



デジタルウェーブガイドメッシュを用いた3次元MRIデータからの声道断面積関数の抽出

井上 謙次 (大阪工業大院・情)

竹本 浩典 (ATR-Promotions/ATR-CIS)

北村 達也 (甲南大)

正木 信夫 (ATR-Promotions)

中嶋 鴻毅 (大阪工業大・情)



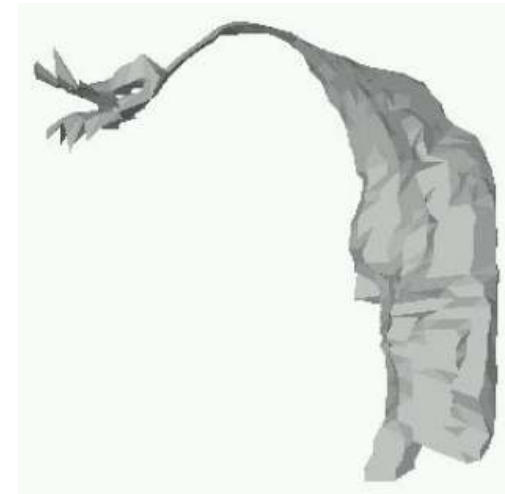
目次

- 研究背景
- 研究目的
- 提案手法
- 実験
- 結果と考察
- 結論

研究背景

- Articulatory synthesis -

- 高品質な音声の合成技術として期待
- ヒトの音声生成過程・調音機構をモデル化
 - 声道形状を利用
- 近年の磁気共鳴画像法 (MRI) の発展による精微な3次元声道モデル

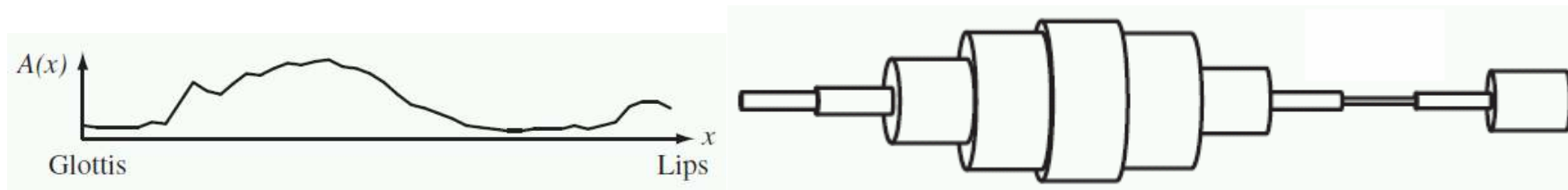


母音 /i/ の3次元声道形状
[Kröger 2000]

研究背景

- 声道断面積関数 -

- 3次元声道モデル内の音波伝播の支配方程式(波動方程式): 実時間では計算が困難
- 3次元 声道形状 → 1次元 声道断面積関数



母音 /i/ の声道断面積関数(左)とその円筒音響管アナロジー(右), 左端が声門, 右端が口唇 [Mullen 2006]

研究背景

- 声道断面積関数の推定手法 -

- 1次元の声道断面積関数
 - 声道中を平面波が伝播すると仮定
 - 声門から口唇へのその平面波の面積の変化
- 既存の推定手法
 - ほとんどが幾何学的に求めている
(音波伝播を考慮していない)
 - 音波伝播を考慮した有限要素法(FEM)を用いた方法
[中井他 2002, Mochizuki et al. 2007]
 - 周波数領域で求めているためやや複雑
 - 正中矢状面(2次元)のみを利用



研究目的

- 音波伝播を計算し, その伝播経路に沿って声道断面積関数を求める
 - 時間領域で計算
- 既存の幾何学的手法との比較
 - 領域二分法 [Story et al. 1996]
 - マンハッタン距離による方法 [Takemoto et al. 2006]
 - エッジ間最短距離点による方法 [Referred in Mochizuki et al. 2007]

提案手法

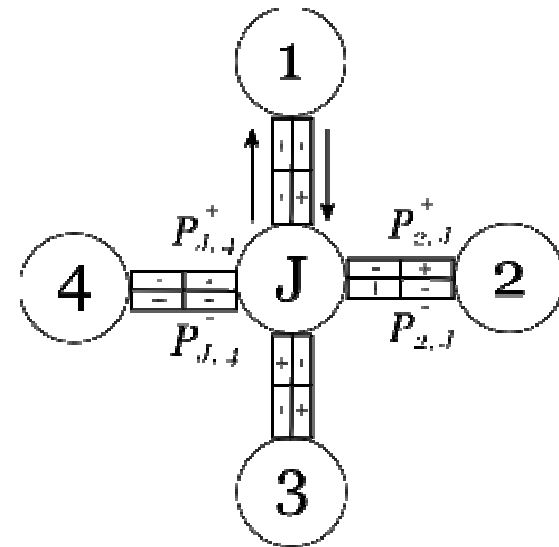
- デジタルウェーブガイドメッシュ -

- 3次元声道形状に対して、デジタルウェーブガイドメッシュ(DWM)を用いて支配方程式(波動方程式)を時間領域で離散化

- 境界条件:
声道内部(空気)と声道壁では特性インピーダンスが異なる

- 結果:
$$p_k^- = r_k p_k^+ + \sum_{i=1, i \neq k}^M (1 + r_i) p_{J,i}^+$$

$$r_k = \frac{2 \frac{1}{Z_k} - \sum_{i=1}^M \frac{1}{Z_i}}{\sum_{i=1}^M \frac{1}{Z_i}}$$



2次元DWMにおける
ジャンクションJ
(ジャンクション数 M=4)

提案手法

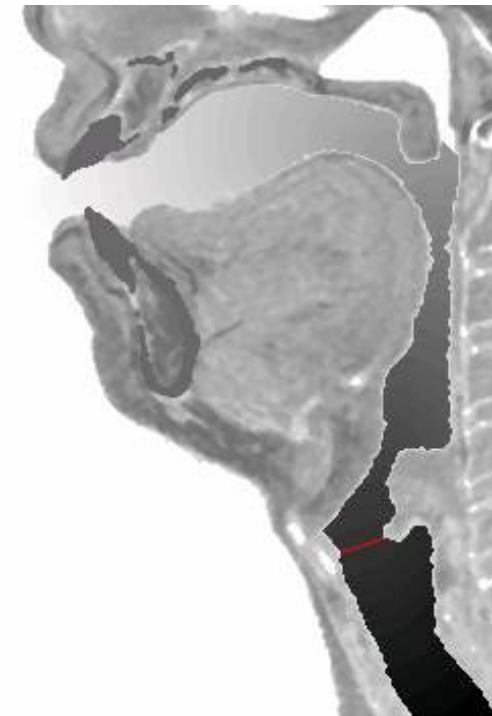
- シミュレーション条件 -

- パラメータは初期圧力とインピーダンス
- 初期状態として声門全体に総音圧 1.0 を付与
- 各ボクセル(画素)にジャンクションを設定
 - 空気と声道壁には別々のインピーダンスを設定
 - 今回は簡単のため声道壁の振動を省略
 - 空気から声道壁への圧力は吸収される(反射は存在)ように設定

