

同一楽曲に対する多数の歌唱の 基本周波数推定値分布の可視化

伊藤 貴之^{1,a)} 中野 倫靖^{2,b)} 深山 覚^{2,c)} 濱崎 雅弘^{2,d)} 後藤 真孝^{2,e)}

概要: 2次創作やソーシャルメディア環境の普及にともない、同一楽曲に対する多数の歌唱を人々が楽しめるようになった。そうした歌唱群の音響データを分析して比較することで、同一楽曲に対する各歌唱者の癖や個性の違いを理解することが可能になる。本報告では、そうした理解の一助として、同一楽曲に対する多数の歌唱群の音響データからそれぞれの音高（基本周波数：F0）の推移を抽出し、その分布を可視化する手法を提案する。本手法では、時刻および周波数の対数値を2軸とする2次元ヒストグラム画像を生成し、これに適応2値化・ラベリングといった画像処理手法を適用することで、音高の特徴的な分布を強調表示する。本報告では2024人の歌唱者による同一楽曲の音高分布を可視化した例を示し、本手法の実用の可能性について議論する。

1. はじめに

近年の2次創作の流行に伴い、同一楽曲に対して、様々な歌唱者の歌声を誰でも容易に楽しめるようになった。また近年の歌声合成技術の発達に伴い、同一楽曲に対して異なる制作者が異なる声色で歌声合成する事例も出てきている。さらに、これらを組み合わせて斉唱・輪唱・合唱などの音源を制作するようなN次創作も見られるようになった。

このような多数の歌唱者・制作者による同一楽曲の歌唱データを分析・比較することは、学術的な観点からも興味深い。例えば、多数の歌唱者による同一楽曲の歌唱を比較することで、どのような癖や個性を有する歌唱者が多い傾向にあるか、楽譜上の音高からどのようにずれた表現をする歌唱者が多い傾向にあるかを分析できる。また、プロの歌唱者と他の歌唱者の歌唱を比較することで、どのような表現はプロの歌唱者の模倣が容易か困難か、といった議論も可能になる。あるいは歌声合成による歌唱データの場合には、どのような表現手段を用いているデータが多いか、あるいはどのような表現手段を用いている歌唱が注目される傾向にあるか、といったパターンを分析できる。

歌唱の個性や違いを測る基本的な指標の一つに音高（基本周波数：以下F0と称する）があげられる。歌唱を録音した音響データの各時刻における音高を推定し、楽譜から得られる音高と比較することで、音高をあえて外す意図的な歌唱、あるいは音高が大きく逸脱している歌唱などを発見できる [1]。あるいはビブラート等の歌唱テクニックをF0の推移から検出したり [2,3]、オーバーシュートなどの動的変動 [4] に着目したりして、歌唱の個性を分析できる可能性がある。

音高の推移は一種の時系列データとして捉えることができる。時系列データの分類や特徴検出には従来から多数の研究がなされており [5,6]、それらを適用することで歌唱の違いや逸脱を検出することは可能である。一方で歌唱の分析には主観的な解釈を要する場面もあり得る。例えば同一楽曲の特定の瞬間の音高に個人差が見られた際に、意図的な歌唱技法として音高をずらしているのか、技量不足や練習不足により意図せずに音高がずれているのか、といった点を解釈するには分析者の理解や判断が必要な場合がある。そこで分析者が主観的・定性的に歌唱データを観察するツールとして、本研究では可視化手法の適用を検討する。可視化は大規模なデータの理解の目的で多用されており、本研究の問題設定である「多数の歌唱データの理解」にも合致している。また鑑賞するのに一定の時間を要する音楽のようなデータをより短時間で理解できるという意味において、可視化は歌声の音高を分析する際に有用な手段であるとも考えられる。時系列データに対する汎用的な可視化手法は従来から多数研究されており、これを歌唱データの

