

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。本年度から研究後期にさしかかることから、取り組みの全体像を表すような成果の確保を目指し、研究成果の集約化に力を入れた。

具体的には、一般にもわかりやすい音楽系アプリケーション領域として1) 歌唱デザインシステムの開発、音楽情報科学研究領域への貢献として2) 音楽デザイン転写に基づくコンテンツ制御、広く音楽学研究全体への貢献として、3) 音楽デザインのための **Directable-User-interface** の開発を主要研究目標とした。このうち、1) 歌唱デザインシステムの構築・統合化については、河原グループ、後藤グループ、片寄グループ、奥乃グループによる共同研究プロジェクト、2) 音楽デザイン転写に基づくコンテンツ制御については、片寄グループ、嵯峨山グループとの共同研究プロジェクトとして研究を展開した。

これらの取り組みのうち、後藤グループが開発した「VocaListener」がヤマハ株式会社に技術移転されるなど実用化に関する成果も上がっている。また、9月に神戸で開催された音楽情報検索に関する国際会議 ISMIRのサテライトワークイベントとしてCrestMuse Workshopを開催し、上記、3テーマでのデモセッションを執り行った。CrestMuse Workshopには内外から 80 名の参加者があり、その模様が、JSTのサイエンスニュース¹としてネット上で公開された。

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

- ・ 音楽情報処理統合研究環境の整備
 - CMX の設計と API 開発 (嵯峨山グループとの共同研究)
 - 演奏表情 DB のバージョン 3.0 の制作と配付

¹ <http://sciencenews.jp/index4.html>

- ・ 音楽デザインインタフェースにおける **Directability** の実現
 - 表情付けシステム
 - 歌唱モーフィングインタフェース(河原グループとの共同研究)
- ・ 歌唱デザインシステムの構築・統合化(河原グループ、後藤グループ、奥乃グループとの共同研究)
 - システムの統合化マネージメント
 - 歌唱 DB の制作(河原グループとの共同研究)
- ・ CrestMuse Workshop の運営マネージメント

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

- ・「歌唱デザイン」のための歌声、歌詞に着目した「歌声情報処理」の研究
- ・歌唱デザインシステムの構築・統合化
(主に河原グループ、片寄グループとの共同研究)
- ・音楽音響信号中の音楽的要素の推定と「能動的音楽鑑賞インタフェース」のデザイン
- ・音楽デザイン転写用インタフェースの研究
- ・音楽音響信号中の各パートの分離・編集を可能にする音源分離技術の研究
(奥乃グループとの共同研究)

河原グループ:聴覚情報表現に基づく音楽デザイン転写

- ・ 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(片寄グループ、後藤グループ、奥乃グループとの共同研究)
- ・ 歌唱デザインシステムの構築・統合化(片寄グループとの共同研究)。
 - 歌唱における歌手の声質の転写アルゴリズムの開発
 - 歌唱における歌手の歌い回しの転写アルゴリズムの開発
 - 歌唱デザイン転写の操作インタフェース
 - 歌唱における演奏表現データベースの構築

嵯峨山グループ:数理的アプローチに基づく音楽デザイン転写

- ・ 歌詞の韻律に基づく歌唱曲の自動作曲の高機能化(楽曲階層構造、和声自動設計、打楽器パート付与、リズムスタイル指定、特殊音階サポートなどの検討と実現)
- ・ 作曲理論(和声学)に基づく楽曲の和声の自動解析技術開発、和声ラベル付与データ作成
- ・ 伴奏つき音楽信号からの歌声の抽出および除去技術(自動カラオケ作成)の開発
- ・ 合奏音楽信号からのリズムパターン抽出と楽曲のジャンル分類への応用

- ・ キーボード演奏への自動伴奏システムの開発

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)
- ・ 新しい楽器音演奏法を提供する音パレット
- ・ ベース音高を組み合わせた和音進行認識技術
- ・ 実時間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)
- ・ 歌唱デザインシステムの構築・統合化(河原グループ、片寄グループ、後藤グループとの共同研究)

§ 2. 研究実施体制

(1)「片寄」グループ(関西学院大学理工学部)

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((独)産業技術総合研究所)

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(和歌山大学システム工学部)

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(東京大学大学院情報理工学系研究科)

