

和音進行を用いた音楽ハイライト検出の一手法と、組み込みシステムへの応用 A Method for Detecting Highlights of Music with Chord-Proceedings and Its Application on an Embedded System

荻山真一†
Shinichi Gayama

1. まえがき

大容量HDDの普及に伴い、カーナビゲーションを始めとした組み込みシステムにおいても、多くの音楽を蓄積し、楽しむことができるようになった。そうした背景をもとに、使い勝手を向上させるための様々な検索手段や再生方法が提案されている。中でも音楽の印象的な部分、すなわちハイライトを検出して音楽を再生する手法は、多くの音楽を短時間で試聴することを可能とし、検索時間の短縮のみならず、新たな音楽の楽しみ方を提供するものとして期待されている。音楽ハイライト検出に係わる研究としては、後藤によるサビ区間検出手法[1][2]が新しく、サビを含む音楽の繰り返し構造を分析し、任意の部分を再生する手法およびシステムを発表している。

一方、組み込みシステムへの応用を考えた場合、他の機能との兼ね合いが重要となるため、演算速度とメモリ量の制約が極めて厳しくなる。そのため、従来手法では十分な性能を確保することが困難であった。

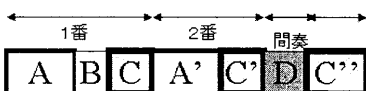
本研究の目的は、音楽の和音進行を抽出し、組み込みシステムにおいて実用可能なハイライト検出手法を開発することである。和音進行は、音楽の基本的な構造を成すことで知られている。また、和音変化の差分値列は、旋律の有無や編曲の違いに対して感度が低く、かつ類似する曲調を表現するため、楽曲構造を検出する目的においては効率的な演算が可能である。

そこで本稿では、著者の和音進行抽出手法[3]を用いた音楽のハイライト検出手法を提案し、アルゴリズムの概要を説明する。また、本手法を実装した組み込みシステムのパフォーマンスと、ポップス音楽を対象としたハイライト検出性能について報告する。

2. 音楽ハイライト検出

2.1 概要

日本で流行しているポップスやロック音楽の多くは、複数のコーラスの組み合わせで構成されている。図1に一例を示す。A、A' (以下 A 群) や C、C'、C'' (以下 C 群) では、同様の曲調が繰り返されている。後藤[1]が指



- ・最も繰り返し回数が多い部分
→ 作曲者が聴かせたい部分
→ 特徴部分(ハイライト)
- ・2回繰り返し → 通常フレーズ
- ・繰り返さない部分 → 前奏、間奏、最後

摘している通り、最も繰り返される C 群がいわゆるサビの集合である可能性が高い。本稿においても同様の発想で、最多繰り返し区間 C 群を検出し、その一つをハイライトとして捉える。実際には、

図1 繰り返しを持つ音楽構造の一例

C 群の中でも旋律の動きや伴奏、リズムが異なっている場合が多く、単純に楽音信号やスペクトルの比較を行っても繰り返しの検出はできない。しかしながら、人間が C 群を同様なフレーズとして認識できるのは、各部分の曲調が類似しているからであり、その共通の曲調を表現できれば、人間の理解に近い形で C 群を検出することができる。

そこで著者は、共通の曲調を表す音楽的要素として和音進行に注目した。本章では、和音進行を用いて音楽ハイライトを検出するアルゴリズムについて説明する。

まず、最初に音楽の楽音信号から和音を抽出する。次に一曲全体の和音進行と、所定の位置から切り出した部分和音進行の相関演算によって、曲調の類似する最多繰り返し区間の検出を行う。そして最後に、繰り返し区間の中からハイライト位置を選択する処理を行う。以下、各々について説明する。

2.2 楽音信号からの和音進行抽出

PCM形式の楽音信号を解析し、和音と信号パワーを抽出した後、補正処理によって和音進行を生成する[3]。和音進行は、音楽の速さに対して比較的緩やかに変化する傾向がある。本処理によって抽出した和音進行のデータ量も非常に少なく、次に述べる和音進行の相関演算など、各種演算処理の負荷を軽減する効果が期待できる。

2.3 最多繰り返し区間の検出

図2に示すように、一曲全体の和音進行の先頭から 8~13 和音長の部分和音進行を切り出し、全体の和音進行との相関演算を行う。以降、切り出し位置を時間方向にずらしながら同様の処理を継続する。演算結果の一例を図3に示す。対角線上の縞以外に出現する縞が、曲調の類似する繰り返し区間を示しており、対角線を含む出現回数が繰り返し回数を意味する。

なお、和音進行の相関演算においては、絶対和音の比較ではなく、前後の和音間に生ずる度数の差分値列と属性列(メジャー、マイナー)すなわち相対的な和音の動きを用いる(以降、和音差分値列とする)。また、両者の和音進行の相関距離として、各和音変化点における和音差分値列の距離和を用いるが、同一変化点の距離を演算すると同時に、1~2 つ先の変化点との距離も計算する。そして、現時点の距離よりも先の和音差分値との距離が近い場合には比較位置を先に移動し、それ以降は新たな比較位置を基準に、互いの和音差分値列の距離計算を継続する。

以上の工夫によって、部分的なテンポの変化や転調している繰り返し区間も容易に検出することができる。

2.4 ハイライト位置の選択

類似する曲調の最多繰り返し区間に対して、各区間の

† パイオニア(株) 研究開発本部総合研究所

信号パワーの平均値を求め、それが最も大きい区間の開始位置をハイライト位置として選択する。本来、どの区間もハイライトの候補として考えることができるが、多くの音楽を連続試聴した際の楽しさを演出するために、このような指針を導入した。

