

多話者におけるリアルタイムMRI動画への歯列補填の精度分析*

☆常盤朔也, 竹本浩典 (千葉工大), 前川喜久雄 (国語研)

1. はじめに

われわれは調音音声学の精緻化のためにリアルタイム MRI を用いて正中面の調音運動を記録した動画のデータベース (rtMRIDB) を構築し, 公開している[1]。その動画を用いることにより, 調音運動を定量的に分析できるが[2], MRI では歯列と口腔が同じ輝度値 (黒) であるため, 歯茎摩擦音などの分析で問題となっている。そこでわれわれは, 動画の各フレームに画像処理により歯列を補填する手法を考案して報告した[4]。しかし, その精度の検討は1名の話者に対してのみであった。そこで本研究では, 複数の話者のMRI動画へ歯列を補填してその精度を検討したので報告する。

2. 材料と方法

2.1. 材料

rtMRIDB に登録されている話者 10 名 (s1, s4, s5, s7, s11, s12, s17, s18, s26, s27) の動画に歯列を補填した。動画にはキャリア文「これは〇〇型」による 2 モーラ語の約 20 発話などが記録されている[1]。各動画は 512 フレームで構成され, 色深度は 16 bit, 画像サイズは 256×256 ピクセル, 空間解像度は 1×1×10 mm である。また, 撮像速度は約 14 フレーム毎秒または約 27 フレーム毎秒である。なお, 撮像は ATR-Promotions に設置された Siemens 製 MAGNETON Prisma fit 3 で行った。

2.2. 歯列データの補填手順

前報[4]で詳細を記述したので, ここではその概略を述べる。まず, 話者ごとに1本の動画から上・下顎の歯列が舌や口唇などの軟組織で囲繞されて明確に観察できるフレームを決定した。これを歯列抽出フレームとする。そして, 上・下顎とその歯列をトレースして抽出し, それぞれ上・下顎歯列データとした。なお, 本稿では歯列のみの補填を目的としているので前報[4]で含めていた口蓋骨は除外して歯列データを作成している。

次に, 歯列抽出フレームと補填する動画の

第1フレームに上顎では鼻の先端付近, 下顎では下顎骨下部を内包する矩形領域を設定する (Fig.1)。そして, 歯列抽出フレームと補填する動画の第1フレームの上顎と下顎の対応する矩形領域を位置合わせする剛体変換行列を画像の強度に基づく勾配降下法で求める。最後に, 歯列データを前述の剛体変換行列で写像して第1フレームに歯列を補填する。なお, 下顎は顎の動きが大きいので, 求めた剛体変換行列を用いて歯列抽出フレームの矩形領域を移動させて第2フレームの矩形領域とした。これを第2フレーム以下に順次適用し, MRI 動画へ歯列を補填する[4]。

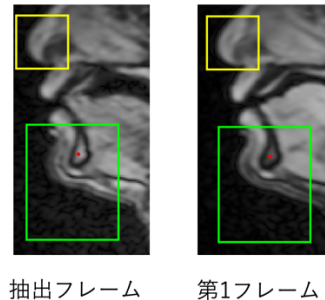


Fig.1 矩形領域 (黄: 上顎, 緑: 下顎)

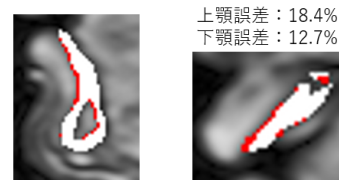


Fig.2 補填した歯列 (白) とはみ出した部分 (赤). 左: 下顎, 右: 上顎

2.3. 補填精度の評価

補填精度の評価も前報[3]に準じる。すなわち, 歯列を補填したフレームで上・下顎歯列データを構成するピクセルがどの程度の割合で周囲の軟組織にはみ出したかを表す誤差面積率で評価する (Fig.2)。また, 歯列抽出フレームで歯列データを上下左右に 1 ピクセル平行移動, あるいは±1 度回転移動させた際の誤差面積率を算出し, その最大値を誤差の参考値とする。各話者で誤差面積率が参考値を超えたフレーム数を, 歯列を補填した動画の

* Accuracy of tooth superimposition on real-time MRI videos for multiple speakers, by TOKIWA, Sakuya, TAKEMOTO, Hironori (Chiba Institute of Technology), and MAEKAWA, Kikuo (NINJAL)

