

26 話速変化情報抽出のための口唇形状変化分析に関する検討

角丸貴洋* 広重真人 荒木健治 栃内香次

北海道大学大学院工学研究科†

1 はじめに

日本語音声の音韻的特徴のひとつとして、話速がほぼ一定であることがあげられる。これは朗読音声に顕著な特徴であるが、日常会話音声においてはしばしば大きな話速変化が観測される。このことから会話などにおいて、局所話速変化の個所には話者の意図や強調などの表現が含まれていると考えられる。我々は、そのような個所を捉えるために、局所話速変化検出の研究を進めている [1]。

一方、人間の口の動きは音声に深い関わり合いがあると考えられている。これまでも、口の動きのみから話している言葉を推測しようとする読唇や、口のカラー動画像から音韻を自動認識する音韻認識など、口唇画像を音声処理に応用するさまざまな研究が進められている [2],[3]。筆者らは、話速分析へ口唇画像を利用することを考え、視覚的情報である口唇形状の変化を、話速変化を捉えるための補助情報として利用するための解析手法を検討している。従来より、PARCOR 係数から声道断面積比を求め、それを利用して口唇端の断面積比を求める方法が知られているが [4]、この手法では単一時刻の断面積比が求まるのみであり、連続的な面積変化の算出には困難を伴う。筆者らはこのことから画像を併用した解析手法の可能性を検討している [5]。

本稿では、3名の話者について口唇面積の変化を口唇画像を用いて測定し、話速変化との関連や話者による相違について比較・検討を行う。

2 処理の流れ・測定方法

口唇面積の変化を求める手順を述べる。まず、口唇全体を正面から家庭用ビデオカメラで撮影し、同時に DAT

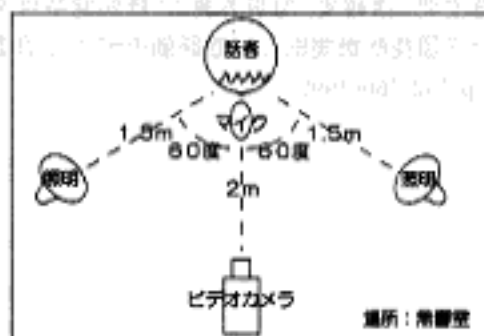


図1: 撮影時の位置関係

で音声を録音する。この際、口唇内は暗くなるように照明をあてる。カメラ・マイク・話者・照明の位置関係を図1に示す。ビデオカメラで撮影した映像から画像サイズ 640 × 480(pixel)、フルカラー (1677 万色)、フレームレート 30(fps) で口唇画像を取り込み口唇の面積測定を行い、DAT で 48kHz サンプルングで録音した音声を 12kHz にダウンサンプルングしたものを用いて、音素区切りおよび文節区切りごとに求めた平均話速の時間変化を求める。ビデオカメラの音声を基にしてビデオカメラの画像と DAT の音声の同期を取る。

2.1 顔画像からの口の面積の測定方法

口唇画像の口の面積の測定方法を以下に示す。フルカラーで取り込んだ画像をグレースケールに変換する。取り込んだ画像から処理対象となる口唇周辺領域を手動で抽出し、この領域について処理を行う (図3参照)。



