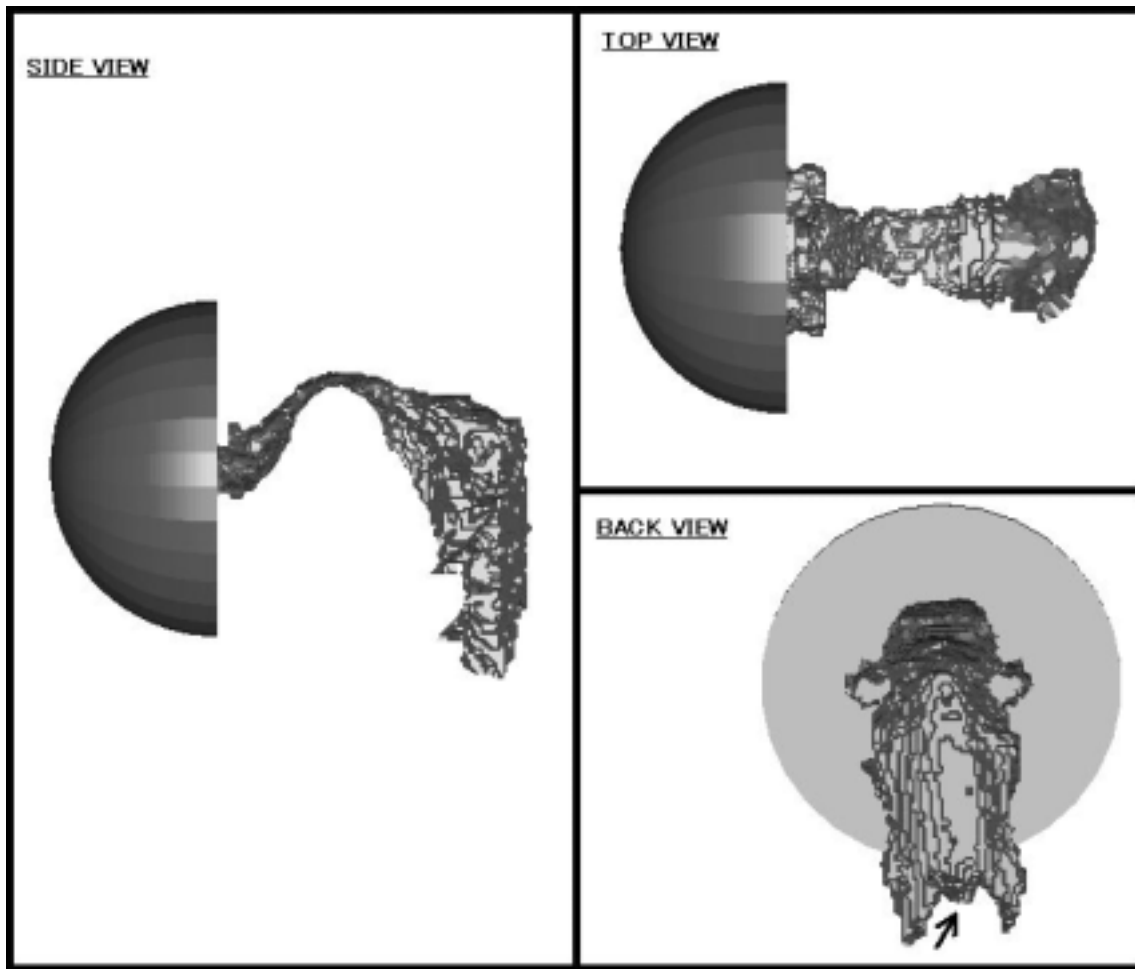
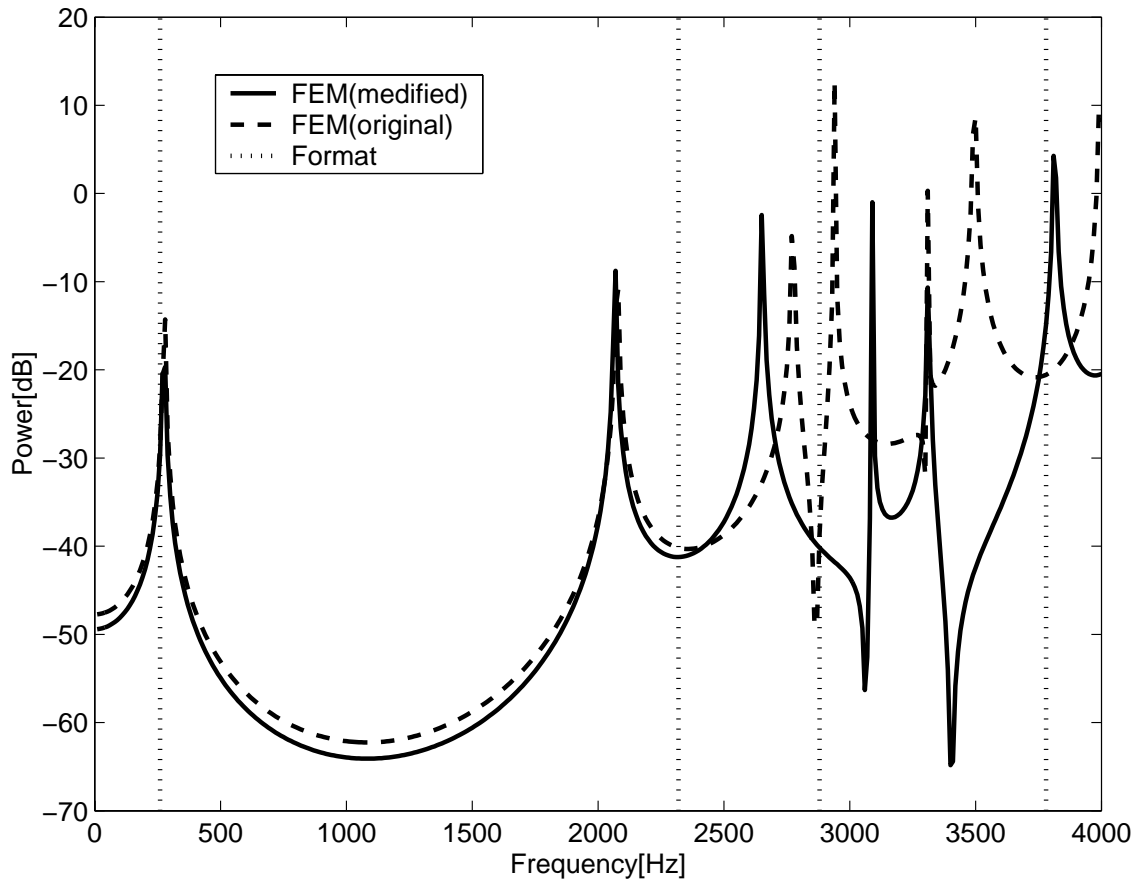


図 4.10: A Normal-A vocal tract model uttering Japanese vowel /i/. This model is cut at ventricular fold.



⊠ 4.11: A transfer function of a Normal-A vocal tract excited at the ventricular fold.

4.5.2 梨状窩を除去したモデル

梨状窩は咽頭室に対して背中側に声道管を挟むようにして左右に存在する対になったロート状の管である．音源である声帯に近いため音声中に影響を及ぼしやすく，声道の分岐となるために，音声スペクトル中の $4000[Hz]$ 付近に零点を与える．ここでは，声道モデルの梨状窩を取り除くことによる声道伝達特性の変化を調べる．

今回は 4.4.3 節で示したモデルの梨状窩を除去したモデルを構築した．これを 図 4.12 に示す．有限要素モデルの要素数は 94381，節点数は 22406 である．入力面は声帯部であり，出力面は放射球面の先端とした．このモデルの FEM によるシミュレーション結果を 図 4.13 に示す．また，シミュレーションの結果得られる声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 4.6 に示す．

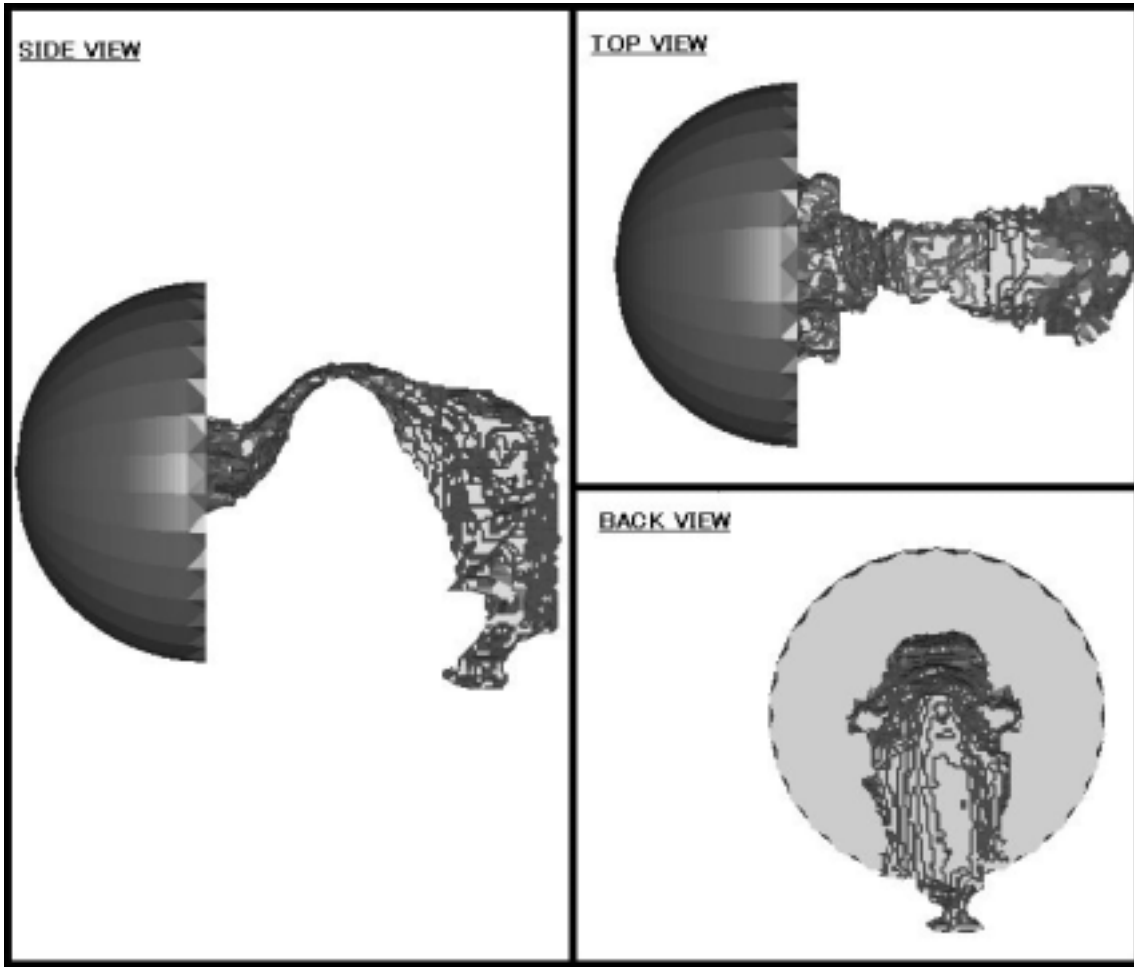
表 4.6: Peak frequencies of estimated vocal tract function and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	290	260	11.5
F_2	2350	2320	1.3
F_3	2700	2880	6.3
—	3350	—	—
F_4	3840	3780	1.6