¹ お茶の水女子大学
東京都文京区大塚 2-1-1

² 産業技術総合研究所
茨城県つくば市梅園 1-1-1

a) itot@is.ocha.ac.jp

b) t.nakano@aist.go.jp

c) s.fukayama@aist.go.jp

d) masahiro.hamasaki@aist.go.jp

e) m.goto@aist.go.jp

分析に適用することも可能であると考えられる。

本報告では、多数の歌唱者が歌唱した同一楽曲の音高の推移分布を可視化する一手法を提案する。本手法では一定時刻ごとに推定した基本周波数を時系列情報とみなし、横軸を時刻・縦軸を周波数の対数値とする格子の上にプロットする。格子上の各長方形領域におけるプロット回数を集計することで、2次元ヒストグラムを構成する。この2次元ヒストグラムをグレースケールの画像として表示し、一定以上の明度を有する連続領域をラベリングすることで、多くの歌唱者が同様に用いる音高の推移を可視化する。

本報告では2024人の歌唱者による同一楽曲の音高分布を可視化した例を示し、今後の展望や実用の可能性について議論する。

2. 関連研究

2.1 多数の歌唱データの活用

同一楽曲に対する多数の歌唱データがあれば、歌唱技法の傾向を分析することが可能になる。この点に着目した研究の例としてWilkinsら[7]は、20人のプロ歌唱者による10時間以上の歌唱録音データベースを構築し、ビブラートやトリルといった歌唱技法を分析した結果を示している。

また、同一楽曲に対する多数の歌唱を活用した2次創作によって、合唱や輪唱などの新しい作品を制作することができる。歌唱作品の制作環境の一例として都築ら[8,9]は、そうした同一楽曲に対する複数歌唱を組み合わせて合唱作品を制作する過程を支援するツールを提案している。

2.2 音高推移の可視化

歌唱データおよびそれに限定しない演奏データにおける音高推移の分析や観察に可視化を用いた研究事例は、既にいくつか報告されている[10]。例として、Nakanoら[2]のMiruSinger, Shiraishiら[11]のHAMOKARA, Moschosら[12]のFONASKEIN, Mayorら[13]の歌唱採点手法では、歌唱の練習成果を正解楽譜(ピアノロールなど)と比較可視化する機能を搭載している。複数の歌唱を対象とした例として、Nakanoら[14]のVocaRefinerは複数の歌唱録音の編集により楽曲を制作するための対話的環境を構築しており、この中で音高の可視化も採用している。

また演奏情報の中から基本周波数およびその時間推移の適切な同定を支援するための可視化[15,16]や、周波数情報から推察される調性の可視化[17]などの事例がある。また、歌唱の音高推移から歌唱スタイルを理解するために可視化を用いる手法として、音高とダイナミクスを2軸とした可視化[18]や、音高と音高差分を2軸とした可視化[19]が報告されている。Wilkins[7]らによる歌唱技法の同定結果はスペクトログラムとして可視化されている。

しかし我々が調査する限り、数百人・数千人単位の多数の歌唱データを一画面に可視化する研究事例は見当たらない。

ない。

2.3 時系列データの可視化

歌唱の音高の推移は時系列データとして扱うことが可能であり、汎用的な時系列データ可視化手法を適用することが可能である。

ここで n 個の標本がそれぞれ m 個の時刻における実数値を有する時系列データがあるとすると、このようなデータに関する多くの可視化手法は以下のいずれかのアプローチを有する。なお、以下での「実数値」はF0値に、「密度」は近い音高を有する歌唱者の人数に対応する。

- (1) 一方の座標軸に m 個の時刻、他方の座標軸に実数値を割り当てた折れ線グラフ[20,21]や散布図。
- (2) (1)の折れ線や点群を密度関数に置き換えて、密度を各画素の明度や色相に変換したヒートマップで表現したもの[22]。
- (3) 一方の座標軸に m 個の時刻、他方の座標軸に n 個の標本を割り当てたマトリクスに対して、実数値を各画素の明度や色相に変換したヒートマップで表現したもの[23,24]。

