

音符表現によらない旋律の木構造表現の予備検討

Preliminary Consideration of Non-notewise Tree-based Melody Representation

北原 鉄朗

Tetsuro Kitahara

日本大学文理学部

College of Humanities and Sciences, Nihon University

This paper describes a method for extracting a melodic contour using the discrete Wavelet transform. The advantage of the use of the discrete Wavelet transform is that we obtain the parameter set of a melodic contour as a binary tree. We make an attempt to extracting the repetition structure of the contour by detecting low-distance pairs in the subtrees of this binary tree.

1. はじめに

旋律とその背後にある構造を計算機上で表現する方法を検討することは、音楽情報科学における主要な研究課題の1つである。最も基本的で広く使われている旋律表現は、楽譜である。楽譜は、音高と音価からなる音符の系列で、計算機ができる前から音楽家同士の旋律の伝達手段として広く用いられている。旋律の構造については、よく用いられる表現形式に Generative Theory of Tonal Music (GTTM) [Lerdahl 83] がある。これは、旋律を二分木で表したものであるが、各ノードから伸びる2つの枝に優勢な枝かどうかを表す情報が載っており、優勢でない方の枝を削除することで旋律の簡約などが行える。

旋律の認知表現を考えたときに、音符列を出発点とし、その上に二分木を構築するアプローチは本当に妥当なのだろうか。人が旋律をどのように聴いているかはその人の音楽経験などによって様々であると予想されるが、個々の音符を個別に聴く方法の他に、旋律の全体的な「流れ」を聴く方法も考えられる。このような考え方から、個々の音符に関する情報をあえて隠し、旋律の全体的な「流れ」を表現する「旋律概形」を提案した [土屋 13, 北原 14]。旋律概形は、旋律の音高の時系列に対してフーリエ変換を行い、低次のフーリエ係数のみを残して逆フーリエ変換を行うことで得られる。しかし、旋律概形に含まれる階層的な構造（たとえば繰り返し構造）を木構造として表す方法については未検討であった。

本稿では、離散ウェーブレット変換を用いて旋律概形の抽出方法を再定義することで、旋律概形に含まれる階層的な構造の抽出を検討する。離散ウェーブレット変換は、再帰的に繰り返すことでパラメータを二分木として得ることができる。[北原 14] でも離散ウェーブレット変換を用いた旋律概形の抽出については簡単に報告したが、本稿ではそれに加え、二分木として得られるパラメータに対して距離を定義することで、旋律概形間の類似度計算を実現するとともに、部分二分木に対して類似度を計算することで繰り返し構造の検出を試みる。

2. 離散ウェーブレット変換を用いた旋律概形抽出

旋律概形は、旋律の全体の流れを曲線として表したものである。[土屋 13] では次の手順により得られるものと定義した。

連絡先: 北原 鉄朗, 日本大学文理学部, 東京都世田谷区桜上水 3-25-40, kitahara@chs.nihon-u.ac.jp

1. 旋律を音高の時系列（一定の時間分解能でその時々々の音高を表した系列。音高は基本周波数やノートナンバーなどどんな表記法でも構わない。音がない箇所は便宜的にその直前の音が続いているとみなす）として表す。
2. 音高の時系列を周期信号とみなし、フーリエ変換を行う。これにより、旋律の大まかな流れを表す低次のフーリエ係数と、旋律の細かな動きを表す高次のフーリエ係数に分離することができる。
3. 低次のフーリエ係数のみを残して逆フーリエ変換を行う。これにより、細かな動きの成分が削除された音高の時系列が得られる。

