

母音合成のための手の動きを利用した声道形状設定ツールの開発

Development of vocal tract shape setting tool using hand gestures for vowel synthesis

松村 洸平[†] 緒方 公一[†]
Kohei Matsumura Kohichi Ogata

1. はじめに

本研究室では、より柔軟な音声の生成を目的として、音声の生成過程に基づいた音声合成システム^[1]の開発を進めてきた。このシステムでは、生成に関与する声道形状を介して音声を合成するが、形状を表現する音響管の断面積を直感的な方法で設定できるツールの開発は、容易な操作や応用インターフェースの創出に有効と考えられる。本稿では、手の動きを利用した声道断面積パラメータ設定ツールの開発について報告する。

2. 音声合成システム

ここでは、本研究室で開発されている音声合成システムについて説明する。本システムでは、Sondhi と Schroeter により提案された声道シミュレータ^[2]を用いている。このシミュレータは、音声の生成過程を模擬しており、図 1 に示すように、声帯モデル、声道モデル、口唇放射モデルの 3 つから構成されている。

声帯モデルは、Ishizaka と、Flanagan の 2 質量モデルを用いて時間領域で表現されたモデル^[3]である。声道モデルは、図 2 に示すように、断面積が異なる等長円筒管を 20 個縦続接続することで、声道を 1 本の音響管として近似している。

本システムによる合成音声の生成は、図 3 に示すように、図 1 で示した声道断面積や声道長などのパラメータ

の値が記述されたパラメータファイルを、音声合成システムに入力することにより行われる。このファイルのうち、声道断面積パラメータとして、各円筒管の断面積の時系列データが記述されており、これらの値により声道形状の変化を表現する。音韻は声道形状により変化するので、声道断面積パラメータの変化により、音韻の変化を表現することができる。

3. 声道形状設定ツール

ここでは、本稿で開発した声道形状設定ツールについて説明する。声道形状設定ツールとは、前章で説明した音声合成システムにおいて、声道断面積パラメータを手の動きにより設定するツールである。指の曲げ具合をデータグローブにより数値化し、一定周期ごとにサンプリングを行い、さらに、スケール変換した各指の数値を、各指と対応した円筒管の断面積とすることで声道断面積パラメータの設定を実現している。

データグローブには 5DT 社製の 5DT Data Glove 5 Ultra を使用した。このサンプリング周波数は 60Hz で、指の曲げ具合は 12bit で A/D 変換される。また、センサ数は各指に 1 つずつの計 5 個である。

今回は、図 4 に示すように、各指を、円筒管 20 個のうち等間隔に選択した 5 個の円筒管と対応させている。また、指と対応付けされていない円筒管の断面積は、対応させている円筒管の断面積を用いて線形補間により求めた。なお、声門部の円筒管の断面積はあまり変動がないので固定値(0.81cm²)とした。この固定値は、本研究室で導出した、成人男性を想定した平均的声道形状のデータ^[4]の声門部の断面積のうち、他の母音に比べて大きな値を示した/u/を除いた母音についての平均値となっている。

指の曲げ具合から声道断面積パラメータへのスケール変換には式(1)の線形変換を用いた。



図 3 音声合成システムによる音声合成の流れ

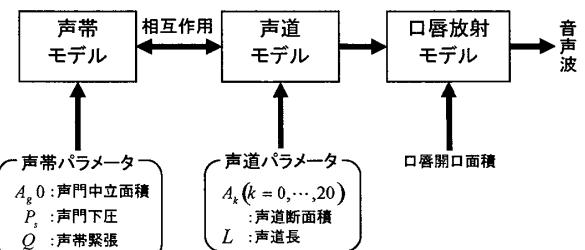


図 1 声道シミュレータの構成

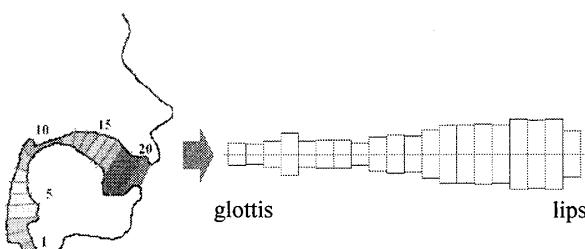


図 2 声道モデルの表現

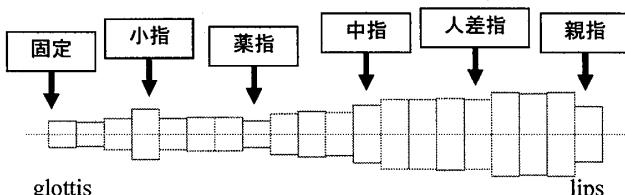


図 4 各指と円筒管との対応関係

[†]熊本大学大学院自然科学研究科 Graduate school of Science and Technology, Kumamoto University

$$A = \frac{x_{\max} - x}{x_{\max} - x_{\min}} A_{\max} \quad (1)$$

ここで、 x は指の曲げの測定値、 A は声道断面積の値を表す。システム起動後、手を開いた状態と閉じた状態とで指の曲げ具合を測定しており、それぞれ x_{\min} と x_{\max} の値として採用している。また、 x の範囲はデータグローブの仕様により [0, 4095] である。 A の範囲は [0, A_{\max}] である。 A_{\max} は、本研究室で導出した、成人男性を想定した平均的な声道形状のデータをもとに定めた。今回は母音合成を行うので、前述の母音のデータのうちでもっとも大きい/o/の17番目の断面積の値(7.38cm²)を用いた。

