

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 6 日現在

機関番号：13301

研究種目：若手研究(B)

研究期間：2011～2013

課題番号：23700197

研究課題名(和文) 歌声の生理・音響的特徴の操作に基づく高度な歌声合成

研究課題名(英文) Singing voice synthesis based on control of acoustical and physiological characteristics

研究代表者

齋藤 毅 (Saitou, Takeshi)

金沢大学・電子情報学系・助教

研究者番号：70446962

交付決定額(研究期間全体)：(直接経費) 2,400,000円、(間接経費) 720,000円

研究成果の概要(和文)：本研究では、多様で表情豊かな歌声を合成可能にする歌声合成技術の構築を目指し、以下に示す4つの成果を達成した。(1)オペラ・ポピュラー歌唱者25名による歌声500トラックから構成される歌声データベースの構築。(2)歌声の裏声・地声(声区)の生理的・音響的特徴の操作に基づく声区識別・変換手法の構築(3)オペラ歌唱固有の音響特徴の制御に基づくオペラ歌唱合成手法の構築。(4)歌手の個性を規定する音響特徴量の検討。

これらの成果は、従来の歌声合成で実現できなかった豊かな声質、高度な歌唱技量に基づく歌唱スタイルを表現可能にするものである。

研究成果の概要(英文)：The aim of this study was to develop a singing voice synthesis method that can synthesize artistic and expressive singing voice. Research achievements of the study were followings. (1) A large database which contains 500 singing tracks sung by 25 operatic or pop singers was constructed. (2) A vocal register conversion method between a modal and falsetto registers based on modification of acoustical and physiological characteristics was developed. (3) A operatic singing voice synthesis method based on controlling acoustical characteristics unique to operatic singings was proposed. (4) Acoustic characteristics that define singer individuality (vocal timbre, and way of singing) was examined. These achievements contribute to development of the singing voice synthesis.

研究分野：知能情報処理・知能ロボティクス

科研費の分科・細目：音声情報処理

キーワード：歌声合成 歌声知覚・生成 歌声固有の音響特徴 オペラ歌唱 声区 物真似歌唱 歌声の類似度

1. 研究開始当初の背景

計算機で人工的に歌声を作る歌声合成技術は、素片接続型音声合成や統計的音声合成(例えば HMM 合成)といったコーパススペースの手法に基づいたものが主流となっている。これらの手法は、話し言葉を対象とした音声合成研究で長年に渡り培われた実用性の高い技術であり、YAMAHA の VOCALOID に代表されるような市販の歌声合成ソフトウェアでも用いられている。しかし、合成される歌声は、機械的な音質で自然性が高いとは言いがたい。さらには、様々な音色の歌声を合成する為には、膨大な歌声データの収録、或いは統計モデルを異なる音色毎に学習する必要がある。

一方、豊かな音色を持つヒトの歌声は、声帯や声道といった発声器官から生み出される音事象に他ならない。その為、自然で表情豊かな歌声を合成するためには、従来のコーパススペースの合成方法ではなく、歌唱における生理的・音響的特徴の操作に基づいた合成方法を検討する必要がある。

2. 研究の目的

本研究では、歌声固有の生理的・音響的特徴に着目し、表情豊かな歌声の合成を試みる。具体的には、歌唱(歌声の発声)における発声器官の運動に起因する音響的特徴を抽出し、それら特徴の聴感印象に与える影響の調査することで、歌声の多様な表情・音色を規定する音響的特徴を明らかにする。そして、この音響的特徴を操作することで、従来のシステムでは表現できない多様な歌声を合成可能なシステムを構築する。また、研究期間を通じて、大規模な歌声データベース(歌声 DB)の構築を行う。

