音楽情報処理最前線！

未来の音楽の楽しみ方、作り方はどう変わるのか？
コンピュータは音楽を理解できるようになるのか？
コンピュータを使って音楽を研究する「音楽情報処理」
という研究分野が、世界的に注目を集めています。
本連載では、そうした最先端の研究事例を紹介していきます。

楽器名で音楽が探せる　「楽器認識技術」が叶える音楽の新しい聴き方・探し方

「ピアノ曲が聴きたい」、「ブラスセクションの入った曲が欲しい」等々、
楽器を指定して音楽を聴きたいと思ったことはないでしょうか。
今回はこんなときに便利な、検索技術を紹介します。

1. サックス入りの楽曲を探せ！

例えば、あなたがいま友人と曲作りに励んでいるとしよう。あなたと友人は、ある程度カタチになってきた自作曲を聞き返している。すると、2回目のブリッジ部で、友人はあなたにこんなことを言った。

「ここにサックスのソロを入れてみようよ」

しかし、あなたはサックスらしい旋律なんて知らない…。

「仕方ない、サックスの旋律をいろいろ聴いてみて勉強するしかないか」

そう思い、CDをあさってみたが、参考になりそうな楽曲はそれほど多くなさそうだ。

こんなとき、楽曲のデータベースからサックス入りの曲だけを探し出し、さらに曲中でサックスが演奏されている区間まで表示してくれたらどんなに便利だろう。今回ご紹介するのは、こんなことを実現してくれる技術なのである。

2.「内容に基づく音楽情報検索」とは？

音楽検索と言えば、曲名やアーティスト名で探すのが一般的で、他には歌詞やメロディのデータベースを検索するサービスくらいしかない。複数の楽器音が混在する波形データから情報を抽出するのはやはり難しいのだ。しかし人間はCDを聴くだけで「この音はピアノだ」などと判断できる。つまりプログラムさえ書ければコンピュータでもできるはず！　こう考えた研究者は世界中にいた。波形データから情報を取り出して検索する研究は「内容に基づく音楽情報検索」と呼ばれ、実はすでに10年近い歴史があるのだ。

3. 波形から楽器を認識させるには？

音楽CDの波形から、コンピュータにどうやって楽器を認識させるのか。図1を見てほしい。まず楽器ごとに見本となる音(学習データ)をたくさん用意する必要がある。次に用意した多数のサンプルから、各楽器の音の特徴を数値化(特徴抽出)するのだ。そして、楽曲が入力されたら、同じように特徴抽出を行ない、特徴量の傾向が「学習データのうち、どの楽器に近いか」で楽器名を判定すればよい。

実際の特徴抽出は「調波構造」というものを頼りに行なう。楽器音には基本となる周波数成分の他に、周波数が2倍や3倍などの周波数成分(倍音成分)も一緒に含まれている。例えばクラリネットでは偶数次倍音(2倍、4倍…)が奇数次倍音(3倍、5倍…)より小さいという特徴がある。こういった性質を特徴抽出に用いるわけだ。

しかし、多くの楽曲では異なる楽器が同時にいくつも鳴っていて、そもそもこの調波構造の発見自体がなかなかうまくいかない。そこで我々が考えたのは、調波構造を発見して特徴抽出を行なうのではなく、「調波構造があらゆるところにあると仮定して、しらみつぶしに特徴抽出を行なう」という方法である。

ある平面を考えてみよう。この平面を「Instrogram」と呼ぶ。横軸には時刻の目盛りが、縦軸には周波数の目盛りが付いている。

この平面は色の濃い部分と薄い部分があり、色の濃い部分は、その時刻、その周波数にピアノの音が存在する確率が高いことを表す。この平面は楽器の種類数だけ用意されている(図2)。