総フレーム数で除した値を参考値超過率とし、この値が 5%未満であれば補填精度が高いと評価した。

3. 結果と考察

Table1, Table2 は話者ごとの上顎と下顎の歯列補填の精度を示す。上段から、誤差面積率の平均値, 最大値, 参考値, 歯列を補填した総フレーム数, 参考値超過率である。

Table 1 上顎の歯列補填の精度

	s1	s4	s5	s7	s11
平均値	5.9	3.8	2.5	10.0	5.0
最大値	30.4	27.3	16.5	52.3	34.4
参考値	15.7	8.9	10.3	18.5	11.9
総フレーム数	28,160	27,140	30,727	30,720	29,188
参考値超過率	2.0	3.3	1.4	6.2	5.0
	s12	s17	s18	s26	s27
平均値	11.0	3.4	12.0	4.4	10.1
最大値	39.6	40.9	47.9	26.2	40.9
参考値	21.8	13.5	21.9	14.0	26.2
総フレーム数	29,701	28,676	30,208	27,648	27,136
参考値超過率	5.2	2.8	5.3	0.8	0.7

Table 2 下顎の歯列補填の精度

	s1	s4	s5	s7	s11
平均値	21.4	6.8	2.5	20.6	11.3
最大値	50.0	21.1	17.9	42.5	31.3
参考値	39.6	24.7	28.8	35.1	30.2
総フレーム数	28,160	27,140	30,727	30,720	29,188
参考値超過率	0.4	0.0	0.0	0.1	0.0
	s12	s17	s18	s26	s27
平均値	11.9	9.9	12.4	9.2	7.5
最大値	59.9	33.2	45.0	26.5	27.8
参考値	47.7	37.5	36.5	35.5	45.1
総フレーム数	29,701	28,676	30,208	27,648	27,136
参考値超過率	0.1	0.0	0.2	0.0	0.0

Table 1 と Table 2 で示すように、誤差面積率の平均値は参考値を下回った。しかし参考値超過率は、下顎では全ての話者で 5%を下回ったが上顎では話者 s7, s11, s12, s18 で 5%を超過した。

これを改善する方法を検討する。まず、その原因の一つは矩形領域の大きさが適切でないことであると考えた。そこで上顎の矩形領域を上下に 1.5 倍に拡大した (Fig.3)。次に、発話に伴う上顎の顎の動きは小さいと仮定して上顎では位置合わせの矩形領域を固定していたが、これらの話者では上顎の動きも大きい考え、下顎と同様に 1 フレームごとに矩形領域の位置を移動させて歯列補填を行った。

Table 3 はその結果で、全ての話者で参考値超過率が 5%未満に減少した。よって、上顎で参考値超過率が大きかった原因は、矩形領域が小さかったことと、上顎の動きが小さいと仮定したことであった。しかし、下顎に比べ

て上顎では参考値超過率が大きい傾向は変わらなかった。



オリジナルの矩形領域 拡大した矩形領域

Fig. 3 上顎の矩形領域を拡大

Table 3 矩形領域の大きさとフレームごとに位置合わせを実施した場合の参考値超過率

	s7	s11	s12	s18
参考値超過率	4.0	4.7	2.9	3.6

4. まとめ

本研究では前報[4]で開発した MRI 動画へ歯列を補填する手法を用いて、日本人話者 10 名に歯列補填を行い、その精度を検討した。その結果、各話者の上顎と下顎の歯列を補填した際の誤差面積率の平均値は参考値を下回った。しかし、参考値超過率は下顎では全ての話者において 5%を下回ったが、上顎では一部の話者で 5%を超えた。その原因は、矩形領域の大きさが小さかったことと、上顎の動きが小さいと仮定して位置合わせに用いる矩形領域をフレームごとに更新しなかったことであった。これらを改善したところ、上顎でも全ての話者で参考値超過率は 5%を下回った。よって、全ての話者の 95%のフレームで位置合わせの誤差は、平行移動・回転移動の成分ではそれぞれ 1 ピクセル程度・1 度程度以内であった。しかし、全体として下顎より上顎で参考値超過率が大きい原因と、上顎でさらに参考値超過率を減少させる要因を明らかにすることはできなかった。

謝辞

本研究は JSPS 科研費 20H01265 の支援を受けた。

参考文献

- [1] 前川ら, rtMRIDB_v1, <https://rtmridb.ninjal.ac.jp/>.
- [2] 天野ら, 音講論 (春), 1013-1014, 2022.
- [3] H. Takemoto *et al.*, AST., 468-474, 2004.
- [4] 常盤ら, 音講論 (春), 597-598, 2023.