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(京都大学大学院情報学研究科)

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

§ 3. 研究実施内容

(文中に番号がある場合は(4-1)に対応する)

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

研究活動の後半に入ることに伴い、グループ内の研究の深化を進めるとともに、チーム全体としての成果のまとめあげの準備を課題とした。グループ内の研究としては、CMX Library 利用による予測型セッションシステム[5]、歌唱モーフィングインタフェースの実時間処理化、演奏デザインインタフェース mixtract の特許出願(2010 年2月初旬出願)、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 の ver2.3 のリリース(全80曲)、EC (Entertainment Computing)シンポジウムの特別イベントとしての Rencon の開催などの成果が上がっている。

チーム全体としての成果のまとめあげについては、1)音楽デザイン転写に基づくコンテンツ制御、2)音楽デザインのためのDirectable-User-interfaceの開発、3)歌唱デザインシステムの開発、の3項目に絞り込んで、各グループを超えて共同研究を展開し、粒度の大きい成果として各種イベントでのデモ展示を執り行った。本年度9月に神戸で開催されたISMIRのサテライトワークイベントとしてCrestMuse Workshopを開催し、上記、3テーマでのデモセッションを執り行った。

CrestMuse Workshopには内外から 80 名の参加者があり、その模様が、JSTのサイエンスニュース²としてネット上で公開された。



CrestMuse Workshop の様子

音楽デザイン転写に基づくコンテンツ制御については、ベイジアンネットワークに基づくハーモナイゼーションシステムと嵯峨山グループによる自動作曲システム Orpheus を統合し、連携動作させるためのアプリケーションを開発した。音楽デザインのための Directable -User-interface の開発については、演奏の表情付けデザイン方法論に関連してフレーズ構造の分析支援と頂点(演奏表現上での重要音)を提示する機能を実装した。また、指揮システムについては、指揮者の持つヒューリスティクスの導入と楽曲スタイルに応じたスケジューラを構成した。この他にも iPhone や wii 等を利用したメディア演奏システムの開発を進めて来た結果、メディアアートコンペティション BAKA-JA のネットアート部門の最優秀賞を受賞するなどの成果が上がっている。歌唱デザインシステムの開発については、歌唱モーフィングを時系列上で実施するインタフェースの実装と全体のとりまとめを担当した。

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。

研究実施項目・概要

本年度は、前年度に引き続き「歌唱デザイン」に関する研究に注力し、歌声、歌詞に着目した「歌声情報処理」に関する研究開発を様々な角度から進めた。その具体的な成果として、例えば、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」、話声を歌声に変換するシステム「SingBySpeaking」等を研究開発した。特に VocaListener はヤマハ株式会社に技術移転し、実用化活動に取り組んでいる。さらに基礎研究として、楽曲の混合音中の歌声に対する

² <http://sciencenews.jp/index4.html>

音素認識手法や、歌唱指導による音響特徴の変化とその歌唱力評価への影響を明らかにするための検討も進めた。

さらに、これまで研究開発してきた、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」(active music listening interfaces)の議論を発展させ、新たな研究アプローチとして「音楽理解力拡張インタフェース」(augmented music-understanding interfaces)の考察を深め、この我々の独自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等によって高めた。他にも、音楽に同期したコメントを自動生成するシステムの研究や音楽推薦システムの研究等に取り組んだ。さらに、奥乃グループと共同で、楽器音イコライザのための音源分離手法[1]や、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等にも取り組んだ。

