

# 音楽情報処理が実世界と結び付く： あなたも使える音楽情報処理

## Linking Music Information Processing with the Real World: Music Information Processing for End Users

後藤 真孝

Masataka Goto

産業技術総合研究所

National Institute of Advanced Industrial Science and Technology (AIST)

あらまし 実世界と結び付いた音楽情報処理技術には様々なものがあるが、本招待講演では、音楽家でないエンドユーザが直接恩恵を受けるような新しい技術を中心に、研究事例を紹介する。近年、音楽情報検索が急速に注目を集め、既に、携帯電話に向かって鼻歌を歌うだけで音楽を入手できる技術や、街中で流れている音楽に携帯電話を向けるだけで曲名がわかる技術が実用化されて、エンドユーザが利用可能となっている。そして、こうした音楽情報検索を支える基盤技術として、音楽の音響信号を人間のように理解できる計算機システムを実現する重要性がますます高まってきた。さらに、そうした音楽音響信号理解研究の新しい応用例として、従来の受動的な鑑賞とは違う、能動的な音楽鑑賞を可能にするエンドユーザ向け音楽インタフェースを実現する研究アプローチも提案されている。エンドユーザのための音楽情報処理は、これから大きく発展していくことが期待される。

### 1 はじめに

一万曲を携帯型音楽プレーヤに入れて、いつでもどこでも手軽に持ち歩けるようになった今日、音楽家でないエンドユーザが直接恩恵を受けるような新しい音楽情報処理技術が次々と登場している<sup>1),2)</sup> <sup>1</sup>。その背景として、計算機やネットワークの性能向上と共に、1992年に規格化されたMP3 (MPEG Audio Layer-3) が1990年代後半から普及し、2000年代前半にはMP3以外にもOgg Vorbis, MPEG-4 AAC, WMA (Windows Media Audio) 等の圧縮方式が提案され、誰もが容易に音楽を計算機や携帯型機器上で扱えるようになってきたことが挙げられる。携帯電話で着メロや着うたで音楽を楽しむことは日常化し、インターネット経由で音楽配信をするサービスも続々と出現している。ダウンロード型の音楽配信サービスに加え、聴き放題の定額制音楽配信サービスも登場し、好きな曲を好きなときに

好きな場所で好きなだけ聴ける時代が到来しつつある。

そのため、インターネットとWWWの普及に伴ってテキストに対する情報検索技術(WWWサーチエンジン等)が不可欠になったのと同様に、音楽においても情報検索技術の重要性が高まっている。例えば、以下のような「欲しいけれど曲名がわからなくて聴けない」問題が考えられる。

- メロディーは思い出せるけれど、曲名がわからなくて聴けない。
- 街中やテレビ、ラジオで何気なく聴いて気に入った曲があるのに、曲名がわからなくて聴けない。

さらに、定額制音楽配信サービスが普及してくると、以下のように「自分の好みの曲がどれかわからなくて聴けない」問題も起きると考えられる。

- ある好みの曲があるときに、それに似た他の曲を聴きたくても、どれかわからなくて聴けない。

<sup>1</sup> デモンストレーションを交えた招待講演をする関係上、本稿では筆者の研究事例を中心に紹介する。より広い範囲の文献紹介は、文献1),2)等の解説記事を参照されたい。

## 2 音楽情報検索

このような問題を解決するために、従来の曲名やアーティスト名に基づくテキストベースの音楽情報検索に加えて、主に以下の三種類の検索技術が研究されてきた。

### 1. ハミング検索 (QBH: Query by Humming)

聞いたことのある曲を「ラララー」等のように口ずさむと、その曲名を検索できる方法である。メロディーの歌唱やハミングを検索キーとして、そのメロディーを持つ楽曲を検索する。研究段階<sup>3)</sup>を経て、携帯電話に向かって鼻歌を歌うだけで音楽を入手できるサービスは既に国内で実用化され、エンドユーザが利用できるようになった。

### 2. 断片を含む楽曲の検索

街中で流れている音楽の曲名を知りたいときに、携帯電話でその一部分を録音すると、曲名を検索できる方法である。楽曲の断片を検索キーとして、その断片を含む楽曲を検索する。研究段階<sup>4),5)</sup>を経て、こうした携帯電話を向けるだけで曲名がわかるサービスも、既に国内外で実用化されている。

