

## 研究課題別評価

### 1. 研究課題名 :リアルタイム音楽情景記述システムの構築

### 2. 研究者氏名 :後藤 真孝

### 3. 研究の狙い :

本研究は、音楽音響信号を人間のように理解できる計算機システムを工学的に実現することをねらっている。人間は音楽を聞いて、メロディーを口ずさんだり、サビ(曲中で盛り上がる部分)がわかったり同じフレーズが繰り返されていることに気付いたりできる。しかし、こうした音楽理解能力を備えた計算機システムの構築は、メディア検索や音楽編集等の様々な用途で望まれているにも関わらず、困難な課題であるために従来実現されていない状況であった。本研究は、その状況を打破すべくコンパクトディスク(CD)等に収録されている実世界の複雑な音響信号を入力として、リアルタイムに音楽理解が可能なシステムを実現することを目的としている。

### 4. 研究成果 :

本研究では、複数種類の楽器音や歌声を含む音楽音響信号を入力とし、それに対応した記述を理解結果として出力するリアルタイム音楽情景記述システムを構築した。音楽情景記述とは、音楽演奏中の刻一刻と変化する情景を分析・理解した結果を記述する処理過程である。ここで問題となるのは、「何ができれば音楽を理解したといえるのか」という問題設定である。本研究では、この問いに答えること自体も研究対象と考え、具体的な提案をした上で実際に実現した点が重要である。具体的には、この問いに対する現段階での答えとして、メロディー、ベース、繰り返し区間、サビ区間の4種類の音楽的要素の記述を得る音楽情景記述システムの実現に取り組んだ。これらの音楽的要素は、いずれも楽曲の情景を表す本質的な要素であるが、従来は推定が困難だと考えられていたものであり、その実現は重要な課題である。

以下では、本研究の主要な四つの成果である

1. リアルタイム音楽情景記述システムの実現
2. 応用システム「サビ出し機能付き音楽試聴機」の実現
3. RWC 研究用音楽データベースの構築
4. 音楽情景記述エディタの開発と記述のラベリングの実施

の概要を述べる。

#### 4.1. リアルタイム音楽情景記述システムの実現

メロディーとベースの推定手法 PreFEst (Predominant-F0 Estimation Method) とサビ区間と繰り返し区間の推定手法 RefraiD (Refrain Detecting Method) を提案し、それに基づいて音楽情景記述システムを構築した。本システムは、複数の楽器音が混在したモノラルの音楽音響信号に対し、4種類の音楽的要素の記述をリアルタイムに得ることができる。

メロディーとベースの推定手法 PreFEst

PreFEst は、帯域制限された混合音中で最も優勢な音高を推定する手法である。メロディーは中高域において最も優勢な高調波構造を持ち、ベースは低域において最も優勢な高調波構

造を持つことが多いため、それぞれに適した帯域に制限して適用すれば、メロディーとベースの音高を推定できる。本手法が対象とするCD等による実世界の音響信号は、事前に音源数を仮定することが不可能な混合音であり、ここでは、周波数成分が頻繁に重複する上に、基本周波数成分がないような音も存在する。しかし、従来の音高推定手法の多くは、少数の音源数を仮定し、周波数成分を局所的に追跡したり、基本周波数成分の存在に依存したりしていたために、そのような実世界の混合音には適用できなかった。それに対して本手法は、音源数を仮定せず、周波数成分の局所的な追跡もおこなわず、基本周波数成分の存在を前提としないという特長を持つ。

#### サビ区間と繰り返し区間の推定手法 RefraiD

RefraiD は、ポピュラー音楽の音響信号に対して、楽曲中に出現するすべてのサビの区間を網羅的に検出する手法である。従来、楽曲の音響信号中に何度も出現するサビのどこか一箇所を、指定した長さだけ切り出して提示する研究はあったが、サビ区間の開始点と終了点はわからず、サビの転調も扱えなかった。本手法は、音楽 CD 等による実世界の複雑な混合音に対して、各サビの区間の開始点と終了点の一覧を求めることができるだけでなく、転調を伴うサビを検出することも可能である。さらに、楽曲全体の中での様々な区間の繰り返しに基づいてサビを検出するため、その中間結果として、繰り返し区間の一覧も得ることができる。