提案手法

- 距離マップの算出 -

- 声門からの距離マップを算出
(非定常状態における等音圧面)
 - 通過エネルギーの総量が
閾値 $T_{dm}(t)$ を超えた
ステップ時刻 t を記録
 - 2次元: $T_{dm}(t) = 0.01/N_g(t^2 + 1)$
 - 3次元: $T_{dm}(t) = 0.1/N_g(t^3 + 1)$
- Ng: 声門のジャンクション数



声門上部が正
声道下部が負

提案手法

- 中心線の算出 -

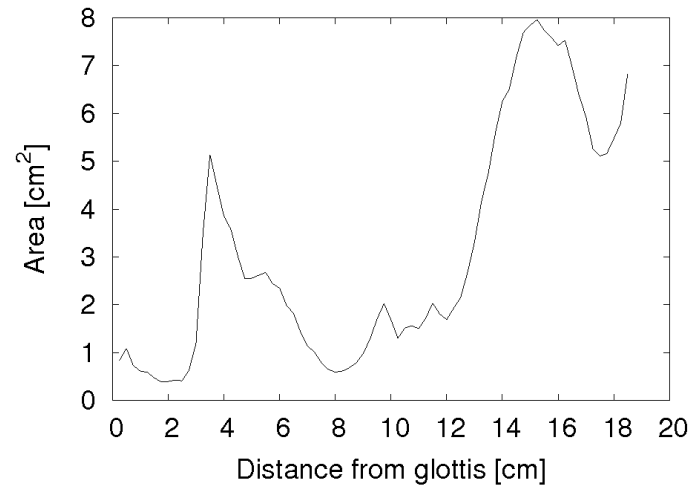
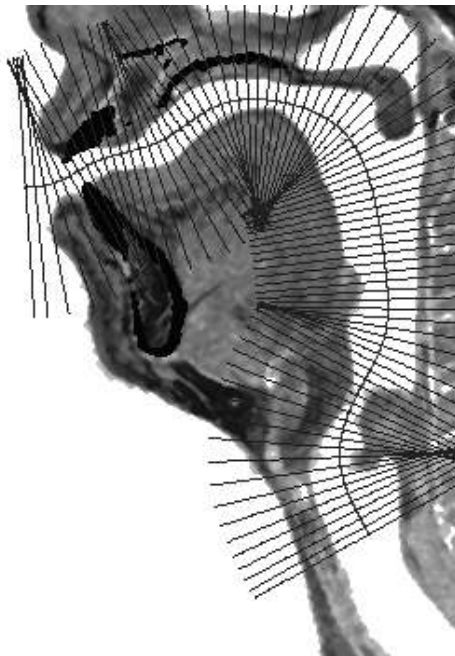
- 各ステップ距離(等距離面)における点集合の重心を計算
- スムージング: 一定距離ごとに再度重心を計算



提案手法

- 断面積関数の抽出 -

- 中心線に沿って断面を取る
- 断面上の空気領域のボクセル数から面積を求める





実験

- 本提案手法の有効性を示すために2つの実験を実施
 - 実験1: 有効なDWM法のパラメータの調査
 - 実験2: 既存手法とDWM法との比較

実験

- 使用データ -

- 使用した3次元MRIデータ
 - 日本語5母音 /aeiou/ (成人男性1名)
 - “ATR母音発話MRIデータ” (ATR-Promotions)
- 声道模型の音響特性と比較可能 [北村他 2007]
 - 上記MRIデータを基に製作



母音 /a/ の声道模型 [北村 2007]

実験1

- DWM法のパラメータ -

- DWM法のパラメータ
 - 2次元(正中矢状断面)および3次元
 - 距離マップの閾値: 固定
 - 空気のインピーダンス = 1.0 : 固定
 - 声道壁のインピーダンス =
1(吸収), 100(中間), 10000(凡そ実際)
- 等音圧面が算出可能かどうか

