

学位論文題名

音声生成過程における音響特徴量抽出と
実体的声道モデルに関する研究

学位論文内容の要旨

我々が、音声合成の高品質について考える時、そのファクタとして、「了解性」、「自然性」、「個人性」、「情緒性」といったものが挙げられる。前者の2つは、主に韻質に関係するもので、いわゆる「聴き易い」音声を指すものである。一方、後者の2つは音質に関係している。音声には、誰の声であるのかということ、どのような心的状態であるのかといった情報をも含むことが可能であり、このような大きな利点を十分生かしてこそ高品質の音声合成が実現すると言える。

人間の音声は、アクセントやイントネーションといったことを論じる以前の、単なる発声の段階で既に個人性は生じている。発声は、基本的には音波の伝搬といった物理現象であるから、発声の過程を追っていく際に出会う、例えば声帯の振動状態、あるいは声道形状、声道長といった要素がそのまま個人の特徴パラメータ量となり得るわけである。

声道アナログ・モデルは、人間の発声器官である声道を模擬するもので、音声の生成機構を声道内の音波の伝搬までさかのぼってシミュレートすることで、合成音が得られるモデルとして知られている。このモデルは、音響管の伝搬問題を1次元の分布定数線路もしくはデジタル・フィルタで近似表現することで、コンピュータによる音声合成器が実現される。

しかしながら、声道アナログ・モデルが実体的モデルであると言っても、近似モデルである以上、いくつかの現象については簡単化がなされている。現在、簡単化のため近似、あるいは考慮されていない現象のいくつかは、合成音のより高品質化への重要な特徴量となり得るという指摘がなされるようになった。例えば、口唇部の境界条件もしくは音源生成部における空力学的な考察がなされていないことによって、高周波数において物理モデルとの整合性がとれないこと、または鼻腔を単に音響管モデルとして扱うのみでは鼻音の実測結果をうまく説明できないことなどが挙げられる。本研究で主に取り扱う「声道壁インピーダンス」もそのひとつであり、声道壁の振動によって起こる声道中音波の伝搬損失が、音声生成過程に複雑かつ効果的に影響していると考えられ、十分な検討を必要とするテーマである。本論文では、この声道壁振動について、従

来の音響管モデルの簡単化によって高品質化の妨げとなっている諸問題点を挙げるとともに、改良を行うために必要な人間の物理的音響特徴量の取得を試みる。

本論文は9章より構成されている。1章の緒論に続く各章の概要を以下に述べる。

第2章では、人間の音声生成機構についての概説と、声道アナログモデルの概説を行う。音声生成機構モデルを「音源部」、「声道部」、「放射部」に分離し、それぞれについて「物理的モデル」、「電気的モデル」、「デジタル・モデル」と移行していく段階を示す。また、3章以後の考察で多く使用される声道伝達特性の計算法についても述べる。

第3章では、本論文の全般に渡って考察される声道壁インピーダンスについての概説を行う。ここでは、従来より研究されている声道壁インピーダンスの報告例を示し、それらがどのようにして得られたかを解説する。また、声道壁からの放射を考慮した声道モデルの、電気的等価回路について述べる。

第4章では、人間の頬を音圧で駆動することでその音圧反射特性を実測し、そこから頬の音響インピーダンスを導出する。音響管法を用いれば、その音響管を口唇部より口腔内に挿入することによって、声道壁を直接測定試料として用いることが可能である。実験で得られた頬の音圧反射係数の周波数特性から、声道壁インピーダンスを導出する際、単純なRL素子によるパラメータで同定することは困難であることを示し、測定周波数の低減、高域で異なった値を持つ2種類のインピーダンスによって近似を行う。

第5章では、弾性体を壁とする均一音響管を用いて、管内の音圧分布の測定結果から減衰定数の同定を行う。管壁を粘弾性体モデルによって表現し、この時得られるインピーダンスを壁インピーダンスとして音響管モデルへ導入すれば、管壁の共振による伝達特性の変化も1次元の伝搬モデルによって十分近似可能となることを示す。

第6章では、MRI-CTを用いて声道形状の抽出を行い、そこから得られる断面積関数、周長関数および声道形状の潰れ度のファクタから、母音による特徴量の比較検討を行う。実際の声道形状データの取得は、音声生成過程における個人性のパラメータを獲得するという点で有効であり、第4、5章で得られた結果について、より実体的な考察が可能となる。ここで得られる「潰れ度関数」は、声道断面積関数と同様に非常に特徴的な概形を持つことを示す。また、ここでは同時に鼻腔の断面積関数の抽出も行い、鼻腔を考慮した声道伝達特性を求める。

第7章では、声道壁インピーダンスの粘弾性体モデルによる表現と、声道モデルへの導入について述べる。第3、4章で得られた声道壁インピーダンスの結果を参照し、さらに第5章で議論した粘弾性体モデルを声道壁に適応させて、実測結果を近似することが可能な声道壁モデルを構

築する。このモデルは、壁の軟らかさを壁厚のパラメータを変化させることによって調節することが可能である。第6章で抽出した声道断面積関数を用い、さらにMR画像から声道内の壁厚分布を視察することで、より実体的な声道内減衰特性を持つ声道モデルとしての検討が可能となる。得られた声道伝達特性は、実音声スペクトルとほぼ同様な特性を示し、良好な近似が可能となることを述べる。声道内で声道壁インピーダンスが様々な値で分布している場合、最も軟らかい、すなわち影響の大きい部分がどこに位置するかによって声道伝達特性は変化し、またその影響の度合は母音の調音状態によって異なることが確認できる。