ピアノの音が存在する確率は、「何かの音が存在する確率」と「何かの音が存在するならそれがピアノの音である確率」の積で表

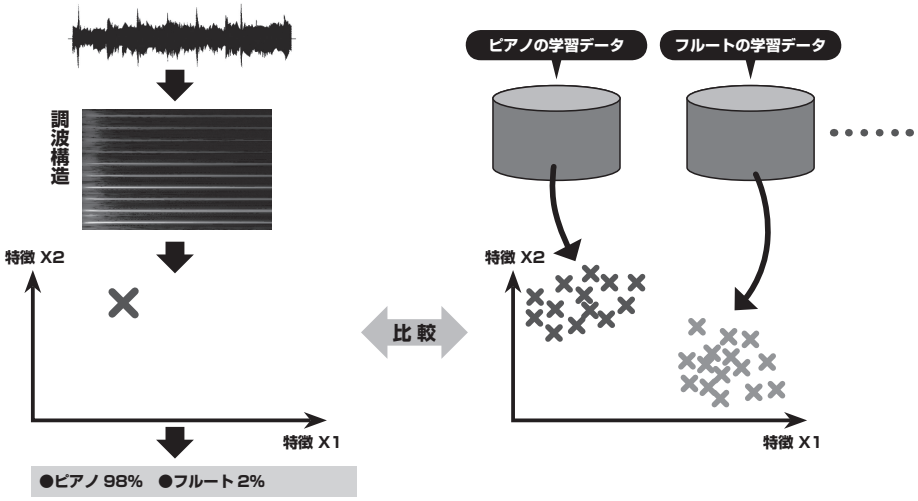

図1　楽器認識の基本的な考え方。あらかじめ用意した楽器名が既知のデータと、楽器を認識したい音波形の両方から特徴抽出(特徴の数値化)を行なって、どの楽器の学習データに特徴が近いかから楽器名を判定する。

北原 鉄朗
（きたはらてつろう）

2002年東京理科大学理工学部卒業。
2007年京都大学大学院情報学研究科博士後期課程修了。
博士（情報学）。
日本学術振興会特別研究員（DC2）を経て、現在、科学技術振興機構 CrestMuse プロジェクト研究員。人はどうして音楽を聴き、楽しみ、創ることができるのか、機械にこれらをさせる技術の研究開発を通じてこの謎に迫る。
http://ist.ksc.kwansei.ac.jp/~kitahara/

「音楽情報科学研究会」へ参加してみませんか？

情報処理学会 音楽情報科学研究会（SIGMUS）は、コンピュータと音楽とが関わり合うあらゆる場面を活動対象とする学際的研究会で、年5回の研究発表会を開催しています。研究会に会員登録すると、参加できなかった研究発表会の論文集の郵送、過去の全研究発表会の論文のダウンロードなどの特典があります。研究会の登録方法や研究発表会の開催に関する最新情報などは http://www.ipsj.or.jp/sigmus/ をご覧ください。

すことができる。前者は信号処理技術を、後者は本節冒頭の方法を使えば求めることができる。

ここで重要なのは、本節冒頭の方法では「調波構造がないところ」から特徴抽出を行なうとでたらめな結果が返ってくるのだが、今回の考え方なら、本来「調波構造がないところ」は「何かの音が存在する確率」が0になるので掛け算すれば結果はちゃんと0になるという点だ。だから安心して（本当に調波構造があるかどうかを気にせずに）、あらゆる時刻のあらゆる周波数に調波構造があると仮定して、処理を進めることができるわけだ。

4. アプリケーション

この楽器認識技術を使って作ったアプリケーションをいくつか紹介しよう。

●楽器構成の可視化（図3）

「Instrogram」の表示を見ることで、どの楽器がいつからいつまで演奏しているか一目で見ることができ、ある楽器の存在確率が濃くなりだしたところをクリックすれば、その楽器が演奏を始める部分から聴くこともできる。

その他、楽曲再生に同期して動く棒グラフやアニメーションで、いま演奏している楽器を知ることもできる。

●楽器名を検索キーにした楽曲検索

指定された楽器名の存在確率ができるだけ「1」に近く、それ以外の楽器の存在確率が「0」に近い楽曲を探してくる検索方式。

●類似楽曲検索

ユーザがある楽曲（WAVファイルなど）を指定すると、データベース中の楽曲と同様に「Instrogram」の計算を行ない、各楽器の存在確率ができるだけ似通った楽曲を探してくる。

楽器構成は、その曲の雰囲気に密接に関係している。楽器構成を比べることで、雰囲気の似た楽曲を探すことが可能だ。

●プレイリスト生成機能付き音楽プレイヤー（図4）

類似楽曲検索をさらに1歩進めたもので、たとえば10曲聴きたいと思ったら、最初と最後とあと何曲かを指定したら、雰囲気（楽器構成）が滑らかに変わるように残りの楽曲を埋めてくれる。

