

MRI に基づく音声生成の研究手法の概要

Overview of Methods and Techniques Used in MRI-based Speech Production Study

竹本浩典 北村達也

Abstract

MRI は発話中の声道形状を計測できるので、音声生成の研究に有用なツールである。近年、様々な撮像法が開発されて誰もが利用できるようになったが、日本では MRI を用いた研究はまだ少ない。その一つの理由は、撮像法と画像処理に関する情報が不足しているためであると考えられる。そこで本稿では、これから MRI を用いて音声生成の研究を開始する研究者のために、MRI の撮像法と画像処理について概説し、声道の代表的な音響解析手法を紹介する。

キーワード：音声生成, MRI, 撮像法, 画像処理

1. はじめに

音声は、声帯振動及び声道^(用語)内に生じた乱流雑音などの音源が、声道形状に固有の音響的修飾を受けて口唇や鼻孔から放射される現象である。そのため、声道の形状とその変化を知ることは、音声生成研究において極めて重要である。しかし、声道は直接的に観測できないため、その形状計測には特殊な装置が必要である。磁気共鳴画像法 (MRI: Magnetic Resonance Imaging) はその一つで、核磁気共鳴を利用して断層画像を得る。1986年、Rokkaku ら⁽¹⁾は MRI で初めて声道を計測した。それから 25 年の間に装置や撮像法が改良され、MRI で計測した声道形状に基づく様々な音声生成の研究がなされてきた (例えば文献 (2)~(4))。これらの研究では、MRI で撮像した声道形状を声道断面積関数^(用語)に変換し、縦続音響管モデル^(用語)で声道形状と音響特性との関係を分析している。

本誌では、①文部省 (文部科学省) 学術用語集電気工学編、②本誌編の改訂電子情報通信用語辞典、③本誌編のエンサイクロペディアハンドブックに基づき用語を統一している。日本磁気共鳴医学会用語委員会編 MR 用語辞典では、英語の static magnetic field に「静磁場」との訳語を当てているが、本誌では①に従い、「静磁界」の訳語を当て、「静磁場強度」ではなく「静磁界強度」と記載する。

竹本浩典 独立行政法人情報通信研究機構ユニバーサルコミュニケーション研究所

E-mail takemoto@nict.go.jp

北村達也 甲南大学知能情報学部知能情報学科

E-mail t-kitamu@konan-u.ac.jp

Hironori TAKEMOTO, Nonmember (Universal Communication Research Institute, National Institute of Information and Communications Technology, Kyoto-fu, 619-0288 Japan) and Tatsuya KITAMURA, Nonmember (Faculty of Intelligence and Informatics, Konan University, Kobe-shi, 658-8501 Japan).

電子情報通信学会誌 Vol.94 No.7 pp.585-590 2011年7月

©電子情報通信学会 2011

日本では、2000 年から (株) ATR-Promotions 脳活動イメージングセンタ (ATR-BAIC) で声道の様々な撮像法が検討され、多くのノウハウが蓄積されてきた。ATR-BAIC では、専属の技師が操作する MRI 装置や実験設備を 90 分単位で貸し出しており、誰でも利用できる (<http://www.baic.jp/>)。

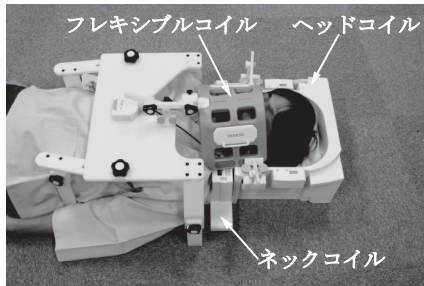
しかし、日本では MRI を用いた音声生成の研究はまだ少ない。その一つの理由は、MRI を使えば何を撮像できるのか、画像をどのように処理するのかに関する情報が不足しているためであると考えられる。そこで本稿では、これから MRI を用いて音声生成研究を始める研究者のために、2011 年 1 月現在の ATR-BAIC の実験設備を例にとり、音声生成の研究で用いられる代表的な撮像法を解説し、実験者や被験者が気を付けるポイントについて触れ、MRI 特有の画像処理について述べる。そして、声道形状の代表的な音響解析手法を紹介し、MRI を用いた音声生成研究の今後の展望について述べる。