これらのアプローチの各々にはいくつかの問題点がある。(1)に示した折れ線グラフや散布図には、画面上の描画物の過密状態が引き起こすVisual Clutteringと呼ばれる視認性の低下が避けられない。また、可視化結果からのデータ読み取りにおいて色の識別能力は高くない[25]ことが知られており、その点で(3)に示したヒートマップでは実数値を正確に読み取れない可能性がある。以上により本研究では(2)に示す「密度関数のヒートマップ」というアプローチをとることにする。

本研究の目的の一つとして「歌唱における音高推移のパターンを発見する」という点がある。時系列データの可視化においても、クラスタリングや部分頻出パターン検出などの汎用的な手法を用いている事例がいくつかある[26]。ここで音高推移の部分パターンには「同一の瞬間にも異なる時間長のパターンが同時に出現する」という特徴があり、この点に着目した時系列データ可視化手法はまだ多くない。一方で、任意のタイミングで併合や分離を繰り返し、異なる長さのクラスタが同時に生じるような時系列データにおいては、Sankey GraphやStoryline[27]の適用が有効な場合が多い。しかしこれらの表現では、クラスタ間の標本の入れ替えが頻繁に発生するようなデータにおいて表示内容が複雑に絡まるため、視認性の低下が避けられない。一方で本研究が採用する「密度関数のヒートマップ」は画像の一種である。このような画像に領域分割手法を適用することで、密度関数のヒートマップから密度の濃い領域を抽出し、これをクラスタとみなすことが可能である。本研究ではこのアプローチによって、歌唱における音高推移のパターンを可視化する。

3. 基本周波数分布の可視化

本章では基本周波数 (F0) の可視化のための画像処理的なアプローチについて、その処理手順に沿って論じる。

3.1 F0 値の同定

本章では歌唱者集合 S を構成する各歌唱者の音高の推移を以下のように表記する。

$$S = \{s_1, s_2, \dots, s_n\}$$
$$s_i = \{p_{i1}, p_{i2}, \dots, p_{im}\} \quad (1)$$

ここで s_i は i 番目の歌唱者による歌唱の音高系列、 n は歌唱者の総数、 p_{ij} は i 番目の歌唱者の j 番目の時刻における F0 値の対数、 m は基本周波数推定の対象区間における標本化された時刻の総数 (各音高系列の F0 値の個数) である。なお休符に相当する無音部分には、便宜上、F0 値の対数にゼロを代入した。

なお本研究では、全ての歌唱が同じ長さ・同じタイミングで収録された上で、同一の時刻における F0 を推定することを前提としている。この前提が成立しないデータが与えられた際には、DP マッチング等による歌唱間の F0 系列のアライメントが前処理として必要である。

3.2 グレースケール画像の生成

本手法では、時刻を横軸、周波数の対数を縦軸とした長方形領域を設定し、これを格子状に分割する。歌唱の開始時刻および終了時刻をそれぞれ $t_{\text{start}}, t_{\text{end}}$ として長方形領域 R の左右端にわりあて、この区間を N 個に分割する。また可視化の対象となる周波数領域の上限と下限を設定し、各々の対数をそれぞれ $p_{\text{max}}, p_{\text{min}}$ として R の上下端にわりあて、これを M 個に分割する。なお、以下の記述では $t_{\text{start}} = t_1, t_{\text{end}} = t_{N+1}, p_{\text{min}} = f_1, p_{\text{max}} = f_{M+1}$ とする。