最初にブラス系の曲を聴いて最後はピアノソロ曲でしっとりと、なんてことが可能だ。

5. 楽器認識技術は人間に追いつけるのか

以上、楽器認識とそのアプリケーションについて述べてきたが、現時点での楽器認識精度はまだまだ人間には追いついていない。

3～4種類の楽器の混合であれば何とかなるが、10～20の楽器が同時に鳴ったら完全にお手上げという状態だ。やはり人間の耳はすごいのであって、コンピュータがそれに追いつけるのはもう少し先の話になりそうだ。

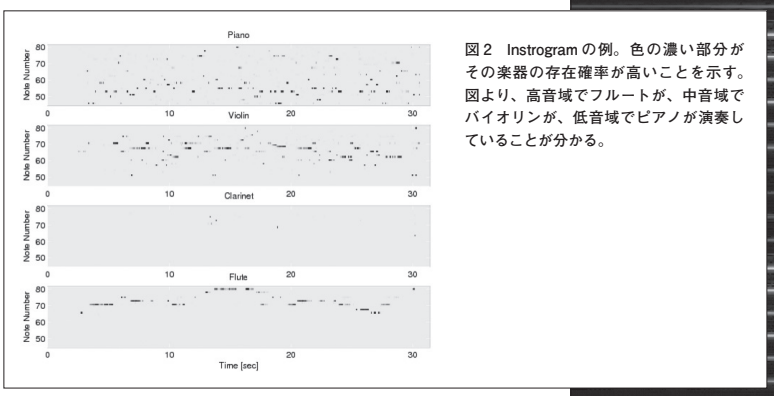

図2　Instrogramの例。色の濃い部分がその楽器の存在確率が高いことを示す。図より、高音域でフルートが、中音域でバイオリンが、低音域でピアノが演奏していることが分かる。

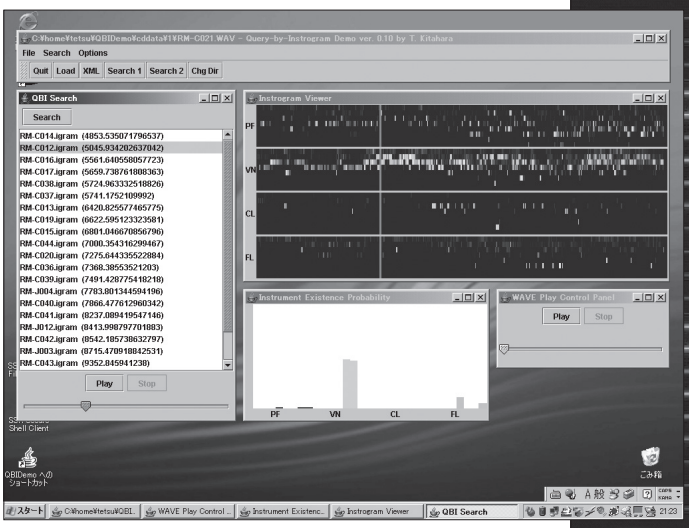

図3　Instrogramを表示しながら再生する音楽プレイヤー。再生中の楽器構成がInstrogramと棒グラフの両方で表示されている。左のウィンドウには類似楽曲検索の検索結果が表示されている。

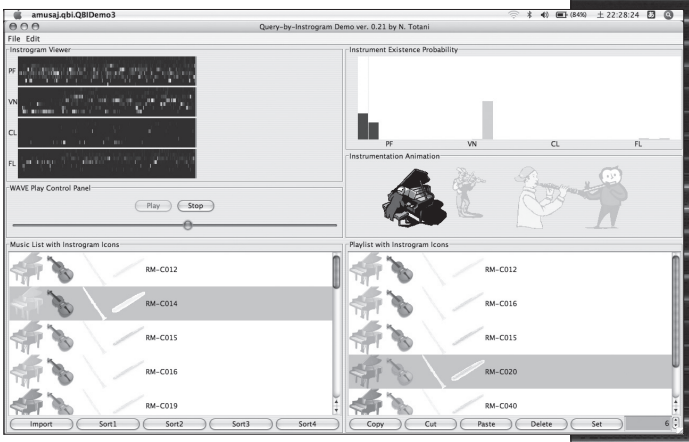

図4　プレイリスト生成機能つき音楽プレイヤー。左下の楽曲リストから右下のプレイリストにドラッグ＆ドロップして最後に「Set」ボタンを押すとプレイリストの空き項目に楽曲を埋める。各楽器の存在確率が楽器アイコンの色の濃さに反映されており、聴かなくても楽器構成を知ることができる。