2. MRI 撮像法

ATR-BAIC が保有する声道撮像用の MRI 装置は、Siemens 製 MAGNETOM Verio (3T) である (図 1(a))。この装置の静磁界強度は 3.0 T で、ガントリ (中央の穴) の内径は 70 cm である。撮像時には、被験者を仰向けに寝かせ、頭頸部の周囲にコイル (図 1(b)) を配置してガントリに入れる。このコイルは、画像データを得るために必要なパルスの送信と体内からの信号の受信 (MRI データ収集) を行う。このとき、コイルの振動で



(a)



(b)

図1 MRI装置(a)とコイル(b)

100 dB を超える騒音が発生する。なお、重力は仰向けに寝た被験者の体に対して前後方向に作用するため、MRI で得られる声道形状は、立位や座位のものとは異なる⁽⁵⁾。

2.1 被験者の選定

MRI 画像の良否は以下に述べる被験者の能力に依存する。MRI の騒音は、気導音による聴覚フィードバックを阻害するため調音がひずみやすい。そのため、骨伝導音を聞くことで調音する能力が被験者に求められる。また、後述する発声同期撮像法や動画撮像法では、発声や発話の繰り返しが必要なため、正確な発話運動を同じタイミングで繰り返す能力も必要である。また、歯列に金属製の補てつ物(詰め物)がない被験者が望ましい。こ

用語解説

声道 声門(左右の声帯の隙間)より上部にあって、音声波が体外に放射されるまでに通過する体内の空間。喉頭腔、咽頭腔、口腔、鼻腔が含まれる。

声道断面積関数 声門からの距離に対する声道の断面積の関数。

縦続音響管モデル 声道を断面積の異なる円筒形の音響管の連なりで表現したモデルで、各種音響特性を計算できる。なお、個々の音響管は平面波伝搬のみを仮定している。

矢状断面 身体を左右に分割する断面。特に正中を含む面を正中矢状断面と呼ぶ。

ボクセル解像度 ボクセルは断面の厚さを考慮したピクセルのこと。例えば、 256×256 mm の方形の領域(厚さ 2 mm)を MRI で 256×256 ピクセルの画像にすると、ボクセル解像度は $1 \times 1 \times 2$ mm となる。

れは、補てつ物によっては磁界がひずみ、画像が乱れるためである。

2.2 静止画撮像法

静止画撮像法は、歯列と持続発声可能な母音や幾つかの子音の声道形状を計測するために用いる。声道は左右方向の幅が最も小さいため、声道を一連の矢状断面^(用語)で撮像して三次元再構築し、声道を抽出して音響解析を行う。このとき注意しなければならないのは空間解像度である。MRI は撮像時間を短縮するために様々な補間を行うため、補間しない場合の空間解像度、すなわち実効的な空間解像度を確認する必要がある。必要とされる実効的な空間解像度は解析対象と方法によって異なる。例えば、成人男性の母音を対象とする場合、縦続音響管モデルで 5 kHz 以下の音響特性を解析するのであれば、ボクセル解像度^(用語)が $2 \times 2 \times 4$ mm、断面の間隔が 4 mm 以上の実効的な空間解像度があればよい⁽⁶⁾。一方、三次元の音響解析手法を用いて 10 kHz 以下の音響特性を解析するのであれば、ボクセル解像度が $1 \times 1 \times 2$ mm、断面の間隔が 2 mm 以上の実効的な空間解像度が必要である⁽⁷⁾。

2.2.1 歯列撮像法

MRI は体内の水素原子の分布を画像化する。一般に、水素原子が多い組織(脂肪や筋肉など)が高い輝度値(白)、少ない組織(骨や歯など)や空気が低い輝度値(黒)となる。そのため、画像では歯列と口腔とを区別できない。しかし、声道形状を正確に計測するためには、何らかの方法で歯列を撮像し、画像処理で補填する必要がある。そこで歯列の輝度値が低いことを逆に利用して、歯列を輝度値の高い物質で囲んで撮像し、輝度値を反転すれば歯列を抽出できる。