3. 研究の方法

表情豊かな歌声合成を実現する為、以下の3項目の課題を設定した。

- (1) 歌声固有の声質制御
- (2) 高度な歌唱技巧が要求される歌唱スタイルの合成
- (3) 歌手固有の歌唱表現の制御。

(1)については、声区の変換・制御に着目した。声区とは、声の高さの変化に伴う声質の変化であり、地声や裏声は代表的な声区である。世界中の様々な歌唱スタイルにおいて、地声と裏声の間で声区を変換する歌唱技巧(声区変換)は頻繁に使用され、声質の多様性を生み出すだけでなく、個々の歌唱スタイルを規定する重要な要素となっている。そのため、歌声合成においても声区の違いを表現できれば、表情豊かな歌声が合成可能となる。地声と裏声の差異は、声帯の振動様式の違い、更にはその様式の違いによって生じる生理的・音響的・知覚的(声質)な違いとして定義される。そこで、地声・裏声発声を対象に、ステレオ高速度カメラを用いた声帯振動様

式の計測と、音響分析を行い、両声区の生理的・音響的な差異を調査した。また、その調査結果に基づき、地声・裏声を変換するための制御方法、更には地声・裏声を識別する技術を構築した。

(2)に関しては、代表的な歌唱芸術であるオペラ歌唱に着目した。一般に習得が難しい高度な歌唱技巧によって生み出されるオペラ歌唱を合成できれば、歌声合成技術の実用性・有効性は高まることが期待できる。オペラ歌唱を科学的に扱う研究は、古くから行われている。しかし、その音響的性質については、未解明なことが多く、既存技術で合成されるオペラ歌唱の音質も十分なものとは言えない。特にオペラ歌唱固有の特徴であるヴィブラートと歌唱ホルマントは、オペラ固有の声質を規定する重要な要素であるものの、その生理的・音響的な理解は不足している。ヴィブラートとは、声の高さの4~8Hz程度の準周期的な振動である。また、歌唱ホルマントとは、スペクトルの3 kHz 近傍に生じる顕著なピーク成分である。本研究では、オペラ歌唱データベースを作成し、これらのデータを対象に、2種の特徴の音響的特性を詳細に分析した。また、分析結果に基づき、各特徴を精密に制御に基づく高品質なオペラ歌唱合成を目指した。

(3)に関しては、歌手の個性に着目した。歌声の多様性を生み出す要因の1つに、歌手それぞれの歌唱表現(個性)がある。歌声合成技術によって任意の歌手の個性を表現できれば、歌声合成技術による表現の可能性を飛躍的に広げることができる。とりわけポピュラー歌唱の合成を考えた場合、歌手の個性の表現は、歌声を特徴付ける重要な要素となる。しかし、歌手の個性を対象にした研究は行われていない。そこで、ポピュラー歌唱を対象に、歌手夫々の声質・歌い方の特徴(個性)を対象に研究を進めた。プロ歌手の歌声を研究対象にするのは困難なため、素人歌唱者がプロ歌手の歌声を真似る際、どのような特徴を真似ているかを調査した。真似る際に積極的に制御する音響特徴を調査することにより、歌手の個性を音響的に定義することを目指した。

4. 研究成果

(1) 声区の変換・制御について

代表的な2種の声区である地声・裏声、及びこれら声区間の変換時の生理的、音響的分析を行った。高速度ステレオカメラを用いた各声区の声帯振動様式計測の結果、地声発声の方が裏声発声に比べ、声帯全体が大きく振動し、且つ声帯の上下(正中矢状面において)間に明確な位相差があることを確認した。変換時においては、変換速度に関わらず、声帯振動様式が連続的に遷移することが明らかとなった。また、音響分析の結果、声帯振動の

