

rtMRI動画から抽出した発話器官の輪郭データに基づく 平均声道の生成と分析*

☆後藤翼, 天野沢海, 竹本浩典 (千葉工大),
北村達也 (甲南大), 能田由紀子, 前川喜久雄 (国語研)

1 はじめに

近年, MRI 撮像技術の進展により発話運動をリアルタイムの動画 (rtMRI 動画) として記録することが可能となった。そこで, われわれは日本語調音音声学の精緻化のために, rtMRI 動画のデータベースを構築してきた。[1]。2020 年現在, このデータベースには 25 名の被験者によるそれぞれ 676 種類のバイモーラ発話などを記録した多数の動画が収録されている。各動画はフレームレート 14 fps で約 37 秒撮像された 512 フレームからなり, 約 20 種類の発話を含んでいる。

rtMRI 動画から発話器官の輪郭を点群データとして抽出すれば, 発話運動の定量的な分析が可能である。しかし, rtMRI 動画は発話器官の輪郭が不鮮明なため, 画像処理で自動的に輪郭を抽出することは困難である。この問題を解決するため, われわれは機械学習を導入した。その結果, 被験者ごとに約 20 フレームから発話器官の輪郭を手動で抽出して学習器を生成すれば, その被験者の全ての動画から輪郭を抽出することが可能となった[2]。

そこで, 本稿では, データベースから抽出した 18 名の日本人被験者の安静呼吸時の輪郭から平均声道を計算し, その分析を行ったので報告する。

2 材料と方法

2.1 輪郭

材料は, 日本人成人男性 12 名 (M1~M12), 女性 6 名 (F1~F6) の発話前の安静呼吸時の動画フレームから手動で抽出した輪郭である。この輪郭は, 舌, 口唇・下顎, 軟・硬口蓋, 咽頭後壁・披裂部, 喉頭蓋・声帯の 5 つの部位からなり, それぞれ 40 点, 40 点, 30 点, 30 点, 30 点から構成されている。なお, 各輪郭線を構成する点群は等間隔に配置した。

2.2 平均声道

前述したように, 全ての被験者において, 同じ部位の輪郭は等間隔で等しい数の点で構成されているので, 個々の点の幾何重心をとることで平均の輪郭を生成することが可能である。そこで, 全ての被験者の輪郭を前鼻棘が原点となるように平行移動した後, 硬口蓋が水平になる角度に調整した。そして, 5 つの部位の輪郭を構成する全ての点で幾何重心を求めて得られる輪郭を平均声道とした。

また, 男性のみ, 女性のみ輪郭を用いてそれぞれの平均声道を求め, 男性平均声道, 女性平均声道とした。そして, 薄板スプライン[3]を用いて, 各被験者の声道および男性・女性平均声道のそれぞれが平均声道を基準としてどのように変形しているかを可視化した。

2.3 声道形状の分析

声道形状の個人差を検討するために, 全被験者の輪郭を用いて主成分分析と最小分散法 (ward 法) によるクラスタ分析を行った[4]。

3 結果・考察

Fig.1 は平均声道, 男性平均声道, 女性平均声道と, 平均声道で設定した直交格子が男性・女性平均声道にどのように写像されたかを示している。平均声道と比較すると, 男性平均声道は前下方へ拡大するように変形している。一方, 女性平均声道は後上方に縮小するように変形している。これは, 男性に比べて女性では声道が小さく, 喉頭が高く, 声道が狭いことを反映していると考えられる。また, 変形量は全体的に男性で小さかった。これは, 被験者の内訳が男性 12 名, 女性 6 名であり, 平均声道は男性の声道形状の影響をより大きく受けたため, 結果として平均声道と男性平均声道の差が小さくなったと考えられる。

* Analysis of mean vocal tract shape based on edge lines extracted from rtMRI movies, by GOTO, Tsubasa, AMANO, Takumi, TAKEMOTO, Hironori (Chiba Institute of Technology), KITAMURA, Tatsuya (Konan Univ.), NOTA, Yukiko, and MAEKAWA, Kikuo (NINJAL)

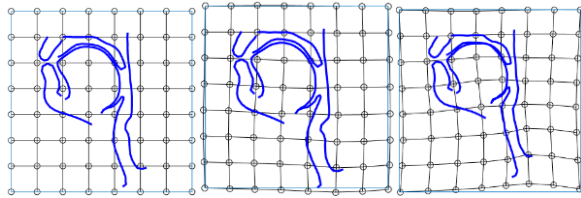


Fig. 1 平均声道 (左), 男性平均声道 (中央), 女性平均声道 (右)

