

歌声の習熟度と LPC スペクトル包絡線から得られる フォルマントとの関係に関する考察

A Consideration of the Relationship between Singing Proficiency and the Formant of the LPC Spectral Envelope

寺内 大晴^{*1}, 富田 孝幸^{*2}, 山下 泰樹^{*3}, 舘 伸幸^{*2}, 永井 孝^{*2,4}, 香山 瑞恵^{*2}
Taisei TERAUCHI^{*1}, Takayuki TOMIDA^{*2}, Yasuki YAMASHITA^{*3}, Nobuyuki TACHI^{*2},
Takashi NAGAI^{*2,4}, Mizue KAYAMA^{*2}

^{*1}信州大学大学院 総合理工学研究科, ^{*2}信州大学 工学部

^{*3}長野工科大学 情報エレクトロニクス学科, ^{*4}ものづくり大学 技能工芸学部

^{*1}Graduate School of Science & Technology, Shinshu University

^{*2}Faculty of Engineering, Shinshu University

^{*3}Department of Information and Electronics, Nagano Prefectural

^{*4}Department of Manufacturing Technologists, Institute of Technologists

Email: 24w6050a@shinshu-u.ac.jp

あらまし: 本研究では、音響特徴量に基づき、声楽における歌声の習熟度を測る評価指標の提案を目的としている。これまでの研究で、歌声特有の周波数応答である Singer's formant を、周波数成分の割合やフォルマントの鋭さを用いて定量的に評価してきた。本稿では、ソプラノ音域のプロ歌手と声楽学習者を対象に、LPC スペクトル包絡線から得られる Singer's formant の候補となるフォルマントの周波数と相対的な強度を比較検討し、歌声の習熟度を測る音響特徴量としての可能性を見出した成果について述べる。

キーワード: 歌声, 周波数特性, 歌声評価指標, 音響特徴量, 定量化, Singer's formant

1. はじめに

声楽の研究では、歌唱力と音響特徴量の関係が長年研究されてきた^[1]。その中で我々は、歌声の周波数特性に現れる「Singer's formant (SF) ^[2]」に注目し、それを基にした音響評価指標を提案・検証してきた^[3,4]。

本稿では、LPC (Linear Prediction Coding) スペクトル包絡線から得られるフォルマントに基づいた音響特徴量について、プロ歌手と声楽学習者の分布の比較と学習者の指導前後の分布の比較から、歌声の習熟度を表す音響特徴量として妥当か検証する。

2. 音響特徴量

本研究では、声道の共振特性を表す LPC スペクトル包絡線に基づき、音響特徴量を抽出する。なお、LPC スペクトル包絡線の導出は山下らの提案したシャープ分析における分析条件及び式を用いる^[3]。また、図 1 に歌声のスペクトル (灰実線) と LPC スペクトル包絡線 (赤実線) の例を示す。本研究では、3 章で示す解析条件において 3kHz 付近に見られる LPC スペクトル包絡線の 2 番目のフォルマント (以下, SP: Secondary Peak) に着目する。具体的には、0~2 kHz の区間に 1 番目のフォルマント (以下, PP: Primary Peak) が得られた LPC スペクトル包絡線において、PP の周波数から 6 kHz の区間にあるフォルマントを SP として抽出する。本研究では、この SP に基づく特徴量として、SP の周波数と相対強度を用いる。なお、相対強度として 2 凸^[3]を用いる。2 凸とは、SP と SP の直前の谷とのパワーレベル差である。以降は相対強度に表現を統一する。

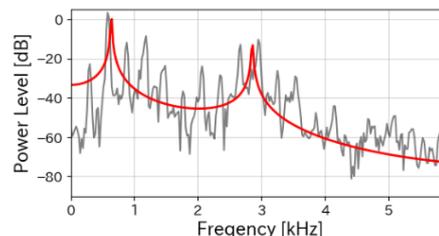


図 1 LPC スペクトル包絡線の例

3. 実験条件

本実験における被験者は、プロ歌手 (以下, pro) 5 名, 音楽大学声楽科在籍の学習者 (以下, mu) 6 名, 教育学部在籍で音楽教育を専修する学習者 (以下, ed) 5 名である。なお、全ての被験者において声種はソプラノである。学習者の各被験者については、各大学において歌唱指導を開始した直後の 2022 年 4 月~6 月に収録した歌声データ (以下, 指導前/bf) と、歌唱指導を定期的に受けた後の 2024 年 1 月~3 月に収録した歌声データ (以下, 指導後/af) を分析対象とした。

実験に使用する楽曲は、「Caro mio ben」(作詞: 不明, 作曲: Tommaso Giordani) とし、21 小節目「tanto」区間を解析対象とした。収録は音の反響しない静かな部屋で行いレコーダは LS-P2 (OLYMPUS) を使用した。被験者にはレコーダから 2 m 離れた正面立位で歌唱させた。音声はサンプリング周波数 44.1 kHz, 量子化ビット数 16 bit で収録した。