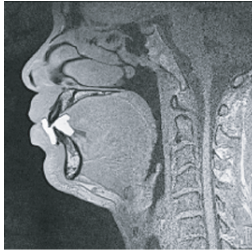
最も実用的で簡便な方法は、仰向けで唇をしっかり閉じ、舌を歯の裏側に密着させて撮像するものである(図 2(a))。この方法は道具が不要であり、声道の撮像に引き続いて撮像できる。しかし、画像処理の際に、上下の歯が接触する部分を目視で分離する必要がある。その他に、セプトン製のマウスピースを装着する方法⁽⁸⁾(図 2(b))、ブルーベリージュースを口に含む方法⁽⁹⁾(図 2(c))がある。これらの方法では、歯列をより正確に抽出できるが、被験者の負担が大きく、時間がかかる。なお、いずれの方法でも実効的な空間解像度は、ボクセル解像度が $1 \times 1 \times 1$ mm、断面の間隔が 1 mm である。

2.2.2 短時間撮像法

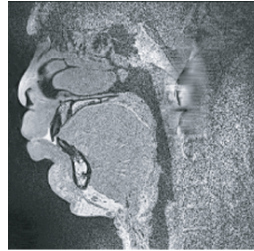
この撮像法は、一息の持続発話で声道全体を撮像できるため、被験者によらずほぼ確実に声道の三次元形状が得られる。15 秒間の持続発話による実効的な空間解像度は、ボクセル解像度が $2 \times 2 \times 3$ mm、断面の間隔が 3



(a)



(b)



(c)

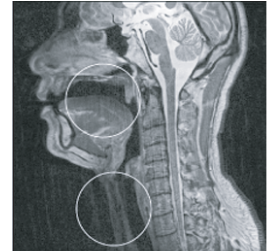
図2 歯列撮像法 画像全体のコントラストや輪郭に注意。(a)舌を歯に密着させて撮像、(b)セプトン製マウスピースを装着して撮像、(c)ブルーベリージュースを口に含んで撮像。



(a)



(b)



(c)

図3 MRI 静止画撮像 (a)短時間撮像 (15秒), (b)短時間撮像 (28秒), (c)発声同期撮像。

mmである。断面の間隔が広いと、狭い部分や声道細部の形状がやや曖昧になる(図3(a), 円で囲んだ部分)。

持続発声の時間を長くすれば、実効的な空間解像度が向上する。例えば、持続発声の時間を28秒にすれば、実効的な空間解像度は、ボクセル解像度が $2 \times 1 \times 2$ mm, 断面の間隔が2 mmとなる(図3(b))。しかし、長時間発声するために、被験者によっては喉頭腔の上部を狭めて呼気流を抑制する傾向が現れる。なお、撮像途中で息継ぎをすると、吸気に伴って軟口蓋が挙上して左右の声帯が離れる運動が生じ、軟口蓋と喉頭で声道形状が不正確になる。

2.2.3 発声同期撮像法

ある程度以上の長さの持続発声は不可能であるため、発声と同期した撮像を繰り返すことで実効的な空間解像度を高める。これが発声同期撮像法である。例えば、5秒間の発声を33回繰り返すと、ボクセル解像度が $1 \times 1 \times 2$ mm, 断面の間隔が2 mmの実効的な空間解像度が得られる(図3(c))。

この撮像法では、発声とMRI撮像を同期させるために外部トリガボックスを用いる。これは被験者には吸気と発声を促す聴覚刺激を呈示し、MRI装置には撮像を開始するトリガパルスを出力する装置である⁽¹⁰⁾。聴覚刺激は声の高さを一定に保ちやすくするための複合正弦波である。その最も低い周波数は被験者にとって自然な声の高さに調整可能で、例えば男性なら120 Hz, 女性なら200 Hzなどに設定する。