### 3. 楽曲間の類似度に基づく検索

ある曲が気に入っているときに、それに似た曲調の曲を探ることができる方法である。楽曲を検索キーとして、それに類似した楽曲を検索する。そのためには、楽曲間の類似度が必要だが、そうした類似度は検索以外にも重要であり、類似度に基づく楽曲の自動分類(音楽ジャンル、曲調の分類)等も研究されている。こうした楽曲間の類似度を利用した検索や分類技術も、既に実用化され始めている。

他にも、エンドユーザ向けの技術として、協調フィルタリング(collaborative filtering)や内容に基づくフィルタリング(content-based filtering)を活用した音楽推薦手法も研究されている<sup>1),2)</sup>。

## 3 音楽音響信号理解

こうした音楽情報検索を支える基盤技術として、音楽音響信号理解に関する研究の重要性が高まっている。楽曲同士が似ていることを適切に類似度に反映したり、楽曲の音響的特徴や内

容に基づく適切な楽曲推薦やプレイリスト作成をしたりするには、音楽CDなどによる実世界の複雑な音響信号を理解する必要があるからである。音楽音響信号理解は、単旋律から、単一楽器の多重音、複数種類の楽器による混合音へと、対象の複雑さを徐々に増しながら進歩し、近年の主流は、音楽CDなどの楽曲を対象とした研究に移ってきている。こうした動向は、人工知能の分野でかつて指摘されていた「単純な実験室環境を対象としたプロトタイプシステムの多くは、実際には実世界の環境を扱えるように拡張するのが極めて難しい」というスケールアップ問題とも関連している。実際に、MIDI音源等による単純な合成音を対象として考案された音楽理解手法は、音楽CDを扱おうとすると単なる改良では対処できず、新たな手法を考案しなければならないことが多い。

音楽情報検索の観点からも、最初から音楽家のような深く詳細な音楽理解を実現しようとして音楽CDが扱えないよりは、音楽CDから楽曲の内容を反映した何らかの有用な情報が得られることが求められている。そこで、例えばパワースペクトルの局所的な特徴量(セントロイド、ロールオフ、フラックス等)の楽曲全体に渡る平均と分散や、音高・ビートに関連した周期性を反映した特徴量等が抽出され、活用されてきた。しかし、これらは音楽理解というよりは、曲調を反映した特徴抽出に位置付けられる。よりの確に音楽演奏中の情景を分析・理解した結果を記述する研究アプローチとして、例えば「音楽情景記述」<sup>6),7)</sup>が提案されている。これは、従来のように音楽音響信号から個々の音符や個々の構成音の分離信号を得ることは目指さずに、音楽的に訓練されていない「しろうと」の音楽理解の実現を目指す研究アプローチである。既に、音楽CDからビート、メロディー(図1)、ベース(図1)、サビ、フレーズの繰り返し等の記述を得る手法が実現されている。

## 4 エンドユーザ向け音楽インタフェース

音楽音響信号理解研究の進歩の恩恵をエンドユーザが受けられるようにするためには、日常的な場面で使える音楽インタフェースの研究が

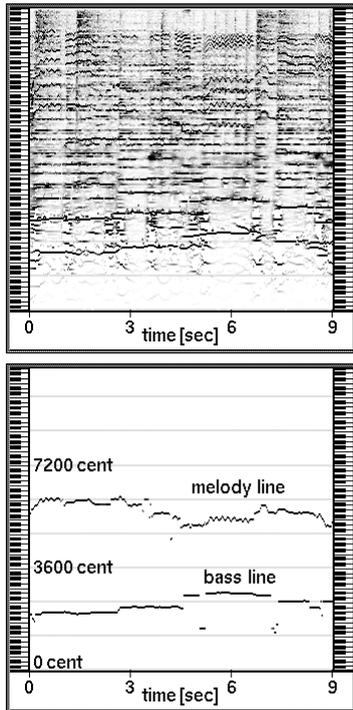


図 1: 音楽音響信号を対象としたメロディーとベースの音高推定手法 PreFEst<sup>7)</sup> の画面表示例: 上側のウィンドウはドラムを伴うポピュラー音楽を入力したときの周波数成分で、下側のウィンドウは PreFEst によってリアルタイムに推定されたメロディーラインとベースラインの音高 (F0) を示す。

