

フォルマントの変動を用いたビブラート検出法の提案

榎 真吾[†] 波多野 賢治[†][†]同志社大学文化情報学部

1 はじめに

カラオケは1970年代に登場して以降、一般人の歌唱という行為に対する意識を変化させたことや、公的空間の中にカラオケボックスという私的空間を作ったこと等、日本の社会全体に対して様々な影響をもたらしてきた。現在では、カラオケは日本の娯楽産業における代表的なヒット商品として定着しており、世界中でKARAOKEとして親しまれている。

カラオケがヒットした要因の一つに、歌唱力採点機能の存在が挙げられる。しかし、歌唱力採点機能による歌声の評価は、聴取者による歌声の評価と一致しない。その原因の一つとして、歌声の基本周波数である音声ピッチを基にしてビブラート検出を行っている為に発生するビブラートの誤検出が挙げられる。音声ピッチが正弦波状に変動していても、それが聴取者に対して不快な印象を与える場合がある。ビブラートとは、歌声に規則的な変動を付加することで聴覚印象的に心地良い響きを聴取者に与える歌唱技術であるが、そのような変動は声帯の操作によって容易に作り出すことが可能である為、ビブラートの検出に音声ピッチを用いるのは不適切である。

そこで、本稿では現状の歌唱力採点機能の問題点の一つであるビブラートの誤検出を解消することを目的とし、音声の歌声らしさに関連したフォルマントの変動を用いたビブラート検出法を提案する。

2 関連研究

竹内らは、音声ピッチの周期変化量に対して、自己相関法に基づいた周期性の分析によるビブラート検出手法を提案した[1]。この研究では、自己相関関数の調波構造の特性を利用することで、高精度な音声ピッチの抽出を可能にした。その結果、ビブラートの検出精度を向上させている。

また、齋藤らは歌声らしい音声とは何であるかを検討する為に、歌声特有の音響特徴量と歌声知覚の関係性を考慮した歌声らしさの知覚モデルを提案した[2]。この研究では、歌声らしさと基本的な心理的特徴、基本的な心理的特徴と音響特徴量の関係性についてそれぞれ

調査を行い、歌声らしさの聴覚印象には、ビブラートとそれに同期したフォルマントの振幅変調成分が大きく寄与していることを明らかにした。

3 提案手法

本稿では、齋藤らによる歌声らしさの定義を基に、歌声特有の聴覚印象に大きく寄与するフォルマントの時間変動を用いたビブラート検出手法を提案する。

フォルマントは、声道の周波数成分であるスペクトル包絡のピークに現れる。スペクトル包絡の抽出には線形予測法を用いる[3]。プリエンファシス、ハニング窓処理を行った音声データに対して、レビンソン・ダービンのアルゴリズムを適用させることでLPC係数を算出し、スペクトル包絡を抽出する。そして、スペクトル包絡のピークを抽出することでフォルマントの周波数を求める。本稿では、フォルマントの中でも特に周波数の変動の大きい第1、第2、第3フォルマントの時間変動をビブラート判定に用いる。図1は、スペクトル包絡(縦線)とその頂点であるフォルマント(横線)の時間変動を示している。

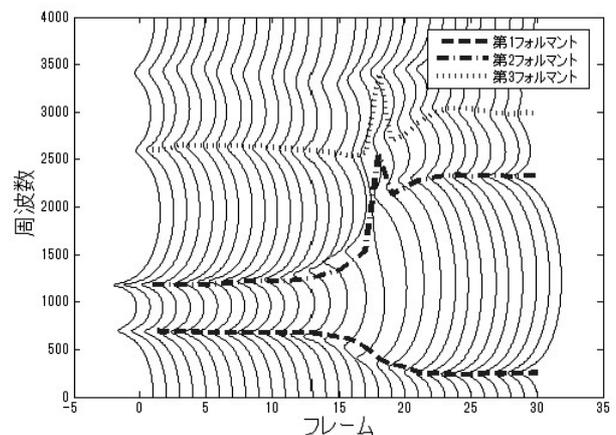


図1: スペクトル包絡とフォルマントの時間変動

次に、抽出した各フォルマントの周波数をそれぞれ時系列毎に並べ、式(1)より周波数変動率 $c_n(i)$ を算出する。ここで、 n はフォルマント番号 ($n = 1, 2, 3$)、 N は切り出す区間の長さ、 t はフレーム、 i はフレームのずらし量である。

A Vibrato Detecting Method using Fluctuation of Formant

[†] Shingo ENOKI (bii0022@mail4.doshisha.ac.jp)[†] Kenji HATANO (khatano@mail.doshisha.ac.jp)

Faculty of Culture and Information Science, Doshisha University (†)

$$c_n(i) = \sum_{t=1}^{N-i-1} \frac{f_n(t+i)}{f_n(t)} \quad (0 \leq i \leq N-1) \quad (1)$$