図2: 取り込んだ口唇原画像

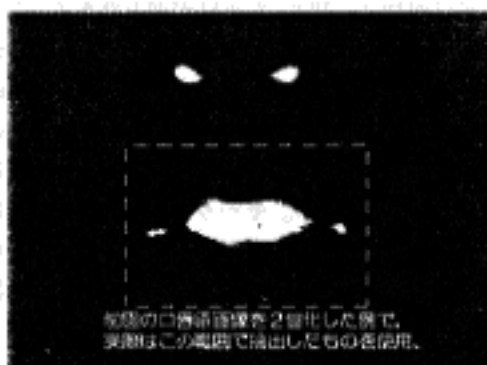


図3: 2値化画像

閾値を設定して画像を2値化し、暗い部分を開口領域として抽出する。ここで、開口領域に値が含まれる可能性を考慮し2段階に分けて抽出する。以下に2つの方法を示す。

*kaku@media.eng.hokudai.ac.jp

†札幌市北区北13条西8丁目北海道大学工学部

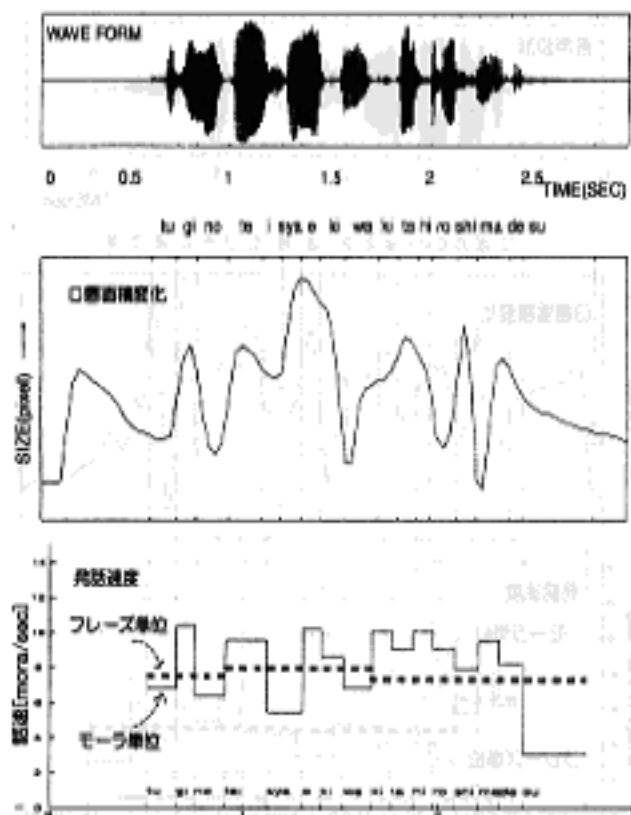


図4: 自然な速さでの発話例(話者S)

取り込んだ口唇原画像の例と、その画像を2値化し暗い開口領域を抽出した画像の例を図2,3に示す。

2.1.1 暗い開口領域の抽出法

暗くなっている口唇内を抽出できるように、閾値を手動で設定して画像を2値化する。閾値を少しずつずらし、口唇原画像と2値化画像とを見比べて最適な閾値を用いて、暗い開口領域を抽出する。歯が出ていない場合はここで抽出した領域が開口領域となる。

2.1.2 開口領域内の明るい領域(歯)の抽出法

暗くなっている口唇内の中の明るい部分(主に歯)を抽出できるような閾値を手動で設定して画像を2値化する。閾値を少しずつずらし、口唇原画像と2値化画像とを見比べて最適な閾値を用いて、明るい領域を抽出する。この領域と2.1.1に述べた方法で抽出した領域を論理和により合わせたものを口唇開口領域としこの面積を口唇面積とする。

3 実験と考察

3.1 実験と結果

実験に用いたデータは3名の男子大学院生によって発話された数文の音声と口唇画像で、話者には自然な口形で発話することと、1つの文につき以下に示す4種類の異なる速さで発話することの2つを指示した。