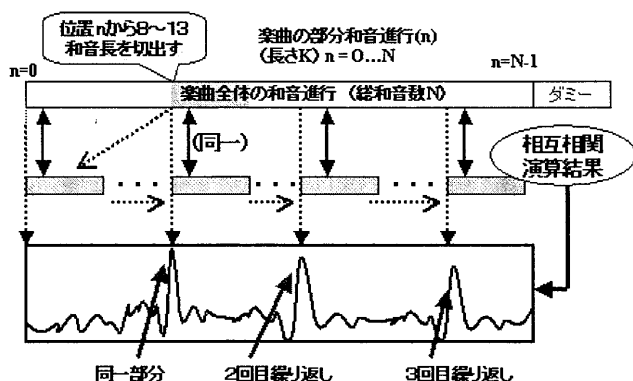


図2 部分と音進行を用いた類似曲調区間検出

3. 組み込みシステムへの実装

3.1 リアルタイム処理

楽音信号からの和音進行抽出は、音楽の録音と同時にされる。まず、入力楽音信号に対してモノラル変換とダウンサンプリング処理を実行する。次に、離散フーリエ変換によって得られた平均律音程のパワーを用いて、時系列の和音候補群と信号パワーを出力する[3]。なお、各和音候補は3つの音程で構成される3和音で表現されている[3]。

3.2 音楽単位の後処理

各音楽が録音された後、リアルタイム処理によって得られた時系列の和音候補群に対する補正処理[3]を行い、和音進行を生成する。その後、和音進行の変化点のみの情報を記憶し、2.3および2.4に記述したハイライト検出処理を実行する。

4 性能評価

4.1 パフォーマンス

3.1で記述したリアルタイム処理を、16ビット固定小数点DSPに実装した。その結果、サンプリング周波数44.1kHzの入力PCM信号のモノラル変換とダウンサ

ンプリング処理に約3 MIPS、和音進行検出および信号パワー算出に約1.1 MIPSで計約4.1 MIPSを要した。所要コードサイズは約9kバイト、ワークメモリは約17kバイトである。

また3.2に記述した、和音進行の補正処理およびハイライト位置の検出処理は、動作周波数400 MHzの組み込みCPUで処理を行った。その結果、単独動作において3分の音楽で平均150 msec、6分の音楽で平均400 msecの処理時間を要した。

DSPとCPUの性能も飛躍的に向上しており、本結果は十分実用レベルであると言える。ただし、実際のシステムにおいて他の優先度の高い機能が動作している場合には、上記所要時間が3~5倍になることも十分に考えられる。そのため、システム全体を安定に動作させるための設計が重要である。

4.2 ハイライト検出性能

ハイライト検出の評価楽曲には340曲のポップス音楽を用いた。本研究におけるハイライト検出の目的は、操作手段の限られた組み込みシステムにおいて、多くの音楽を簡単に試聴することである。そのためには、録音した各音楽のハイライトを15~20秒程度再生できることが望ましい。したがって評価方法としては、ハイライト検出位置から5秒以内でサビを再生できた場合を正当とし、その曲数をカウントした。結果として308曲は良好な結果が得られ、約9割の正当率を確保した。

検出できなかった音楽の特徴として、サビ以上に多く繰り返されるフレーズが存在する、あるいは明確な繰返し構造をもっていないことが挙げられる。また、元々ハーモニー性に乏しい楽曲に対しては、和音進行の検出精度が悪化する。これらが、結果的にハイライト検出精度の低下に影響していると考えられる。

5 むすび

和音進行を用いた音楽のハイライト検出手法を提案し、組み込みシステムへの実装と評価を行った。組み込みシステムに適用する場合には、演算量や使用メモリの削減という課題があり、従来手法では性能を確保することが難しかったが、本手法を実装した結果、十分な動作速度と性能が得られた。なお、本手法は実際にカーナビゲーション製品に搭載されており、楽曲のハイライト再生に用いられている。

参考文献

- [1] 後藤真孝, “リアルタイム音楽情報記述システム: サビ区間検出手法,” 情報処理研究報告, 2002-MUS-47-6, Vol.2002, No.100, pp.27-34, October 2002
- [2] 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, November 2003
- [3] 我山真一, “類似楽曲検索を目的とした楽音からの和音進行抽出手法,” FIT2003, 一般講演論文集, Vol.2, p245-246, 2003

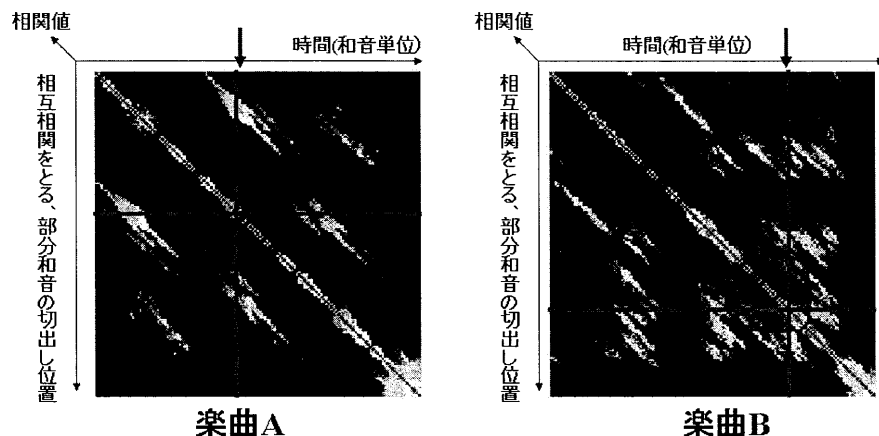


図3 実際の類似曲調演算結果と検出されたハイライト位置