

スペクトル包絡と原音抽出をせずに 喉締めと裏声の度合いを得る方法の分析

宮城研一[†] 鍋木崇史[†]

国際基督教大学教養学部[†]

はじめに：歌手志望の人やカラオケ好きな人など、歌唱力を上げたい人はいつの時代も少なくないだろう。しかし、歌唱力を向上させる事は簡単ではない。金銭的、精神的、または正確さなどの理由から、自身の発声状態のフィードバックを得られにくいからだ。その点正確なフィードバックをしてくれるソフトウェア・アプリがあれば、無料で、安心して、一人で、継続的に使用できると期待できる為、多くの人に役立つと考えられる。現状、音程や音域、リズム、ビブラートなどの表現を評価してくれるアプリは多数あるが、発声状態を推定してくれるものはあまり多くない。そのようなソフトウェア・アプリを開発するには音声から発声状態を推定する方法が必要である。例えば、発声時の喉の締め度合いは第2フォルマント周波数から一定の精度で推定できる事がわかっている[1]。しかし第2フォルマントを求めるには線形なソースフィルタ理論を仮定することが多く[2]、非線形な求め方も提案されてはいるものの一般的ではない[3]。その為、喉締め度を周波数スペクトルからそのまま求める手法を検証する。他にも、どの程度裏声かを推定するには声帯原音のスペクトル重心に注目すれば良い事がわかっている[4]。しかし既に述べた理由から声帯原音を抽出せずに裏声度を推定する手法を検証する。

提案手法：喉締め度合いを周波数スペクトルから推定する手法として、三つの傾斜を検証する。喉締めが起きたということは声道の形が変化したと考えられる。声道の形が変化すれば増幅されやすい周波数帯に変化があるのでと考へ、傾斜を検証することにした。発声は C4 から喚声点付近のまで G4 の音域で主に行い、出来るだけ地声でロングトーンで行うようにした。実験環境としては、iPad Pro (9.7-inch)のボイスメモで録音を行った。傾斜の導出計算方法としては、一つ目はスペクトルの 0 から 15kHz の周波数で 3 点公式で傾斜を計算し、その平均と喉締め指標

との相関を計算するものだ。二つ目はスペクトルの基本周波数付近から同じく 3 点公式で傾斜を求め、喉締め指標との相関を計算した。三つ目はスペクトルの基本周波数と整数次倍音の五つ付近のピークを求め、それらの組み合わせで 15 種類の傾斜を計算し、それぞれと喉締め指標との相関を求めた。なお水平方向の長さは基本周波数を 1 とした。これらの方法を相関で評価することにした。

検証（筋電位の扱い）：発声の際の喉締め度合いを主観的ではなく客観的に判断する必要がある。まずは舌骨筋群の働きを筋電位で観察できるかを調べる為に嚙下の際の筋電位を確認した。嚙下の際に明らかな筋電位の上昇と低下を確認できた為、筋電位で舌骨筋群の働きを観察できる事がわかった。舌骨筋群が介入すると喉締めとなる為[5]、筋電位からの喉締めの定性的評価が可能となった。しかし筋電位は皮膚表面の状態や、姿勢、電極の位置、触るものなどに多大に影響される事が確認できた為、定量的評価は難しいと言えよう。それを可能にする為に、筋電位データをどのように扱えば良いか検証した。具体的には筋電位のばらつきが周期的であると仮定し、周波数スペクトルを確認したが、主観的な喉締めの程度に対応している特徴は特に見つからなかった。筋電位の分散も対応していなかった。発声時の筋電位と安静時の筋電位の平均を求め、その差を主観的な喉締めの程度に照らし合わせた所、ある程度対応している事がわかった。その為今後これを喉締めの指標とし実験を進めた。

検証（本実験）：傾斜の平均、基本周波数付近の傾斜、倍音付近のピークの組み合わせの順に実験をした。傾斜の平均、基本周波数付近の傾斜の実験では/a, i, e/で発声したが、母音の違いや開口母音の開口に伴う舌骨筋の働きの影響を考慮して途中から/i/で発声した。各方法の相関係数を図 1 に示した。3 番目の整数次倍音と 5 番目の整数次倍音の付近のピークの組み合わせの傾斜と喉締め指標との相関が最も強かった。次に 3 番目のものと 6 番目のものの組み合わせ

Methods to infer vocal strain and falsetto without spectral envelope and extraction of glottal source

