

# 病的音声のための GRBAS 尺度の 自動推定に関する検討

松岡 駿平 佐宗 晃  
産業技術総合研究所

## 1. はじめに

疾患の早期発見は、根治させる観点において重要である。現在、発話から声に関する疾患を検査する手法として、GRBAS 尺度に基づいた聴覚心理的評価が行われている。GRBAS 尺度とは、音声の総合的な異常度を表す Grade という尺度と、異常さの内容の尺度である粗糙性(Rough)、気息性(Breathy)、無力性(Asthenic)、努力性(Strained)の頭文字を並べたものであり、それぞれの尺度において 4 段階で評価を行う。この尺度を基準に評価を行うことで発声に関する疾患の診断が可能である[1]。

一方で、GRBAS 尺度をもとにした評価は主観的になるため評価者によって結果に差異が生じやすい。そのため、正確かつ統一的な評価を行うために、発話データから GRBAS 尺度の自動推定を行う研究が活発に行われている。本稿では、推定に有効な新たな特徴量の検討について論じる。

## 2. 関連研究

先行研究で用いられてきた特徴量は、主に振幅スペクトルや音声波形の時系列データから計算されている。振幅スペクトルから計算される特徴量としてはメル周波数ケプストラム係数が、音声波形からは LPC 残差の高次特徴量などが挙げられる。これらの特徴量は単体で用いるより、組み合わせることで使用することによって推定精度が向上することが示されている[2]。本研究では新たに、位相スペクトルが持つ特徴に着目した特徴量を提案し、従来の特徴量と組み合わせることによって識別率の向上を試みる。

## 3. 提案する特徴量の抽出方法

音声波形の位相は、声帯が閉じる瞬間に全ての周波数の位相が一致するが、声帯に何らかの異常が存在すると、正常に声帯が閉じずに、位相に乱れが生じることを事前調査により確認した。この位相の乱れを特徴量とすることによって、振幅スペクトルのみでは捉えきることのできないわずかな異常を検知することが可能であると推定される。

次に特徴量抽出の方法について述べる。1 秒程度の持続母音に対して、短時間フーリエ変換を行う。位相の細かな変化を取得するため、窓関数の幅は 2ms 程度と短く設定する。フーリエ変換後、位相スペクトルを求め、切り出したフレームごとに位相がどのくらい一致しているか計算し、特徴量とする。全てのフレームで計算後、外れ値を除外し、平均値をその音声の特徴量とする。

## 4. 評価実験

帰無仮説を”Grade 間で特徴量に差がない”、対立仮説を”Grade 間で特徴量に差がある”として、Tukey-Kramer の検定による多重比較検定を行った。使用するデータは、Grade が 1 以上と医師に評価された 63 人の音声データで、1 人につき 5 つの母音[a,i,u,e,o]の発声が記録されている。これらの音声から特徴量を抽出し、Grade ごとの特徴量の平均を計算して、母音ごとに検定を行った。また、振幅スペクトルから計算される特徴量である Harmonic-to-Noise Ratio(HNR)[3]を提案特徴量の比較対象とし、同様の検定を行った。表 1 に、それぞれの検定結果を示す。

表 1 特徴量ごとの多重比較検定結果の比較 ( $\alpha=0.05$ )

母音	HNR			提案特徴量		
	G1-G2	G1-G3	G2-G3	G1-G2	G1-G3	G2-G3
a	Accept	Reject	Reject	Reject	Reject	Reject
i	Reject	Reject	Reject	Accept	Reject	Accept
u	Reject	Reject	Reject	Reject	Reject	Reject
e	Accept	Reject	Reject	Accept	Reject	Accept
o	Accept	Reject	Reject	Reject	Reject	Reject

表 1 より、HNR は母音”a”と”o”について、G1-G2 間で母平均に有意な差が得られなかったが、提案特徴量ではこれを相補するように有意な差が得られた。この結果は、提案特徴量は従来の特徴量と組み合わせることによって識別率が向上することを示唆している。

## 5. まとめ

本稿で提案した特徴量は、従来の特徴量では Grade 間で母平均の有意な差が得られにくかった母音について有意な差を得ることができた。今後は、様々な特徴量と組み合わせることで識別器を作成し、評価を行っていく。

## 参考文献

- [1] 日本音声言語医学会, “新編 声の検査法”, 医歯薬出版株式会社, page 45, 2012.
- [2] Zhijian Wang et al., “Automatic Assessment of Pathological Voice Quality Using Multidimensional Acoustic Analysis Based on the GRBAS Scale”, Journal of Signal Processing Systems, vol. 82, no. 2, pp. 241–251, 2016.
- [3] Carole T Ferrand, “Harmonics-to-Noise Ratio: An Index of Vocal Aging”, Journal of Voice, vol. 16, no. 4, pp. 480–487, 2012.