第8章で扱うのは、これまでの議論とは多少異なる現象についてである。第7章までは、声道内の音圧によって生じる声道壁振動についての諸問題を取り扱った。ここでは、声道壁振動の影響を考える時もうひとつ重要となる声帯振動によって直接引き起こされる壁の“機会的”振動について考察する。この章では、合成音の実験として声帯振動の変位が振動波として声道の壁部分を伝搬してゆくような声道シミュレータを構築する。振動波の影響は、定常母音においても分析スペクトルのホルマントの時間的なばらつきが確認される。またこの現象も、/i/で最も大きいといった、調音状態による明確な違いがあることを示す。

第9章は、本論文全体の総括として、本研究で得られた成果を要約する。また、残された課題についての記述を行う。

本研究で得られた音響特徴量のいくつかは、出力端で観測される音声時系列波形を観測するだけでは十分な説明ができない、いくつかの複雑な現象の解明につながり、声道内で起こる多くの現象のうち、どのようなものをパラメータとして取捨選択すれば近似モデルとして有効であるかを示唆することができた。

学位論文審査の要旨

| | | |
|----|-----|------|
| 主査 | 教授 | 永井信夫 |
| 副査 | 教授 | 小川吉彦 |
| 副査 | 教授 | 栃内香次 |
| 副査 | 教授 | 小柴正則 |
| 副査 | 助教授 | 三木信弘 |

音声には、誰の声であるのかということ、どのような心的状態であるのかといった情報をも含むことが可能であり、このような大きな利点を十分生かしてこそ高品質の音声合成が実現すると言える。個人性に関しては、アクセントやイントネーションといったことを論じる以前の、単なる発声の段階で既に生じていると言える。発声は、基本的には音波の伝搬といった物理現象であるから、発声の過程を追っていく際に出会う、例えば声帯の振動状態、あるいは声道形状、声道長といった要素がそのまま個人の特徴パラメータ量となり得るわけである。

本研究で主に取り扱う「声道壁インピーダンス」もそのひとつであり、声道壁の振動によって起こる声道中音波の伝搬損失が、音声生成過程に複雑かつ効果的に影響していると考えられ、十分な検討を必要とするテーマである。本論文では、この声道壁振動について、従来の音響管モデルの単純化によって高品質化の妨げとなっている諸問題点を挙げるとともに、改良を行うために必要な人間の物理的音響特徴量の取得を試みる。本論文は9章より構成されている。1章の緒論に続く各章の概要を以下に述べる。

第2章では、人間の音声生成機構についての概説と、声道アナログモデルの概説を行う。音声生成機構モデルを「音源部」、「声道部」、「放射部」に分離し、それぞれについて「物理的モデル」、「電気的モデル」、「デジタル・モデル」と移行していく段階を示す。

第3章では、本論文の全般に渡って考察される声道壁インピーダンスについての概説を行う。ここでは、従来より研究されている声道壁インピーダンスの報告例を示し、それがどのようにして得られたかを解説する。

第4章では、人間の頬を音圧駆動することでその音圧反射特性を実測し、そこから頬の音響インピーダンスを導出した。音響管法を用いれば、その音響管を口唇部より口腔内に挿入することにより、声道壁を直接測定試料として用いることが可能である。結果より、声道壁インピーダンスは単純なRL素子によって同定することは困難であることを示し、測定周波数の低減、高域

で異なった値を持つ2種類のインピーダンスによって近似を行った。

第5章では、弾性体を壁とする均一音響管を用いて、管内の音圧分布の測定結果から減衰定数の同定を行った。管壁を粘弾性体モデルによって表現し、この時得られるインピーダンスを壁インピーダンスとして音響管モデルへ導入すれば、管壁の共振による伝達特性の変化も1次元の伝搬モデルによって十分近似可能となることを検証した。

第6章では、MRI-CTを用いて声道形状の抽出を行い、そこから得られる断面積関数、周長関数および声道形状の潰れ度のファクタから、母音による特徴量の比較検討を行った。実際の声道形状データの取得は、音声生成過程における個人性のパラメータを獲得するという点で有効であり、より実体的な考察が可能となる。

第7章では、声道壁インピーダンスの粘弾性体モデルによる表現と、声道モデルへの導入について述べた。このモデルは、壁の軟らかさを壁厚のパラメータを変化させることによって調節することが可能である。得られた声道伝達特性は、実音声スペクトルとほぼ同様な特性を示し、良好な近似が可能となる。声道内で声道壁インピーダンスが様々な値で分布している場合、最も軟らかい部分がどこに位置するかによって声道伝達特性は変化し、またその影響の度合は母音の調音状態によって異なることを示した。

第8章では、声道壁振動の影響を考える時もうひとつ重要となる声帯振動によって直接引き起こされる壁の“機械的”振動について考察した。ここでは合成音の実験として声帯振動の変位が振動波として声道の壁部分を伝搬してゆくような声道シミュレータを構築した。振動波の影響は、定常母音においても分析スペクトルのホルマントの時間的なばらつきが確認された。またこの現象も、調音状態による明確な違いがあることが示された。

第9章は、本論文全体の総括として、本研究で得られた成果を要約する。また、残された課題についての記述を行う。

以上のように本論文では、音声合成モデルを発生の物理現象とより整合のとれたレベルで構築を試み、この段階で得られた種々の特徴量は音声の生成過程に重要なファクタとなる新所見が得られている。これらは、高品質音声合成器構築のためのデジタル信号処理および電子工学に寄与するところが多い。よって本論文は学位論文として受理に値するものと認める。