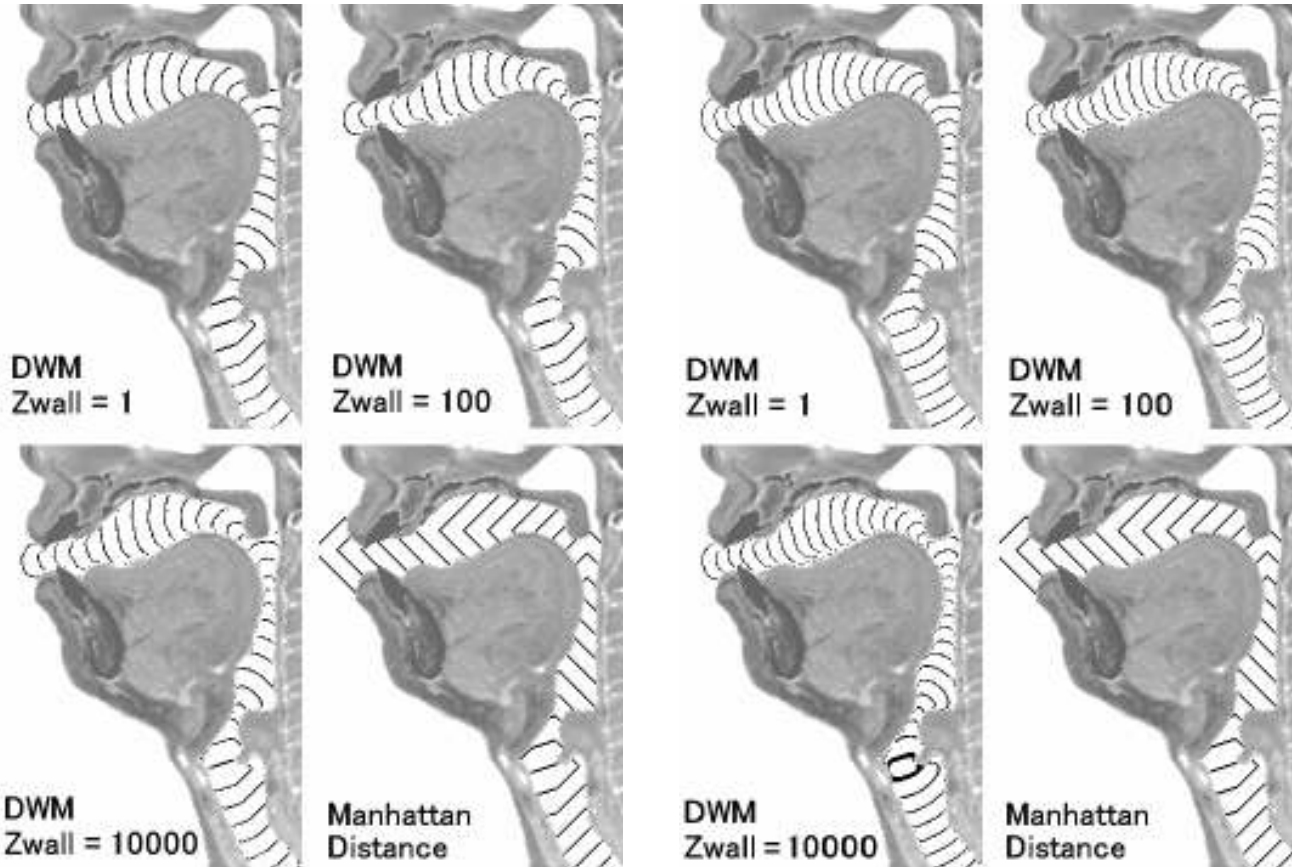


結果と考察

- 実験1: DWM法のパラメータ調査
 - 等音圧面の算出

等音圧面の算出

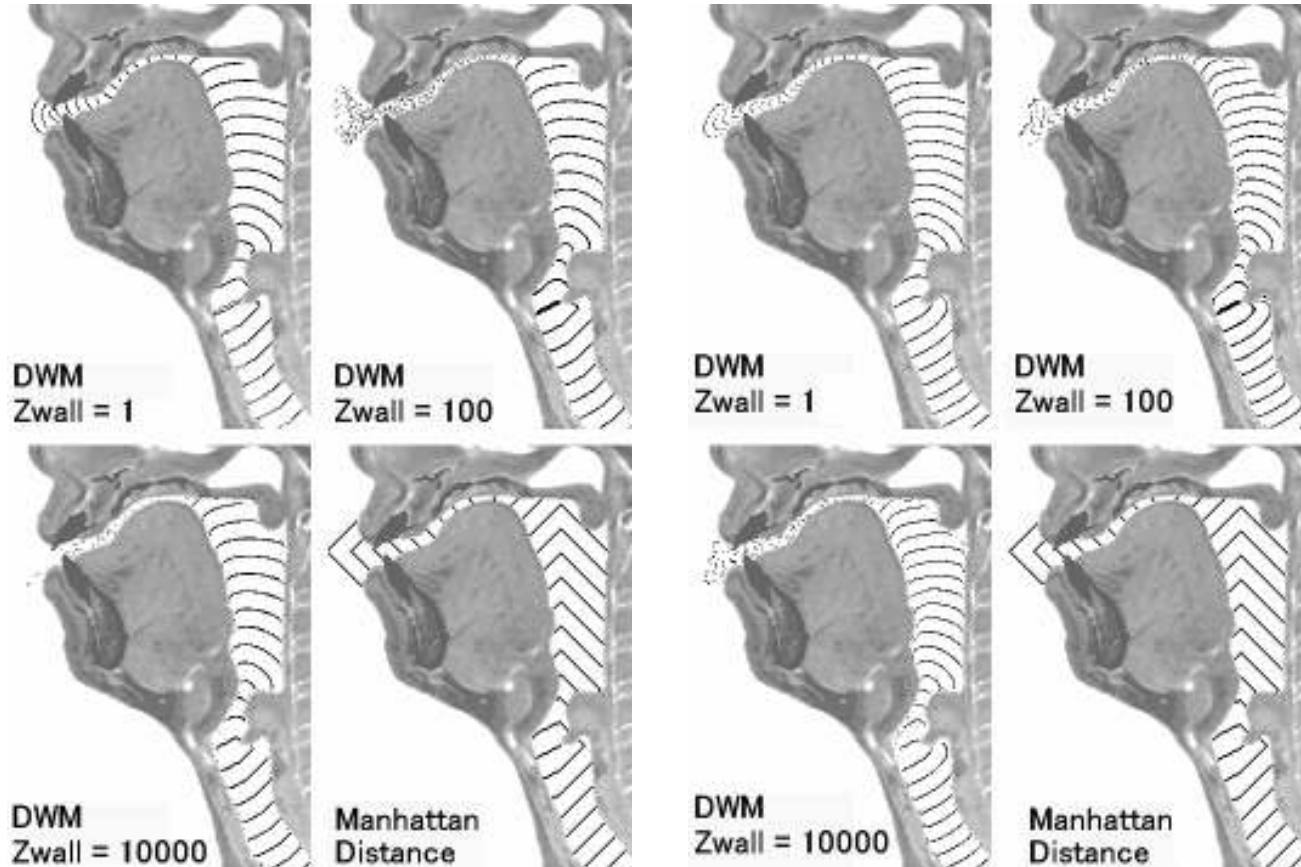
母音 /a/ (/o/, /u/も同様)



2D (正中矢状面)

3D

等音圧面の算出 母音 /i/（/e/も同様）



2D (正中矢状面)

3D

結果と考察

- 実験1(等音圧面の算出) -

- 母音 /e/, /i/ において声道壁インピーダンスが100, 10000の時に等音圧面が算出できないケースがあった
 - 狭窄領域は /i/ で3ピクセル幅(1.5mm), /e/ で4ピクセル幅(2.0mm)
 - シミュレーションのメッシュサイズの問題と考えられる(数値的不安定)
- 実験2では $Z_{wall} = 1$ のみを用いる

実験2

- 他手法との比較 -

- 中心線算出アルゴリズム
 - 本提案手法 (DWM)
 - マンハッタン距離 (MD) [Takemoto et al. 2006]
 - 領域二分法 (BS) [Story et al. 1996]
 - エッジ間最短距離 (NE) [Mochizuki et al. 2007]
- 声道模型とのフォルマント周波数の比較
 - 第1～第4フォルマント (F1-F4) の相対パーセント誤差の和を指標

