

「声の見える化」による歌唱学習支援システムの開発 —録音・再生機能の概要—

Conceptual Design of Singing Training Support System using Singing Voice Visualization - Outline of Record and Replay Functions -

羽賀 翼*1, 香山 瑞恵*2, 伊東 一典*2, 池田 京子*3
橋本 昌巳*2, 大谷 真*2, 山下 泰樹*4, 浅沼 和志*5

Tasuku HAGA*1, Mizue KAYAMA*2, Kazunori ITOH*2, Kyoko IKEDA*3
Masami HASHIMOTO*2, Makoto OTANI*2, Yasuki YAMASHITA*4, Kazushi ASANUMA*5

*1 信州大学大学院理工学系研究科
*1 Division of Science & Technology Master's
Program, Shinshu University

*2 信州大学工学部
*2 Shinshu University,
Faculty of Engineering

*3 信州大学教育学部
*3 Shinshu University,
Faculty of Education

*4 長野県工科短期大学校
*4 Nagano Prefectural Institute of Technology

*5 国立長野工業高等専門学校
*5 Nagano National College of Technology

Email: 13tm536e@shinshu-u.ac.jp

あらまし: 本研究の目的は, 歌声の音響特徴量の変化の可視化する歌唱学習支援ツールを開発することである. ここでは, 音響特徴量として, 音高, 音量, 周波数スペクトルなどに着目する. 本稿では, 上記の音響特徴量の変化をリアルタイムに可視化するツールの概要を示す.

キーワード: 音響特徴量, 見える化, 歌唱学習支援, 声楽発声指導者

1. はじめに

歌唱とは, 声によって行われる音楽的な行為である⁽¹⁾. 従来より, 歌唱や歌声に関する研究がすすめられてきている⁽²⁻³⁾. これらの研究では, 歌声らしさを示す属性や, 習熟度に関連する物理量などの解明が進んでいる. また, 音高や音量といった音響特徴量と人による歌声に対する心理的印象とに関係があることが分かり始めている⁽⁴⁻⁵⁾. これらの成果は歌唱支援システムとしても具体化されてきた⁽⁶⁻⁸⁾. しかし, 歌唱指導者が容易にかつ簡便に利用できるシステムの提案には至っていない.

本研究の目的は, 歌唱指導者が求める方式での「声の見える化」である. ここで見える化の対象とするのは, 音高・音量・周波数スペクトル・ビブラートとする. これらの音響特徴量は, 山邊らの研究⁽⁹⁾で習熟度との関係が示唆されたものである. 声楽発声指導者(以下, 指導者)としては, 大学レベルの声楽科の指導者を想定する. 提案システムの利用場面としては, 対面での個人レッスンにおいて, 学習者および指導者が利用する場合と, 学習者が自主練習する場合である. それゆえ, システムのインターフェースは極力簡略化を, 容易で簡便な操作を実現することを目指す.

2. 音響特徴量

音響特徴量とは, 音声データを分析することにより, 出現する特徴量のことである. 本研究では, 音の高さを表すピッチや音の大きさを表すパワー, ピッチの周期的な変動であるビブラート, 音の響きに

関連するといわれる周波数スペクトルといった音響特徴量を取り上げる. 周波数スペクトルからは Singer's formant と倍音構造を求める.

ピッチの推定にはケプストラム法を用いる. 男声は 60Hz から 590Hz, 女声は 130Hz から 790Hz の区間⁽¹⁰⁾でピッチを推定する. パワーは, 二乗平均法で得られた数値をデシベル単位で利用する. ビブラートは主に音を伸ばすときに見られる音高の周期変動である. 5~8Hz の範囲⁽²⁾の成分の強さを調べる.

周波数スペクトルから得られる Singer's formant とは, 主に男性オペラ歌唱者の母音スペクトルにおける 2k から 4kHz に見られる顕著なピークのことである⁽¹¹⁾. このピークをスペクトル曲線の鋭さを示す数値(Q 値)として評価する. また, 倍音構造として, 当該音の倍音成分に対する Singer's formant に相当する倍音成分の比率も合わせて評価する.

3. 提案システム

3.1 概要

本システムは指導者の意見に基づき設計されている. このシステムでは, 学習者の音声をリアルタイムに分析し, 同時に歌唱音声を録音する. また, 事後に再生と詳細な分析結果を示すことができる. リアルタイム分析はピッチとパワーを対象とし, 詳細分析では上記に加え, ビブラートと周波数スペクトル(Singer's formant と倍音構造)を対象とする.

3.2 機能

本システムには, 主として録音と再生の二機能が

ある。各機能に含む利用モードの説明を以下に示す。

3.2.1 録音モード

録音モードでは歌声のピッチとパワーの分析と録音を行う。これは、歌唱中は歌唱に集中させるべきという指導者の意見により、必要最低限の音響特徴量のみを示すこととした結果である。

歌唱前に男声か女声かを指定させた上で、ピッチを計算している。この時の計算パラメータは窓長 4096 (92ms 相当)、検査回数は 40 回毎秒である。

事前に歌唱の楽譜情報を入力した場合には、ピッチマッチ方式での表示となる。これにより、音取り練習が可能となる。また、同一楽曲の歌唱データ(自己および他者の歌唱)を指定することで、例えば、指導者の歌唱と比較しながらの練習が可能となる。