声道形状設定ツールによる声道断面積パラメータのパラメタファイル生成の流れを図 5 に示す。指の曲げの記録時には、図 5 の中央に示すように、その時点での指の曲げ具合により生成される声道形状が音響管として表示される。なお、目標とする母音の声道形状を点グラフで表示する機能を有しているので、それを参考に手の形を変えて声道形状を変形させることにより、母音の声道断面積パラメータを生成することができる。

4. 単母音合成実験

本ツールを用いて、図 5 の流れで単母音 (/a/, /i/, /u/, /e/, /o/) の声道断面積パラメータ、および、合成音声を生成した。そして、それらと、成人男性を想定した平均的な声道形状のデータ、および、その合成音声とを比較した。

図 6 は母音/a/における声道断面積での比較結果である。これより、本ツールにより作成した声道断面積パラメータは、成人男性を想定した平均的な声道断面積と比較して、大局的に類似した形状となっていることがわかる。他の母音においても同様に大局的に類似した形状となつた。

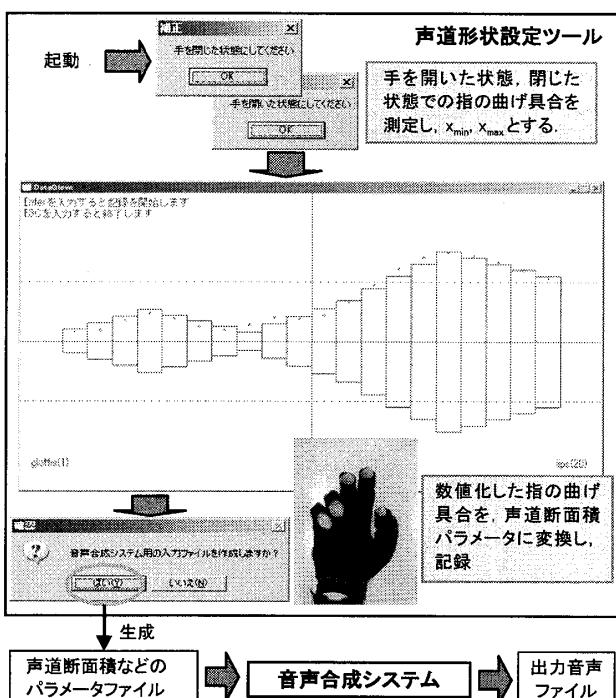


図 5 声道形状設定ツールによる音声合成の流れ

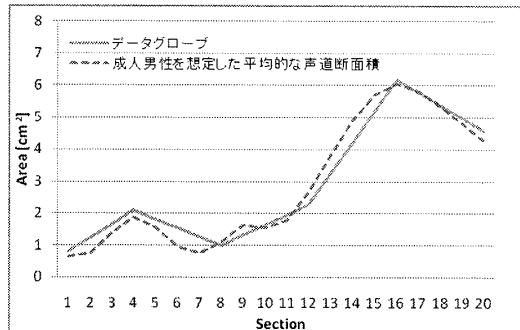


図 6 声道断面積パラメータの比較 (/a/の場合)

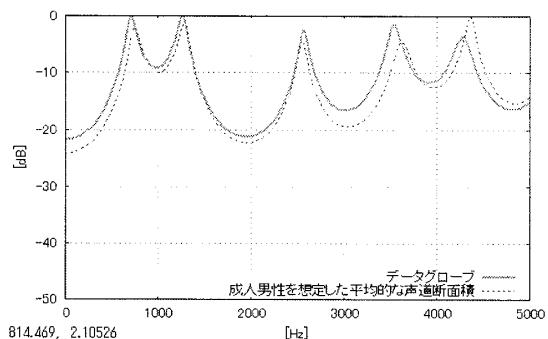


図 7 伝達特性の比較 (/a/の場合)

図 7 は母音/a/における伝達特性での比較である。これより、伝達特性においても、本ツールにより生成した母音と、成人男性を想定した平均的な声道断面積による母音は大局的に類似した特性となっていることがわかる。他の母音においても同様に、大局的に類似した結果となった。

5.まとめ

本稿では、母音合成のための手の動きを利用した声道断面積パラメタ設定ツールを開発した。実験の結果、手の動きにより大局的に声道形状を再現することができ、生成された声道音響管の声道伝達特性も大局的に類似したものとなった。このことから、手の動きにより母音を生成することは可能といえる。今後は、連続母音の生成や、入力方法の改良について検討する。

謝辞

本研究の一部は科学研究費補助金((C)20560398)の援助によることを記し、謝意を表する。

参考文献

- [1] 緒方 公一, 増矢 拓郎, “Web アプリケーションとしての声道音響管モデルに基づく母音合成システムの開発”, 日本音響学会誌, Vol.62, No.3 (2006).
- [2] M.M.Sondhi, J.Schroeter, “A hybrid time-frequency domain articulatory speech synthesizer”, IEEE Trans. Acoust., Speech & Signal Process., Vol.ASSP-35, No.7 (1987).
- [3] K.Ishizaka, J.L.Flanagan, “Synthesis of voiced sounds from a two-mass model of the vocal cords”, Bell Syst.Tech.J., Vol.51, No.6 (1972).
- [4] K.Ogata, B.Yang, “A study of vocal tract area function for speech synthesis system based on transmission line model”, Proceedings of 19th International Congress on Acoustics (ICA2007), Vol.CD-ROM(cas-03-010.pdf) (2007).