続いて本手法では、式 1 に示す p_{ij} の各々が上述の格子構造のいずれの長方形領域に該当するかを算出する。具体的には、左から u 番目、下から v 番目の長方形領域について、

$$t_u < i < t_{u+1}$$
$$f_v < p_{ij} < f_{v+1} \quad (2)$$

が成立するようであれば、 p_{ij} は当該長方形領域に属するとして、変数 r_{uv} に 1 を加算する。

以上の処理による集計結果は 2 次元ヒストグラムを構成するが、本手法ではこれを横 N 画素、縦 M 画素の画像として扱う。長方形領域に包括される p_{ij} の個数を集計した変数 r_{uv} から、以下の式

$$I_{uv} = 1.0 - (\alpha r_{uv})^\gamma \quad (3)$$

によって、左から u 画素目、下から v 画素目の明度 I_{uv} を

求める。ただし α および γ はユーザが調節可能な変数とする。

3.3 ラベリングによる頻出周波数推移領域の特定

頻出する周波数推移を見つけやすくするための一手段として、本手法では上述の画像を閾値 β によって白黒 2 値化し、さらにラベリング処理を適用する。まず以下の式

$$B_{uv} = 1(I_{uv} > \beta)$$
$$B_{uv} = 0(I_{uv} \leq \beta) \quad (4)$$

によって、左から u 画素目、下から v 画素目の画素値 I_{uv} を 1 または 0 のいずれかを有する画素値 B_{uv} に変換する。続いて $B_{uv} = 1$ である画素を 1 個抽出し、隣接画素で $B_{uv} = 1$ であるものを再帰的に探索する。そして、探索が終了するまでに訪問した画素の集合に固有のラベルを割り当てる。この処理を $B_{uv} = 1$ である全ての画素に割り当てることで、一定以上の頻度で現れる周波数推移領域を抽出する。なおラベリング結果は画素の処理順に依存しない。

4. 実行例

本手法を実装した結果を紹介する。著者らは本可視化手法を Java 1.10.0 および JOGL (Java OpenGL) 2.3.2 で実装した。実行例には DAMP-balanced dataset *1 に収録された”Let It Go”の 2024 人の歌唱を用いた。DAMP-balanced dataset には F0 値の推定結果を記述したデータファイルも収録されているが、本報告では我々自身で音響データから STRAIGHT [28] を用いて推定した F0 を入力とした。可視化結果の画素数は $N = 1000, M = 500$ とした。

図 1 は周波数推移分布をグレースケール画像として表示した例である。画像中の上部に黒に近い部位が左右に分布しており、多くの歌唱者が同様な音高をとっていることがわかる。さらに、これと同様な動きが画像中の下部にもうっすら見える。このような部位が見られる理由として、少数の歌唱者が音高を 1 オクターブ低く歌唱していた可能性が考えられる。

図 1 はグレースケール画像を 2 値化してラベリングを適用した例である。この結果では、多くの歌唱者が有する同様な周波数推移に固有の色が割り当てられて表示されている。なお、この図ではマウスオーバーした箇所に該当する 1 人の歌唱者の周波数推移を折れ線グラフとして同時に表示している。この図の楕円 (1) に着目すると、多くの人が同様な歌唱をしていることが紺のラベルで表示されているのに対して、それよりも高い周波数で赤のラベルが存在することがわかる。また楕円 (2) に着目すると、多くの人が同様な歌唱をしていることが緑のラベルで表示されているのに対して、それよりも低い周波数で水色のラベルが存在