本稿では、旋律概形を離散ウェーブレット変換を用いて再定義する。離散ウェーブレット変換は、時系列 $\{x_1, \dots, x_N\}$ をスケーリング関数とウェーブレットを用いて近似する技術である。スケーリング関数は低域の、ウェーブレットは高域の成分しか持たないため、これを行うことで時系列を低域と高域に分離することができる。スケーリング関数で近似した際の係数を $\{c_{-1,1}, \dots, c_{-1,N/2}\}$ 、ウェーブレットで近似した際の係数を $\{d_{-1,1}, \dots, d_{-1,N/2}\}$ と表すとすると、前者に対して改めて離散ウェーブレット変換を行うことで、さらに低域成分 $\{c_{-2,1}, \dots, c_{-2,N/4}\}$ と高域成分 $\{d_{-2,1}, \dots, d_{-2,N/4}\}$ に分けることができる。これを再帰的に繰り返す（**多重解像度解析**と呼ばれる）ことで、最も低域の近似係数 $\{c_{-n,1}\}$ ($n = \log_2 N$) が長さ1の系列として、多段階の近似係数 $\{d_{-1,1}, \dots, d_{-1,N/2}; d_{-2,1}, \dots, d_{-2,N/4}; \dots; d_{-n,1}\}$ が二分木として得られる（元の系列の長さが2のべき乗の場合）。

このようにして得られたパラメータに対して逆離散ウェーブレット変換を行うと、元の時系列を得ることができる。二分木のある深さ以上のノードの値を0に書き換えて逆離散ウェーブレット変換を行うと、その深さに応じて分解能を落とした時系列を得ることができる。このようにすることで、旋律概形を得ることができる。ただし、どのような概形が得られるかは用いるマザーウェーブレットに依存する。たとえば、Daubechies 4 ウェーブレットを用いた場合は連続的な概形が得られるのに対し、Haar ウェーブレットの場合は、このウェーブレットが不連続な直線からなるため、概形も不連続な直線で構成される。本稿では、得られたパラメータが何を表すかが分かりやすいため、Haar ウェーブレットを用いる。他のマザーウェーブレットを用いた場合については今後の課題とする。



図 1: モーツァルト「ピアノソナタ K.331」第 1 楽章の冒頭部。この冒頭 8 小節の主旋律のみを使用。http://imslp.org/より。

3. 旋律概形に対する距離の定義と繰り返し構造の検出

旋律概形は $\{c_{-n,1}\}$ と $\{d_{i,j}\}$ の 2 種類のパラメータセットで表現されるが、前者は鉛直方向の位置を表し、旋律概形の形状は後者が表す。そこで、旋律概形に対する距離尺度を後者のみにより定義する。いま、ある 2 つの旋律概形を表すパラメータの二分木を $T_1 = \{d_{i,j}\}, T_2 = \{e_{i,j}\}$ とすると、これらの間の二乗距離 $\text{dist}^2(T_1, T_2)$ を次のように定義する：

$$\text{dist}^2(T_1, T_2) = \sum_{i=1}^n \left\{ \frac{1}{2^i} \sum_{j=1}^{2^i} (d_{-i,j} - e_{-i,j})^2 \right\}.$$

ここで係数 $1/2^i$ が付いているのは、木の深さによって要素数が異なるがその距離に対する影響度を等しくするためである*1。

この距離尺度は、1 つの二分木の部分木同士に対しても計算することができる。ある旋律概形を表す二分木の各部分木に対して、距離が一定の値以下になる部分木を同じ二分木から検出し、これらを同一視する。同一視された部分木は、片方を削除し、削除された部分木への辺をもう一方の部分木につなぎ直す。これを「部分木を共有する」と呼ぶ。この処理により、旋律の繰り返し構造を階層的に見つけることができる。

4. 実行例

上で述べた手法を実際の旋律に適用した例を示す。用いた旋律は、モーツァルト「ピアノソナタ K.331」第 1 楽章の冒頭 8 小節（図 1）である。この旋律は、前半 4 小節と後半 4 小節のそれぞれ前半分（1~2 小節目と 5~6 小節目）が共通になっており、その中にさらに繰り返し（1 小節目と 2 小節目、5 小節目と 6 小節目）がある形になっている。

この旋律に対して音高の時系列を求めたもの、さらに多重解像度解析により得られた二分木をそれぞれ図 2、図 3 に示す。この二分木のうち深さ 5（または 4）以上のノードの値を 0 に置き換えて逆ウェーブレット変換を行った結果を図 4 に示す。この例では Haar ウェーブレットを用いているので、連続的な曲線にはならず、旋律を 32 個（または 16 個）の区間に均等に分割し、各区内で平均を取ったものと等価になっている。また、各ノードの値は、その 1 つ上の深さまでで近似される旋律概形からの差分になっており、正の値であれば対応する区間において差分が右下がり、負の数であれば右上がりであることを表している。