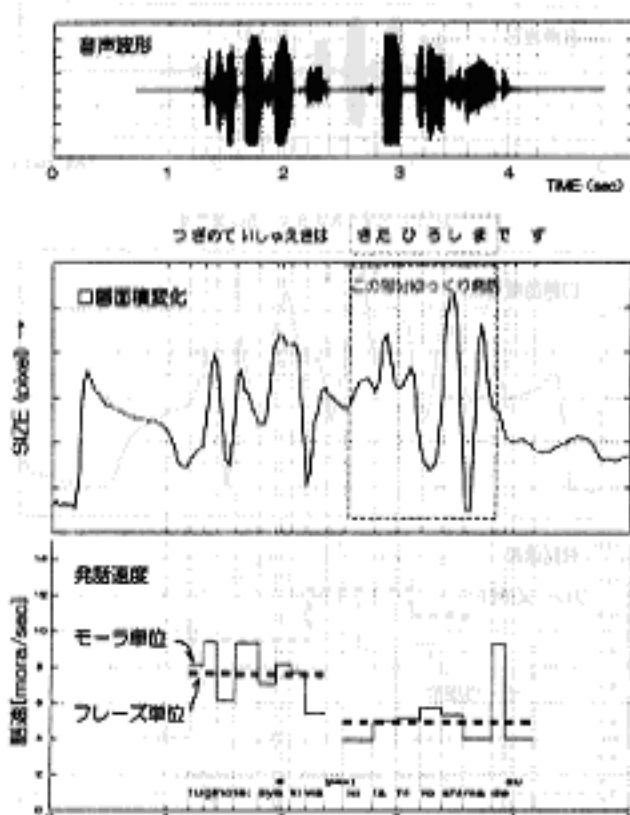


図5: 一部分をゆっくりと発話した例(話者S)

1. 自然な速さで発話する。
2. 全体的に速く発話する。
3. 一部分(指定したところ)をゆっくりと発話する。
4. 全体的にゆっくりと発話する。

このデータに対して前述の方法で画像から口唇面積変化を測定し、話速変化との比較、および話者の違いによる変化の検討を行う。

「次の停車駅は北広島です。」

という持続時間が約3秒間の文に対して、それぞれの話速パターンについて、面積変化と話速変化の例(話者S)を図4,5,6,7に。

「春になって田植えの季節がやってきた。」

という文に対して、「自然な速さで発話」について、面積変化と話速変化の例(話者S)を図8に示す。下線部はゆっくりと発話するように指定した部分である。それぞれの図において上から、音声波形、面積変化、話速変化を示している。「自然な速さで発話」した同一文について、話者Sと話者Mによる面積変化の比較を図9に示す。

これらの結果から、同一話者が同じ文を話していても話速が変化すると口唇面積変化も異なってくるのがわかる。しかし、口を閉じているとみられる顕著に下に凸な箇所はすべてにおいてほぼ同じ回数であることが観測される。話者が異なると同一文をほぼ同じ話速で発話していても口唇面積変化のグラフ形状は異なっている。こ

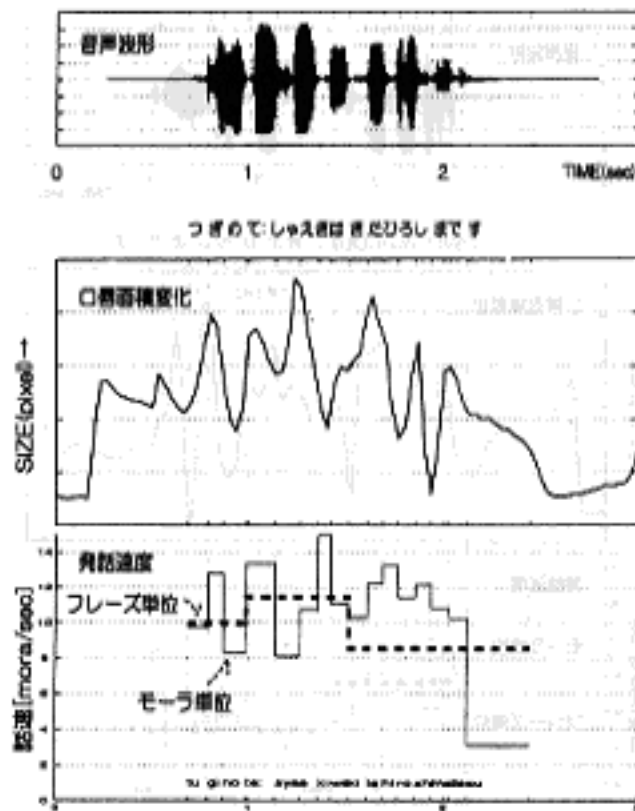


図6: 全体的に速く発話した例 (話者S)