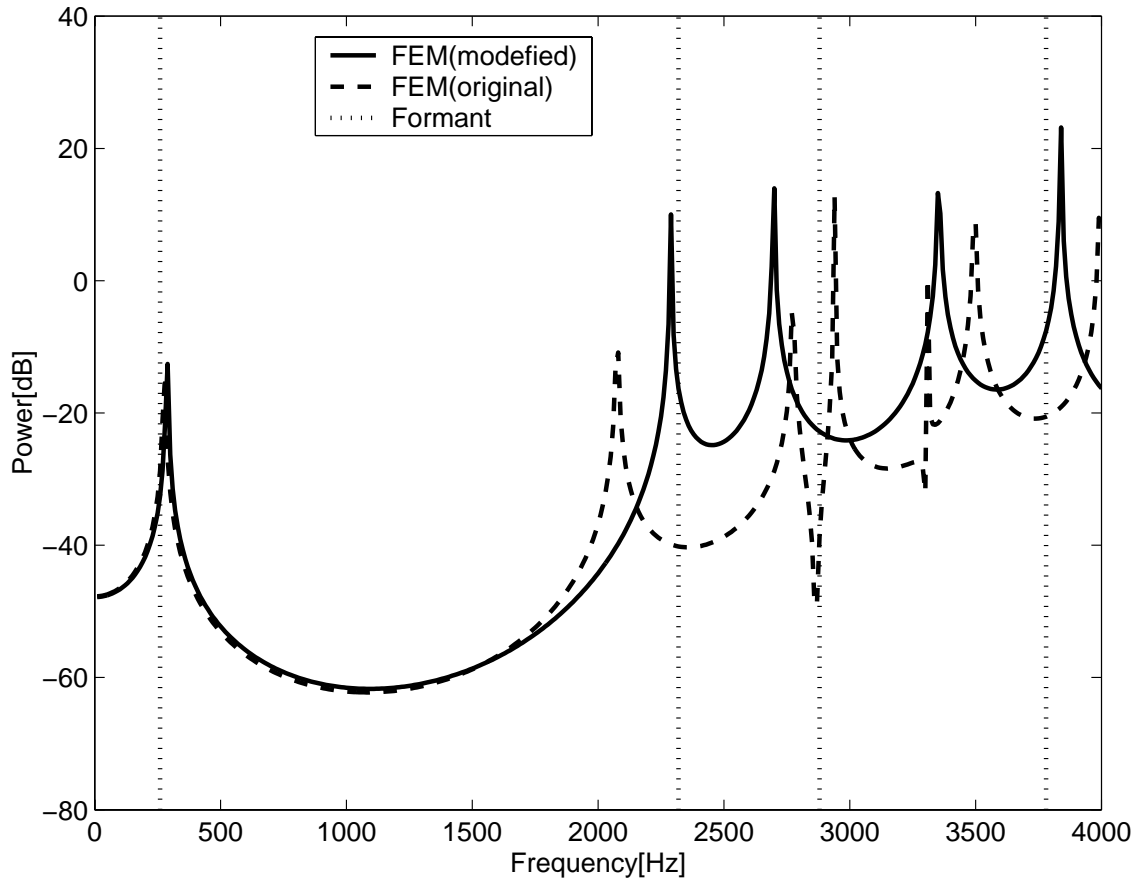
梨状窩を完全に除去した 3 次元声道モデルを用いた場合， F_2 は高周波数方向へ， F_3 は低周波数方向へシフトしている．シミュレーション結果の F_2 が実音声に近づいたことが分かる．

梨状窩は発声中に膨張，もしくは収縮を起こす器官である．また，MR 撮像の姿勢は仰臥位であり，MR 画像には写らないような唾液等の体液により，梨状窩が実際には声道の分岐としての役割を果たしていないことが考えられる．そのため， F_2 ， F_3 への影響を考慮しつつ，梨状窩を適度に切除した状態のほうが F_2 に関して実音声との対応がよくなるという可能性がある．

Normal-A の梨状窩は一般的な健常者よりも大きな体積である．そのため，梨状窩が分岐した声道という役割の他に，後室側の声道の一部の役割を果たしているという事が考えられる．日本語母音 /i/ の場合， F_2 は後室の影響をうけるため，梨状窩をなくしてしまうと F_2 に影響が出たと考えられる．



☒ 4.12: A Normal-A vocal tract model removed pyriform fossas.



⊠ 4.13: A transfer function of a Normal-A vocal tract removed pyriform fossas.

4.6 1次元モデルとの比較

ここでは、3次元声道モデルのFEMによる解析法と1次元近似モデルによる解析結果の比較を行う。使用する症例は Normal-A とし、4.4.3 節で示したモデルを用いた。このモデルのFEMと1次元近似モデルによるシミュレーション結果を図4.14に示す。

図4.14からFEMによる分析結果では、実音声の各フォルマント周波数に対応した形状となっている。しかし、1次元近似モデルによるシミュレーション結果では F_2 、 F_3 は低周波数方向へシフトしているが、だいたいの形状は表現できている。健常例に対する声道伝達関数の推定は、どちらの方法でも表現できることが確認された。

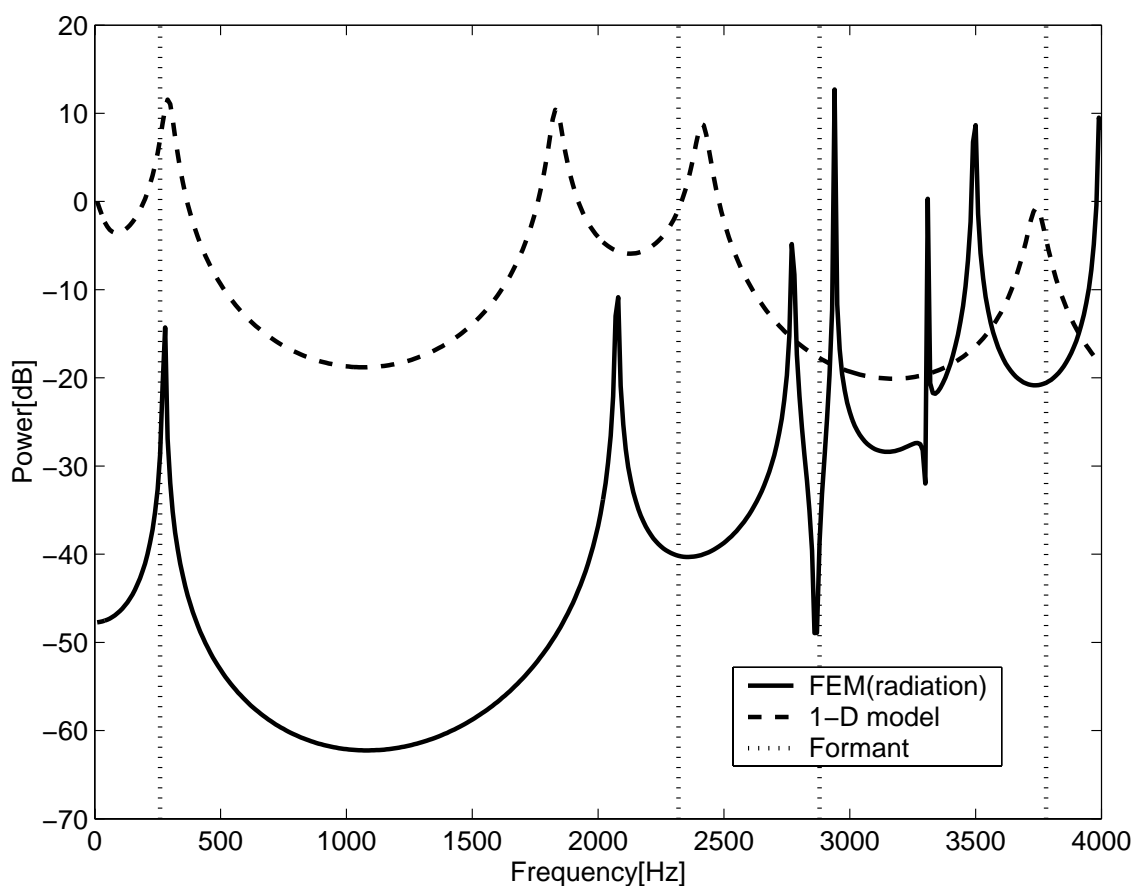


図 4.14: Transfer functions of Normal-A vocal tract estimated by FEM and 1-D model.