この撮像法は最も精密に声道を計測できるが、繰り返

し発声の際に口の形が変われば、声道の輪郭が不鮮明になる。また、コイルの感度むらで画像が局所的に暗くなるという問題もある(図3(c), 円で囲んだ部分)。

2.3 動画撮像法

MRIは時間分解能が低いが、発話運動を定期的に繰り返してストロボスコープに類似した撮像法を用いれば、発話運動を動画として記録できる⁽¹⁰⁾。以下では、一断面における動画撮像法(MRI動画)について解説する。なお、多断面で動画を撮像して動く立体を再構成する手法(三次元MRI動画)は文献(6)を参照された。

MRI動画は、主に正中矢状断面における発話器官各部の運動タイミングや調音結合による発話運動の変化の観測に用いられる。この撮像法でも被験者の発話とMRIの撮像を同期させる外部トリガボックスを用いる。被験者に呈示する聴覚刺激はノイズバーストからなるリズムミカルな信号で、発話タスクに応じて調整する。例えば、/pataka/という発話タスクであれば、1.5秒の長さの3拍の刺激で、最初の拍で息を吸い、次の拍で/pata/, 最後の拍で/ka/を発話する。1回の発話の長さは動画の総フレーム数と、繰り返し回数はボクセル解像度に対応する。また、発話運動のタイミングと軌跡の正確さは、輪郭の明瞭さに影響を与える。なお、ATR-BAICのMRI装置では、フレーム数の上限は128, 時間分解能の上限は120フレーム/sである。

通常、発話を128回繰り返すことによって、ボクセル解像度が $2 \times 2 \times 4$ mm, フレーム間隔(時間分解能)が33 msの動画を得る。しかし、発話回数が多いと疲労で

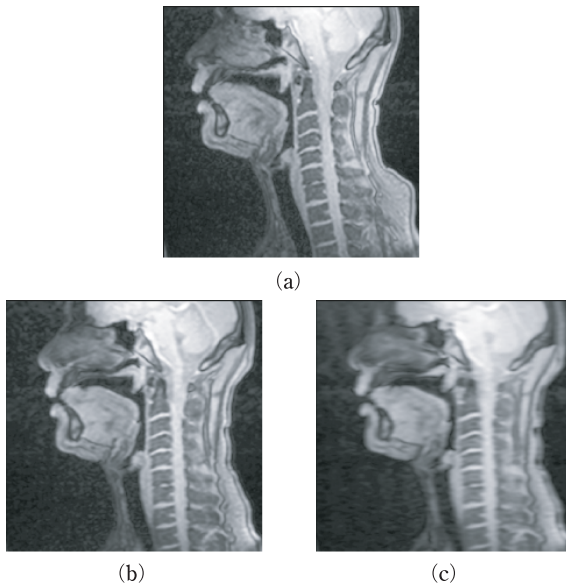


図4 MRI 動画による/pataka/発話時の/t/のフレーム ボクセル解像度を犠牲にして発話の繰返し回数を減らしたときの画質の変化. 発話の繰返し回数は, (a)128回, (b)96回, (c)64回.

発話運動が不正確になり, かえって輪郭が不明瞭になる. そこで, ボクセル解像度または時間分解能のどちらかを犠牲にすることにより, 発話の繰返し回数を減らす方法が提案されている. どちらの方法で発話の繰返し回数を減らすかは, 発話タスクに応じて判断する.

図4は, ボクセル解像度を犠牲にしたときの/pataka/の/t/のフレームの変化である. ボクセル解像度を $2 \times 2 \times 4$ mm から $2.7 \times 2.7 \times 4$ mm, $4 \times 4 \times 4$ mm へと低下させると, 128回の繰返し回数を96回, 64回に減らすことができる. しかし, 画像全体のコントラストや輪郭が曖昧になる.