*1 <https://ccrma.stanford.edu/damp/>

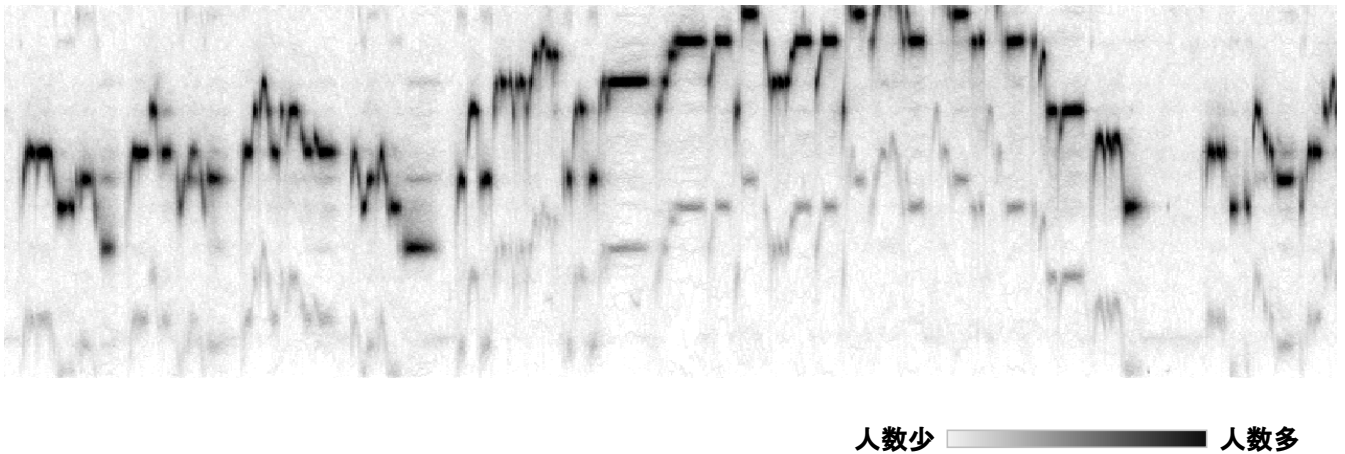


図 1 音高推移分布をグレースケール画像として表示した例. 黒に近いほど多くの歌唱者が同じ音高推移をとっていることを示している.

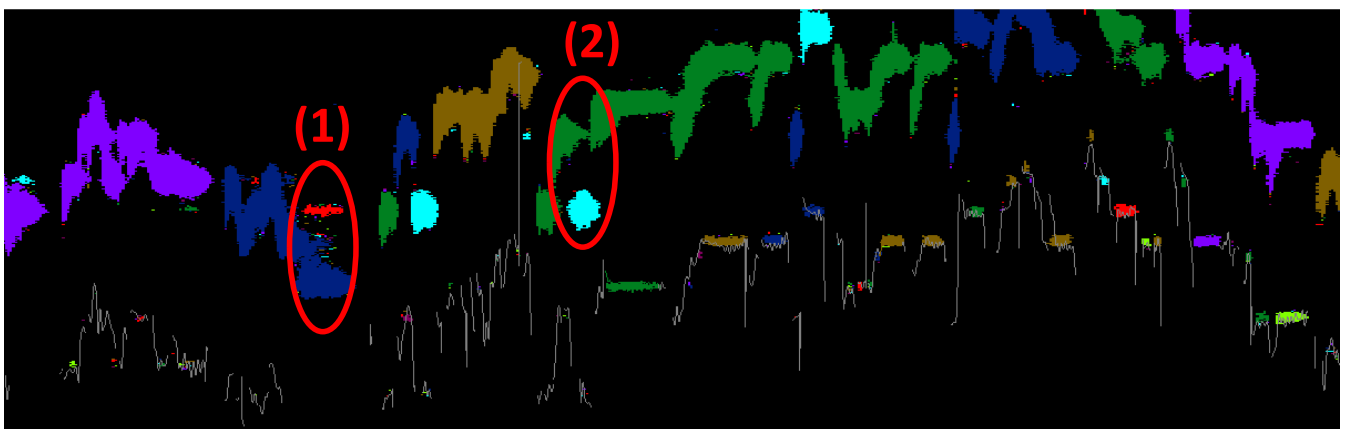


図 2 グレースケール画像を 2 値化してラベリングした例. ラベルごとに異なる色が割り当てられている. ただし現時点での実装では所定の 10 色のいずれかがラベルに付与されているため, 複数のラベルに同一色が割り当てられている箇所がある.

