

平成 21 年 6 月 11 日現在

研究種目：基盤研究(B)

研究期間：2006～2008

課題番号：18300064

研究課題名（和文） 多元観測信号を用いた音信号の予測及び復元に関する研究

研究課題名（英文） Study on prediction and restoration of an original sound signal by using plural observed signals

研究代表者

板倉 文忠 (ITAKURA FUMITADA)

名城大学・理工学部・教授

研究者番号：30168299

研究成果の概要：本研究では、多元観測信号を用いて、変化・劣化した音信号から元の信号や所望の信号を予測・復元する実用的方法について検討を行った。まず、本研究の理論的基盤となる新しい音声のスペクトル表現を開発した。これにより、高品質な音声の再合成や、音声モーフィングなどの様々な変換に対する自然な音色変化の再現が可能となった。この理論や既存の枠組を利用し、様々な音声信号・音響信号の分析を行った。具体的には、話速変換処理を施した音声と早口で発声した音声の比較検討や、様々な表現を含む歌唱音声の解析、アナログレコードに含まれる歪みの数値化、室内の音響特性の場所による変化の分析、マイクロホン特性のモデル化、車内騒音のモデル化などを行い、数多くの知見が得られた。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	5,600,000	0	5,600,000
2007年度	4,500,000	1,350,000	5,850,000
2008年度	4,500,000	1,350,000	5,850,000
年度			
年度			
総計	14,600,000	2,700,000	17,300,000

研究分野：音声・音響処理

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：音声信号処理、音響信号処理、スペクトル、声道断面積関数、モデル化、頭部伝達関数、歌唱音声合成

1. 研究開始当初の背景

音信号は、様々な要因によって変化・劣化する。その要因としては、例えば、伝送路・通信路、収録機器による歪み、発声者の感情や意図などが挙げられる。また、どのような音信号を収録したいかは、応用の種類や収録者の意図などによって異なる。

本研究では、多元観測信号を用いて、変化・劣化した音信号から元の信号や所望の信号を予測・復元する実用的方法を開発する。

さらに、応用の種類や収録者の意図を考慮し、目的に応じた予測・復元を行うシステムの構築を目指す。

2. 研究の目的

本研究では、音声における話速・感情・意図による変化や、歌唱音声における唱法等による変化、雑音の重畳や収録機器・通信経路等に起因する歪みによる劣化、音響機器や楽器の構造等による音色の変化、受聴者・音源

の空間的な位置による、両耳までの伝達特性の変化などを対象とする。

このような変化・劣化した音信号から、元の音信号を予測・復元する方法を確立することが本研究の目的である。上記のいずれの場合も、変化・劣化要因、すなわち収録音の歪みや、音声の感情や話速などを、段階的に制御してデータ収集することが可能である。本研究での予測処理は、このように段階的に変化した音信号を用いて所望の信号を求めることに相当する。これを実現するには、音声・音響信号における変化のモデル化と内挿・外挿を含めた高度な補間手法の確立が必要であり、これらの理論基盤の構築を行う。

また、歪みなどによる劣化を回復する場合には、歪みを完全に除去しようとする、意図しない歪みが発生したり、異種の歪みに対処できなかつたりすることが考えられる。伝達特性の場合にも、完全な復元を目指した場合には、わずかな位置の違いや温度変化などによる影響を受けやすくなる。同様の現象は他の変化・劣化の場合にも起こり得るものである。すなわち、完全な復元を狙った場合には、環境の変化に脆弱であったり、意図しない劣化が発生する可能性があったりするなど弊害も多い。このことを踏まえ、本研究では、どの程度の復元が必要かをユーザが指定できるシステムの構築を目指す。

3. 研究の方法

まず、スペクトル予測・復元における理論基盤の構築と、この理論に基づくスペクトル表現方法（STRAIGHT スペクトル）の改良を行う。最初は、収録した多元観測信号において、要因の変化によってスペクトルがどのように変化・劣化するかを徹底した観察に基づいて数理的にモデル化する。観察によって明確な傾向が見出せない場合は、特徴量についての再検討や、主成分分析などの分析方法を用いた別の角度からの観察の実施、必要に応じて統計的モデルを併用するなどの手段を講ずる。