4.7 まとめ

本章では健常例 Normal-A, Normal-B の日本語母音 /i/ 発話時の 3 次元声道モデルでのシミュレーションを行い、実音声とシミュレーション結果の比較を行った。その結果、MR 画像から抽出する声道形状情報を $0.5[mm]$ 間隔でサンプリングしたものにすれば、音声の揺れや MR 撮像による計測法の誤差から本シミュレーションの妥当性が言えた。

次に、駆動面を声帯部にした 3 次元声道モデル、及び梨状窩を除去した 3 次元声道モデルによるシミュレーションを行った。その結果、仮声帯を駆動面に指定した 3 次元声道モデルでは声道伝達特性の F_1 , F_2 に変化は見られなかったが、それ以上の周波数帯では、ピーク周波数の位置に変化が見られた。音源の位置を変えることで F_3 以上の周波数帯でのフォルマント位置に変化がある事が分かった。梨状窩を除去した 3 次元声道モデルでは、 F_2 の実音声との対応の良さが向上した。 F_2 , F_3 への影響を考慮しつつ、梨状窩を適度に切除した状態のほうが F_2 に関して実音声との対応がよくなるという可能性がある。

さらに、本手法による分析結果と 1 次元近似モデルによる分析結果を比較した。その結果、健常者の場合は 1 次元近似モデルでも表現できるという事が示された。

第 5 章

舌・口底切除症例の 3 次元声道形状モデルによるシミュレーション

5.1 はじめに

口腔疾患症例である舌・口底切除患者 (Patient-A) と, 口底切除患者 (Patient-B) の MR 画像がある. 本章では, それより抽出された声道形状情報より 3 次元声道モデルを構築し, その FEM によるシミュレーション結果を示す. また, Patient-A の声道形状に変形を与えたモデルでのシミュレーション結果を示す.

5.2 Patient-A によるシミュレーション

5.2.1 PAP 非装着時のモデル

Patient-A は術後の機能障害に対して PAP が適用されており, PAP 装着, 非装着の場合で, それぞれ撮像を行った. 本節では PAP 非装着時の MR 画像を使用し, それにより構築される 3 次元声道モデルによるシミュレーションを行った. また, 放射球面の有無による声道伝達特性の変化を見るため, 放射球面がある声道モデルと放射球面がない声道モデルを用意し, それぞれに関してシミュレーションを行った.

なお, それぞれのモデルの口唇部は口角までであり, 歯の形状は考慮されていない. 放射球面がない Patient-A の 3 次元声道モデルを図 5.1 に, 放射球面がある Patient-A の 3 次元声道モデルを図 5.2 にそれぞれ示す. 放射球面のない有限要素モデルの要素数は 92365, 節点数は 22586 であり, 入力面は声帯部であり, 出力面は口唇部としている. 放

射球面がある有限要素モデルの要素数は 105038 , 節点数は 25488 である . 入力面は声帯部であり , 出力面は放射球面の先端としている . それぞれのモデルの FEM によるシミュレーション結果を図 5.3 に示す . また , シミュレーションの結果得られる声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 5.1 , 表 5.2 に示す .

表 5.1: Peak frequencies of Patient-A (PAP(-)) vocal tract function without hemispherical surface of radiation and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	340	340	0
F_2	1940	1450	33.7
F_3	2280	2300	0.9
F_4	3450	3920	12.0

表 5.2: Peak frequencies of Patient-A (PAP(-)) vocal tract function with hemispherical surface of radiation and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	330	340	3.0
F_2	1700	1450	12.8
F_3	2170	2300	5.6
F_4	3410	-	-
F_4	3540	3920	9.7

Patient-A (PAP(-)) の実音声の分析結果から , PAP が無い場合は構音が一定しないため , フォルマント周波数が一定せず , その影響は F_2 に大きく現れることが分かっている . そのため , ここでは音響分析結果との厳密な比較は困難であるが , Patient-A (PAP(-)) の特徴である 1500[Hz] 周辺にできる F_2 の存在との対比はできる . その点では F_2 が Patient-A (PAP(-)) の特徴通り表現できていると言える .

図 5.3 から , 放射球面を取り付けることによって , 声道伝達関数のピーク位置のシフトが見られる . これは , 声道形状が左右非対称であり , 特に舌・口底を切除したために口腔内での声道の形状が声道の軸に対して左右非対称になっているため , 口唇部での音圧分布

が場所により異っているためと考えられる．そのため，実際の音声の伝搬をモデル化するためには，口唇部からの放射を近似できるモデルである，放射球面付きの3次元声道モデルを使用する必要があることが分かる．

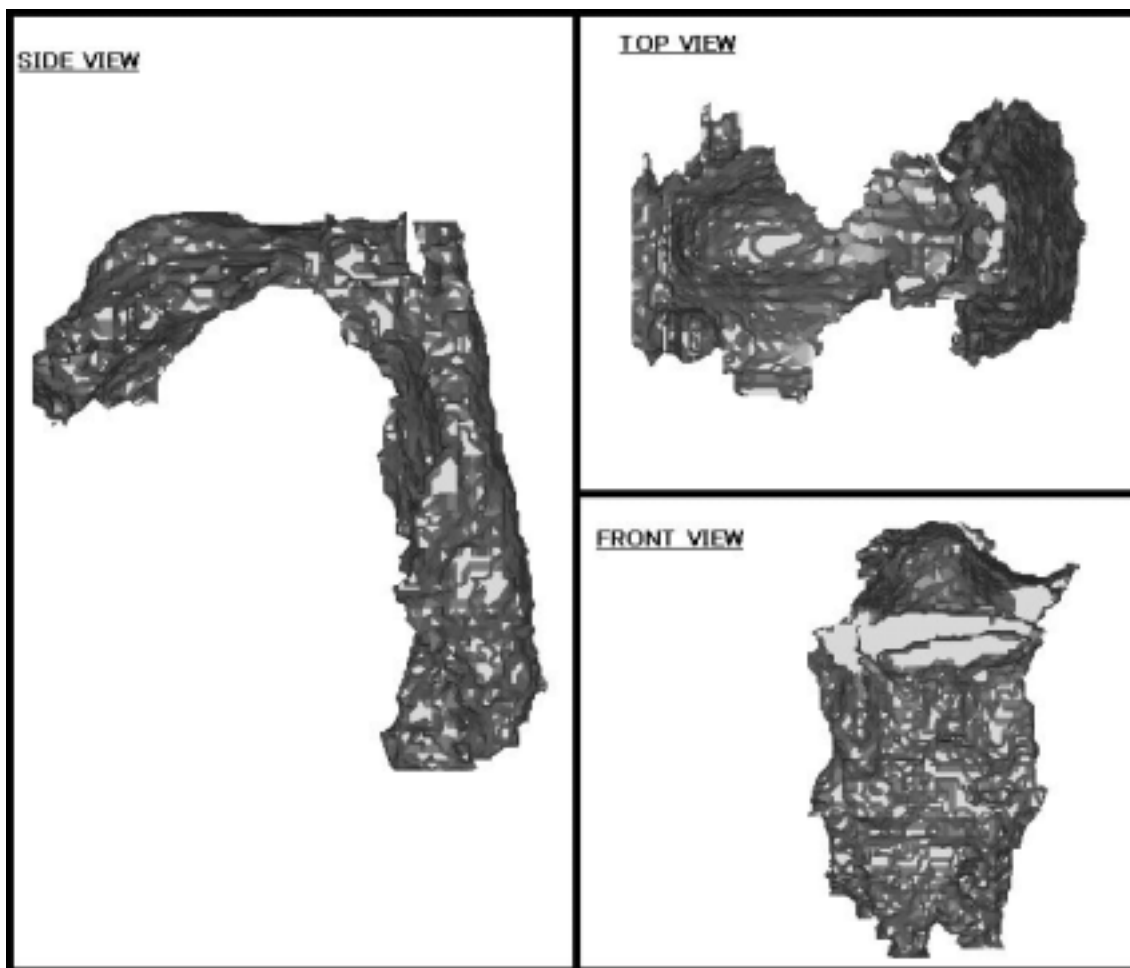
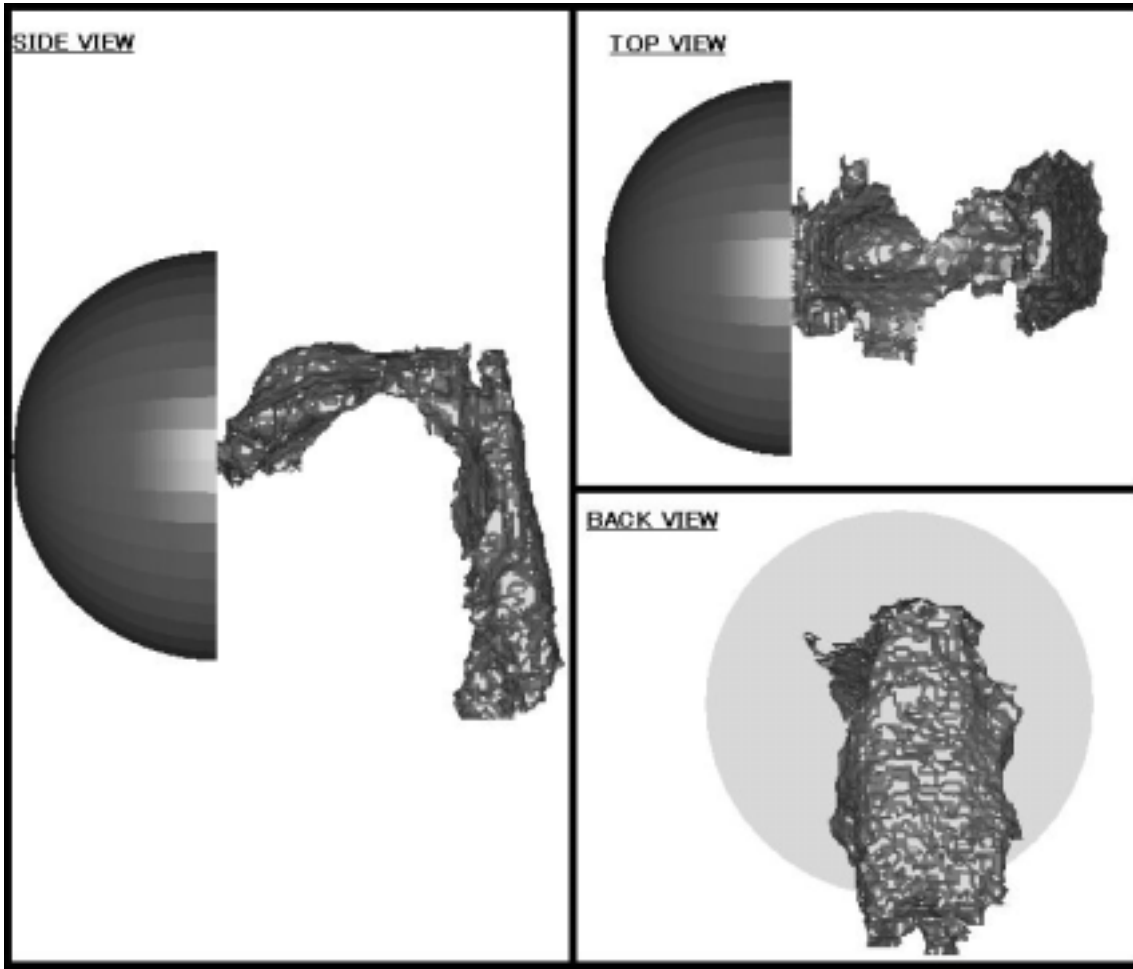
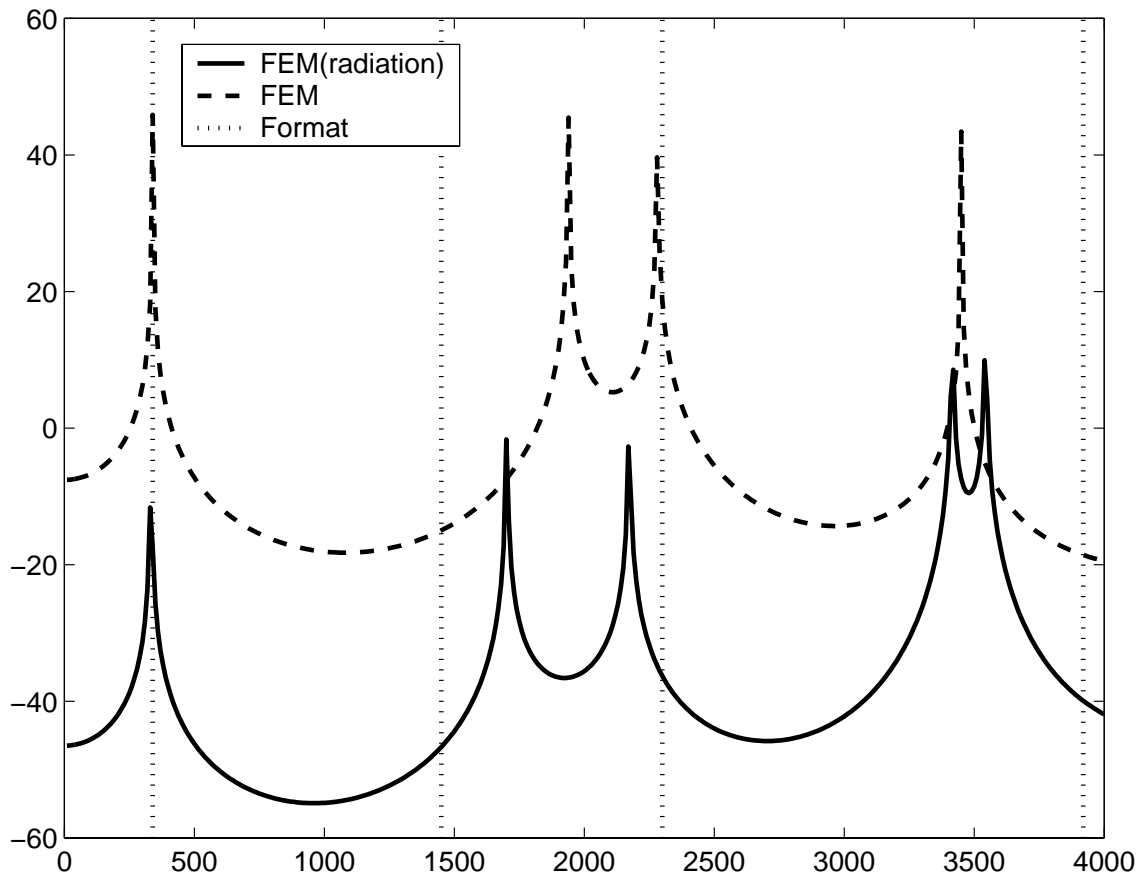