河原グループ:聴覚情報表現に基づく音楽デザイン転写

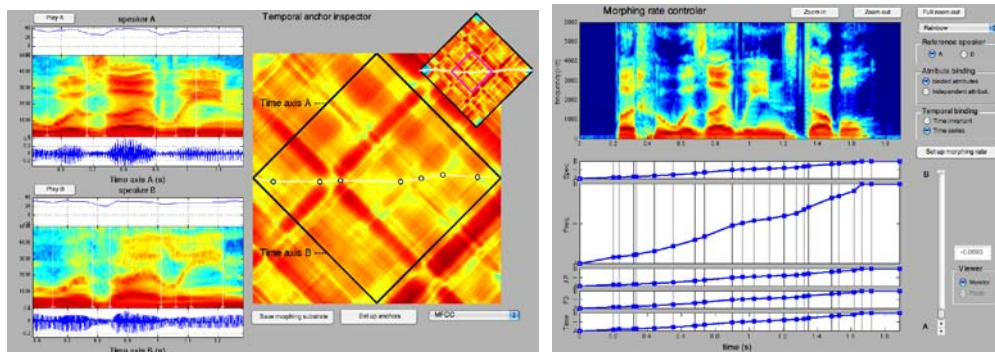
研究の目的および内容

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

基盤技術である STRAIGHT の基礎アルゴリズムの本質的な進歩(TANDEM-STRAIGHT)と新たに開発した高速な音源情報抽出法[8]を活用することにより、インタラクティブに歌唱デザインの操作を行うための一連のアルゴリズムを、実時間処理に適した形で整備することに成功した。本年度は、これらの成果と、様々なレベルでの歌唱デザインの操作性(directability)の検証を行うためのテストベッドとして開発したインタフェースおよびプロトタイプシステムに基づく知見を統合して、インタラクティブに歌唱デザインの様々な属性を操作できる歌唱インタフェースの開発を進めた。また、これらの開発を支えるため、研究インフラとしての TANDEM-STRAIGHT に基づくモーフィングアルゴリズムの拡張と要素アルゴリズムの最適化を進めることにより、プログラム基盤の整備を行った。また、歌唱の重要な要素である演奏表現の操作のため、データベースの拡張を進めた。

具体的には、演奏表現として用いられる多様な発声法による特異な声質・歌い回しを表現できる音源情報表現を TANDEM-STRAIGHT を基盤技術として開発した。また、片寄グループ、後藤グループおよび奥乃グループとの共同研究により、時間軸変形と基本周波数軌跡の変形による歌い直し転写を行うための特徴の分析方法と合成方法の開発とインタフェース実装を執り行った。



TANDEM-STRAIGHT を基盤とした音源情報表現

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

本グループは、数理的な手法を独自に開拓して行くことにより、既存作曲事例および演奏事例中のデザインを抽出し、かつそれを自動作曲や自動演奏に転写する技術の開発を中心課題として研究を行っている。任意の日本語歌詞の韻律に基づく自動作曲、楽曲の自動ジャンル分類、能動的な音楽鑑賞のためのイコライザを3つの柱として以下のような研究を進めた。

・歌詞の韻律に基づいた歌唱曲の自動作曲と歌唱演奏合成のシステム

日本語歌詞入力に基づく自動作曲システム Orpheus において、楽曲の長さの制約をなくし、打楽器パートを付加してリズムスタイルの指定できるように、また、作曲理論をより厳密に反映したアルゴリズムを開発した。

・作曲理論に基づく楽曲の和声の自動解析技術開発、和声ラベル付与データの作成

動的計画法を用い、非和声音を含めた自動和声解析を行うアルゴリズムを開発した。また、専門家の協力を得て和声解析の正解ラベルを付与した楽曲データを作成した。これを、音楽モデルの機械学習と和声認識に用いた。

・調波音と打楽器音のブラインド分離技術の応用研究

調波音と打楽器音のブラインド分離技術をより一般化した理論を構築し、CD 等におさめされたポピュラー音楽音源から、歌声のみを抽出するあるいは歌声のみを消去する技術を開発した。また、打楽器パターンを抽出し、それを音声認識におけるパターン学習のアルゴリズムと組み合わせることで楽曲のリズムパターン要素を抽出する画手法を開発した。この技術を応用し、楽曲ジャンルやムードの自動分類処理を実現した。

- ・キーボード演奏への自動伴奏システムの開発

キーボード演奏への自動伴奏技術の追従性と精度を高め、完成度を向上し、演奏者の音楽的な表現を自動伴奏にも転写できるシステムを開発した。具体的には、演奏者のテンポと打鍵強度と打鍵タイミングが、伴奏にも反映されるようにした。企業とともに製品化の検討を行った（製品化は未決定）。