図5は, 時間分解能を犠牲にしたときの/pataka/の/t/のフレームの変化である. 通常は, フレーム間隔ごとに1回のMRIデータ収集しか行わないが, 時間分解能を犠牲にする場合には, フレーム間隔ごとに複数回のデータ収集を行い, これらをまとめて一つのフレームを構成する. これはマルチエコー法と呼ばれる. 図5はフレーム間隔ごとに3, 5, 7回のデータ収集を行い, 発話回数をそれぞれ128回の $\frac{1}{3}$ (43回), $\frac{1}{5}$ (26回), $\frac{1}{7}$ (19回)にしたものである. ボクセル解像度はいずれも $2 \times 2 \times 4$ mmであり, 一見画質に差はない. しかし, 時間的に幅のあるデータで一つのフレームを構成するため, 素早い運動では輪郭が不明瞭になる.

3. MRI 画像処理

本章では, MRIの画像処理を実際の流れに沿って解説する. 各処理の詳細は, 文献を参考にして頂くことにし, ここではその概要を記述する.

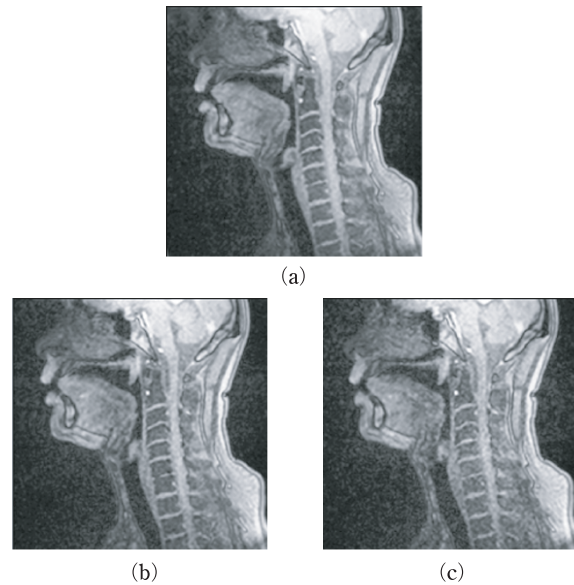


図5 MRI 動画による/pataka/発話時の/t/のフレーム 時間分解能を犠牲にして発話の繰返し回数を減らしたときの画質の変化. 発話の繰返し回数は, (a)43回, (b)26回, (c)19回.

3.1 ファイル形式

MRI画像は, 医療画像の標準であるDICOM形式で出力される. この画像には, 被験者の氏名, 年齢, 性別等の個人情報が含まれているため, 取扱いには注意を要する. また, 画像は16 bit階調であるため, コンピュータの標準的な形式の8 bit階調のファイル形式に変換する必要がある.

3.2 しきい値決定

声道形状を抽出するには, 空気と体組織とを分離するしきい値が必要である. しかし, MRI画像は元々雑音が多く, コイルとの距離により輝度値が変化するなどの問題がある. そのため, 妥当なしきい値を決定することは難しい. 我々は, 平均隣接数に着目してしきい値を求めている⁽¹¹⁾. この方法が最適とは限らないが, この方法で決定したしきい値により抽出した声道断面積関数から計算したホルマントは, 被験者の音声から抽出したホルマントとよく一致した⁽⁶⁾.

3.3 歯列補填

ここでは歯列補填の手法⁽⁹⁾の概要を述べる. まず, 2.2.1で紹介した手法で歯列の撮像を行い, 上下の歯列を抽出する. 次に, 上下の歯列にそれぞれ四つの参照点を決め, 補填対象となるMRIデータにも対応する点を求める. そして, これら四つの参照点の座標を相互に対応付けるアフィン変換行列を計算して歯列を補填する(図6).この行列は, 補填した歯列が周囲にはみ出す部分を焼きなまし法により最小にすることで最適化できる.

