

「デジタルメディア作品の制作を支援する基盤技術」
平成 17 年度採択研究代表者

片寄 晴弘

関西学院大学理工学部・教授

時系列メディアのデザイン転写技術の開発

1. 研究実施の概要

本チームは、デザイン支援、能動的鑑賞、新エンタテインメントをキーワードに掲げ、研究開発を進めている。平成 19 年度の取り組みの過程で、「デザイン転写」技術の実利用を進めるためには抽象化レベルでの操作性 (=directability) が重要であることを再認識し、研究の前半を締めくくる平成 20 年度は、「デザイン転写」技術の集積と「directability」を有するインタフェースの実装を主たる研究課題とした。

具体的には、能動的音楽鑑賞に関する成果として、a) ユーザの評価と楽曲の内容併用による音楽検索・推薦システム、b) 音響信号とその楽曲の楽譜を利用した楽器別音楽イコライザ、c) 調波構造の和音性モデルに基づく楽曲ムードビジュアライザの開発に至った。また、fNIRS を用いた脳機能計測により、d) 聞き方のモードによって脳の賦活状況が変化することを示唆するデータを取得した。デザイン転写に関する成果として、e) 与えられた歌詞の韻律に基づく自動作曲システム、f) フレーズとパルスモデルに基づく演奏の表情付けシステムを開発したほか、g) MusicXML 準拠の階層的な音楽データ共通記述方式(CMX: CrestMuseXML)を策定するとともに、ピアノを対象とした h) 演奏表情データベース CrestMuse PEDB Ver.2.0 の一般公開に至った。また、他の応用システムの実現を支える i) 多声音楽信号分析技術や歌唱デザインインタフェースのプラットフォームとなる j) STRAIGHT の高性能化等、信号処理に関する成果も確保された。

メディア系の研究領域のアウトリーチ活動を考える際、キラー領域、キラーアプリケーションの設定は不可欠である。今年度より、一般ユーザにとってもなじみ深い「歌唱・歌声」を重点研究テーマとして位置づけ、各グループの力を結集する形でのシステム開発に着手した。k) 伴奏を含む音響データを対象とした似た歌声楽曲検索システムや、l) SingBySpeaking: 話声を歌声に変換するシステム、m) VocaListener や v.morish 等の歌唱デザインインタフェースの開発を進めた。以上の成果をビデオとして取りまとめ、一般公開した。

2. 研究実施内容(文中にある参照番号は 4.(1)に対応する)

各グループの進捗状況を以下に示す。

片寄グループ: 認知的音楽理論に基づくデザイン転写と全体総括

研究の目的および内容

音楽には、複数のデザイン記述レベルにおいて、1) 未来の elaboration は、観測された当該 elaboration の振る舞いと因果性を持つ、2) elaboration は、reduction の階層的な組み合わせに基づいて付与されるという性質がある。これらの性質に基づく音楽情報処理システムを構成することを目的として、認知的な視点を取り入れた音楽理論の計算モデルを構成する。

研究実施項目・概要

音楽のデザイン転写に関する研究の一貫として、ベイジアンネットワークを利用したハーモナイゼーション[4]、予測型伴奏システムのシステム化を進めるとともに、音楽情報処理研究のフレームワーク Java クラスライブラリ CMX Library¹の整備を進めた。また、音響として存在する名演奏の演奏制御情報を抽出・アーカイブした演奏表情 CrestMuse PEDB ver2.0[5]²を一般公開した。

事例活用型デザインインタフェースの構成 directability 基軸に関するものとして、歌い回しと声質の2軸で歌唱の実時間モーフィングを可能とするインタフェース v.morish[65]、CrestMuse PEDB ver2.0 に準拠したフレーズ単位の演奏デザインインタフェース mixtract のデザインを実施した。また、指揮型の演奏インタフェースとして、iFP.MJ, AIIM を開発した。

能動的音楽鑑賞の評価に関連した研究項目としては、fNIRS を用いた脳機能計測実験により、音楽境界の判別タスクにおいて聞き方のモードによって脳の賦活状況が変化することを示唆するデータを取得した[128]。