のことからも、口を閉じているとみられる回数はほぼ同じであることが観測される。

3.2 考察

本稿では、以下の3項目について検討を行った。

- 同一話者による同一文で話速の変化による比較
- 同一話者による異なる文での比較
- 異なる話者による同一文での比較

口を閉じているとみられる顕著に下に凸な箇所はすべてにおいてほぼ同じ回数であることが観測されることから、口を閉じていなければならない音素を発話している箇所は、話速パターンが変化しても本稿の手法で検出可能である。話者が異なる場合細部において面積変化が異なっていることから、このような細部の変化を取り除き大きな変化のみを抽出する、話者間に依存しない面積変化の検出方法の検討が必要である。

4 おわりに

本稿では、局所的な話速変化検出の補助情報として口唇面積変化を利用することに着目し、口唇面積変化と話速との比較を行ない、さらに話者の違いによる相違を調査することにより口唇面積情報の可能性について検討した。

話者の違いによる口唇面積変化の共通点や相違点の検討が必要であり、さらには口唇面積の測定方法の改良、音声パワーなど他の音響パラメータとの関連、面積変化

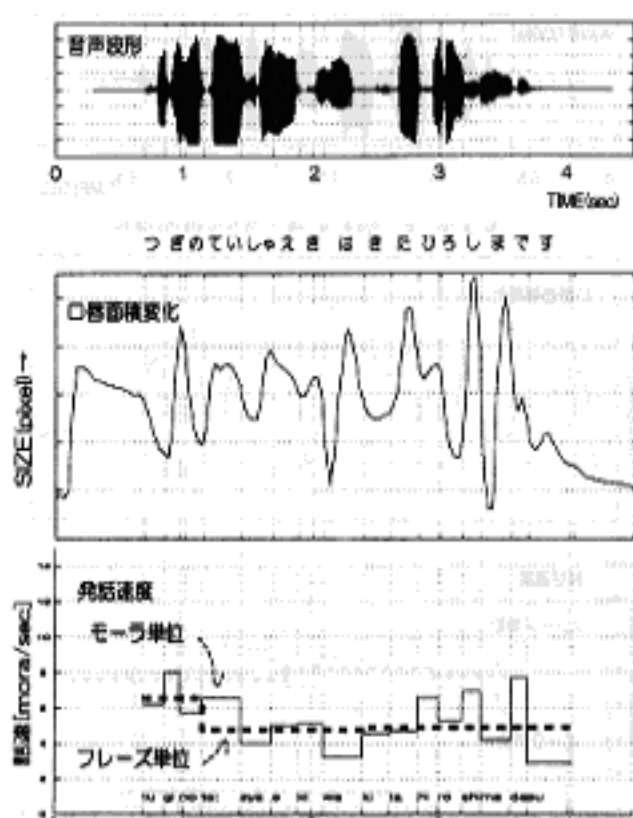


図7: 全体的にゆっくり発話した例 (話者S)