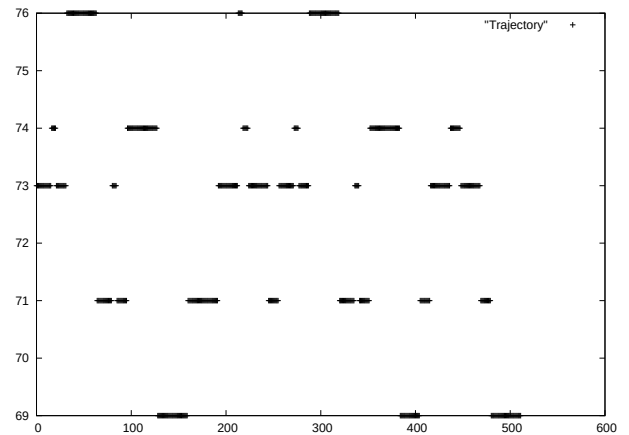
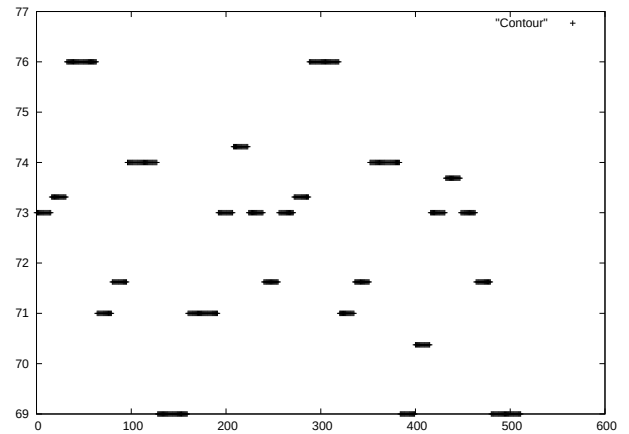
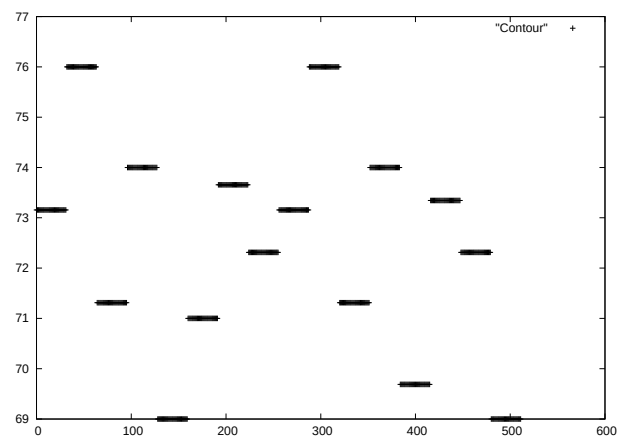