CrestMuse プロジェクトにおける取り組みをヒューマンインタフェース学会、Audio Engineering Society (AES 日本支部)、CEDEC 等の招待講演で発表した[21][23]。また、今年度は、プロジェクト中間期にあたるため、内外からのゲスト講演者を招聘し、シンポジウムを開催した。そこでのデモ内容を素材として中間成果ビデオを制作して一般公開するなど、アウトリーチ活動につとめた。

後藤グループ: 音楽デザイン転写・音響信号理解に基づく音インタフェース

研究の目的および内容

音楽デザイン転写と音響信号の分析・理解に関する研究を実施すると共に、それらの研究成果に基づいて具体的な音インタフェースの事例を提案し、ユーザが実際に操作可能なインタフェースシステムを研究開発する。

¹ <http://www.crestmuse.jp/cmz/>

² <http://www.crestmuse.jp/pedb/>

研究実施項目・概要

本年度は、「歌唱デザイン」に関する研究に注力し、歌声や歌詞を対象とする研究領域を、学会誌解説記事[13]において我々が新たに「歌声情報処理」と名付けて、様々な研究開発を実施した。その具体的な成果として、例えば、混合音中のボーカルパートに対する音高推定・モデリング手法[2][20]や、複数の楽曲の歌詞中に共通して登場するフレーズ間へのリンク作成手法[78][135]、ユーザ歌唱を真似る歌声合成パラメータを自動推定するシステム「VocaListener」[9][61][85][152]、話声を歌声に変換するシステム「SingBySpeaking」[14][153]等を実現した。さらに基礎研究として、歌声に含まれる個人性知覚に寄与する音響特徴[75]や、歌唱力評価に寄与する音響特徴[86][143]を明らかにするための検討も進めた。

さらに、前年度に引き続き、音楽家でないエンドユーザが容易に能動的な音楽の聞き方を体験できる「能動的音楽鑑賞インタフェース」の研究開発も進め、我々の研究アプローチの重要性と主要な成果に対する対外的認知度を基調講演 (TELECOM ParisTech Workshop)[19]や招待講演 [28]等によって高めた。他にも、音響的特徴に基づく楽曲のサムネイル画像生成手法の研究[79][53]や、奥乃グループと共同で、楽器音イコライザによる音色の類似度に基づく楽曲検索システムの研究[76][134]等に取り組んだ。今後の歌唱デザイン研究のためのデータベースについても、河原グループ、片寄グループ等と共同で検討した。

河原グループ:聴覚情報表現に基づく音楽デザイン転写

研究の目的および内容

人間の聴覚における音声・音響信号の情報表現の本質的理解に基づいて構築された高品質音声分析変換合成システム STRAIGHT を主な研究手段として、歌唱を中心とする音楽のデザイン転写を行うシステムを研究開発する。

研究実施項目・概要

本年度は、他のグループにおける歌唱研究のインフラとなる TANDEM-STRAIGHT に基づくプログラム基盤の整備および技術の高度化を進めた[8][37][42][54][101]。この過程で、歌い回しと声質の2軸で歌唱を実時間でモーフィングするインタフェース v.morish[65]の制作に協力するとともに、その基盤となるモーフィング率の時間軸を含めた時系列化の拡張を可能にする方式を確立した[100]。また、実行時間の大きな部分を占める基本周波数の抽出に関しても、実時間処理に適した基本周波数抽出法を開発した[57][71]。さらに、歌唱デザインの重要な要素である歌手性の制御に関して、母音のみに基づいて声質を変換する方式の開発を進めた[58]。これらの大きな技術的進歩に加え、TANDEM-STRAIGHT および、拡張したモーフィングをより利用し易くするためのインタフェースを新たに開発し、インフラとして提供した[111]。これらのインフラは、他のグループの研究を促進するとともに、聴覚の脳機能を解明するための重要な方法論として利用され始めている[1]。さらに、これらの成果の普及を図るため、招待講演 [18][25][27][29]、雑誌記事[10]等を通じた広報活動を積極的に進めた。