ここで、理想的なビブラート周期 ($17 \leq i \leq 33$) における周波数変動率が最大値をとる地点におけるフレームのずらし量を p とし、周波数変動率の極値がずらし量 $\frac{p}{2}$ 毎に逆転し、且つ変動率の差が1%以上であることをビブラートの条件として設定する。これにより、各フォルマントの周波数変動率がこの条件を満たした区間に対し、ビブラートを含む区間として検出する。

図2は、ビブラートを含む区間における周波数変動率である。等間隔のずらし量で極値が逆転し、且つ変動も大きいことが分かる。

一方、図3はビブラートを含まない区間における周波数変動率である。変動率の大小関係がばらばらで、且つ変動も小さいことが分かる。

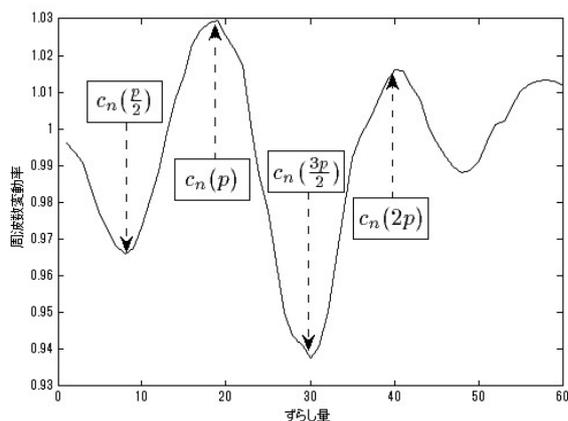


図2: ビブラートを含む区間における周波数変動率

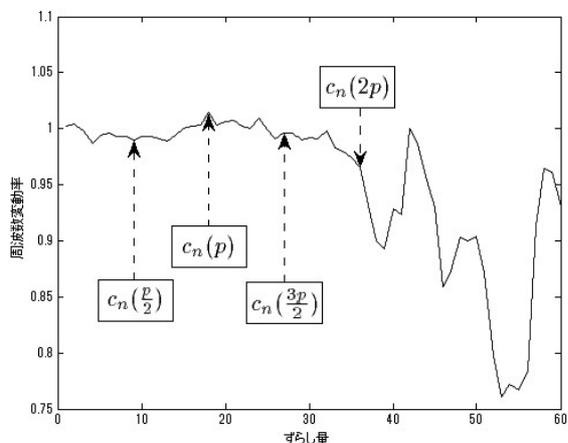


図3: ビブラートを含まない区間における周波数変動率

4 評価実験と考察

本手法の評価実験用音声データとして、歌唱者4名(男性2名, 女性2名)の、音声ピッチの変動が発生するあらゆる状況を想定した何種類かの歌唱音声データを使用した。これらの音声データに対して、ボイストレーナーにビブラートを含むか否かを判定してもらい、音声ピッチの変動を用いた従来手法によるビブラート検出と、フォルマントの変動を用いた提案手法によるビブラート検出をそれぞれ行った。

ボイストレーナーの判定によるビブラートを正解¹として、情報検索の分野で評価の際によく利用される再現率、適合率、F値の算出を行ったところ、表1のようになった。

表1: F値の算出結果

	再現率	適合率	F値
従来手法	0.875	0.824	0.848
提案手法	0.889	0.941	0.911

よって、従来手法よりも提案手法の方がF値が大きく、提案手法の有用性が確認された。

特に、提案手法は頭を振りながらの歌唱、ジャンプしながらの歌唱、腹筋を振動させながらの歌唱に対してビブラートの誤検出を抑制することができ、提案手法の有用性を確認できた。

5 おわりに

本稿では、ビブラートの誤検出を解消する為に、フォルマントの時間変動を用いたビブラート検出法の提案を行った。その結果、ビブラートの誤検出を抑制させることが可能となった。

実際のカラオケではリアルタイムでビブラートの検出を行う為、今後の方針として、提案手法のカラオケでの実用性について検討する必要がある。

参考文献

- [1] 竹内英世, 保黒政大, 梅崎太造. 人の主観評価に近いカラオケ採点法. 電気学会論文誌, pp.1042-1053, 2010.
- [2] 齋藤毅, 辻直也, 鷗木祐史, 赤木正人. 歌声らしさの知覚モデルに基づいた歌声特有の音響特徴量の分析. 日本音響学会誌 64(5), pp.267-277, 2008.
- [3] 青木直史. デジタル・サウンド処理入門. CQ出版社, 2006.

¹ ボイストレーナーは通常、ビブラートによる歌唱法は容認していない為、ビブラートはノイズであると判断される。その為、正解のビブラートとはボイストレーナーがノイズと判断した音声データを指す。