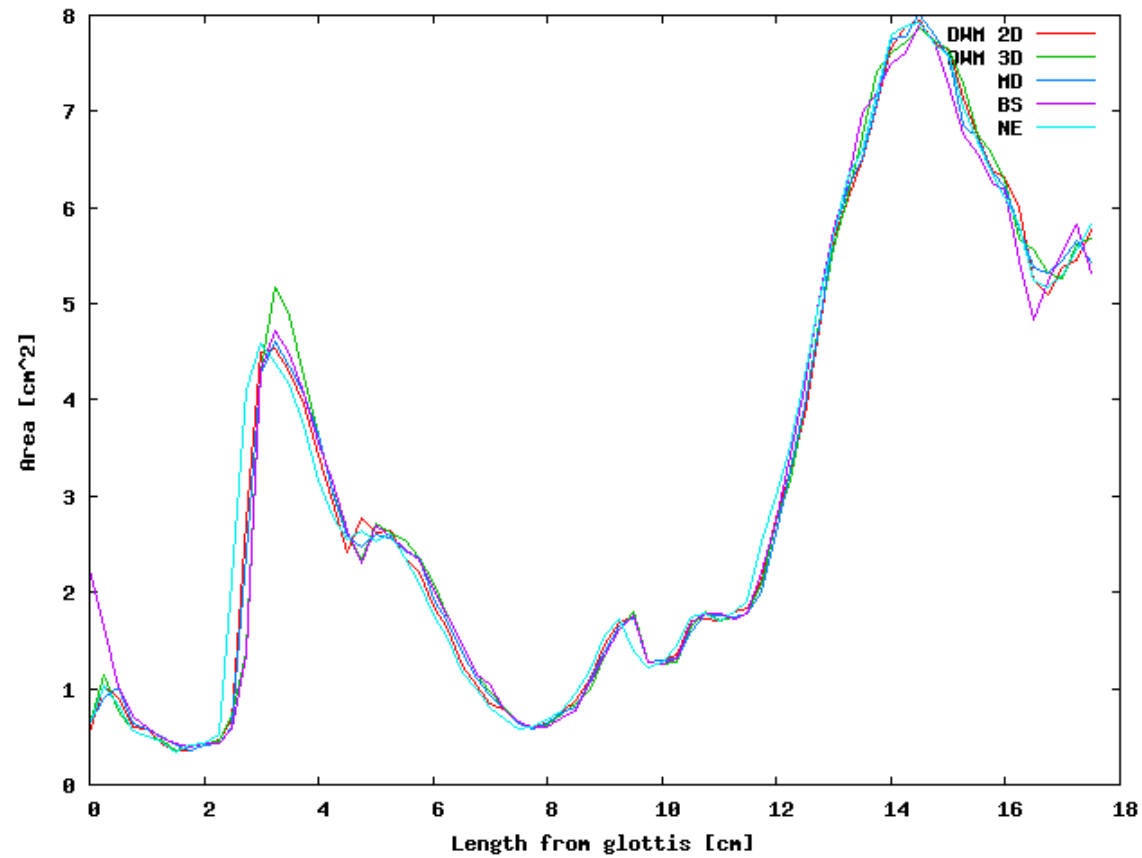


結果と考察

- 実験2: 他手法との比較
 - フォルマント周波数の比較

実験2

母音 /a/ の声道断面積関数



実験2

母音 /a/

	声道長	F1	F2	F3	F4	$\Delta 1$	$\Delta 2$	$\Delta 3$	$\Delta 4$	$\Sigma \Delta$
声道模型		450	1070	2407	2696					
DWM 2D	17.75	539	1318	2531	2865	19.8	23.2	5.2	6.3	54.4
DWM 3D	17.75	533	1318	2514	2889	18.4	23.2	4.4	7.2	53.2
MD	17.75	533	1318	2496	2877	18.4	23.2	3.7	6.7	52.0
BS	17.75	527	1301	2092	2865	17.1	21.6	13.1	6.3	58.1
NE	17.75	527	1301	2625	2941	17.1	21.6	9.1	9.1	56.8

DWM : デジタルウェーブガイドメッシュ [提案手法],

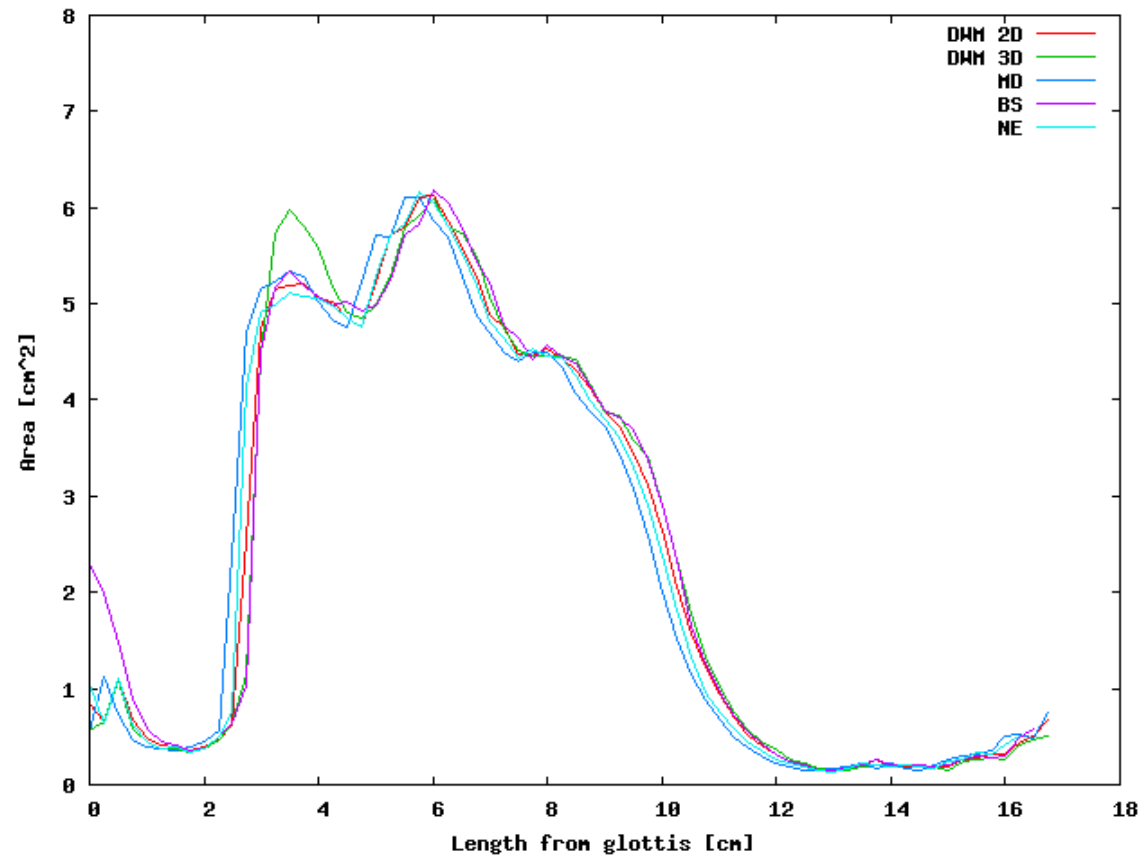
MD : マンハッタン距離 [Takemoto et al.]

BS : 領域二分法 [Story et al.],

NE : エッジ間最短距離点 [Mochizuki et al.]