理論は、実環境下で収録された多元観測信号において評価・検証される。収録においては、歌唱音声における唱法等による変化、雑音の重畳や収録機器・通信経路等に起因する歪みによる劣化、音響機器や楽器の構造等による音色の変化、受聴者・音源の空間的な位置による、両耳までの伝達特性の変化などを捉えるよう条件を変えて収録する。収録された信号に理論を適用し、その結果をフィードバックすることにより、理論を磐石なものにしていく。実環境での評価を重ねることにより、机上の空論に終わらない真に実用的な理論の構築を目指す。

その他、実用的な応用システムの構築や、研究開発用ツールの整備・公開も随時行い、

研究成果の社会への還元も進めていく。

4. 研究成果

(1) 音声のスペクトル表現

本研究では、多様な音声信号に対し、高品質な再合成が可能で、声質変換・音声モーフィングなどの応用において、様々な変換を加えても劣化が少ないこと、変換を加える際の操作と音色の変化が直感的に対応することなどの条件を満たす音声のスペクトル表現について検討してきた。その結果、STRAIGHT スペクトルに対して、研究代表者が以前より検討している PARCOR 分析・LSP 分析を組み合わせた線形予測分析に基づく声道断面積関数を計算する手法を開発し、これらの条件を満たすことに成功した。

まず、我々は従来から検討されている線形予測分析に基づく声道断面積関数抽出手法を STRAIGHT スペクトルに対して計算し、音声分析・声質制御に適用することを検討した。この線形予測分析に基づく声道断面積関数抽出手法においては、適応逆フィルタを用いて抽出精度を高める方法が提案されており、ここでもこれを採用している。ところが、この手法は、高品質な音声の再合成に不可欠な高サンプリング周波数の信号に対しては、効果が薄いか、高域のスペクトルの影響により、特に声道断面積関数に変形を加えて再合成した際に極端な劣化を発生させ得ることが実験的に確かめられた。線形予測分析に基づく声道断面積関数抽出手法は、スペクトルが平坦であるという条件を入力信号が満たす必要があるが、高サンプリング周波数の信号においては、この条件を満たしていないことが原因の1つとして挙げられる。そこで、STRAIGHT スペクトルを低域と高域とに分離し、それぞれに帯域において、周波数領域のパラメータである LSP 係数を求め、それを結合した後に、PARCOR 係数へと変換し、声道断面積関数を求める手法を開発した。ここでは、低域は適応逆フィルタに基づく手法を用い、高域はケプストラムリフタリングに基づくスペクトル平坦化を予め行っており、その後自己相関関数へと変換し、LSP 係数を求めるという手順となっている。

現在、歌唱音声の音量による声道特性の変化を、この手法により求めた声道断面積関数を用いて分析し、従来手法では観察できなかった違いが観測できるなどの成果が得られている。また、この手法に基づく声質制御手法については、従来法を上回る品質が得られることが確認されている。

(2) 話速・感情による音声の変化

話速による音声の変化に関する検討では、特に、話速変換処理を行った音声（話速変換音声）と、実際に話速を変化させて発声した自然音声の特徴の違いについて分析した。そ

の結果、話速変換音声と比較した場合に、自然音声の時間長の変動が少ない音素と大きい音素に分類できることや、長時間スペクトルの500~1500Hz程度に強調が見られること、スペクトル変動が小さい箇所が少ないことなどが明らかとなった。これらの特徴を考慮した話速変換システムを構築し、話速変換率が大きくなるに従って既存のシステムよりも品質が向上することが主観評価実験により確かめられた。

また、感情による音声の変化についても、怒りや悲しみなど、いくつかの感情を段階的に変化させた音声を収録し、長時間スペクトルや基本周波数に一定の傾向があることが明らかとなった。

その他、発声者の意図の分析として、意図的に明瞭に発声した音声（以下、明瞭音声）とそうでない音声の特徴の比較を行った。その結果、明瞭音声では音素の開始点でよりスペクトルやパワーの変動が大きいことが確認された。この結果に基づき、スペクトル変動特性の視覚化を可能とする尺度を構築した。この尺度はデルタケプストラムを利用したものであり、これを用いることにより、特に明瞭音声ではホルマントのピークの変動がより強調される傾向にあることが明らかとなった。