することがわかる. この可視化結果は, 一定数の歌唱者に同様な癖があることが示唆される.

2.3 節で論じた通り, 時系列データ可視化の主要な手法には「折れ線グラフ」「密度分布のヒートマップ」「実数値のヒートマップ」があり, 本研究では「密度分布のヒートマップ」を採用している. ここで, 本章で示したデータを「折れ線グラフ」および「実数値のヒートマップ」で可視化した例を図 3 および図 4 に示す.

折れ線グラフでは本手法と同様に, 横軸に時刻, 縦軸に F0 値を割り当てている. 画面上の折れ線の過密状態によって Visual Cluttering が生じていることがわかる. また描画処理時間が画面解像度よりも折れ線の数に依存するという問題もある.

実数値のヒートマップでは, 横軸に時刻を割りあて, 縦軸に沿って歌唱者を並べ, 画素値の色相で F0 値を表現している. この表現では音高推移の個人差を読み取ることが困難であることがわかる.



図 3 図 1,2 と同一のデータを折れ線グラフで表示した例. 画面上の折れ線の過密状態による Visual Cluttering を回避することが難しい.

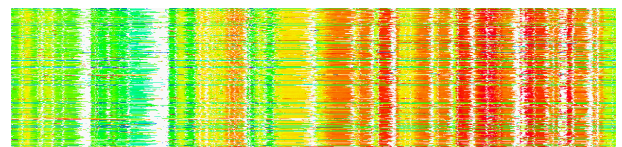


図 4 図 1,2 と同一のデータを実数値のヒートマップで表示した例. 音高推移の個人差を読み取ることが困難である.

5. 今後の展望

本研究に対する今後の展望として, 以下のようなことを

検討している。

- 正解データとの比較可視化

歌唱の音高推移を可視化する際に、楽譜から得られる音高推移、あるいは、プロの歌唱者の録音物から得られる音高推移を正解データとして同時に表示しても有用である。現時点での本手法では、特定の音高推移を正解データとして強調表示するような機能はまだ実装されていない。これを実装することで、歌唱者群がどの程度正確な歌唱を実現できているか、あるいは逆に楽譜上の音高との差分を有する音高を意図的に歌唱する唱法がどこに用いられているか、またプロの歌唱者とどの程度近い音高推移を有する人が多いか、といった点を観察できるようになる。

- 付随情報の同時表示

歌唱者の属性と音高推移の相関に興味深い傾向がある場合も考えられる。例えば歌唱者の年齢や歌唱経験年数、あるいはソーシャルメディア等で公開されている歌唱であればその評価や再生数、といった属性を各歌唱データに付与させることが可能である。これらの属性と音高推移との相関が発見できれば興味深いであろう。あるいは、歌唱の各時刻における歌唱技法を数値化することも可能である。音高推移から評価と数値化がある程度可能な歌唱技法の例として、ビブラート [3] やオーバーシュートがあげられる。これらの歌唱技法と音高推移の相関を可視化することも考えられる。これらを総合すると、周波数以外にも変数を有する時系列データとして歌唱群を可視化する、ということが今後の課題の一つとして考えられる。

- 可視化手法の改善

図1からもわかるように、密度関数のヒートマップを適用した可視化では、個々の標本（1人の歌唱の音高推移）を鮮明に視認することが難しい。この解決方法として、密度関数のヒートマップをラベリング処理だけのために用いて、個々の歌唱の音高推移を折れ線グラフや散布図で表示する、ということが考えられる。一方で2章で議論した通り、折れ線グラフや散布図には Visual Cluttering という問題が避けられない。これを解決する手段として、ネットワークや多次元データの可視化において既に多用されている Bundling(束化) [29] という手法を用いることが考えられる。さらに別の問題として、現状の実装では、頻出パターンから外れた例外的な音高推移を発見するのが難しいという点もあげられる。この点についても、ラベリング処理から外れた音高推移を折れ線グラフや散布図で表示するという形での解決が考えられる。