図 5.1: A 3-D vocal tract model of Patient-A(PAP(-)).



☒ 5.2: A 3-D vocal tract model of Patient-A(PAP(-)) with with hemispherical surface of radiation.



⊠ 5.3: A transfer function of the Patient-A(PAP(-)) vocal tract model with hemispherical surface of radiation and without hemispherical surface of radiation.

5.2.2 PAP 装着時のモデル

ここでは Patient-A の PAP 装着時の MR 画像を使用し，それにより構築される 3 次元声道モデルによるシミュレーションを行った．また，5.2.1 節と同様に放射球面の有無による声道伝達特性の変化を見るため，放射球面がある声道モデルと放射球面がない声道モデルを用意し，それぞれに関してシミュレーションを行った．

なお，それぞれのモデルの口唇部は口角までであり，歯の形状は考慮されていない．放射球面がない Patient-A の 3 次元声道モデルを図 5.4 に，放射球面がある Patient-A の 3 次元声道モデルを図 5.5 にそれぞれ示す．放射球面のない有限要素モデルの要素数は 83487，節点数は 21067 であり，入力面は声帯部であり，出力面は口唇部としている．放射球面がある有限要素モデルの要素数は 104698，節点数は 25553 である．入力面は声帯部であり，出力面は放射球面の先端としている．それぞれのモデルの FEM によるシミュレーション結果を図 5.6 に示す．また，シミュレーションの結果得られる声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 5.3，表 5.4 に示す．

表 5.3: Peak frequencies of Patient-A(PAP(+)) vocal tract function without hemispherical surface of radiation and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	320	310	3.2
F_2	1870	2100	11.0
F_3	3040	—	—
F_4	3500	3420	2.3

図 5.6 から，5.2.1 節と同様に放射球面を取り付けることによって，声道伝達関数のピーク位置のシフトが見られる．これも，声道形状が左右非対称であり，特に舌・口底を切除したために口腔内での声道の形状が声道の軸に対して左右非対称になっているため，口唇部での音圧分布が場所により異っているためと考えられる．また，放射球面を取り付けることによって実音声との整合がよりよくとれていることが分かる．

表 5.4: Peak frequencies of Patient-A(PAP(+)) vocal tract function with hemispherical surface of radiation and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	320	310	2.9
F_2	1860	2100	11.4
F_3	2730	—	—
F_4	3210	3420	6.1
F_4	3730	—	—