(3) 歌唱音声における唱法等による変化

歌唱音声においては、単母音を用いて、音高や音量、ビブラートなどのバリエーションを収録し、周波数特性の分析及びモデル化を行った。本研究で開発した声道断面積関数抽出手法を用いて分析を行うなどした結果、ある程度のモデル化が可能であることが確認された。また、様々な子音を含むスキットにおいて、テンポによるパワー変動特性の違いの分析及びモデル化を行った。パワーの時系列に対し2つの時間窓を組み合わせた窓を用い、時間方向の平滑化の程度を制御する方法を開発し、主観評価実験により一定の効果が得られることが確認された。そして、これらの知見を導入したリアルタイムに歌唱音声合成を行うシステムを構築した。これは、MIDI キーボードからの入力に応じてリアルタイムに合成するシステムである。

(4) 雑音の重畳や収録機器・通信経路等に起因する変化・歪み

雑音の重畳については、様々な音圧レベルの車内騒音を収録し、分析及びモデル化を行った。モデル化した騒音を用いて音声認識システム及び被験者を用いた書き取りによる認識実験を行い、認識率の傾向に違いがあるかを検討した。その結果、認識率の傾向が実際の騒音とモデル化したもので極めて類似しており、モデル化した騒音により、車内騒音をほぼ近似できることが明らかとなった。

また、音楽信号の楽音符号化による劣化や、

アナログレコードの歪みなどについて分析・モデル化を行った。特に、アナログレコードの時間的なゆらぎ特性のモデル化を行い、ゆらぎを含むものと含まないものの聴覚的な距離を適切に計算できる尺度の構築を行った。この尺度により、楽音符号化された信号やアナログレコード等に含まれる劣化を数値化することに成功している。

(5) 音響機器や楽器の構造等による音色の変化

様々なマイクロホンの特性を測定し、その違いについて分析を行った。その結果、マイクロホンが特有の周波数特性を持つだけでなく、特有の非線形歪みを持っていることが明らかとなった。この非線形歪みのある程度モデル化することが可能となった。この成果により、マイクロホンの個性を忠実に再現できるシミュレータの構築につながる可能性が開かれたと言える。

また、楽器に関しても、ボディの大きさや形状の異なるアコースティックギターや、アルトサクソ・テナーサクソなど大きさの異なるサクソを対象として分析・モデル化を行った。アコースティックギターについては、接触面に振動を与えて音を再生する特殊なスピーカを用い、ボディの響きに関するインパルス応答を測定した。そして、それを、サイレントギターを用いて別途収録した弦の振動成分と畳み込むことにより、ギター音を合成できる。これにより、ボディの響きを変更することのできるギター音合成システムの構築が可能となった。

サクソについては、長時間スペクトルや立ち上がり部分がサクソの種類によって異なることが明らかとなった。この知見を取り入れ、アルトサクソ音からテナーサクソ音へと変換するシステムの構築を行った。

(6) 受聴者・音源の空間的な位置による、両耳までの伝達特性の変化

頭部伝達関数の測定において、マイクとスピーカの距離を変化させて測定を行った。分析の結果、距離によってスペクトルに若干の違いは表れるものの、明確な傾向は見られなかった。

また、室内伝達特性における収録位置による特性の変化についても検討を行った。その上で、頭部伝達関数と空間伝達特性を組み合わせ、実際の部屋での受聴環境をシミュレートするシステムを構築した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

1. Hideki Banno, Hiroaki Hata, Masanori Morise, Toru Takahashi, Toshio Irino and

Hideki Kawahara

Implementation of realtime STRAIGHT speech manipulation system, Acoustical Science and Technology, 査読有, 28 (3), pp.140-146, 2007

[学会発表] (計 32 件)