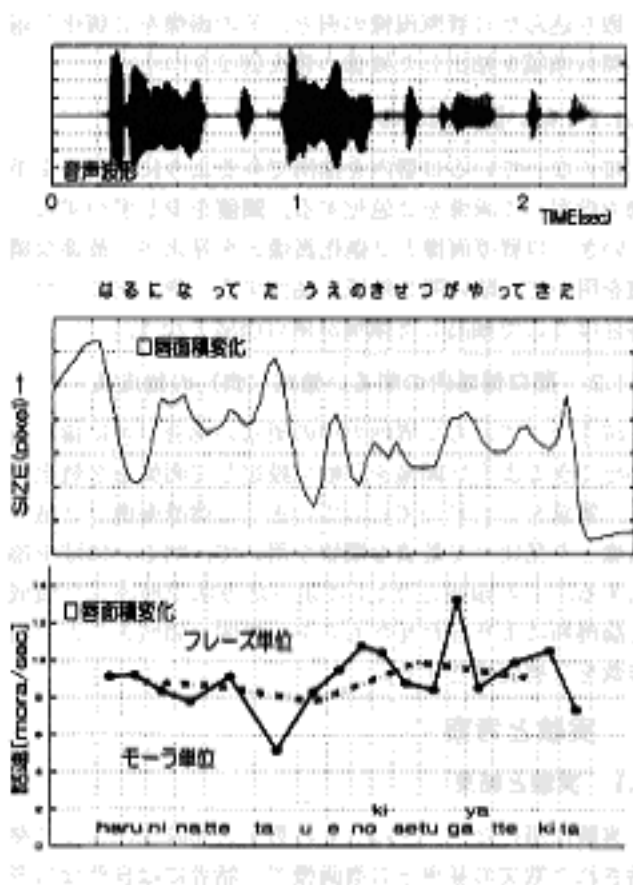


図8: 異なる文において、自然な速さで発話した例 (話者S)

からの話速変化パラメータの抽出などを今後進めていく予定である。

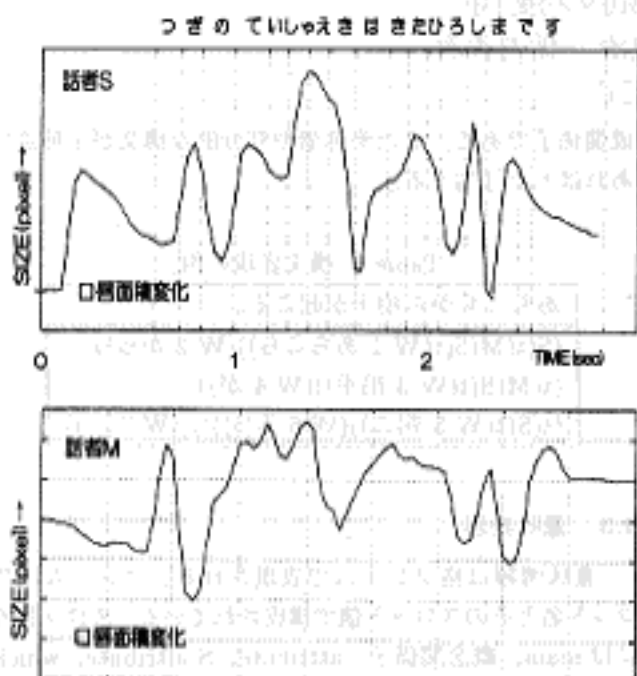


図9: 2人話者による自然な速さで発話した例での比較 (上図: 話者S, 下図: 話者M)

参考文献

- [1] 佐々木洋一, 広重真人, 宮永喜一, 板内香次: "局所的な話速変化検出を指向した簡易セグメンテーション手法と実環境音声への適応について", 信学技報, SP98-142, pp.47-54(1999-02).
- [2] 積山薫, 城和貴, 梅田三千雄: "読唇による単音節の認識", 信学技報, IE87-108, pp.33-40(1988).
- [3] 寺田賢治, 山中理聖子, 大恵俊一朗: "口のカラー動画像を用いた音韻認識", 電学論 D, Vol.119-D, No.1, pp.37-43(1999).
- [4] 古井貞照: "デジタル音声処理", 東海大学出版会, (1985).
- [5] 角丸貴洋, 広重真人, 荒木健治, 板内香次: "話速変化情報抽出のための口唇形状変化分析に関する予備的検討", 電気関係学会北海道支部連合大会講演論文集, p.300(1999).

の抽出を目的として、この論文では話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。まず、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。

本論文は、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。まず、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。

この論文は、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。まず、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。

この論文は、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。まず、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。

この論文は、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。まず、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。

この論文は、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。まず、話者の自然な発話速度と、その速度変化の検出と、その検出結果を用いた音声処理の手法について検討する。