- 可視化結果からの集合知の抽出

多数の楽曲の歌唱データに対して本手法を適用することで、歌唱群の音高推移に対する多数の可視化結果を

集めることができる。昨今の VOCALOID 等の歌声合成技術の流行により、同一楽曲に対して異なる制作者が歌声合成技術を用いた歌唱作品を制作して動画共有サービス等で公開する機会が生じており、それらを可視化して比較することは興味深い。そのようにして得られた複数の可視化結果に対して機械学習を適用することで、「このような音高推移が見られる多くの楽曲に対して、このような音高推移での歌唱がよくみられる」といった一般性の高い歌唱分析が可能になるであろうと期待する。

- 制作支援ツールとしてのケーススタディ

歌唱者が自分の歌唱技術や表現力を向上させるための支援ツールとして、あるいは歌声合成技術を用いる制作者が他者の制作技術を参考にするための支援ツールとして、本手法がどのように貢献できるかを実証したい。

6. まとめ

本報告では同一楽曲に対する多数の歌唱データに対して推定した基本周波数の推移を時系列データとみなし、画像処理的なアプローチによって可視化する手法を提案し、その実行例を示した。今後は前章で示した展望に沿って、さらに研究開発を進めたい。

参考文献

- [1] Wager, S., Tzanetakis, G., Sullivan, S., C. Wang, J. S., Kiml, M. and Cook, P.: Intonation: A Dataset of Quality Vocal Performances Refined by Spectral Clustering on Pitch Congruence, *Proc. IEEE ICASSP 2019*, pp. 476–480 (2019).
- [2] 中野倫靖, 後藤真孝, 平賀 譲: 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] Driedger, J., Balke, S., Ewert, S. and Müller, M.: Template-Based Vibrato Analysis in Complex Music Signals, *Proc. ISMIR 2016*, pp. 239–245 (2016).
- [4] 後藤真孝, 齋藤 毅, 中野倫靖, 藤原弘将: 歌声情報処理の最近の研究, *日本音響学会誌*, Vol. 64, No. 10, pp. 616–623 (2008).
- [5] Last, M., Kandel, A. and Bunke, H.: *Data Mining in Time Series Databases*, World Science Publishing (2004).
- [6] Liao, T. W.: Clustering of Time Series Data – A Survey, *Pattern Recognition*, Vol. 38, pp. 1857–1874 (2005).
- [7] Wilkins, J., Seetharaman, P., Wahl, A. and Pardo, B.: VocalSet: A Singing Voice Dataset, *Proc. ISMIR 2018*, pp. 468–474 (2018).
- [8] Tsuzuki, K., Nakano, T., Goto, M., Yamada, T. and Makino, S.: Unisoner: An Interactive Interface for Derivative Chorus Creation from Various Singing Voices on the Web, *Proc. Joint ICMC SMC 2014 Conference*, pp. 790–797 (2014).
- [9] 都築圭太, 中野倫靖, 後藤真孝, 山田武志, 牧野昭二: Unisoner: 様々な歌手が同一楽曲を歌った Web 上の多様な歌声を活用する合唱制作支援インタフェース, *情報処理学会論文誌*, Vol. 56, No. 12, pp. 2370–2383 (2015).