また、筋電などの生体情報や歌唱姿勢などを示す動画を入力することで、歌唱情報とこれらの情報を同期させて示すことができる。

音声録音時のパラメータは、サンプリング周波数 44.1kHz、サンプリングビット数 16bit、1 チャンネル (モノラル) であり、wav 形式で保存される。

3.2.2 再生モード

再生モードでは、自身の歌声を聞きながら、その音響特徴量を詳細に分析した結果を確認できる。ここでは、直前に録音した歌声の聞き直しや過去に録音した音声の聞き直し、そして過去と現在の歌声の比較再生が可能である。再生時には、停止・一時停止・指定区間のみ再生ができる。

音響特徴量の詳細分析は、再生対象となる歌声区間に対してなされる。ピッチ推定、パワー、ビブラート検出、そして周波数スペクトルからの Singer's formant の同定と倍音構造の計算がなされる。ピッチ推定のパラメータは録音モードよりも細かく、窓長 8192 (183ms 相当)、検査回数は 100 回毎秒である。

図 1 に再生モードを利用した比較を行った例を示す。例では学習者と指導者の音声を本システムに読み込み、詳細分析の結果を出力した。なお、見やすさを考慮し、表示しているのは周波数スペクトルのみである。図は左側は学習者、右側が指導者である。学習者と比較して、指導者の 2k から 4kHz の周波数領域に強い成分が含まれているのがわかる。

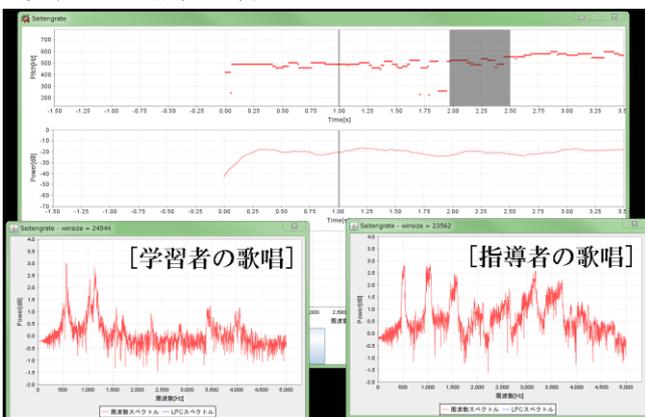


図 1 再生モードを用いた比較

また、本システムでは、同時に複数の歌声に対する詳細分析が可能である。この機能により、例えば、学習者の訓練前後での音響特徴量を歌声をそれぞれ再生しながら確認することができる。

4. おわりに

本研究では、歌声の音響特徴量を見える化し、学習者にフィードバックできるシステムを構築している。このシステムは対面レッスンおよび自主練習での利用を想定している。提案システムの特徴は、指導者の意図に基づいた機能設計にある。

しかしながら、現行の機能では音響特徴量を見える化しているのみである。さらなる歌唱学習支援を目指して、指導者による歌唱指導の様子を分析し、指導語を用いての人的なフィードバックを実現するための機能の実装を検討している。

本システムはプログラミング言語 Java を使用して開発を行っている。また、本システムのグラフの表示には API JFreeChart を利用している。

参考文献

- (1) 山川直治: “日本音楽の流れ”, pp.38-40, 日本之友社 (1990) .
- (2) 中野倫靖, 後藤真孝, 平賀謙: “楽譜情報を用いない歌唱力自動評価手法”, 情報処理学会論文誌, Vol.48, No.1, pp.227-236 (2007) .
- (3) 米沢義道, 伊東一典, 平野主蔵: “音痴の音高感覚に関する検討”, 電子情報通信学会論文誌 A, Vol.17, No.8, pp.1532-1538 (1988) .
- (4) 辻直也, 赤木正人: “歌声らしさの要因とそれに関連する音響特徴量の検討”, 日本音響学会聴覚研究会資料 H-2004-8, Vol.34, No.1, pp.41-46 (2004) .
- (5) 斎藤毅, 後藤真孝, 鶴木祐史, 赤木正人: “SingBySpeaking: 歌声知覚に重要な音響特徴を制御して話声を歌声に変換するシステム”, 情報処理学会研究報告音楽情報科学研究会, Vol.2008, No.12, pp.25-32 (2008) .
- (6) 三浦雅展, 下石坂徹, 斉木由美, 柳田益造: “和声学学習のためのバス課題実施システムのユーザインタフェースとその評価”, 電子情報通信学会技術研究報告教育工学研究会, Vol.100, No.113, pp.33-40 (2000) .
- (7) 中野倫靖, 後藤真孝, 平賀謙: “MiruSinger: 歌を「歌って/聴いて/描いて」”, 情報処理学会インタラクティブ 2007 論文集, pp.195-196 (2007) .
- (8) 中西将, 香山瑞恵, 岡部真美, 浅沼和志, 伊東一典, 為末隆弘, 橋本昌巳: “歌声の音響特徴量を指標とした指導者知識に基づく合唱学習支援”, 人工知能学会全国大会論文集(CD-ROM), 第 23 巻, 352 (2009) .
- (9) 山邊大貴, 池田京子, 香山瑞恵, 大谷真, 橋本昌巳, 伊東一典, 山下泰樹, 浅沼和志: “歌声の心理的印象と音響特徴量との対応付けによる歌唱の熟練度評価に関する基礎的検討”, 電子情報通信学会技術研究報告応用音響研究会, Vol.112, No.266, pp.61-66 (2012) .
- (10) 米沢文明: “声と日本人”, pp.54-55, 平凡社, 東京 (1990) .
- (11) J.Sundberg, "Articulatory interpretation of the 'singing formant'," J.Acoust.Am, Vol.55, No.4, pp.838-844 (1974) .