4. 解析結果

図2に“tanto”の“ta”区間、図3に“tanto”の“to”区間において、それぞれSPの周波数と相対強度のヒストグラムをもとにカーネル密度推定(scipy.stat.gaussian_kde, バンド幅:default)を行った結果を示す。橙色が pro, 水色が mu 指導前, 青色が mu 指導後, 黄緑色が ed 指導前, 緑色が ed 指導後を表す。なお, SP が検出できなかった解析フレームについては周波数, 相対強度ともに 0 として集計した。

pro - mu 指導前 - ed 指導前の3群間において, SPの周波数にはF検定による分散の比較, SPの相対強度にはMann-Whitney U検定(ともにボンフェローニ補正後0.42%)による代表値の比較を行った。検定結果を表1に示す。SPの周波数については“to”区間のmu指導前 - ed指導前を除くすべての組み合わせで有意な差が見られ, SPの相対強度については“to”区間のpro - mu指導前を除くすべての組み合わせで有意な差が見られた。

mu指導前後とed指導前後の分布について, pro分布とのWasserstein距離を導出した結果を表2に示す。muにおいては, toのSPの周波数と相対強度で指導後にWasserstein距離が減少した。edにおいては, taのSPの周波数及びtoのSPの周波数と相対強度で指導後にWasserstein距離が減少した。

5. 考察

本章では, 前章で得られた結果をもとに, 提案評価指標が歌声の習熟度を表す音響特徴量としての妥当性を検討する。

SPの周波数について, プロ歌手のSPの周波数は学習者の分布と比較して有意に散らばりが小さく, 3 kHz付近に集中していることが得られた。学習者間でも習熟度の初期値の違いによって分布の散らばりが異なり, 習熟度が低くなるほど散らばりが大きくなる傾向が得られた。

SPの相対強度について, プロ歌手のSPの相対強度は学習者の分布と比較して有意に大きいことが得られ, 20~40 dBに分布する割合が大きくなる傾向が得られた。

学習者の指導前後の比較においては, mu, edに共通して, “to”区間において特徴量の分布がproの分布に近づく傾向が得られた。“ta”区間については, edにおけるSPの周波数の分布でのみ同様の傾向が得られ, “to”区間の方が習熟度の向上による特徴量の変化が顕著であることが得られた。これらのことから, SPの周波数と相対強度が歌声の習熟度を表す音響特徴量として妥当であることが示された。

ただし, “ta”区間と“to”区間における傾向の要因については, 他の区間を用いた解析や対照実験によって議論される必要がある。

6. おわりに

本稿では, LPCスペクトル包絡線から得られるフ

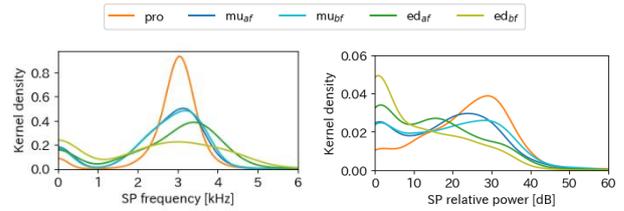


図2 “ta”区間のSPの周波数(左)と相対強度(右)の分布

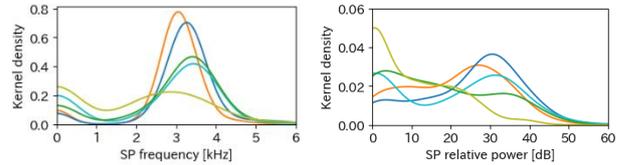


図3 “to”区間のSPの周波数(左)と相対強度(右)の分布

表1 検定から得られたp値(*: p < 0.0042)

組み合わせ	SP 周波数		SP 相対強度	
	ta	to	ta	to
pro - mu _{bf}	*	*	*	0.2488
pro - ed _{bf}	*	*	*	*
mu _{bf} - ed _{bf}	*	0.6079	*	*

表2 pro分布とのWasserstein距離

群	SP 周波数		SP 相対強度	
	ta	to	ta	to
mu _{bf}	445.75	701.14	4.65	12.34
mu _{af}	467.31	250.16	5.28	8.82
ed _{bf}	1118.57	1186.94	3.72	11.59
ed _{af}	647.03	486.65	3.89	4.72

オルマントに着目した音響特徴量について, 習熟度の違いによる音響特徴量の分布の違いを検証した。その結果, 提案特徴量が歌声の習熟度を表す音響特徴量として妥当であることを示した。

今後は, SPの周波数と相対強度を組み合わせた歌声の習熟度を評価する指標を検討し, 主観的な評価との比較分析を行う。

謝辞 本研究は科研費22K00237により支援された。

参考文献

- (1) W.T.Bartholomew, A Physical Definition of “Good Voice- Quality” in the Male Voice, J. Acoust. Soc. Am., Vol.6, No.1, pp.25-33 (1934)
- (2) J.Sundberg, Articulatory interpretation of the “singing formant”, J. Acoust. Soc. Am., Vol.55, Issue 4, pp.838-844 (1974)
- (3) 山下泰樹他, 声楽の習熟度に関連する周波数特性に基づく音響特徴量分析条件の検討, 電気学会論文誌 C, Vol.144, No.12, pp.1197-1208 (2024)
- (4) 寺内大晴他, LPC スペクトル包絡線から導出される Singer’s formant 中心周波数の声種及び母音種別分布, 情処研報音楽情報科学, Vol.2024-MUS-140, No.27, pp.1-5(2024)