- [10] Hoppe, D., Sadakata, M. and Desain, P.: Development of Real-time Visual Feedback Assistance in Singing Training: A Review, *Journal of computer assisted learning*, Vol. 22, No. 12, pp. 308–316 (2006).
- [11] Shiraiishi, M., Ogasawara, K. and Kitahara, T.: HAMOKARA: A System for Practice of Backing Vocals for Karaoke, *Proc. SMC 2018*, pp. 511–518 (2018).
- [12] Moschos, F., Georgaki, A. and Kouroupetroglou, G.: FONASKEIN: An Interactive Software Application for the Practice of the Singing Voice, *Proc. SMC 2016*, pp. 326–331 (2016).
- [13] Mayor, O., Bonada, J. and Loscos, A.: Performance Analysis and Scoring of the Singing Voice, *Proc. AES 35th International Conference* (2009).
- [14] Nakano, T. and Goto, M.: VocaRefiner: An Interactive Singing Recording System with Integration of Multiple Singing Recordings, *Proc. SMC 2013*, pp. 115–122 (2013).
- [15] Klapuri, A.: A Method for Visualizing the Pitch Content of Polyphonic Music Signals, *Proc. ISMIR 2009*, pp. 615–620 (2009).
- [16] Jure, L., Lopez, E., Rocamora, M., Cancela, P., Sponton, H. and Irigaray, I.: Pitch Content Visualization Tools for Music Performance Analysis, *Proc. ISMIR 2012*, pp. 493–498 (2012).
- [17] Gomez, E. and Bonada, J.: Tonality Visualization of Polyphonic Audio, *Proc. ICMC 2005* (2005).
- [18] Lin, K. W. E., Anderson, H., Agus, N., So, C. and Lui, S.: Visualising Singing Style Under Common Musical Events Using Pitch-Dynamics Trajectories and Modified TRACCLUS Clustering, *Proc. ICMLA '14* (2014).
- [19] Kako, T., Ohishi, Y., Kameoka, H., Kashino, K. and Takeda, K.: Automatic Identification for Singing Style Based on Sung Melodic Contour Characterized in Phase Plane, *Proc. ISMIR 2009* (2009).
- [20] Uchida, Y. and Itoh, T.: A Visualization and Level-of-Detail Control Technique for Large Scale Time Series Data, *Proc. IV09*, pp. 80–85 (2009).
- [21] Perin, C., Vernier, F. and Fekete, J.-D.: Interactive Horizon Graphs: Improving the Compact Visualization of Multiple Time Series, *Proc. ACM CHI 2013*, pp. 3217–3226 (2013).
- [22] Wang, Y., Han, F., Zhu, L., Deussen, O. and Chen, B.: Line Graph or Scatter Plot? Automatic Selection of Methods for Visualizing Trends in Time Series, *IEEE Trans. on Visualization and Computer Graphics*, Vol. 24, No. 2, pp. 1141–1154 (2018).
- [23] Imoto, M. and Itoh, T.: A 3D Visualization Technique for Large Scale Time-Varying Data, *Proc. IV10*, pp. 17–22 (2010).
- [24] Oliveira, G., Comba, J., Torchelsen, R., Padilha, M. and Silva, C.: Visualizing Running Races through the Multivariate Time-Series of Multiple Runners, *Proc. SIB-GRAPI 2018*, pp. 99–106 (2013).
- [25] Mazza, R.: *Introduction to Information Visualization*, Springer (2009).
- [26] van Wijk, J. J. and van Selow, E. W.: Cluster and Calendar based Visualization of Time Series Data, *Proc. InfoVis'99*, pp. 4–9 (1999).
- [27] Tanahashi, Y., Hsueh, C.-H. and Ma, K.-L.: An Efficient Framework for Generating Storyline Visualizations from Streaming Data, *IEEE Trans. on Visualization and Computer Graphics*, Vol. 21, No. 6, pp. 730–742 (2015).
- [28] Kawahara, H., Masuda-Katsuse, I. and de Cheveigné, A.: Restructuring Speech Representations Using a Pitch Adaptive Time-frequency Smoothing and an Instantaneous Frequency Based on F0 Extraction: Possible Role of a Repetitive Structure in Sounds, *Speech Communication*, Vol. 27, pp. 187–207 (1999).
- [29] Holten, D.: Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data, *IEEE Trans. on Visualization and Computer Graphics*, Vol. 12, No. 5, pp. 741–748 (2006).