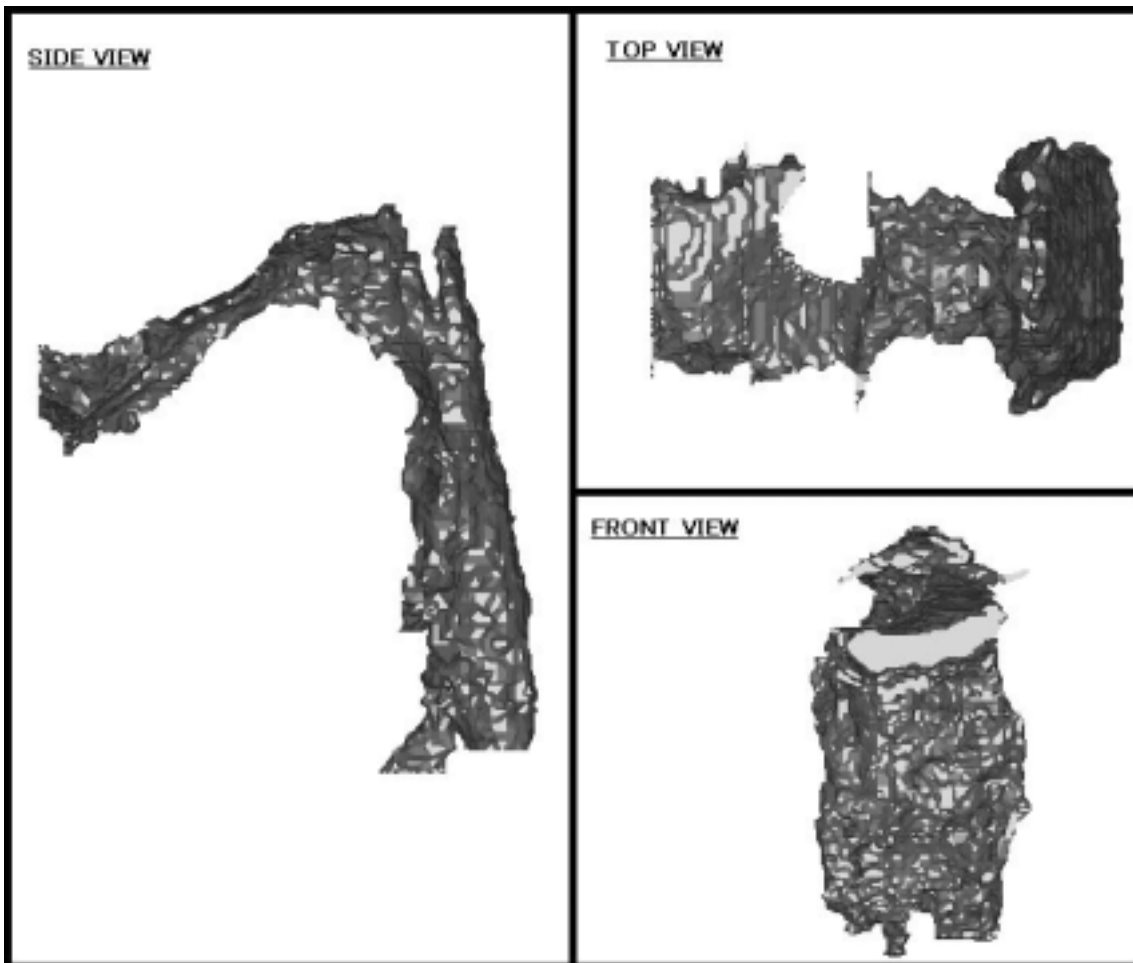
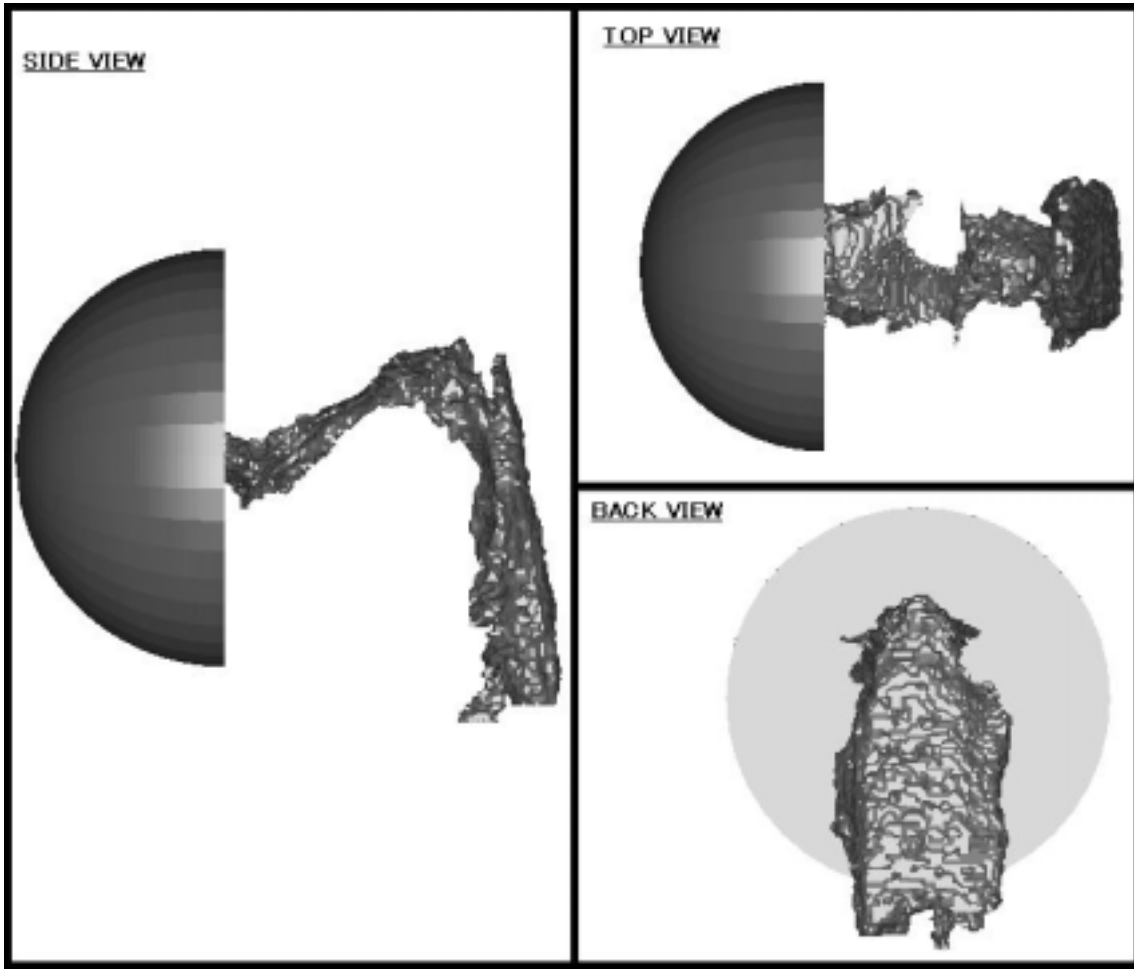


图 5.4: A 3-D vocal tract model of Patient-A(PAP(+)).



☒ 5.5: A 3-D vocal tract model of Patient-A(PAP(+)) with with hemispherical surface of radiation.

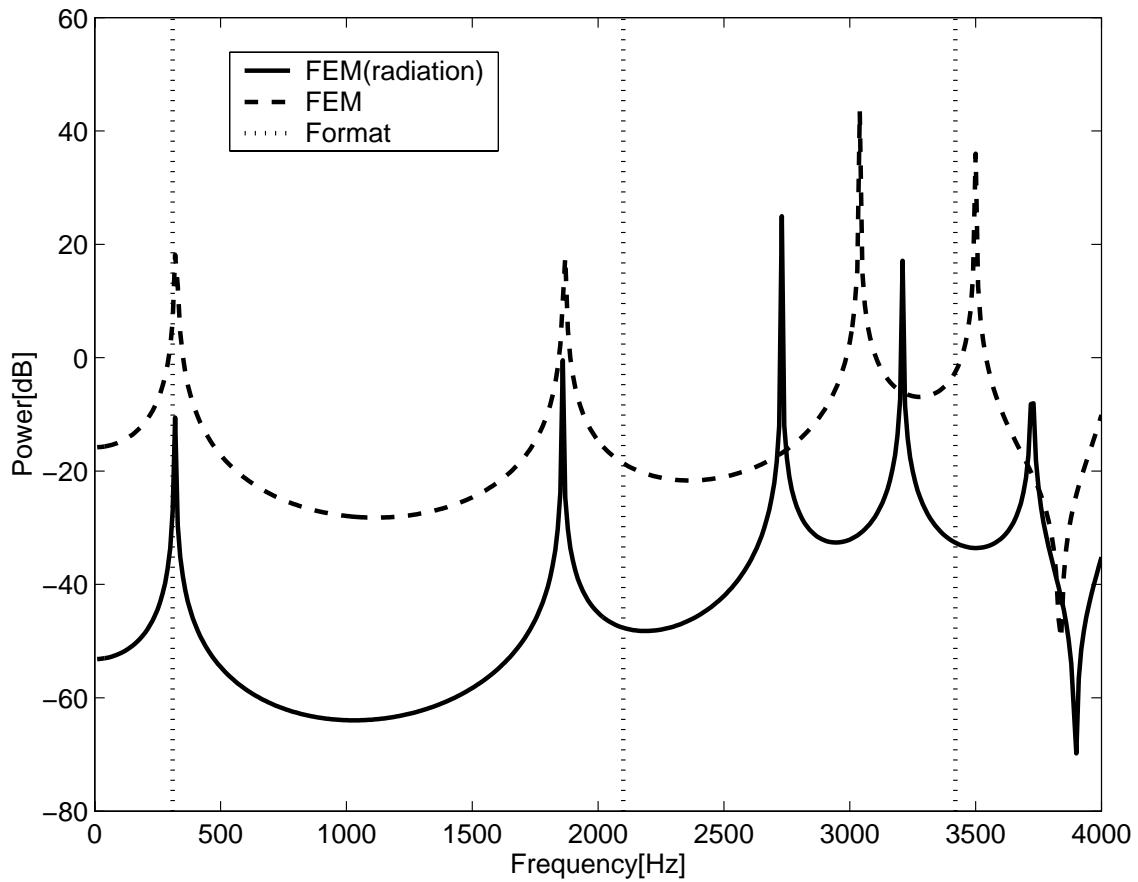


Figure 5.6: A transfer function of Patient-A(PAP(+)) vocal tract model with hemispherical surface of radiation and without hemispherical surface of radiation.