実験2

母音 /i/ の声道断面面積関数



実験2

母音 /i/

	声道長	F1	F2	F3	F4	$\Delta 1$	$\Delta 2$	$\Delta 3$	$\Delta 4$	$\Sigma \Delta$
声道模型		-	2038	2853	3134					
DWM 2D	17.0	170	2104	2690	3492	-	3.2	5.7	11.4	20.4
DWM 3D	17.0	164	2109	2707	3481	-	3.5	5.1	11.1	19.7
MD	16.5	170	2191	2906	3609	-	7.5	1.9	15.2	24.5
BS	17.0	170	1746	2467	3475	-	14.3	13.5	10.9	38.7
NE	16.75	170	2109	2695	3539	-	3.5	5.5	12.9	21.9

DWM : デジタルウェーブガイドメッシュ [提案手法],

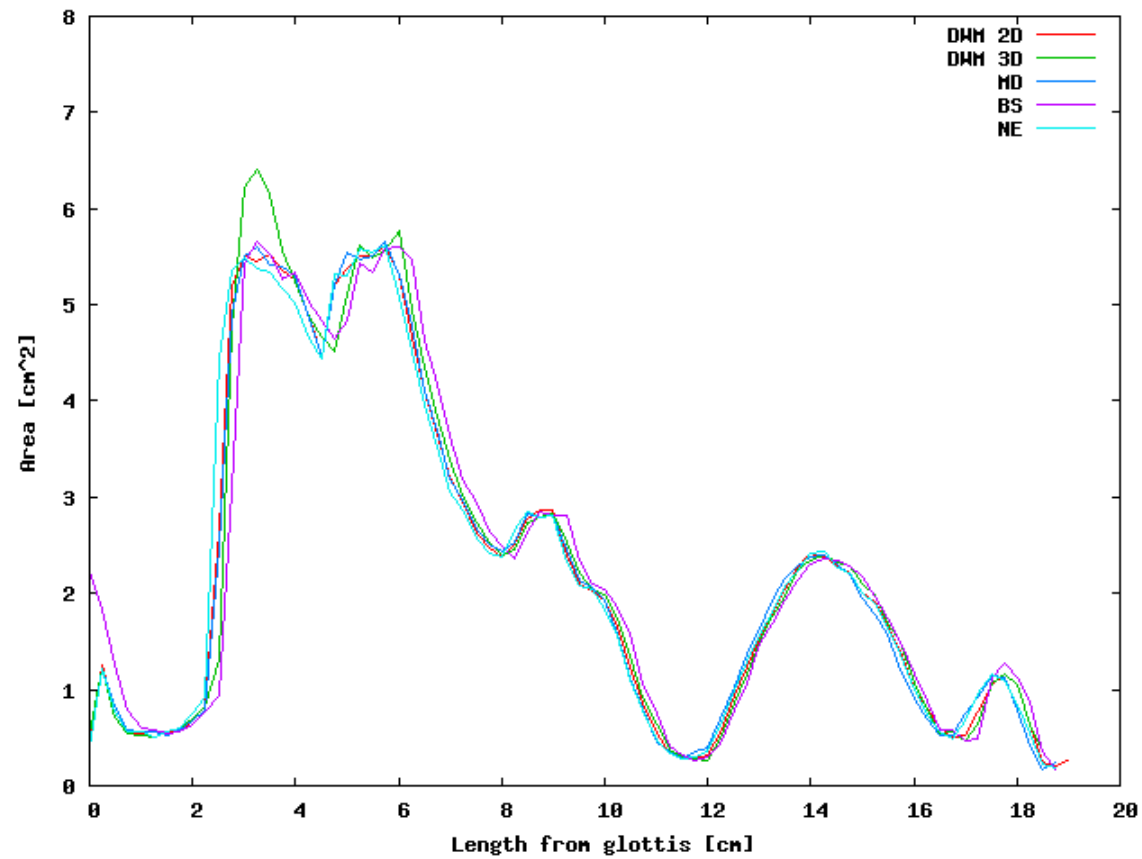
MD : マンハッタン距離 [Takemoto et al.]

BS : 領域二分法 [Story et al.],

NE : エッジ間最短距離点 [Mochizuki et al.]

実験2

母音 /u/ の声道断面面積関数



実験2

母音 /u/

	声道長	F1	F2	F3	F4	$\Delta 1$	$\Delta 2$	$\Delta 3$	$\Delta 4$	$\Sigma \Delta$
声道模型		196	934	2136	3091					
DWM 2D	19.0	234	1078	2291	3182	19.4	15.4	7.3	2.9	45.0
DWM 3D	19.0	234	1084	2279	3135	19.4	16.1	6.7	1.4	43.6
MD	19.0	234	1084	2315	3176	19.4	16.1	8.4	2.7	46.6
BS	19.25	229	1061	2063	2654	16.8	13.6	3.4	14.1	48.0
NE	19.0	234	1066	2285	3258	19.4	14.1	7.0	5.4	45.9

DWM : デジタルウェーブガイドメッシュ [提案手法],

MD : マンハッタン距離 [Takemoto et al.]

BS : 領域二分法 [Story et al.],

NE : エッジ間最短距離点 [Mochizuki et al.]



考察1

■ 全体

- 声道模型のフォルマントが実音声より低い
→声道壁が剛体に由来
- F1が高く出る: 梨状窩の影響か
- F2が高く出る: 口唇部の断面積の取り方の影響及び梨状窩の影響(特に母音/i/)か

■ 領域二分法(BS)

- F3が低く出る: 喉頭腔上部の取り方の影響か
(実装の問題)



考察2

- 既存手法と比較して、2D版、3D版ともに良好な結果
 - しかし、たいして差はない
- 2D版と3D版の結果に顕著な差異は見られなかった



結論

- $Z_{wall} = 1$ の条件の時, 2D版, 3D版ともに安定して等音圧面がDWMで計算可能
- 音波伝播に基づくDWM法と幾何学的方法に基づく既存手法は同程度に有効
- 声道形状の中心線の計算は, 正中矢状断面のみを用いても充分



