

「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

本プロジェクトでは、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、音楽に代表される時系列メディアのハンドリング技術の開発を進めている。平成 19 年度から 20 年度にかけては、プロジェクトタイトルにも謳っている「デザイン転写」を体現する音楽システムの実例の提示、そして、当該領域全般の研究環境の整備を中心課題としてプロジェクトを運営している。

本年度の具体例な成果として、Web 上のアーティストを表す言葉を手がかりした音楽検索推薦インタフェース、楽音（音響信号）中のベース音成分に着目した音楽検索推薦インタフェース、和音性を利用したムードビジュアライザ、歌詞入力に対してメロディ（歌声）合成を行うシステム等を開発した。また、基礎研究として、歌声の個人性知覚に寄与する音響特徴の検証、楽譜情報を利用して混合音中の調波成分と非調波成分を分離するモデルの構築を進めた。環境整備に関連した事項として、音楽情報科学研究用共通 Java クラスライブラリ CMX Library、演奏表情データベース CrestMuse PEDB ver1.0 を一般公開した。

平成20年度は、これらの成果をデモビデオでの形で集積し、アウトリーチ活動を実施するとともに、各グループ成果を結集するものとして、「歌声デザインシステム」の具体例を示す予定である。

### 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

認知に関連した研究項目として、fMRI を用いた実験環境を立ち上げ、ミニマルミュージック聴取および記憶に関する予備実験を執り行った。これにより、能動的音楽鑑賞の評価研究に向け、fNIRS、fMRI、脳波の3手法併用による脳機能計測環境が整った。また、「和音性の評価モデル」に利用し、音楽(MIDI)のムードをリアルタイムで表示するムードビジュアライザを実装した。

システム関連では、音響信号から抽出したベースパート情報を利用して楽曲検索推薦を行うインタフェース、ベイジアンネットワークを用いた予測型伴奏システム、ハーモナイゼーションシステムの各プロトタイプを実装した。また、事例活用型デザインインタフェースの構成に関して Directability 視点の基軸を打ち出し、表情付けシステムやサウンドインスタレーションの形で具体化した。

環境整備に関連した事項として、MusicXML をベースとした音楽の記述形式の標準化とその標準 API の開発を進め、Java クラスライブラリ CMX Library として一般公開した。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 ver1.0 を一般公開した。

アウトリーチ活動に関連し、代表者が中心となって活動してきたデジタルメディア領域全般での事例を利用したデザイン支援研究をプロモートする取り組みが、5年連続で人工知能学会全国大会審査付きオーガナイズドセッションとして採択され、満了となった。

### 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、音楽家でないアマチュア(一般のエンドユーザ)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」を研究開発した。その具体的な成果としては、Web を用いて自動ラベリングしたアーティストを表す言葉を手がかりに、能動的にアーティストの推薦を受けられるインタフェース「MusicSun」があげられる。音楽理解技術によって音楽の聴き方を豊かにする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、二件の招待講演(IEEE 国際会議、アメリカ音響学会大会)等によって高めた。

歌詞や歌声に着目した研究開発も進め、歌声の個人性知覚に寄与する音響特徴を明らかにし

たり、歌声の類似度に基づく音楽情報検索を実現したり、混合音中の歌声と歌詞との時間的対応付け手法を改良したりした。他にも、ドラム音の類似度の計算モデルの研究[3]や、奥乃グループと共同で、ユーザが混合音中の楽器音の音量や定位を変化させることができる能動的音楽鑑賞インタフェース[5]、音楽推薦手法[4]等に取り組んだ。

## 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

昨年度の研究により明らかになった「歌唱音声の歌手の個人性の判断において声質が支配的な重要性を有している [6] 」という知見の適用範囲を検証するとともに、歌唱デザイン転写システムとしてまとめる際に不可欠な歌い回しの転写のための基礎技術の構築を進めた。さらに、演奏表現の転写の研究のための基盤となるデータベースの構築に向け、正弦波モデルおよびイベントに基づく表現を統合できる基盤(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と合成)、テキストチャッピングに基づく音声変換の枠組み(歌唱デザイン転写システムの開発)を構築した。これらの基盤系の成果を利用し、片寄グループとの共同研究で時間軸方向での歌唱モーフィングインタフェースのプロトタイプを実装した。関連して、複数の歌手による多様な演奏表現の歌唱データベースの構築を進めた。なお、本課題の基盤として用いている音声分析変換合成技術 STRAIGHT のアルゴリズムに根本的な発明があり、速度・精度・拡張性が大きく向上したことが特記される。この発明の成果は、平成 20 年度の計画に生かされる予定である。