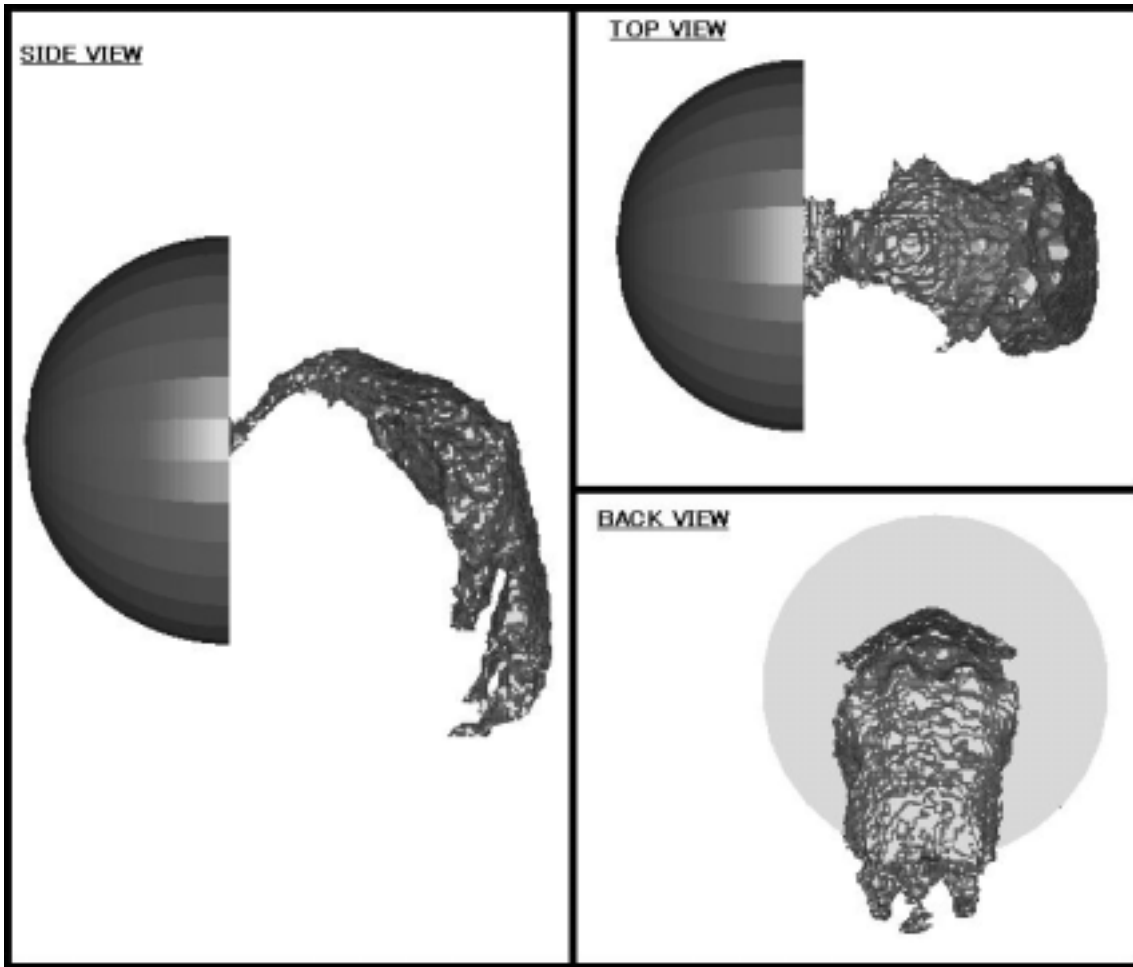
5.3 Patient-B のモデルによるシミュレーション

ここでは Patient-B の 3 次元声道モデルによるシミュレーションを行った。口唇部の放射を近似するために、放射球面を取り付けた声道モデルを用意し、それぞれに関してシミュレーションを行った。なお、このモデルの口唇部は口角までであり、歯の形状は考慮されていない。Patient-B の 3 次元声道モデルを図 5.7 にそれぞれ示す。放射球面のある有限要素モデルの要素数は 142307、節点数は 33890 であり、入力面は声帯部、出力面は放射球面の先端としている。この 3 次元声道モデル FEM によるシミュレーション結果を図 5.8 に示す。また、シミュレーションの結果得られる声道伝達特性のピーク周波数と実音声のフォルマント周波数の比較を表 5.5 に示す。

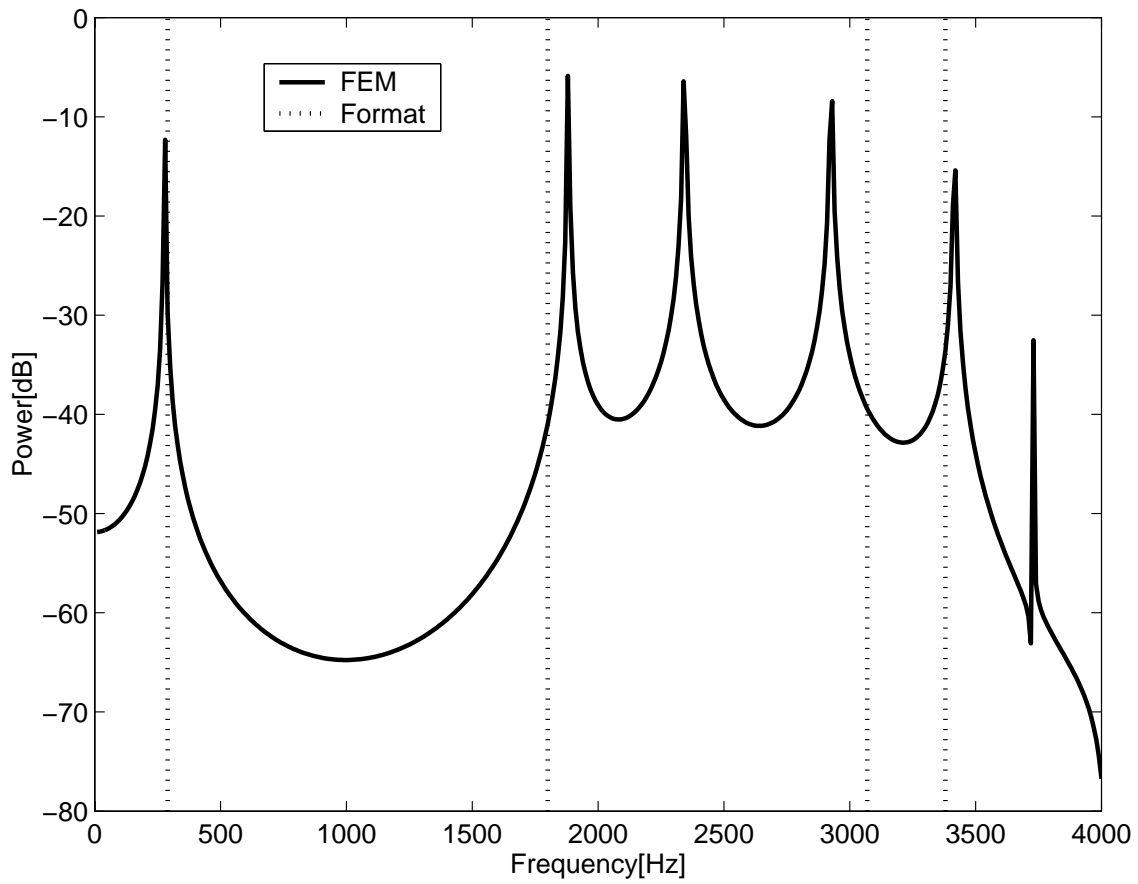
表 5.5: Peak frequencies of Patient-B vocal tract function with hemispherical surface of radiation and formant frequencies.

	Molel[Hz]	Formant[Hz]	Error[%]
F_1	280	290	3.4
F_2	1880	1860	1.1
F_3	2340	—	—
F_4	2930	3070	4.8
F_4	3420	3380	1.2
—	3730	—	—

シミュレーションにより推定された声道伝達特性の F_1 , F_2 に関して言えば、MR 画像の計測誤差と実音声のフォルマントの揺れを考慮すると誤差の範囲内であると言える。しかし、それ以上の周波数帯域ではフォルマント周波数と声道伝達特性のピーク周波数が一致していない。これは、Patient の場合、PAP が無い場合では構音が一定しないため、撮像時間内に声道形状が変化してしまっている可能性がある。また MR 撮像時は仰臥位であるが、音声収録時の立位とは異なっており、それによる影響が出ている可能性がある。



☒ 5.7: A 3-D vocal tract model of Patient-B.



⊠ 5.8: A transfer function of a Patient-B vocal tract model with hemispherical surface of radiation and without hemispherical surface of radiation.

5.3.1 1次元近似モデルとの比較

ここでは, Patient-1(PAP(-)) のデータを用いて 1次元近似モデルによる分析を行い, その結果と, 5.2.1節の FEM による分析結果と比較した.

図 5.9より, 音響分析結果では $0 - 4000[\text{Hz}]$ の間にフォルマントは 4 つ存在するが, 本手法により得られた声道伝達特性は同じ周波数帯域にピークが 4 つ存在していることが分かる. 3章で述べたとおり, 実音声のフォルマント周波数が一定にならないということ, そして本手法での MR 画像の計測誤差は $\pm 5[\%]$ である [6] ということ を考慮すれば, 誤差の範囲内であり, FEM による分析は実音声との対応が良くとれていると言える. 一方, 1次元近似モデルでは実音声に見られる F_2, F_3 近辺にピークが 1 つしかない. これは, 口腔内の異常な形状, 特に舌による狭めの位置のずれや前室の水平方向の広がり起因する音声の歪みを表現できていないためと考えられ, 本手法の優位性が示された.