今後の課題

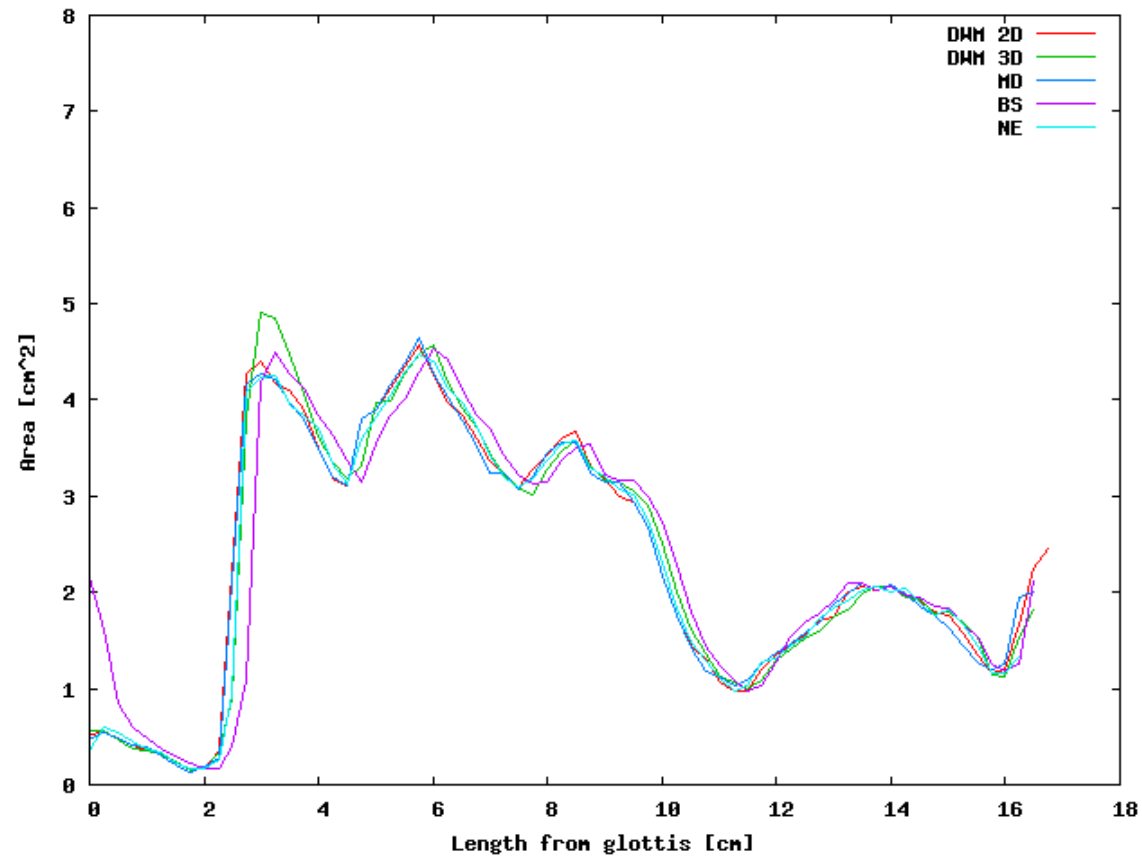
- 評価モデルに梨状窩などの分岐管を組み込む
- 3D版DWMにおいて、中心線を求めずに等音圧面から直接面積を求めた場合の実験と評価
- 3D版DWMを用いれば、口唇からの放射モデルを声道断面積関数に組み込めるため、その実験と評価



ご清聴ありがとうございました

実験2

母音 /e/ の声道断面面積関数



実験2

母音 /e/

	声道長	F1	F2	F3	F4	$\Delta 1$	$\Delta 2$	$\Delta 3$	$\Delta 4$	$\Sigma \Delta$
声道模型		370	1629	2193	2684					
DWM 2D	16.75	422	1852	2420	2854	14.1	13.7	10.4	6.3	44.4
DWM 3D	17.0	416	1828	2391	2813	12.4	12.2	9.0	4.8	38.5
MD	16.5	428	1869	2414	2865	15.7	14.7	10.1	6.7	47.2
BS	16.75	422	1529	1981	2783	14.1	6.1	9.7	3.7	33.5
NE	16.75	422	1846	2385	2848	14.1	13.3	8.8	6.1	42.2

DWM : デジタルウェーブガイドメッシュ [提案手法],

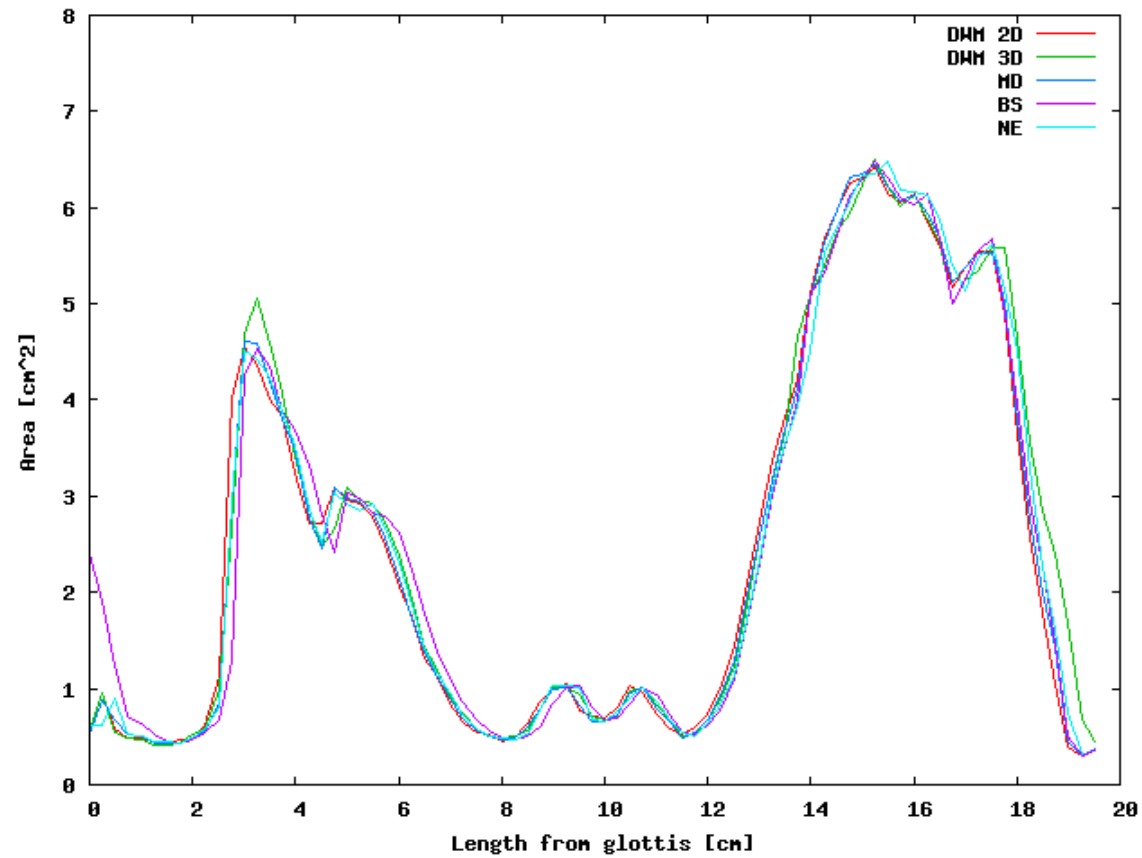
MD : マンハッタン距離 [Takemoto et al.]

BS : 領域二分法 [Story et al.],

NE : エッジ間最短距離点 [Mochizuki et al.]