重要となる。例えば、音楽家でない一般のエンドユーザは、音楽を創作するのではなく、受動的に音楽を鑑賞することが多い。そこで、従来の受動的な鑑賞とは違う、能動的な音楽鑑賞を可能にするエンドユーザ向け音楽インタフェースを実現する研究アプローチ<sup>8),9)</sup>が提唱された。本稿ではこれを「能動的音楽鑑賞インタフェース」(Active Music Listening Interface)と名付ける。これは「音楽音響信号理解研究が音楽の聴き方をどのように豊かにできるか」を追求する研究アプローチである。こうした能動的鑑賞において重要なのは、ユーザが意図を容易に指定でき、かつ、その意図が適切に鑑賞体験に反映されることである。既に「サビ出し機能付き音楽試聴機 SmartMusicKIOSK」<sup>8),9)</sup>(図 2)を筆頭に「楽曲を流してくっつけて並べることのできる新たな音楽再生インタフェース Musiccream」<sup>10),11)</sup>(図 3)、「ドラムイコライズシステム INTER:D」<sup>12)</sup>等の音楽音響信号理解に基づく具体的なインタフェースも提案されている。

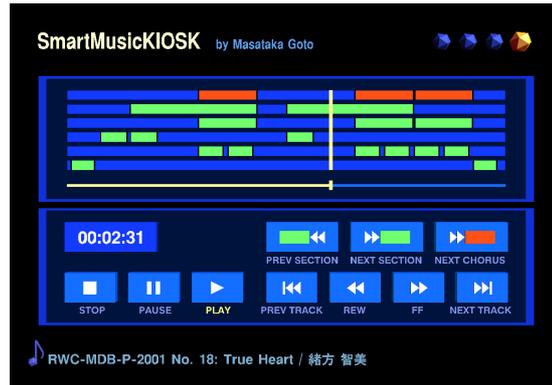


図 2: サビ区間検出手法に基づく音楽再生インタフェース SmartMusicKIOSK<sup>8),9)</sup> の画面表示例 (RWC 研究用音楽データベース<sup>13)</sup> RWC-MDB-P-2001 No. 18 に対する自動サビ区間検出結果): 上側のウィンドウは楽曲内容の視覚化結果 (音楽地図) を表す。横軸は時間軸で楽曲全体を表示しており、最上段がサビ区間の一覧、その下の 5 段が繰り返し構造、その下の横棒が再生位置スライダーを表す。下側のウィンドウは再生操作ボタン群を表し、通常の CD プレーヤー相当のボタンに加え、サビ区間の頭出しをする「サビ出しボタン」等を含む。ユーザはボタンひと押しでサビを聴くことができ、音楽地図を見ながら能動的に聴きたい場所を探ることができる。

## 5 おわりに

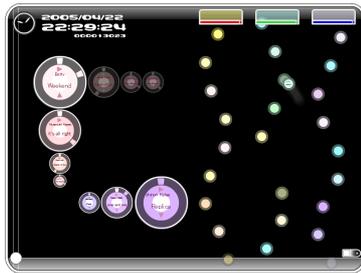
実世界と結び付いた音楽情報処理の研究分野では、欧米を中心に次々と大型プロジェクトが生まれ、多くの研究者が参入して研究が活性化している。関連する研究成果が発表される国際会議も、ISMIR, ICASSP, WASPAA, ICMC, ISMA, ICMPCC, ICA, WEDELMUSIC (AXMEDIS), ICMAL, DAFx, NIME, ICEC, ICME, IJCAI, ACM Multimedia 等の多岐に渡り、特に 2000 年代に入ってから増えてきた。国内でも、情報処理学会 音楽情報科学研究会、日本音響学会 音楽音響研究会、日本ソフトウェア科学会 インタラクティブシステムとソフトウェアに関するワークショップ (WISS)、情報処理学会シンポジウム インタラクション、情報科学技術フォーラム (FIT) 等の会議で、活発に成果が発表されている。基礎研究として魅力的であり、かつ、応用研究としてもエンドユーザに従来とは異なる音楽体験を可能にする研究が、次々と登場してくることを期待したい。