#### 4.2. 応用システム「サビ出し機能付き音楽試聴機」の実現

サビ区間と繰り返し区間の推定手法 RefraiD を応用し、試聴に適した新たな音楽再生インターフェース SmartMusicKIOSK を実現した。CD 販売店の店頭で音楽を短時間試聴する際には、通常の音楽鑑賞における受動的な聴き方と異なり、試聴者は早送りを何度も繰り返しながらサビを探ることが多い。しかし、こうした聴き方に対する支援は従来なかった。本研究では、サビの区間や楽曲中で繰り返される区間の先頭へジャンプする機能と、それらの区間の楽曲中での配置を視覚化する機能を提供する。これにより、試聴者が手探りでサビを見つける煩わしい作業を不要にし、試聴者が能動的に聴きたい場所を探す作業を容易にする。このような、インタラクティブに楽曲中の再生位置を変更しながら所望の箇所を見つけられるインターフェースは、試聴に限らず、音楽を選んで利用する一般的な目的で有用である。従来の音楽再生インターフェースでは、楽曲単位でしか興味のない音楽を飛ばせなかったのに対し、SmartMusicKIOSK によって初めて、楽曲内部の興味のない箇所も容易に飛ばすことが可能になったといえる。

#### 4.3. RWC 研究用音楽データベースの構築

RWC (Real World Computing) 知的資源 WG (ワーキンググループ) 内に、RWC 音楽データベースサブWG という組織を主査として立ち上げ、その活動を通じて RWC 研究用音楽データベースを構築した。本データベースは、共通利用の自由、学術利用の自由が確保された世界初の楽曲 DB であり、音楽情報処理研究全般に広く貢献する成果である。音楽情報処理の研究分野全体の発展のためには、世界中の研究者が自由に利用可能な音楽データベースが必要であるが、従来、共通楽曲データベースや大規模な楽器音データベースは存在していなかった。

そこで、RWC 研究用音楽データベースとして、「ポピュラー音楽データベース」(100 曲)、著作権切れ音楽データベース」(15 曲)、クラシック音楽データベース」(50 曲)、ジャズ音楽データベース」(50 曲)、音楽ジャンルデータベース」(100 曲)、楽器音データベース」(50 楽器、29.1

Gbytes)の六つを構築した。上記の全 315 曲は、すべて新規に演奏、歌唱、録音した。この内、著作財産権の保護期間が終了した 96 曲以外の 219 曲は、新規に作曲、作詞、編曲した。さらに全 315 曲に対し、音響信号、標準 MIDI ファイル、歌詞のテキストファイルを用意した。一方、50 楽器に対しては、音域全体を半音間隔で収録した単独演奏音を用意した。これにより、研究者は、着想(問題発見)、問題解決、実装、評価、発表の各段階で、著作権等の制約を受けない楽曲を使用することが可能となった。

#### 4.4. 音楽情景記述エディタの開発と記述のラベリングの実施

楽曲の音響信号に対して音楽的要素の記述のラベリングを施すことができる音楽情景記述エディタを開発した。RWC 研究用音楽データベースを活用して音楽情景記述システムを構築したりシステムの出力を評価したりするためには、本エディタによる音楽的要素の記述のラベリングが重要となる。実際にラベリング作業も実施し、前述の RWC 研究用音楽データベースのポピュラー音楽データベース 100 曲に対し、メロディーの音高とサビ区間のラベリング作業を完了している。

#### 5. 自己評価：

さきがけ研究 21 「情報と知」領域 研究員として研究を開始したことで、予算的な裏付けを持ちながら様々な研究展開が可能となり、充実した 3 年間で多様な成果を上げられた。

- 1) 音楽的要素の記述の推定手法については、局所的な構造としてメロディーとベース、大局的な構造として繰り返し区間、サビ区間の 4 種類の推定に成功した。特に、繰り返し区間とサビ区間の推定に関しては、研究課題提案当初は階層的なビート構造を利用しながら実現する計画であったが、そうした局所的な記述を使わずにこれを実現できたこと、つまり、大局的な構造を音響信号から直接推定することが達成できたことは、自分自身にとっても大きな驚きであった。なお、以下に述べる応用システムやデータベースのような新たな研究の展開にも取り組むために、研究課題提案当初よりも音楽的要素の対象を絞ったものとなったが、最大限の研究成果を出す上で、正しい選択であったと考えている。今後は、既の実現した手法の推定精度をより向上させていくと共に、新たな音楽的要素の記述の推定にも取り組んでいきたい。
- 2) 応用システムとして実現した「サビ出し機能付き音楽試聴機」は、学会発表や各種デモンストレーション等で、すぐに実用化されそうな完成度を持ち、実際に使って便利なシステムだと高い評価を受けた。本成果は、情報処理学会シンポジウムインタラクション 2003 でベストペーパー賞を受賞し、採択率 4.64 倍の国際会議(ACM UIST2003)に採択され、日本経済新聞朝刊(2003 年 8 月 1 日発行)でも紹介された。本研究の重要性を一般に広くアピールする上で意義のある成果であった。
- 3) 本研究は、研究課題提案当初の見込みを大きく上回る RWC 研究用音楽データベースの構築という成果に結び付いた。これは、さきがけ研究 21 の採択が発端の一つとなって、RWC 音楽データベースサブ WG を発足させて主査として活動することが可能となったことによる。音楽情報処理の研究分野に広く貢献する成果として国内外から高く評価されており、今後の研究分野全体の発展のための基礎が築けた。
- 4) 音楽情景記述エディタの開発と記述のラベリングの実施は、地道な作業であり、通常は光が当たらない部分であるが、着実な研究の進展のために多くの時間を割いてきた。これらは今