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。

研究実施項目・概要

- ・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）

商用音楽CDに対して、ユーザが音楽中の楽器音を自由に制御できる「能動的音楽鑑賞」のための技術開発の高機能化を行い、編集機能付きのオーディオプレイヤーの高性能化と類似楽曲検索に展開した。具体的には、既開発の楽器音の調波・非調波混合モデルに inharmonicity を導入して、より現実に近い音響的な揺らぎを有する楽器演奏に対応した[1]の分析を通じて、「楽器音の音量の組み合わせにより、同一楽曲であってもジャンルが変化する」という仮説の検証を行い、提案手法の有効性を確認した。さらに、残響も含めて、調波・非調波・残響という3成分から構成されるモデル化を行い、より精緻な音楽音響信号分析法への道筋を明らかにできた。

- ・新しい楽器音演奏法を提供する音パレット

音高・音長に加えて、音色の操作を開発し、多重奏演奏中の特定楽器を別の楽器で置き換えるシステムを開発した。

- ・ベース音高を組み合わせた和音進行認識技術

音楽情報検索の重要な要素技術である和音進行認識技術を、メロディの一部を担っているベース音高と組み合わせることにより、昨年度開発したシステムの高性能化を行った。

- ・実時間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）

楽譜情報を利用して商用音楽CDのビート追跡の高性能化という課題に対して2種類のカルマンフィルタを用いる手法を提案した。また、ロボットの姿勢が歌声に反映される歌声モーフィング技

術を開発した。

・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)

演奏表現として用いられる多様な発声法による特異な声質・歌い回しを表現できる音源情報表現をTANDEM-STRAIGHTを応用して、音パレットの音高・音長・演奏法の変換技法を確立するとともに、音楽信号の新しいモーフィング技術を開発した。

§ 4. 成果発表等

(4-1) 原著論文発表

● 論文詳細情報

1. Katsutoshi Itoyama, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno: Parameter Estimation for Harmonic and Inharmonic Models by Using Timbre Feature Distributions *IPSJ Journal*, Vol.50, No.7 (Jul. 2009) 1757-1767, IPSJ. *Journal of Information Processing*, Vol.17 (2009) 191-201, IPSJ.
2. Osamu Fujimura, Kiyoshi Honda, Hideki Kawahara, Yasuyuki Konparu and Masanori Morise, Noh Voice Quality, *J. Logopedics Phoniatrics Vocology*, Vol.34(4), pp.157-170, 2009. DOI: 10.1080/14015430903002288
3. Hiromasa Fujihara, Masataka Goto, Tetsuro Kitahara Hiroshi G. Okuno: A Modeling of Singing Voice Robust to Accompaniment Sounds and Its Application to Singer Identification and Vocal-Timbre-Similarity-Based Music Information Retrieval, *IEEE Transactions on Audio, Speech and Language Processing*, Vol.18, No.3 (March 2010) 638-648, IEEE. doi: 10.1109/TASL.2010.2041386
4. Tetsuro Kitahara: "Mid-level Representations of Musical Audio Signals for Music Information Retrieval", *Advances in Music Information Retrieval*, Springer, Volume 274, 2010, pp. 65-91, 2010.3. doi:10.1007/978-3-642-11674-2_4
5. 北原 鉄朗、戸谷 直之、徳網 亮輔、片寄 晴弘: BayesianBand: ユーザとシステムが相互に予測し合うジャムセッションシステム, *情報処理学会論文誌*, Vol.50, No.12, pp.2949-2953 (2009.12)
6. 平野 砂峰旅, 赤山 仁, 井上 信太, 片寄 晴弘: メディアパフォーマンスシステム "Phenakistoscope Player" とその作品, *情報処理学会論文誌*, Vol.50, No.12, pp.2954-2963 (2009.12)
7. 八田原慎悟, 藤井叙人, 古屋晋一, 風井浩志, 片寄晴弘: テレビゲーム熟達者の脳活動に関するケーススタディ, *情報処理学会論文誌*, Vol.50, No.12, pp.2782-2795(2009.12)
8. 森勢 将雅、河原 英紀、西浦 信敬: 基本波検出に基づく高 SNR の音声を対象とした高速な F0 推定法、*電子情報通信学会論文誌 D*, vol.J93-D, no.2, pp.109-117, Feb. 2010.

(4-2) 知財出願

- ① 平成 21 年度特許出願件数(国内 3件)
- ② CREST 研究期間累積件数(国内 17件)