[†]Kenichi Miyagi, Takashi Kaburagi · International Christian University

のものが強かった。傾斜の平均と喉締め指標との相関は 0.25 であった。基本周波数付近の傾斜と喉締め指標との相関は最も強く 0.2 であった。

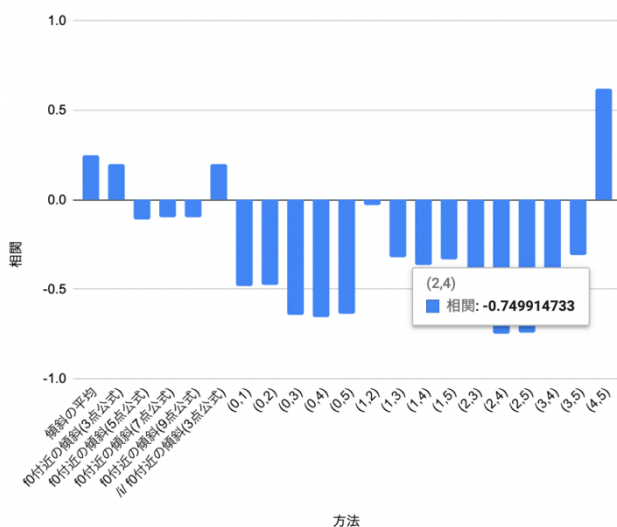


図1 各方法と喉締め指標との相関係数

考察：相関が最も強い3次倍音と5次倍音付近のピークの傾斜と次に強い3次倍音と6次倍音のものは共通して3次倍音付近のピークを持つ。その為3次倍音もしくはその付近に、喉締めによって形が変わった声道が影響を及ぼす周波数帯があると考えられる。

裏声の手法：声帯原音を抽出せずに裏声の程度を推定する方法を検証する。地声と裏声と同程度の力強さで発声できるピッチを探し、そのピッチで3秒地声発声し、すぐに3秒裏声発声する。これを逆でも行い、録音した音声のスペクトル重心を計算し、3秒地点で明らかな変化があったかを目測で判断する。明らかな違いがあった場合は裏声の程度を定性的に判断できたとする。定量的な評価は裏声の程度の正解データを用意するのが難しい為、検証は定性的評価までとした。

検証：図2と図3のように、地声発声から裏声発声と裏声発声から地声発声の両方の切り替え時点でスペクトル重心に明らかな違いが見られた。

考察：地声発声時と裏声発声時の両方でスペクトル重心の細かな変動が確認できた為、スペクトル重心そのままの値では定量的評価に用いることができないと考えられる。その為一定時間で区切り平均を計算するなどの処理が必要にな

るだろう。



図2 裏声から地声発声にした際のスペクトル重心

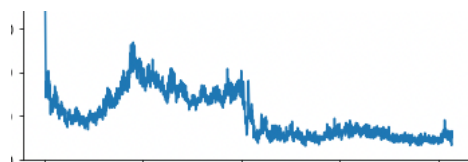


図3 地声から裏声発声にした際のスペクトル重心

あとがき：既存の方法とは別の方法で喉締め度合いと裏声度合いを推定する方法を評価した。しかし相関は最強で-0.75と正確な推定ができるとは言い難い。その為他の手法が必要になるだろう。また、「い」母音以外の母音でも実験する必要がある。「い」母音は口を開かない為、筋電センサーの電極の位置がズレにくく、計測しやすかったが、その他の母音では難しい。その為喉締めを他の方法で計測する必要がある。他にも、今回の実験者は歌唱経験があまりない男性のみだった為、女性や、子供、歌手の人などのデータがあると良いだろう。

参考文献：

- [1] 平山他 “ポピュラー歌唱における高音域の声区と発声状態の判別手法”, 情報学第74回全国大会, 2S-10, 2012
- [2] Titze, Ingo R. “Nonlinear source-filter coupling in phonation: Theory.” The Journal of the Acoustical Society of America 123.4 (2008): 1902-1915.
- [3] 今井聖, and 阿部芳春. “改良ケプストラム法によるスペクトル包絡の抽出.” 電子情報通信学会論文誌 A 62.4 (1979): 217-223.
- [4] 浅野翔大, “裏声判別指標を用いたボイスレーニングソフトウェア”, 三重大学大学院工学研究科研究集録, 2013
- [5] フースラー, フレデリック (1987) 『うたうこと 発声器官の肉体的特質-歌声のひみつを解くかぎ』 (須永義雄訳) 音楽之友社.

謝辞：本研究は JSPS 科研費 16K16392, 22K12276 の助成を受けたものです。