歌唱の重要な要素である歌唱データベースに関しては、感情表現の操作のための収録データの拡張を進めるとともに、上記の技術的進歩を反映させ、要素的な音高遷移やビブラート、歌唱の年齢表現などへの拡張を行った。これらに関して、今後の歌唱デザイン研究のためのデー

データベースについても、後藤グループ、片寄グループ等と、共同で検討を進めた。

嵯峨山グループ: 音楽デザイン転写・音響信号理解に基づく音インタフェース

研究の目的および内容

確率統計と学習理論に基づく数理的アプローチにより、音楽音響信号中の構成音の音高と音長を推定する技術を研究する。また、音楽の和声付けや伴奏自動生成において、事例中に内在する音楽デザインを転写する研究を行なう。

研究実施項目・概要

歌唱曲の自動作曲とその自動歌唱演奏のシステム: 任意の日本語の漢字仮名混じり文章入力から、その歌詞固有の韻律を旋律化することによる自動作曲する技術を開発し、それを自動作曲システム Orpheus[74]として実現し、web ベースのサービスとして公開した。和声進行、リズムパターン、低音進行、伴奏音型などが既存の音楽からのデザイン転写によって与えられ、新たな任意の歌詞に対して妥当な新作曲が自動で行える。大きな反響があり、すでに千曲を超える作曲利用があり、作品は YouTube やニコニコ動画などに多数掲載され、TV や報道で取り上げられ、教育現場での利用実績もできた。

確率モデルに基づく楽曲の和声の自動解析技術開発、音楽デザイン自動学習: 多声音楽・合奏音楽を、隠れマルコフモデル (HMM) および文脈自由文法 (CFG) によりモデル化し、その背後にある和声進行を推定する理論と技術を開発した。特に、音楽信号を入力とする和声推定では、その国際コンテストである MIREX2008 において 1 位の成績を獲得した。

その他に、楽譜と音楽演奏信号との詳細な整合技術[82]、調波音と打楽器音のブラインド分離技術に基づく能動的音楽鑑賞イコライザ技術の開発を行い、前者については片寄グループでの指揮システムに応用利用された。

奥乃グループ: AI アプローチに基づく音楽デザイン転写

研究の目的および内容

音楽音響信号は複数の楽器音や音声から構成される混合音であるので、その中に含まれる楽器音や歌声を分離認識し、記号レベル表現を生成する。まず、楽器音、リズム、和声、メロディなどの基本音楽要素を認識し、MusicXML に基づいた記号表現を生成する。このような音楽音響信号から記号表現への変換においては、音楽的な高次構造に基づく予測を行うことにより、全データを参照することなく、逐次的に処理するオンラインアルゴリズムを設計する。このアルゴリズムは本研究課題である elaboration の実時間認識の基礎技術として展開していく。

奥乃グループ: AI アプローチに基づく音楽デザイン転写

音楽音響信号の分析・理解のための基盤技術として、平成 19 年度に開発した技術の高度化に取り組むと共に、それらの成果を具現化するためにデモシステムの開発にも取り組んだ。本期間中に開発した技術、システム開発事例を以下に示す。

音楽音響信号中の各パートの変換に基づく能動的音楽鑑賞(後藤グループと共同研究): 調

波・非調波混合モデルの洗練化に取り組み、楽器音分離にもとづいた楽器音イコライザを開発し、さらに、「楽器音の音量の組み合わせによって、同一楽曲であっても、ジャンルが変わる」という仮定に基づいた類似楽曲検索技術に応用をした[134].

ドラムパターン認識技術を通じたリズム認識に基づく踊り歌うロボット(後藤グループ・河原グループと共同): 実時間ビート追跡システムを開発し、得られた拍子情報から次のビートを予測し、リズムに合わせて「1, 2, 3, 4」とカウントするロボットを開発した [46] . ビート速度に合わせた発声は種となる音声を STRAIGHT により音長を変えて作成している.