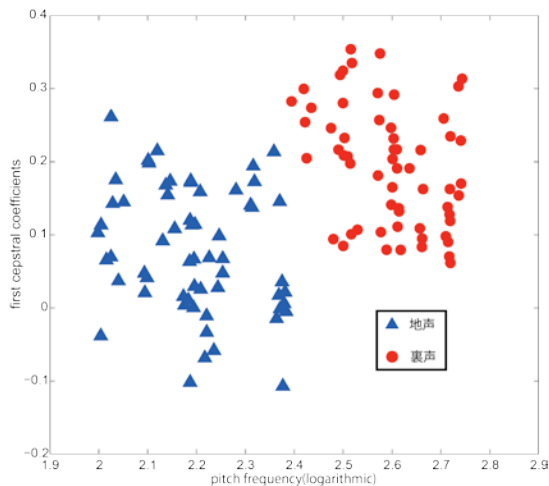


図 1 F0 とケプストラム 1 次項の関係

様式の差異は、音響的に音源スペクトル（音声スペクトル）の傾斜の違いとして現れることを確認した。また、スペクトル傾斜を表す特徴として、ケプストラムの低次項（とりわけ 1 次項）が有効であることを確認した。声の高さ (F0) は、裏声の方が地声に比べて高く、F0 が高いほど（裏声ほど）、ケプストラム 1 次項の値が大きくなる（傾きが急峻になる）関係があることを確認した。図 1 に F0 とケプストラム 1 次項の関係を示す。

これらの分析結果を基に、地声・裏声間の変換手法を構築した。地声・裏声間の切り換え合成においては、F0 の増減、及びそれに伴い図 1 の関係に従ったケプストラム 1 次項の増減を行うことで、自然な声区変換を実現できることが明らかとなった。また、この制御方法を応用し、“あたり” と呼ばれる短時間の地声-裏声変換を含む歌声を合成可能なことを示した。更には、声区変換と同じ特徴を用いて、サポートベクターマシンによって声区を自動識別する手法を提案した。声区変換は、ポピュラー歌唱をはじめ、多くの歌唱スタイルで頻繁に使われる歌唱技巧である。また、あたりは、民謡な長唄といった日本における伝統歌唱だけでなく、世界中の民族歌唱において用いられる歌唱技法である。本研究で提案した声区変換手法は、歌声合成音における声質の幅を広げるだけでなく、様々な歌唱スタイルを表現可能にするものである。

(2) オペラ歌唱の合成について

4 種の声域（バリトン、バス、テノール、ソプラノ）のオペラ歌唱 150 トラックを対象に、ヴィブラートと歌唱ホルマントの分析を行った。

ヴィブラートの音響的な定義は、F0 における 4~8Hz 程度の準周期的な変動成分である。その特性は、揺れの速さ（周波数）と大きさ（振幅）の 2 種のパラメータで記述できる。従来の歌声合成におけるヴィブラートは、2 種のパラメータを一定として正弦波で記述されるのが一般的であった。しかし、その特性は、1 つのヴィブラートサイクル（一つの音

高区間で生じたヴィブラート)において、時間的に変動することが知られている。そこで本研究では、ヴィブラートの速さと大きさの時間変動について分析した。その結果、ヴィブラートの速さは、ヴィブラート終盤において平均で 1.5Hz 増加する（速くなる）ことが確認された。また、大きさについては、徐々に増加した後に減少する山なりパターンが顕著に観測された。

歌唱ホルマントについては、スペクトルピークの大きさと、ピークの帯域幅を分析した。ピーク周波数は歌唱者、発声音高に依存し変化する一方で、ピークの大きさは平均 17dB、帯域幅は 1000Hz 程度で、大きな変動が無いことを確認した。

以上の分析結果を基に、オペラ歌唱合成に適したヴィブラートと歌唱ホルマントの制御方法を検討した。プロのオペラ歌唱者を対象に評価実験を行った結果、両特徴について分析で得られた特性を精密に制御することで、実音声と同等のオペラ歌唱を合成可能なことを確認した。また、各特徴の特性の差異が音質に与える影響が大きいことも明らかとなり、分析結果から僅かに異なる特性の制御した場合でも、その音質が大きく変化することが明らかとなった。これらの結果は、オペラ歌唱合成におけるヴィブラートと歌唱ホルマントの重要性を示すだけでなく、本研究で明らかにした両特徴の特性とその制御の有効性を示すものである。