実験2

母音 /o/ の声道断面積関数



実験2

母音 /o/

	声道長	F1	F2	F3	F4	$\Delta 1$	$\Delta 2$	$\Delta 3$	$\Delta 4$	$\Sigma \Delta$
声道模型		293	616	2175	2924					
DWM 2D	19.5	352	721	2385	2988	20.1	17.0	9.7	2.2	49.0
DWM 3D	19.75	363	762	2379	2947	23.9	23.7	9.4	0.8	57.8
MD	19.75	346	703	2385	2936	18.1	14.1	9.7	0.4	42.3
BS	19.75	334	703	2027	2531	14.0	14.1	6.8	13.4	48.4
NE	19.75	346	709	2350	2912	18.1	15.1	8.0	0.4	41.6

DWM : デジタルウェーブガイドメッシュ [提案手法],

MD : マンハッタン距離 [Takemoto et al.]

BS : 領域二分法 [Story et al.],

NE : エッジ間最短距離点 [Mochizuki et al.]



問題点

- 閉鎖, 分岐に対応できない

付録B.1

- デジタルウェーブガイドメッシュ -

- 3次元声道形状に対して、
時間領域で支配方程式を解くことで
音波の伝播経路を求める
- 音圧 $p(x, y, z)$ に関する3次元の波動方程式

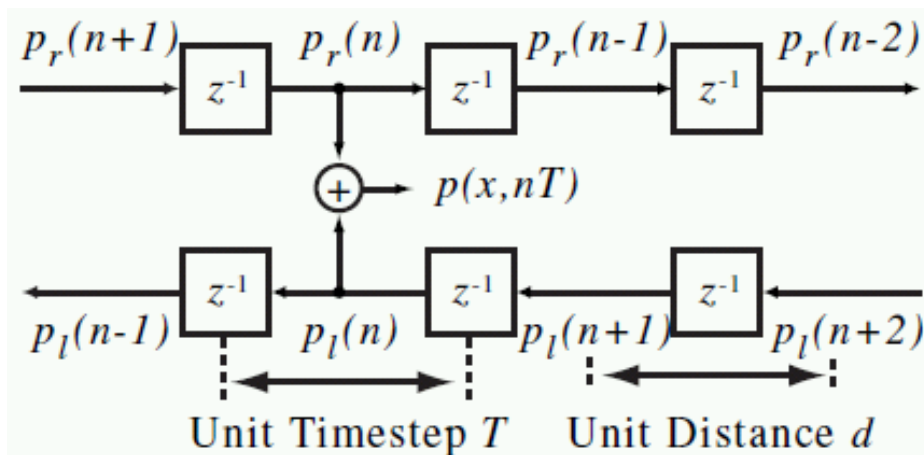
$$\frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2}$$

をデジタルウェーブガイドメッシュ(DWM)
で離散化

付録B.2

- 1次元偏微分方程式の解 -

- 1次元の場合：波動方程式 $\frac{1}{c^2} \frac{\partial^2 p(t, x)}{\partial t^2} = \frac{\partial^2 p(t, x)}{\partial x^2}$
を左方向と右方向の成分に分割した
ダランベールの解 $p(x, nT) = p_r(x - cnT) + p_l(x + cnT)$
を計算 (pは音圧, cは音速, nはステップ時刻)



1次元ウェーブガイド音響管モデルにおける
音圧の構成要素 [Mullen 2006]

付録B.3

- 境界条件とインピーダンス -

- 境界条件: 声道内部(空気)と声道壁では特性インピーダンスが異なる
- DWMジャンクション]において, 連続の法則より

1. $p_i = p_i^+ + p_i^-$

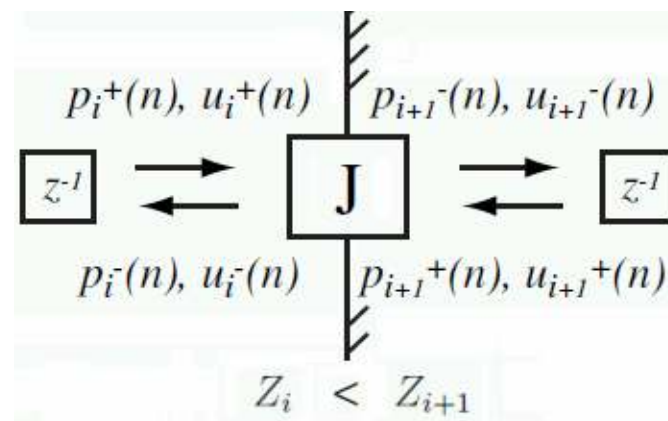
2. $p_i = p_{i+1} = p_J$

3. $u_i = u_i^+ - u_i^-$

4. $u_i + u_{i+1} = 0$

- 媒質のインピーダンス

$$Z_i^+ = \frac{p_i^+}{u_i^+}, Z_i^- = \frac{p_i^-}{u_i^-}$$



インピーダンス不整合点における信号の散乱 [Mullen 2006]

付録B.4

- 多次元への拡張 -

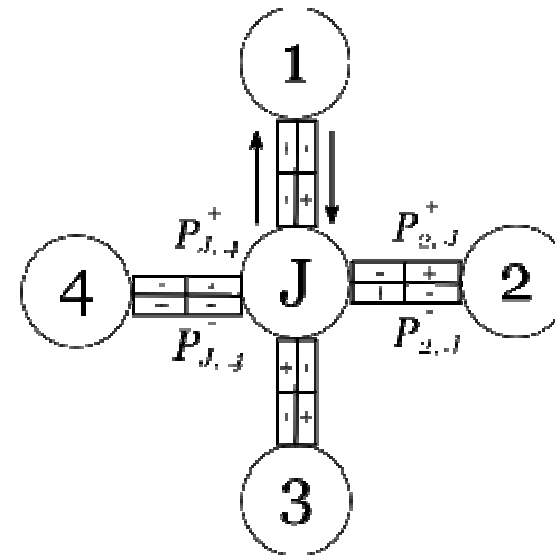
- 接続ウェーブガイド数を M 個 (2次元: M=4, 3次元: M=6) に拡張した場合,

$$p_k^- = r_k p_k^+ + \sum_{i=1, i \neq k}^M (1 + r_i) p_{J,i}^+$$

$$r_k = \frac{2 \frac{1}{Z_k} - \sum_{i=1}^M \frac{1}{Z_i}}{\sum_{i=1}^M \frac{1}{Z_i}}$$