## 参考文献

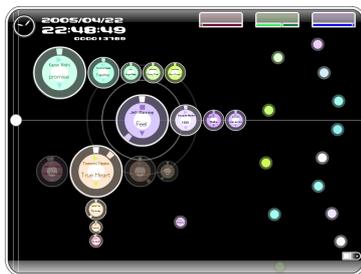
- [1] 後藤真孝, 平田圭二: 解説 “音楽情報処理の最近の研究”, 日本音響学会誌, Vol. 60, No. 11, pp. 675–681 (2004).
- [2] Goto, M. and Hirata, K.: Invited review “Recent studies on music information processing”, *Acoustical Science and Technology (edited by the Acoustical Society of Japan)*, Vol. 25, No. 6, pp. 419–425 (2004).
- [3] Sonoda, T., Ikenaga, T., Shimizu, K. and Muraoka, Y.: The Design Method of a Melody Retrieval System on Parallelized Computers, *Proc. of WEDELMUSIC 2002*, pp. 66–73 (2002).
- [4] Kashino, K., Kurozumi, T. and Murase, H.: A Quick Search Method for Audio and Video Signals Based on Histogram Pruning, *IEEE Trans. on Multimedia*, Vol. 5, No. 3, pp. 348–357 (2003).
- [5] Wang, A.: An Industrial-Strength Audio Search Algorithm, *Proc. of ISMIR 2003*, pp. 7–13 (2003).
- [6] 後藤真孝: リアルタイム音楽情景記述システム: 全体構想と音高推定手法の拡張, 情処研報 音楽情報科学 2000-MUS-37-2, pp. 9–16 (2000).
- [7] Goto, M.: A Real-time Music Scene Description System: Predominant-F0 Estimation for Detecting Melody and Bass Lines in Real-world Audio Signals, *Speech Communication*, Vol. 43, No. 4, pp. 311–329 (2004).
- [8] 後藤真孝: SmartMusicKIOSK: サビ出し機能付き音楽試聴機, 情処学論, Vol. 44, No. 11, pp. 2737–2747 (2003).
- [9] Goto, M.: SmartMusicKIOSK: Music Listening Station with Chorus-Search Function, *Proc. of UIST 2003*, pp. 31–40 (2003).
- [10] 後藤孝行, 後藤真孝: Musicream: 楽曲を流してくっつけて並べることでできる新たな音楽再生インタフェース, WISS 2004 論文集, pp. 53–58 (2004).
- [11] Goto, M. and Goto, T.: Musicream: New Music Playback Interface for Streaming, Sticking, Sorting, and Recalling Musical Pieces, *Proc. of ISMIR 2005* (2005). (accepted, in press).
- [12] 吉井和佳, 後藤真孝, 奥乃博: 実世界の音楽音響信号に対するドラムスの音源同定手法を利用したドラムイコライズシステム INTER:D の開発, 第3回 情報科学技術フォーラム (FIT 2004) 情報技術レターズ LG-005, pp. 143–144 (2004).
- [13] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一: RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース, 情処学論, Vol. 45, No. 3, pp. 728–738 (2004).



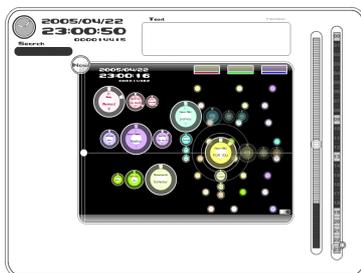
(1)「流し楽曲」機能 流しそめんや回転寿司のように、画面上を次々と楽曲群(円盤)が上から下へ流れ、ユーザは楽曲を何気なく手に取ることで、未知の楽曲に能動的に出会うことができる。



(2)「類似くっつき」機能 流れから取り出した楽曲を持って、流れている他の楽曲に接触させることにより、雰囲気似た楽曲だけを選択的に次々とくっつけることができる(制御可能な類似楽曲検索に相当する)。くっつけて作った楽曲群(円盤の連なり)は、再生順を示すプレイリストになる。



(3)「メタプレイリスト」機能 全プレイリストを画面の上から下へ順番に再生することで、画面全体をさらにメタなプレイリストと捉え直す。プレイリスト(楽曲群)を画面上で自由に並べて曲順を検討できる。



(4)「タイムマシン」機能 Musicream上のすべての操作は記録され、タイムマシンに乗って音楽を聴いている過去の自分に戻るように、時間軸を自在に行き来できる。

図 3: 未知の楽曲と能動的に出会うことができる総合的な音楽再生インタフェース Musicream<sup>10),11)</sup>の画面表示例: Musicreamでは、図中の四つの機能をユーザがシームレスに操作でき、聴き放題の月額制音楽配信サービスで特に有用と考えられる。