今後は、各特徴の制御方式を定式化することを目指す。

(3) 歌手の個性を規定する音響特徴について

本稿では、歌唱時における物真似が音響特徴に与える影響を調査した。3 名のアマチュア歌唱者を対象に、歌謡曲 4 曲を普段通りに歌唱した場合と物真似を意識して歌唱した場合の歌声を収録し、歌い方、声道共鳴特性、声帯振動特性に関する 6 種音響特徴を分析することで、各特徴の物真似歌唱による変化を調べた。その結果、歌い方に関する特徴では、ヴィブラートと音韻継続長に顕著な変化が確認された。声道共鳴特性に関しては、物真似歌唱時に歌唱ホルマントが生起し、更には第 1・第 2 ホルマント周波数の移動が確認された。また、声帯振動特性では、スペクトル傾斜と第 1・第 2 調波の振幅の差が変化していることが分かった。これらの変化は限られた物真似歌唱においてのみ観測されるものであったが、物真似歌唱において、歌声から声質や歌い方に関する様々な特徴を歌手の特徴として抽出、それを真似て歌唱しようとしていることが明らかとなった。

今後は、歌手毎に個性を規定する特徴を抽出し、各特徴の制御方法を構築することで、物真似歌唱合成を実現することを目指す。

上記の課題以外の取組みとして、歌唱データベース（歌声 DB）の構築を行った。この歌声

DB はオペラ歌唱者 10 名, ポピュラー歌唱者 15 名による計 500 トラックで構成される大規模なものであり, 今後の歌声研究の進展に大きく貢献できるものである.

5. 主な発表論文等

[雑誌論文] (計 1 件)

- ① 羽石英里, 齋藤 毅, 城本, 修, ドナ・エリクソン, 岸本宏子, 八尋久仁代, 音楽療法士を対象とした発声訓練プログラム:一開発に向けた予備的研究一, 音声言語医学, 査読有り, Vol. 54, No. 3, 2013, pp. 186-196.
DOI: 10.5112/jjlp.54.186

[学会発表] (計 15 件)

- ① 小島 俊, 齋藤 毅, 三好正人, 歌声解体図を用いた歌声間の類似度評価, 音学シンポジウム 2014, 2014 年 5 月 25 日, 日本大学文理キャンパス(世田谷区).
- ② 寺村恭和, 齋藤 毅, 三好正人, オペラ歌唱におけるヴィブラートと歌唱ホルマントの知覚について, 日本音響学会聴覚研究会・電子情報通信学会騒音・振動研究会, 2013 年 12 月 21 日, 長崎大学(長崎市).
- ③ 森下亮祐, 齋藤 毅, 三好正人, 歌声の地声と裏声の切り替え方法の検討, 日本音響学会聴覚研究会, 2013 年 10 月 10 日, 神戸セミナーハウス(神戸市).
- ④ 齋藤 毅, [招待講演] 歌声の知覚と合成, 日本音響学会聴覚研究会/電気音響研究会/音楽音響研究会・電子情報通信学会応用音響研究会, 2013 年 7 月 18 日, 北海道医療大学(札幌市).
- ⑤ 小島 俊, 齋藤 毅, 中野倫靖, 後藤真孝, 三好正人, 歌声における裏声と地声を識別するための音響特徴量の検討, 日本音響学会電気音響研究会・電子情報通信学会応用音響研究会, 2012 年 10 月 27 日, 牛岳温泉リゾート(富山市).
- ⑥ Takeshi Saitou and Ken-Ichi Sakakibara, [Invited Talk] Abrupt register changing technique "Atari" in traditional Japanese singing, Acoustics 2012, 2012 年 5 月 16 日, Hong Kong.
- ⑦ Takeshi Saitou and Ken-Ichi Sakakibara, A study on perceptual effects of overshoot in singing voice F0 contour, Pan European Voice Conference (PEVOC9), 2011 年 9 月 1 日, Marseille.
- ⑧ 齋藤 毅, 榊原健一, 歌唱時の物真似による音響特徴の変化, 日本音響学会聴覚研究会, 2011 年 7 月 9 日, 東京学芸大学(小金井市).

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

6. 研究組織

(1) 研究代表者

齋藤 毅 (SAITOU, Takeshi)

金沢大学 理工研究域 電子情報学系 助教

研究者番号: 70446942