が導出される

- 初期条件とインピーダンスの設定が必要

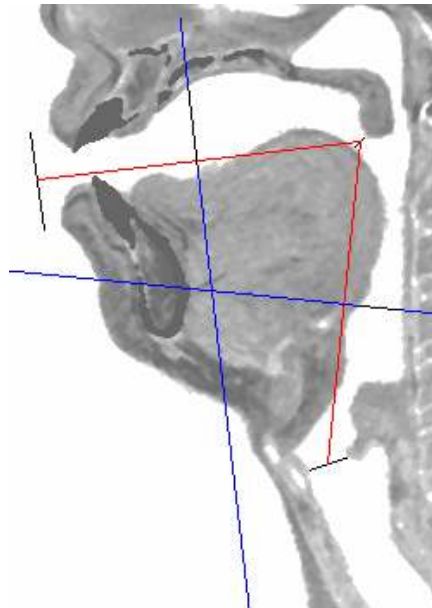
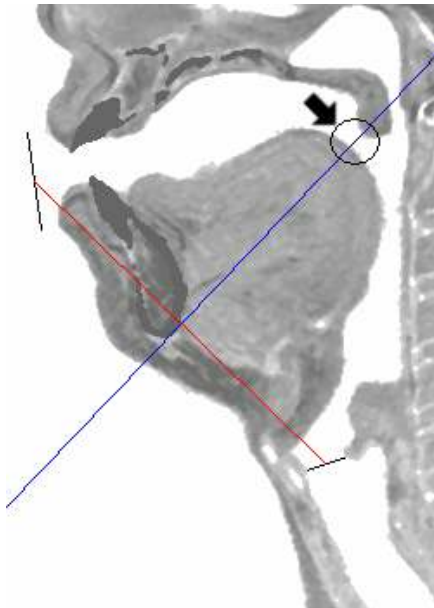


2次元DWM

付録C.1

- 領域二分法 -

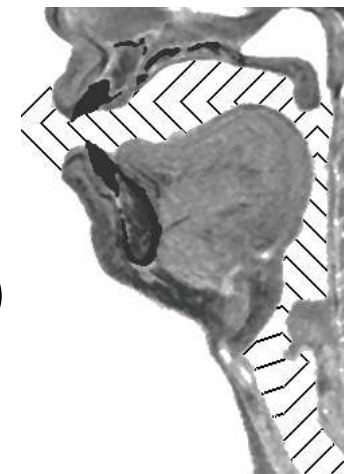
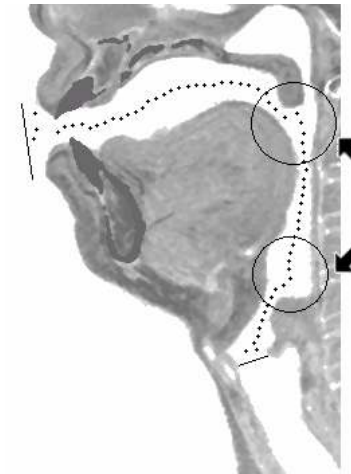
- 既存手法: 領域二分法 [Story 1996]
 - 2点を結んだ線分(赤)に対する直交面(青)との交差領域の中点を取って領域を2分割



付録C.2

- 音波伝播経路 -

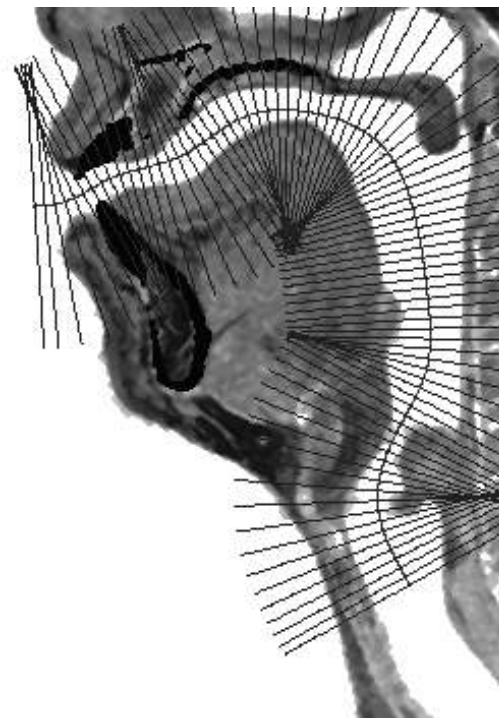
- 音波の伝播経路は必ずしも幾何学的中心を通らない [中井他 1999,2002]
- 伝播経路に沿って断面を算出
 - 有限要素法 (FEM) [中井他 2002, Mochizuki et al. 2007]
 - 周波数領域で計算するため複雑
 - マンハッタン距離による音波伝播の近似 [Takemoto et al. 2006]
- 上記2手法は正中矢状面 (2次元) のみを利用



付録C.3

- 断面積関数の推定 -

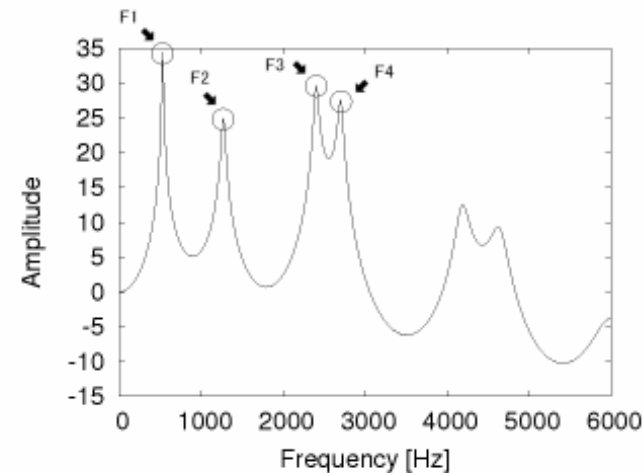
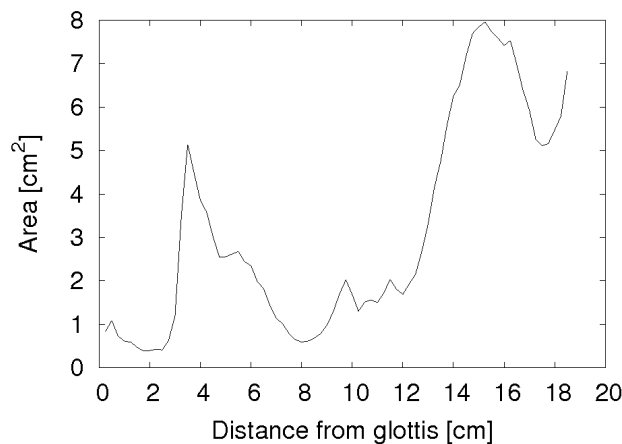
- 3次元MRIデータからの声道断面積関数の一般的な推定手順
 - 1) 声道の中心線を抽出し,
 - 2) 中心線に沿って断面を求める
- 既存の中心線算出手法
 - Story et al. (1996)
 - Kröger et al. (2000)
 - Takemoto et al. (2006)
 - Mochizuki et al. (2007)



付録D.1

- 全体の処理手順 -

1. 声道領域の抽出
2. 中心線の算出
3. 声道断面積関数の計算
4. 伝達関数の計算



付録D.2

- 伝達関数の計算 -

- 伝達関数の計算
 - 伝送線路モデルを用い周波数領域で解析
 - 声道壁は声道模型に合わせて剛体とした
(全体的にフォルマント周波数が低下)
 - 粘性, 熱伝導, 放射のエネルギーロス