1. 片山裕規, 坂野秀樹, 板倉文忠, “アコースティックギター音の分析合成手法に関する検討,” 日本音響学会 春季研究発表会, 2009年3月18日, 東京工業大学, 東京
2. 笠原優也, 坂野秀樹, 板倉文忠, “時間的不整合を含む楽音信号に対する客観評価尺度の提案,” 日本音響学会 春季研究発表会, 2009年3月18日, 東京工業大学, 東京
3. 水野雄介, 坂野秀樹, 板倉文忠, “歌唱音声における音高連続変化音声の分析,” 日本音響学会 春季研究発表会, 2009年3月17日, 東京工業大学, 東京
4. 荒川綾則, 内村佳典, 坂野秀樹, 板倉文忠, “音圧レベル変化を含む歌唱音声の周波数分析,” 日本音響学会 春季研究発表会, 2009年3月17日, 東京工業大学, 東京
5. 西脇涼介, 坂野秀樹, 板倉文忠, “話声から歌声への発話変化に伴う音声特徴量の変化の分析,” 日本音響学会 春季研究発表会, 2009年3月17日, 東京工業大学, 東京
6. 柴田吉輝, 坂野秀樹, 板倉文忠, “意図的な明瞭発声における動的特徴の分析,” 日本音響学会 春季研究発表会, 2009年3月17日, 東京工業大学, 東京
7. 田原拓, 坂野秀樹, 板倉文忠, “早口音声の特徴に基づいた人間らしい話速変換音声の生成に関する分析,” 日本音響学会 春季研究発表会, 2009年3月17日, 東京工業大学, 東京
8. 河原英紀, 森勢将雅, 高橋徹, 坂野秀樹, 西村竜一, 入野俊夫, “音響的イベントの持続時間に基づいた非周期成分の時間構造の制御について,” 日本音響学会 春季研究発表会, 2009年3月17日, 東京工業大学, 東京
9. 坂野秀樹, 吉田翔, 高橋茂樹, 板倉文忠, 森勢将雅, 高橋徹, 河原英紀, “STRAIGHT ライブラリを用いた高品質音声分析合成アプリケーションの開発,” 日本音響学会 春季研究発表会, 2009年3月17日, 東京工業大学, 東京
10. Hideki Kawahara, Masanori Morise, Hideki Banno, Toru Takahashi, Ryuichi Nisimura, Toshio Irino, “Spectral Envelope Recovery beyond the Nyquist Limit for High-Quality Manipulation of Speech Sounds,” Interspeech 2008, 2008年9月24日, Brisbane, Australia
11. Yoshinori Uchimura, Hideki Banno, Fumitada Itakura, Hideki Kawahara, “Study on Manipulation Method of Voice Quality Based on the Vocal Tract Area Function,” Interspeech 2008, 2008年9月23日, Brisbane, Australia
12. 河原英紀, 森勢将雅, 高橋徹, 坂野秀樹, 板垣英恵, 大西荘登, 西村竜一, 入野俊夫, “零周波数フィルタ信号に基づく基本周波数抽出法の評価と応用について,” 日本音響学会 秋季研究発表会, 2008年9月11日, 九州大学, 福岡
13. 田原拓, 坂野秀樹, 板倉文忠, “早口音声と話速変換音声におけるスペクトル変動特性の違いに関する分析,” 日本音響学会 秋季研究発表会, 2008年9月10日, 九州大学, 福岡
14. 荒川綾則, 坂野秀樹, 板倉文忠, “歌唱音声における音圧レベルによる周波数特性変化の個人性,” 日本音響学会 秋季研究発表会, 2008年9月10日, 九州大学, 福岡
15. Hideki Kawahara, Masanori Morise, Toru Takahashi, Ryuichi Nisimura, Hideki Banno, Toshio Irino, “A unified approach for F0 extraction and aperiodicity estimation based on a temporally stable power spectral representation,” ISCA Tutorial and Research Workshop, 2008年6月5日, Aalborg, Denmark
16. Hideki Kawahara, Masanori Morise, Toru Takahashi, Ryuichi Nisimura, Toshio Irino, Hideki Banno, “A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation,” ICASSP 2008, 2008年4月1日, Las Vegas, Nevada, U.S.A.
17. 坂野秀樹, 森勢将雅, 高橋徹, 西村竜一, 入野俊夫, 河原英紀, “リアルタイム STRAIGHT の改良と STRAIGHT ライブラリの実装,” 日本音響学会 聴覚研究会 / 電子情報通信学会 音声研究会, 2008年3月21日, 東京大学, 東京
18. 川添正人, 坂野秀樹, 板倉文忠, “スキヤットにおける声道断面積関数のテンポによる違いの分析,” 日本音響学会 春季研究発表会, 2008年3月19日, 千葉工業大学, 千葉
19. 田原拓, 坂野秀樹, 板倉文忠, “早口音声と話速変換音声における音素長の違いに関する分析,” 日本音響学会 春季