3.4 声道断面積関数の抽出

ここでは、声道断面積関数を MRI データから抽出する方法⁽⁶⁾の概略を述べる。まず、3.3 で求めた歯列を補填した MRI データの正中矢状面において、平面波の経路である声道中心線を求める。次に、この声道中心線に直交する断面を一定間隔で求めて、3.2 で求めたしきい値により声道部分を抽出し、断面積を算出する。図7はこの方法で求めた母音/e/の声道中心線と声道横断面の位置、及び声道断面積関数である。なお、この手法⁽⁶⁾で

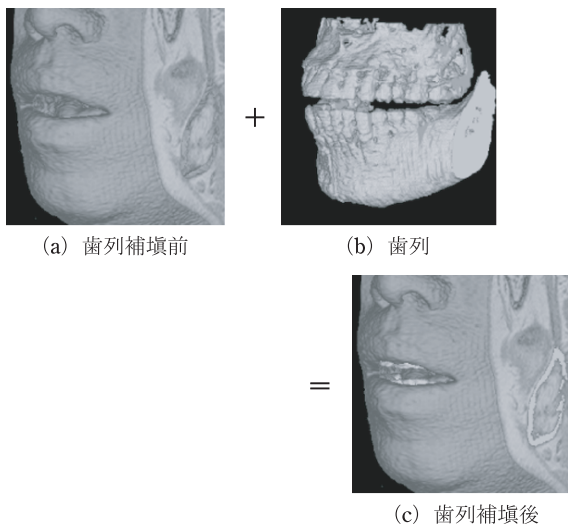


図6 歯列補填の模式図 三次元再構築した発話中の MRI データ(a)に、抽出した上下の歯列(b)を補填する(c)。

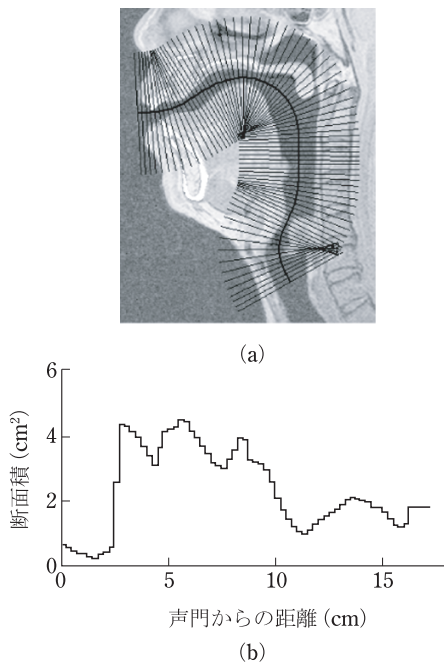


図7 母音/e/の声道中心線と声道断面の位置(a)と声道断面積関数(b)

声道断面積関数を求めるプログラムの公開が予定されている⁽¹²⁾。このプログラムはフリーの画像処理ソフト ImageJ のプラグインで、高速に声道断面積を抽出できる。

4. 音響解析の諸手法

声道の音響解析は、一次元モデルによるものと三次元モデルによるものに大別される。一次元モデルとは、3.4 で抽出した声道断面積関数に基づいて構築した縦続音響管モデル⁽¹³⁾である。このモデルは、平面波伝搬のみを仮定しているため、約 5 kHz 以下の周波数領域が解析対象である。このモデルの長所は、計算が簡単で高速なこと、時間変化する声道断面積関数と音源モデルとを組み合わせることで音声合成器として用いることができることなどである。逆に短所は、断面積関数の抽出の仕方によって計算結果が大きく変わること、梨状窩などの分岐管の相互作用⁽¹⁴⁾を正しく扱えないため、零点に関して誤差が大きいことなどである。

三次元モデルとは、歯列を補填したデータから声道を三次元的に抽出して解析を行うもので、有限要素法⁽¹⁵⁾、境界要素法⁽¹⁶⁾、時間領域差分法⁽⁷⁾などがある。これらの手法の長所は、平面波伝搬を仮定しないため、高い周波数領域(約 10 kHz)まで解析が可能であること、分岐管による零点を正しく扱えることである。逆に短所は、一次元モデルに比べて高い空間解像度の MRI データが必要なこと、三次元の複雑な画像処理が必要なこと、計算に膨大な時間を要することなどである。