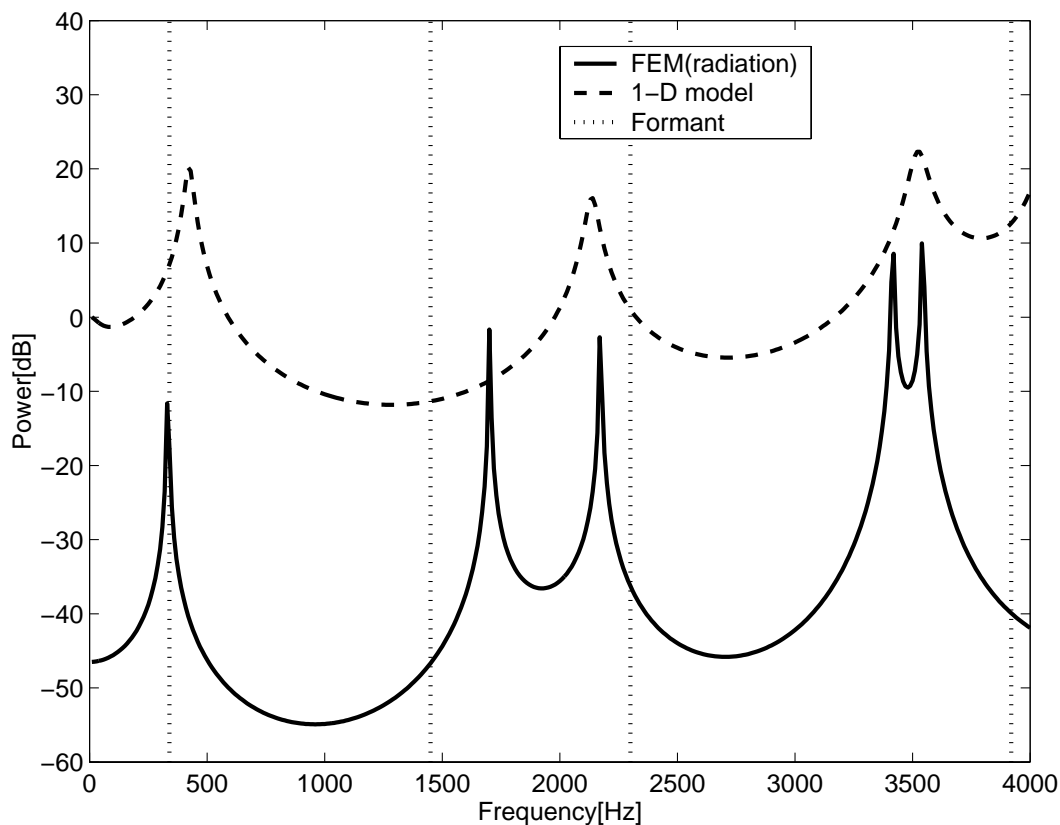


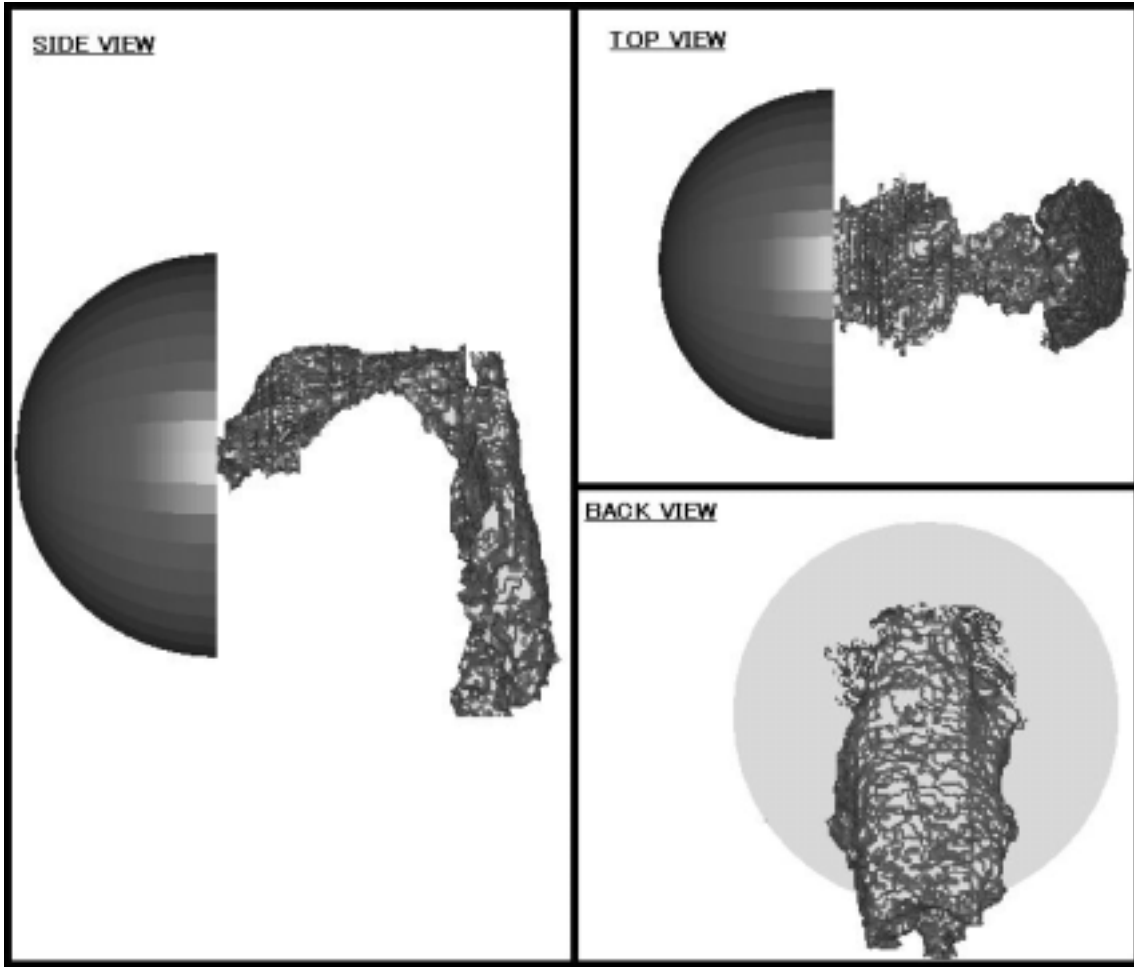
図 5.9: Transfer functions of a Patient-A(PAP(-)) vocal tract model simulated by FEM and 1-D model.

5.4 口腔内の声道形状を変化させたモデルによるシミュレーション

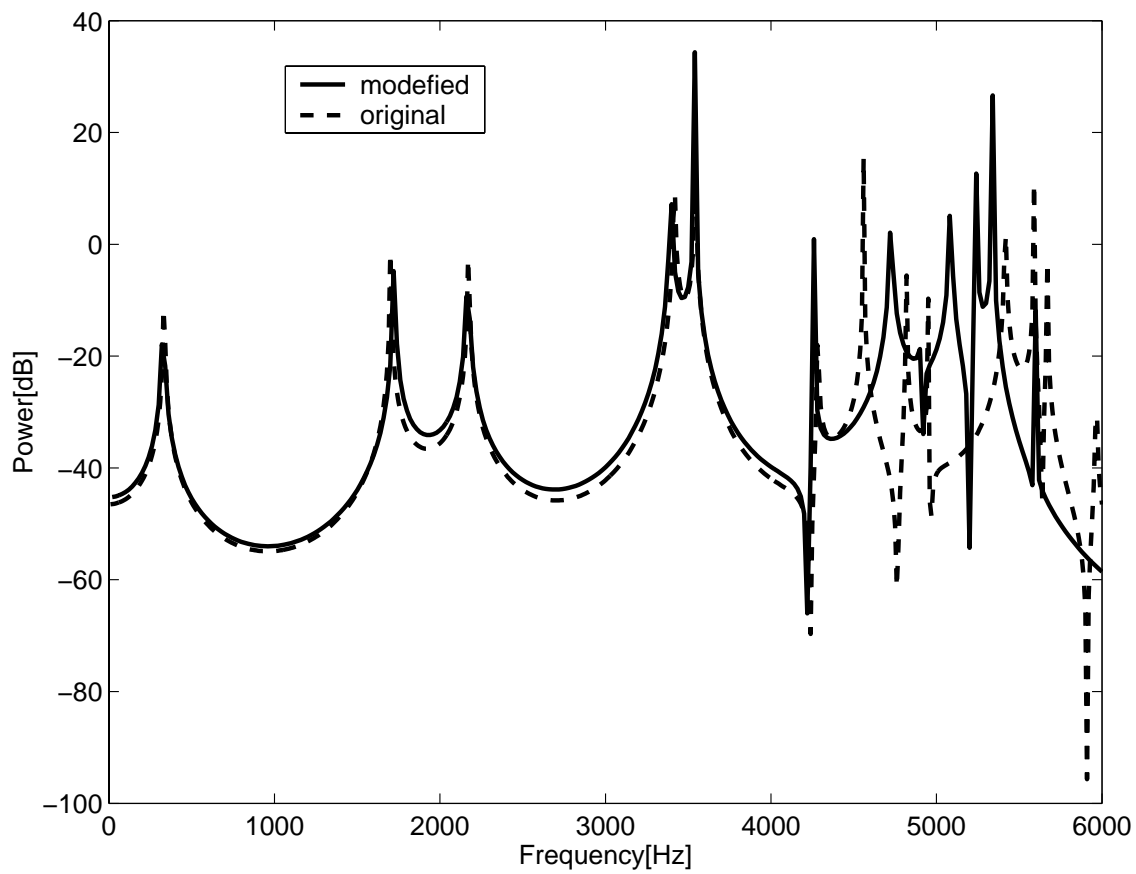
口腔疾患の患者の口腔内の異常な声道形状が音声スペクトルに与える影響を調べるためのシミュレーションを行った。Patient-A(PAP(-))の口腔内の声道管経路は声道の中心軸に対してやや右手側を通っている。そこで、Patient-A (PAP(-))の3次元声道モデル(図5.2)の口腔内に相当する部分に変形を与え、声道の中心軸を基準としてスライドさせたモデルを構築した。それを図5.10に示す。有限要素モデルの要素数は142307、節点数は33890であり、入力面は声帯部、出力面は放射球面の先端としている。また、高周波成分の変化を見るために分析範囲を $10 - 6000[H_z]$ とした。この3次元声道モデルFEMによるシミュレーション結果を図5.11に示す。

推定された声道伝達特性から図5.10から、 $4000[H_z]$ までは余り変化が見られなかった。これは、Patient-A (PAP(-))の口腔内の声道の中心軸に対するずれがあまり大きくなく、モデルの口腔内の声道をずらしても、実質的な声道管経路に変化がないため、声道長にほとんど変化を与えなかったためであると考えられる。一方、 $4000[H_z]$ 以上のオリジナルの声道モデルでのシミュレーション結果との間で違いが見られる。この周波数領域では声道長よりも声道表面の滑らかさ、連続性の影響を受けやすく[18]、変形を与えたモデルでは二セの高次元モードを作り出してしまっていると考えられる。

よって、以上の結果からはPatient-A (PAP(-))のPAPの設計では、声道管経路以外の部分(舌による狭めの幅、狭めを作る位置)に着目した設計が重要であると言える。



☒ 5.10: A modified 3-D vocal tract model of Patient-A (PAP(-)).



⊗ 5.11: A transfer function of a modified vocal tract model of Patient-A (PAP(-)).