Fig.2 は声道形状のクラスタ分析の結果である。まず、女性型 (F1~F6) と男性型 (M3~M9) に大きく分かれた。女性型は声道が小さかった (Fig. 3a)。男性型は4つのクラスタに細分され、M3~M7 は声道が前方に拡大 (Fig. 3b), M4, M12 は声道が下方に拡大 (Fig. 3c), M1, M11 は声道が後方に拡大していた (Fig. 3d)。M8~M9 はその他で、共通した特徴は見られなかった。

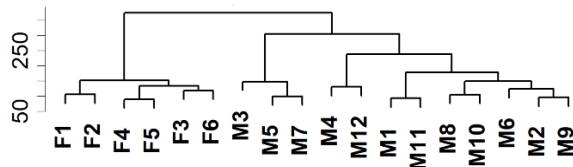


Fig. 2 声道形状のクラスタリング

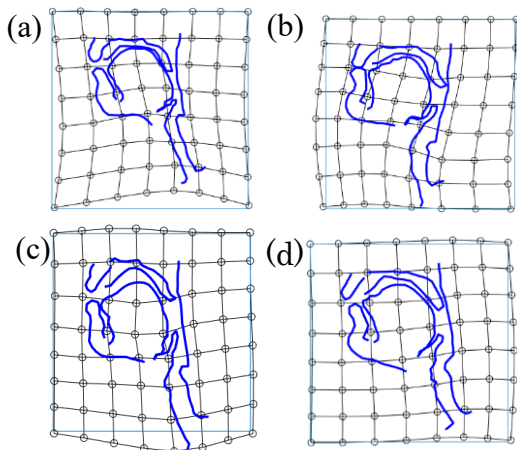


Fig. 3 各クラスタの特徴を含む4名の被験者 ((a): F4, (b): M5, (c): M4, (d): M1)

Fig.4 は主成分分析の結果で、第1~3主成分 (PC1~PC3) による変形 (平均声道から±1標準偏差) を示す。PC1は前方-後上方の変形で、寄与率は0.40であった。これは、男性平均声道と女性平均声道の差 (Fig. 1) とほぼ一致し、Fig. 2の男性のクラスタと女性のクラスタを分離する主要因であると考えられる。PC2は頭部と頸部の角度による変形で、寄与率は0.25であった。これは、MRIによる

撮像は仰臥位で行われるため、発話しやすい頭部と頸部の角度に個人差があることを反映していると思われる。PC3は頸椎と下顎の相対的な位置による変形で、寄与率は0.09であった。これは、頸椎の後湾に伴って下顎や舌が後方に引かれることを示すと思われる。

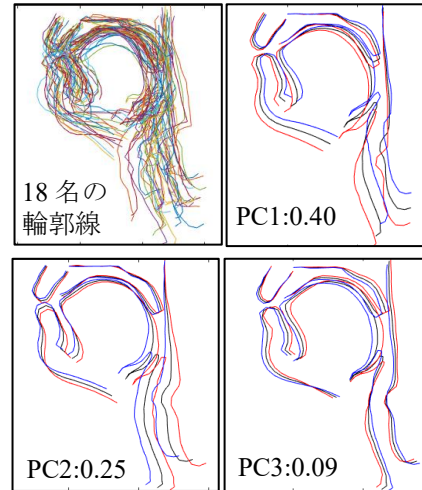


Fig. 4 PC1~PC3による変形 (黒: 平均声道, 赤: 平均声道+1標準偏差, 青: 平均声道-1標準偏差, 数値は寄与率)

4 まとめ

本研究では、rtMRI データベースに含まれる18名の被験者の安静呼吸時における声道の輪郭線を用いて平均声道を生成し、形態分析を行った。その結果、クラスタを形成する形態的な要因などが明らかになった。また、本稿では、薄板スプラインを用いて平均声道から各個人の声道への変形を分析したが、この手法を応用すれば、各個人の発話運動に伴う輪郭の変化を平均声道へ写像することが可能になると思われる。

5 謝辞

本研究は JSPS 科研費 17H02339 の助成を受けた。

6 参考文献

- [1] 前川ら, 音講論 (春), 1247-1248, 2018.
- [2] Takemoto *et al.*, Proc. of Interspeech 2019, 904-908, 2019.
- [3] Ubukata., The Quaternary Research 44(5), 197-313, 2005
- [4] Yoshihara *et al.*, Journal of Surface Analysis Vol.21 No.1, 10-17, 2014