このように、解析手法には一長一短がある(表1)。そのため、研究目的、対象とする周波数領域、MRI データの空間解像度に応じて使い分ける必要がある。

5. ま と め

本稿では、MRI による音声生成時の静止画、動画の撮像法と実験者、被験者が注意する点について述べた。また、画像処理の流れと音響解析の手法も紹介した(図8)。

表1 音響解析手法の比較

	必要な空間解像度	平面波伝搬の仮定	計算速度	計算精度	音声合成器への応用
縦続音響管モデル	低	必要	非常に高速	低い	可
有限要素法	中	不要	非常に低速	非常に高い	不可
境界要素法	中	不要	非常に低速	非常に高い	不可
時間領域差分法	高	不要	低速	高い	不可

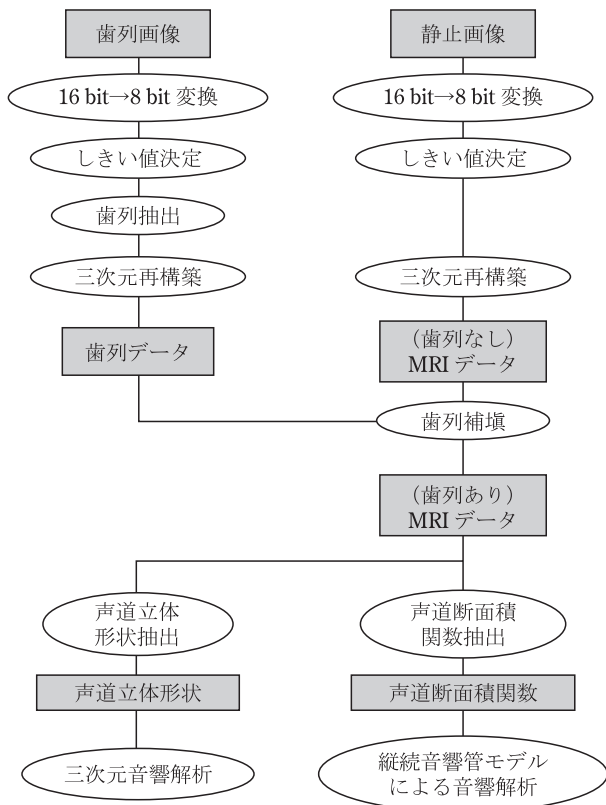


図8 MRI画像の処理の流れと音響解析手法

現在、MRIの静磁界強度は3.0 Tが主流になりつつある。静磁界強度が大きくなると、空間分解能が高くなり、撮像時間も短縮される。これにより、例えば、唇の複雑な筋走行の解明が期待できる。また、短時間撮像により、一息の発話で声道全体の形状が計測できるようになったため、声道形状の個人差の検討が深まることが期待できる。動画撮像に関しては、発話器官各部の運動機序を解析して、時間領域で発話をシミュレートする生理学的調音モデル⁽¹⁷⁾やタイムドメイン合成器⁽¹⁸⁾に反映することで、より自然な音声の生成に貢献すると思われる。

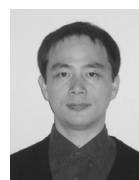
謝辞 本研究の一部は、科学研究費補助金（課題番号21500184, 21300071）により実施した。

文 献

- (1) M. Rokkaku, K. Hashimoto, S. Imaizumi, S. Niimi, and S. Kiritani, "Measurements of the three-dimensional shape of the vocal tract based on the magnetic resonance imaging technique," *Ann. Bull. PILP.*, no. 20, pp. 47-54, 1986.
- (2) T. Bare, J.C. Gore, L.C. Gracco, and P.W. Nye, "Analysis of vocal tract shape and dimensions using magnetic resonance imaging : Vowels," *J. Acoust. Soc. Am.*, vol. 90, no. 2, pp. 799-828, 1991.
- (3) B.H. Story, I.R. Titze, and E.A. Hoffman, "Vocal tract area functions from magnetic resonance imaging," *J. Acoust. Soc. Am.*, vol. 100, no. 1, pp. 537-554, 1996.