新しい楽器音ソースを提供する楽器音パレット: 国際会議[51]で口頭発表を行うとともに、中間シンポジウム用にデモンシステムを開発し、音パレットという新しい概念を提案し、その処理系のデモを通じて、開発手法の有効性を実証した.

歌唱における演奏表現を支える多様な声質と歌い回しの分析・合成・制御(河原グループ、片寄グループ、後藤グループとの共同研究): 今期後半より研究を開始し、歌う姿勢と音響的放射特性について焦点を絞り、無響室での測定を開始した.

3. 研究実施体制

(1)「片寄」グループ

- ①研究分担グループ長: 片寄 晴弘(関西学院大学、教授)
- ②研究項目: 認知的音楽理論に基づくデザイン転写と全体総括

(2)「後藤」グループ

- ①研究分担グループ長: 後藤 真孝((独)産業技術総合研究所、主任研究員)
- ②研究項目: 音楽デザイン転写・音響信号理解に基づく音インタフェース

(3)「河原」グループ

- ①研究分担グループ長: 河原 英紀(和歌山大学、教授)
- ②研究項目: 聴覚情報表現に基づく音楽デザイン転写

(4)「嵯峨山」グループ

- ①研究分担グループ長: 嵯峨山 茂樹(東京大学大学院、教授)
- ②研究項目: 数理的アプローチに基づく音楽デザイン転写

(5)「奥乃」グループ

- ①研究分担グループ長: 奥乃 博(京都大学大学院、教授)
- ②研究項目: AI アプローチに基づく音楽デザイン転写

4. 研究成果の発表等

- (1) 論文発表 (原著論文)

1. Stefan R. Schweinberger, Christoph Casper, Nadine Hauthal, Juergen M. Kaufmann, Hideki Kawahara, Nadine Kloth, David M.C. Robertson, Adrian P. Simpson and Romi Zaeske, Auditory Adaptation in Voice Perception, *Current Biology* 18, 684-688, May 6, (2008).
2. 藤原 弘将, 後藤 真孝, 奥乃 博: 歌声の統計的モデル化とビタビ探索を用いた多重奏中のボーカルパートに対する音高推定手法, *情報処理学会論文誌*, Vol.49, No.10 (Oct. 2008) pp.3682-3693, 情報処理学会.
3. 嵯峨山 茂樹, 亀岡 弘和, “自動採譜技術の展望,” *日本音響学会誌*, 64 巻 12 号, pp.5-720, Dec., 2008.
4. 北原 鉄朗, 勝占 真規子, 片寄 晴弘, 長田 典子: "ベイジアンネットワークを用いた自動コードヴォイシングシステム", *情報処理学会論文誌*, Vol.50, No.3, pp.1067--1078, March 2009.
5. 橋田光代, 松井淑恵, 北原鉄朗, 片寄晴弘: ピアノ名演奏の演奏表現情報と音楽構造情報を対象とした音楽演奏表情データベース CrestMusePEDB の構築, *情報処理学会論文誌*, Vol.50, No.3, pp. 1090-1099, March 2009.
6. 藤澤隆史・谷光彬・長田典子・片寄晴弘: 和音性の定量的評価モデルに基づいた楽曲ムードの色彩表現インタフェース. *情報処理学会論文誌*, Vol.50, No.3, pp.1133-1138, March 2009.
7. 安部 武宏, 糸山 克寿, 吉井 和佳, 駒谷 和範, 尾形 哲也, 奥乃 博: 音色の音高依存性を考慮した楽器音の音高操作手法, *情報処理学会論文誌*, Vol.50, No.3, pp.1054-1066, March 2009.
8. 森勢 将雅, 高橋 徹, 河原 英紀, 入野 俊夫: 分析時刻に依存しない周期信号のパワースペクトル推定法を用いた音声分析, *電子情報通信学会論文誌*, 和文 A 分冊, vol.J92-A, no.3, pp.163-171, 2009.

(2) 特許出願

平成 20 年度 国内特許出願件数 : 4 件 (CREST 研究期間累積件数 : 15 件)