## 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

音楽デザイン自動学習・自動編曲に関連した事項として、和声学の知識を数理的に定式化し、既存の楽曲から確率パラメータを推定することで自動編曲により転写する技術、および、歌詞の韻律に基づいて歌唱曲を自動作曲する技術を開発した。

能動的音楽鑑賞に関連するものとしては、合奏音楽信号から特定の楽器音を除去し、即席カラオケを作成する技術を開発した。打楽器音と調波音(メロディー楽器など)が混合する録音から、そ

これらの混合割合を変化させられる画期的な技術を開発した。

従来から提案していた調波・時間・音色クラスタリングに対して、音色モデルを組み入れ、調波・時間・音色クラスタリング(Harmonic-Temporal-Timbre-structured Clustering; HTTC)として発展させることで、より高精度の音響-MIDI 変換技術を確立した。この考え方をベースとして、和音を構成する個々の音符の発音時刻、消音時刻までも精度良く算出する技術を開発した。この技術は片寄グループで実施している演奏表情データベース制作の支援技術として利用される。

## 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

音楽音響信号の分析・理解のための基盤技術として、平成 18 年度に開発した技術の高度化に取り組むとともに、ハイブリッド型音楽推薦システムの技術開発にも取り組んだ。

基盤技術として、具体的には、後藤 G と共同で、多重奏楽器音認識のために調波・非調波統合モデルのパラメータ推定による音源分離技術を開発した[5]。また、ベース音高を考慮したポピュラー音楽に対する和音進行認識技術を開発した。

応用研究として、上記調波・非調波統合モデルをベースとして、楽譜情報を利用することで、Pops 中の音源分離を実施し、楽器パート毎の音量の制御が可能な能動的音楽鑑賞インタフェースを開発した。また、音楽信号処理に基づいたコンテンツの解析と、ユーザ評価による協調フィルタリングの統合によるハイブリッド型音楽推薦技術を開発した。新エンタテインメントに関連する事項としては、実時間ビートラッキングを実施し、その結果に併せて発声するロボットを製作した。

## 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. 橋田 光代, 野池 賢二, 長田 典子, 片寄 晴弘: 音楽の演奏表現に基づくグループ構造の聴取に関する一検討, 日本感性工学会論文誌, Vol. 7, No. 2, pp. 327-336, 2007.
2. Jonathan Le Roux, Hirokazu Kameoka, Nobutaka Ono, Alain de Cheveigné, Shigeki Sagayama: Single and Multiple F0 Contour Estimation Through Parametric Spectrogram Modeling of Speech in Noisy Environments, *IEEE Trans. on Audio, Speech and Language Processing*, Vol. 15, No.4, pp.1135-1145, May 2007.
3. Elias Pampalk, Perfecto Herrera, and Masataka Goto: Computational Models of Similarity for Drum Samples, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol.16, No.2, pp.408-423, February 2008.
4. Kazuyoshi Yoshii, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol.16, No.2, pp.435-447, February 2008.
5. 糸山 克寿, 後藤 真孝, 駒谷 和範, 尾形 哲也, 奥乃 博: “楽譜情報を援用した多重奏音楽音響信号の音源分離と調波・非調波統合モデルの制約付パラメータ推定の同時実現”, 情報処理学会論文誌, Vol.49, No.3, pp.1465-1479, March 2008.

6. 河原 英紀, 生駒 太一, 森勢 将雅, 高橋 徹, 豊田 健一, 片寄 晴弘: “モーフィングに基づく歌唱デザインインタフェースの提案と初期検討”, 情報処理学会論文誌, Vol.48, No.12, pp.3637-3648, 2007.
7. 森勢 将雅, 高橋 徹, 河原 英紀, 入野 俊夫: “窓関数による分析時刻の影響を受けにくい周期信号のパワースペクトル推定法”, 電子情報通信学会論文誌, Vol.J90-D, No.12, pp.3265-3267, 2007. (研究速報/レター

○解説論文

8. 奥乃 博, 北原 鉄朗, 吉井 和佳: 楽曲の特徴量抽出と検索技術, 電気学会誌, 特集「音響機器は進歩している 6」, Vol.127, No.7 pp.417-420, July 2007.
9. 河原 英紀: “Vocoderのもう一つの可能性を探る—音声分析変換合成システム STRAIGHT の背景と展開—”, 日本音響学会誌, Vol.63, No.8, pp.442-449, 2007.

(2) 特許出願

平成 19 年度 国内特許出願件数:9件 (CREST 研究期間累積件数:16件)