- (4) H. Takemoto, S. Adachi, T. Kitamura, P. Mokhtari, and K. Honda, "Acoustic roles of the laryngeal cavity in vocal tract resonance," *J. Acoust. Soc. Am.*, vol. 120, no. 4, pp. 2228-2238, 2006.
- (5) T. Kitamura, H. Takemoto, K. Honda, Y. Shimada, I. Fujimoto, Y. Syakudo, S. Masaki, K. Kuroda, N. Oku-uchi, and M. Senda, "Difference in vocal tract shape between upright and supine postures : Observations by an open-type MRI scanner," *Acoust. Sci. Technol.*, vol. 26, no. 5, pp. 465-468, 2005.
- (6) H. Takemoto, K. Honda, S. Masaki, Y. Shimada, and I. Fujimoto, "Measurement of temporal changes in vocal tract area function from 3D cine-MRI data," *J. Acoust. Soc. Am.*, vol. 119, no. 2, pp. 1037-1049, 2006.
- (7) H. Takemoto, P. Mokhtari, and T. Kitamura, "Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method," *J. Acoust. Soc. Am.*, vol. 128, no. 6, pp. 3724-3738, 2010.
- (8) 北村達也, 林 俊平, "エラストマー製マウスピースを用いたMRIによる歯列計測法," *音響論集(春)*, no. 3-6-9, pp. 357-358, 2009.
- (9) H. Takemoto, T. Kitamura, H. Nishimoto, and K. Honda, "A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions," *Acoust. Sci. Technol.*, vol. 25, no. 6, pp. 468-474, 2004.
- (10) S. Masaki, M.K. Tiede, K. Honda, Y. Shimada, I. Fujimoto, Y. Nakamura, and N. Ninomiya, "MRI-based speech production study using a synchronized sampling method," *J. Acoust. Soc. Jpn. (E)*, vol. 20, no. 5, pp. 375-379, 1999.
- (11) 笹川耕一, 黒田伸一, 池端重樹, "平均隣接数に着目したしきい値決定法," *信学論 (D-II)*, vol. J73-D-II, no. 3, pp. 360-366, March 1990.
- (12) 成重智則, 北村達也, "ImageJの声道断面積抽出プラグインの開発," *信学技報*, SP2010-71, pp. 13-16, Nov. 2010.
- (13) J.L. Flanagan, *Speech analysis, synthesis, and perception*, Springer-Verlag, New York, 1972.
- (14) 竹本浩典, 足立整治, P. Mokhtari, 北村達也, "左右の梨状窩の音響的相互作用," *信学技報*, SP2010-70, pp. 7-12, Nov. 2010.
- (15) H. Matsuzaki and K. Motoki, "Study of acoustic characteristics of vocal tract with nasal cavity during phonation of Japanese /a/," *Acoust. Sci. Technol.*, vol. 28, no. 2, pp. 124-127, 2007.
- (16) Y. Kagawa, R. Shimoyama, T. Yamabuchi, T. Murai, and K. Takarada, "Boundary element models of the vocal tract and radiation field and their response characteristics," *J. Sound Vib.*, vol. 157, no. 4, 3850403, 1992.
- (17) J. Dang and K. Honda, "A physiological articulatory model for simulating speech production process," *Acoust. Sci. Technol.*, vol. 22, no. 6, pp. 415-425, 2001.
- (18) T. Kitamura, H. Takemoto, P. Mokhtari, and T. Hirai, "An MRI-based time-domain speech synthesis system," *J. Acoust. Soc. Am.*, vol. 120, no. 5, p. 3037, 2006.

(平成23年2月10日受付 平成23年3月18日最終受付)



たけもと ひろあき
竹本 浩典

1993 京大・理卒。1998 同大学院博士後期課程了。2000(株)国際電気通信基礎技術研究所入社。MRIを用いた声道形状の計測と音声生成研究に従事。2006 独立行政法人情報通信研究機構に移り、時間領域差分法を用いた音響シミュレーションに従事。博士(理学)。



きたむら たつや
北村 達也

1997 北陸先端大・情報科学研究科博士後期課程了。1997 静岡大情報学部助手。2002(株)国際電気通信基礎技術研究所。2007 甲南大・理・准教授。2008 改組により同大学知能情報学部准教授。音声生成、音声の個性に関する研究に従事。博士(情報科学)。