5.5 まとめ

本章では口腔疾患を有する患者の3次元声道モデルのFEMにより得られる声道伝達特性と、実音声との比較を行った。その結果、放射空間を近似できる放射球面を声道モデルに取り付ける必要があることがわかった。

本シミュレーション法と1次元近似モデルのPatient-Aによる分析結果の比較を行った。健常者と比較して声道が複雑な構造をしている口腔疾患を有する患者の声道形状をシミュレーションするには1次元近似モデルでは不十分であることが確認された。

Patitent-A (PAP(-)) のモデルの口腔部に変形を与え、それによる声道伝達特性の変化を調査した。その結果、口腔内の形状そのまま音響管の流路だけをずらした場合、 $10 - 4000[Hz]$ の間での声道伝達特性の変化は見られなかった。Patient-A(PAP(-)) の F_2 を健常者のものに近づけるようにするためには、流路をずらす以外の要素が必要であることが分かった。

第 6 章

結論

6.1 本研究の成果

本研究で得られた成果としては以下のことが言える．

- 健常者のシミュレーション結果より，モデルの精度を上げれば実音声の結果と良く一致するようになり， F_1 を実音声と良く対応させるためには特に舌の狭めの部分のサンプリング間隔は $0.5[mm]$ にする必要がある．
- 口腔疾患を有する患者の 3 次元声道モデルを FEM で分析した．その結果得られる声道伝達特性は，音響分析結果と比較すると口唇部からの放射を考慮したモデルにする必要がある．
- 健常者の 3 次元声道モデルに変形を与えた結果，梨状窩は F_2 にも影響を及ぼしているという可能性が示された
- 本手法と 1 次元近似モデルによるシミュレーション法との比較の結果，本手法の優位性が示された．
- 口腔疾患を有する声道モデルに変形を与え，その声道伝達関数を調べた．その結果， $10 - 4000[Hz]$ の間では声道の流路をずらす事による影響が無く， $4000[Hz]$ 以上の周波数帯域で影響を及ぼす．

構音が一定しており，フォルマント周波数の時間的な揺れが少ない健常者を対象とすることで，音響分析結果と比較した結果，本研究で用いた手法の妥当性を示すことができた．また，舌・口底切除症例の場合，1 次元近似モデルと比較した結果，本シミュレーショ

ン法の優位性が示された。これは、本手法が声道の複雑な形状、特に舌・口底切除患者の声道形状が音声スペクトルに及ぼす影響を表現できているためである。また、声道形状に変形を与えることで音声スペクトルの変化が現れることがわかった。舌・口底切除例の3次元声道モデルに対して変形を加えることで、変形を加えた部位が音声に及ぼす影響を調査することができた。これにより、言語障害の治療や PAP の設計への応用が可能であると言える。

6.2 今後の課題

本研究では口腔疾患を有する声道形状と音声スペクトルの関係を調査するための基礎的な研究を行った。今後、更に本研究を発展させるために以下の項目について、更に研究を進める予定である。

- 本研究で用いた3次元声道モデルは口唇部先端が前歯の付け根となっている。そこで、さらに実際の現象に近いシミュレーションを行うために、唇の形状を3次元声道モデルに再現させる予定である。
- 梨状窩を完全に除去することでシミュレーション結果と音響分析結果の差が小さくなることが分かった。そこで、仰臥位での発声時の声道後室の状態を調査し、その結果を3次元声道モデルの作成に反映させていく。
- 口腔疾患を有する患者の声道モデルに変形を与えることで声道伝達特性の変化を見ることができたが、今後は舌による狭めの幅や位置、長さなど他の要素に着目した変形を与え、それによる声道伝達特性の変化を調査する。
- 他の症例の3次元声道形状モデルに変形を与えることで、口腔疾患を有する患者の音声の歪みの原因や、PAP の設計に関する指針が得る。
- 本研究では音声スペクトルと声道伝達特性の關係に着目した。今後は、音圧分布など他のパラメータにも注目したシミュレーションを行い、各構音器官が音声に及ぼす影響を明らかにしていくことを目標とする。そして、声道構造と音声の音響的特徴の關係を導きだし、その一般化を行う。

謝辞

本研究を遂行するにあたり、終始、数多くの貴重な御助言、御指導を頂きました北陸先端科学技術大学院大学 情報科学研究科赤木正人 教授、並びに本学の教官の皆様に深く感謝いたします。本研究を進めるにあたって、MR 画像を提供して頂き、多くの貴重な助言をして頂きました昭和大学 歯学部 鈴木規子 助教授ならびに斎藤浩人 助手に感謝いたします。本研究で用いた声道形状データを提供して頂き、多くの助言をして頂きました静岡大学 情報学部 北村達也助手に感謝いたします。MR 撮像に立ち会わせて頂き、貴重な助言を頂いた昭和大学病院 高橋 俊行 氏に感謝いたします。日頃から大いなる議論と激励を頂きました赤木研究室の皆様、友人、ほか全ての皆様に感謝いたします。

最後に、2年間の研究生生活を暖かく見守っていただいた両親、妹に心より感謝いたします。

参考文献

- [1] Chengxiang Lu, Takayoshi Nakai, Hisayoshi Suzuki, “Finite element simulation of sound transmission in vocal tract,” The journal of the acoustic society of Japan (E), 14, 2, pp63-72, 1993
- [2] 松崎 博季, 三木 信弘, 小川 吉彦, “音響管内の 3 次元の音響的特徴量の有限要素法による解析,” 電子情報通信学会論文誌 A, Vol J80-A, No. 9, pp1351-1359, 1997
- [3] Hiroki Matsuzaki, Nobuhiro Miki, Yoshihiko Ogawa, “FEM analysis of sound wave propagation in the vocal tract with 3-D radiational model,” The journal of the acoustic society of Japan (E), 17, 3, pp163-166, 1996
- [4] Takuya Niikawa, Eri Kawano, Masafumi Matsumura, “Acoustic characteristics of three-dimensional vocal tract shapes measured by MRI during vowel production,” Hokkaido Workshop on speech production, pp8-9, 1998
- [5] 富山 慶一 “有限要素法による 3 次元声道伝達特性の推定に関する研究,” 北陸先端科学技術大学院大学修士論文, 2000
- [6] 齋藤 浩人, 鈴木 規子, 藤田 幸弘, 道 健一, 高橋 俊行, “MR 撮像法を用いた 3 次元声道形状の計測 -研究方法および健常人の声道断面積の検討-”, 日本口腔科学会雑誌, 49, pp.92-101, 2000
- [7] 北村 達也, 鈴木 規子, 齋藤 浩人, 道 健一, 高橋 俊行, “MRI による舌・口底切除患者の 3 次元声道形状の計測の試み,” 日本音響学会講演論文集, pp305-306, 1999
- [8] 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・口底切除患者の声道非対称性の計測,” 日本音響学会講演論文集, 2001

- [9] 和久本 雅彦, 正木 信夫, 党 建武, 本多 清志, 藤本 一郎, 中村 祐二, 島田 育廣, “造影プレートを用いた歯冠部造影法とその構音への影響,” 日本音響学会講演論文集, pp397-398, 1997
- [10] 古市 千枝子, 今井 聖, “対数スペクトルの不偏推定,” 信学論, Vol.J70-A, No.3, pp.471-480, 1987
- [11] 音声信号処理, “音声の性質と聴覚の特性を考慮した信号処理,” pp.152-162, 森北出版株式会社, 1996
- [12] 高橋 信之, 中井 孝芳, 鈴木 久喜, “有限要素法による音響管の要素分割と計算精度,” 静岡大学大学院電子科学研究科研究報告第 18 号, pp41-44, 1996
- [13] Man Mohan Sondhi, JUERGEN SCHROETHER, “A hybrid time-frequency domain articulatory speech synthesizer,” IEEE Transactions on acoustics, speech, and signal processing, Vol. ASSP-35, No. 7, July, pp955-967, 1987
- [14] Man Homan Sondhi, “Model for wave propagation in lossy vocal tract,” The journal of the acoustic society of America, Vol. 55, No. 5, pp 1070-1075, 1974
- [15] Jouji Miwa, “Interactive visualization and auralization of speech production using variable vocal and nasal area function,” ASVA97, pp271-278, 1997
- [16] 日本音声言語医学会, “声の検査法 基礎編,” pp41, 60-62, 医師薬出版株式会社, 1994
- [17] Hiroki Matsuzaki, Nobuhiro Miki, Yoshihiko Ogawa, “FEM analysis of sound wave propagation in the vocal tract with 3-D radiational model,” The journal of the acoustic society of Japan (E), 17, pp.163-166, 1996
- [18] 三木 信弘, “声道のはたらきを見る,” 日本音響学会誌, 56, pp498-502, 2000
- [19] 齋藤 浩人, 鈴木 規子, 北村 達也, 赤木 正人, 道 健一, “舌・口底切除患者における異常構音の音響的特徴-スペクトルピーク分析の試み-,” 信学技報, SP98-149, 25-31, 1999
- [20] 電子情報通信学会, “聴覚と音声,” pp259-274, 株式会社コロナ社, 1980
- [21] 粕谷 英樹, 楊 長盛, “音源から見た声質,” 日本音響学会誌, 51, pp869-875, 1995
- [22] 松田 勝敬, 楊 長盛, 粕谷 英樹 “フォルマント合成における音源に関連した声質の変換法,” 日本音響学会講演論文集, pp193-194, 2000

学会等発表リスト

- 西本 博則, 赤木 正人, 北村 達也, 鈴木 規子, 齋藤 浩人, 道 健一, 高橋 俊行,
“舌・口底切除患者の3次元声道形状モデルの有限要素法による分析,”
音響学会春期発表会, 3-Q-28(2001)
- 西本 博則, 赤木 正人, 北村 達也, 鈴木 規子, 齋藤 浩人, 道 健一, 高橋 俊行,
“舌・口底切除患者の3次元声道形状モデルの有限要素法による分析”
聴覚・音声研究会 (2001)