- 研究発表会, 2008年3月19日, 千葉工業大学, 千葉
20. 荒川綾則, 坂野秀樹, 板倉文忠, “歌唱音声における音圧レベルによる周波数特性の違いの分析,” 日本音響学会 春季研究発表会, 2008年3月19日, 千葉工業大学, 千葉
 21. 内村 佳典, 坂野 秀樹, 板倉 文忠, “声質制御への応用を目的とした、声道断面積関数の話者による違いの分析,” 日本音響学会 春季研究発表会, 2008年3月18日, 千葉工業大学, 千葉
 22. 荒川綾則, 坂野秀樹, 板倉文忠, “線形予測分析に基づくソフトウェア騒音計の開発,” 日本音響学会 春季研究発表会, 2008年3月17日, 千葉工業大学, 千葉
 23. 内村佳典, 坂野秀樹, 板倉文忠, “声質制御への応用を目的とした声道断面積関数の分析,” 音楽情報科学研究会・音声言語情報処理研究会, 2008年2月9日, ホテル暖香園, 静岡
 24. 川添正人, 坂野秀樹, 板倉文忠, “テンポの変化による影響を考慮した歌唱音声合成に関する検討,” 音楽情報科学研究会・音声言語情報処理研究会, 2008年2月8日, ホテル暖香園, 静岡
 25. 河原英紀, 森勢将雅, 高橋徹, 西村竜一, 坂野秀樹, 入野俊夫, “TANDEM-STRAIGHT に基づく基本周波数の抽出と評価について,” 第9回音声言語シンポジウム, 2007年12月21日, NTT 京阪奈ビル, 京都
 26. 河原英紀, 森勢将雅, 高橋徹, 西村竜一, 入野俊夫, 坂野秀樹, “STRAIGHT における時間周波数分析の新しい定式化と実装について,” 日本音響学会 秋季研究発表会, 2007年9月21日, 山梨大学, 山梨
 27. 川添正人, 坂野秀樹, 板倉文忠, “スキヤットにおける, スペクトル変動特性の音韻による違いの分析,” 日本音響学会 秋季研究発表会, 2007年9月19日, 山梨大学, 山梨
 28. 内村佳典, 坂野秀樹, 板倉文忠, “声道断面積関数による声質制御方式に関する検討,” 日本音響学会 秋季研究発表会, 2007年9月19日, 山梨大学, 山梨
 29. Hideki Kawahara, Masanori Morise, Toru Takahashi, Hideki Banno, Toshio Irino, Osamu Fujimura, “Group delay for acoustic event representation and its application for speech aperiodicity analysis,” EUSIPCO 2007, 2007年9月4日, Poznan, Poland
 30. 川添正人, 坂野秀樹, 板倉文忠, “スキヤットにおける, スペクトル変動特性

のテンポによる変化のモデル化,” 日本音響学会 春季研究発表会, 2007年3月13日, 芝浦工業大学, 東京

31. Hideki Banno, Kumi Ohta, Masato Kawazoe and Fumitada Itakura, “Spectral modification technique in conversion of musical notes or tempos for singing voice synthesis system,” 4th Joint Meeting Acoustical Society of America and Acoustical Society of Japan, 2006年11月29日, Hawaii, U.S.A.
32. 河原英紀, 森勢将雅, 高橋徹, 入野俊夫, 坂野秀樹, 藤村靖, “STRAIGHT スペクトルに基づく音源信号の抽出と非周期成分の評価について,” 音声研究会, 2006年11月10日, 産業技術総合研究所, 茨城

[産業財産権]

○出願状況 (計1件)

名称: 帯域分割線スペクトル対による声道断面積関数抽出手法 (仮)

発明者: 坂野秀樹, 板倉文忠, 荒川綾則, 内村佳典

権利者: 同上

状況: 学内処理中

[その他]

STRAIGHT 関連ツール

(STRAIGHT ライブラリ等) URL:

<http://www.crestmuse.jp/cmstraight/>

6. 研究組織

(1) 研究代表者

板倉 文忠 (ITAKURA FUMITADA)

名城大学・理工学部・教授

研究者番号: 30168299

(2) 研究分担者

坂野 秀樹 (BANNO HIDEKI)

名城大学・理工学部・准教授

研究者番号: 20335003

河原 英紀 (KAWAHARA HIDEKI)

和歌山大学・システム工学部・教授

研究者番号: 40294300

西野 隆典 (NISHINO TAKANORI)

名古屋大学・情報メディア教育センター・

助教

研究者番号: 40329769

戸田 智基 (TODA TOMOKI)

奈良先端科学技術大学院大学・

情報科学研究科・助教

研究者番号: 90403328