後様々な研究を展開する上で土台となると考えている。

本研究の成果が、今後の音響信号を対象とした音楽理解研究の発展に大きく寄与していくことを願っている。

#### 6. 研究総括の見解：

本研究は、音楽の理解に対して、従来なかった新しく本質的なアプローチを提案しただけでなく、それを世界初の手法として実現した点が高く評価できる。また、音楽理解の応用例として、過去数十年間変化のなかった音楽再生インタフェースの新たな形態を提案し、産業界にもインパクトのある成果を上げた。さらに、研究用音楽データベースの構築といった研究分野全体としての土台作りにも貢献している。

#### 7. 主な論文等：

##### 論文誌

1. 後藤 真孝: SmartMusicKIOSK: サビ出し機能付き音楽試聴機, 情報処理学会論文誌, Vol.44, No.11, pp.2737-2747, 2003. (推薦論文)
2. Masataka Goto: A Real-time Music Scene Description System: Predominant-F0 Estimation for Detecting Melody and Bass Lines in Real-world Audio Signals, Speech Communication (ISCA Journal), 2004. (accepted)
3. 浜中 雅俊, 後藤 真孝, 麻生 英樹, 大津 展之: 発音時刻の楽譜上の位置を確率モデルにより推定するクオンタイズ手法, 情報処理学会論文誌, Vol.43, No.2, pp.234-244, 2002.
4. 北原 鉄朗, 後藤 真孝, 奥乃 博: 音高による音色変化に着目した楽器音の音源同定: F0 依存多次元正規分布に基づく識別手法, 情報処理学会論文誌, Vol.44, No.10, pp.2448-2458, 2003.

##### 査読付き会議

1. Masataka Goto: A Predominant-F0 Estimation Method for CD Recordings: MAP Estimation using EM Algorithm for Adaptive Tone Models, Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001), pp.V-3365-3368, 2001.
2. Masataka Goto, A Predominant-F0 Estimation Method for Real-world Musical Audio Signals: MAP Estimation for Incorporating Prior Knowledge about F0s and Tone Models, Proceedings of CRAC-2001 (Workshop on Consistent & Reliable Acoustic Cues for Sound Analysis), 2001.
3. Masataka Goto, Music Scene Description: Toward Audio-based Real-time Music Understanding, The Journal of the Acoustical Society of America, Vol.111, No.5, Pt.2, p.2349, 2002. (Invited Paper of The 143rd Meeting of the Acoustical Society of America)
4. Masataka Goto, Hiroki Hashiguchi, Takuichi Nishimura, and Ryuichi Oka: RWC Music Database: Popular, Classical, and Jazz Music Databases, Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR 2002), pp.287-288, 2002.
5. 後藤 真孝: SmartMusicKIOSK: サビ出し機能付き音楽試聴機, 情報処理学会 インタラクシヨ

ン 2003 論文集, pp.9-16, 2003.

6. Masataka Goto: A Chorus-Section Detecting Method for Musical Audio Signals, Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2003), pp.V-437-440, 2003.
7. Masataka Goto: A Chorus-Section Detecting Method for Musical Audio Signals, Proceedings of the 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'03), 2003.
8. Masataka Goto: Music Scene Description Project: Toward Audio-based Real-time Music Understanding, Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003), pp.231-232, 2003.
9. Masataka Goto, Hiroki Hashiguchi, Takuichi Nishimura, and Ryuichi Oka: RWC Music Database: Music Genre Database and Musical Instrument Sound Database, Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003), pp.229-230, 2003.
10. Masataka Goto: SmartMusicKIOSK: Music Listening Station with Chorus-Search Function, Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology (UIST 2003), pp.31-40, 2003.

#### 招待講演

1. Masataka Goto, Music Scene Description: Toward Audio-based Real-time Music Understanding, The 143rd Meeting of the Acoustical Society of America, 2002 年 6 月 4 日.
2. 後藤 真孝: J-POP 革命のための音楽情報処理, DBWeb2002 (データベースとWeb情報システムに関するシンポジウム), 2002 年 12 月 3 日.

#### 受賞

1. 日本音楽知覚認知学会 研究選奨 受賞, 2002 年 11 月.
2. 情報処理学会シンポジウム インタラクション 2003 ベストペーパー賞 受賞, 2003 年 2 月.