図 2: 図 1 の旋律から求めた音高の時系列



(a) しきい値が 5 の場合



(b) しきい値が 4 の場合

図 4: 深さがしきい値以上であるノードの値を 0 にして得られる旋律概形の例

*1 より深いノードの影響度をより小さくするため、ここの係数を $1/2^{\alpha i}$ とし、 α に 1 より大きい定数を与えてもよい。

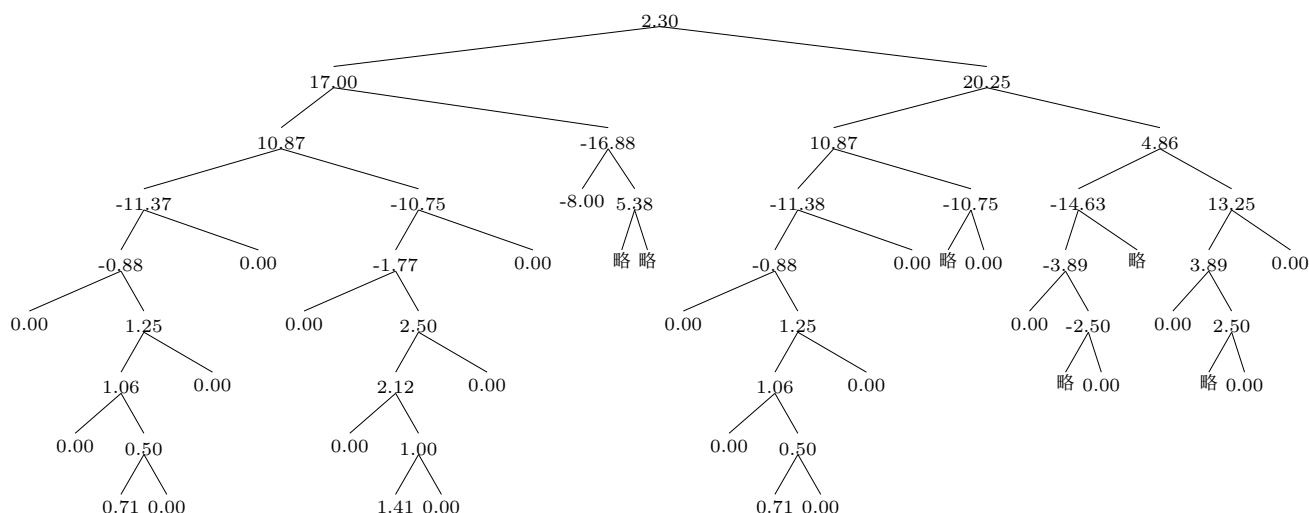
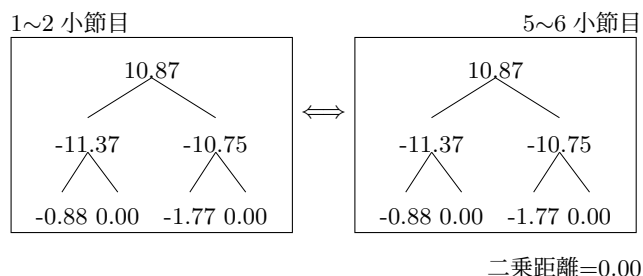
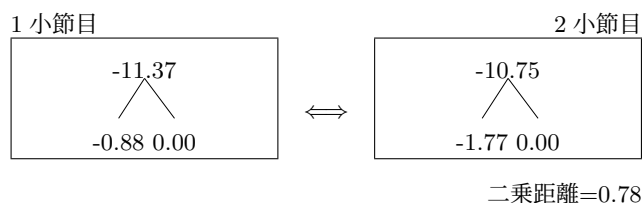


図 3: 旋律概形を表す二分木. 本来完全二分木であるが, 左右両方の子ノードの値が 0 のものは省略した.



(a) 1~2 小節目と 5~6 小節目に対応する部分木間の距離



(b) 1 小節目と 2 小節目に対応する部分木間の距離

図 5: 同一二分木内における部分木間の距離計算例

図 5 (a) は, 1~2 小節目に対応する部分木と 5~6 小節目に対応する部分木の距離, 図 5 (b) は, 1 小節目に対応する部分木と 2 小節目に対応する部分木の距離を表す. 距離の計算では, 深さ 5 以上のノードの値を 0 に置き換えた二分木を用いた. 1~2 小節目と 5~6 小節目は全く同じ旋律であり, 二乗距離の計算結果も 0 であった. 1 小節目と 2 小節目は旋律の形状はほとんど一緒であり, 二乗距離の計算結果は 0.78 となった. さらに, 3 章で述べた方法を用いて部分木を共有した結果を図 6 に示す. 部分木を同一視するかどうかを決める距離のしきい値は 2.0 とした. この結果より, 1~2 小節目と 5~6 小節目が共通で, その中にさらに繰り返しが含まれていることが分かる.

5. おわりに

本稿では, 離散ウェーブレット変換を用いて旋律概形を得る方法, および, その旋律概形に対して距離計算および繰り返し

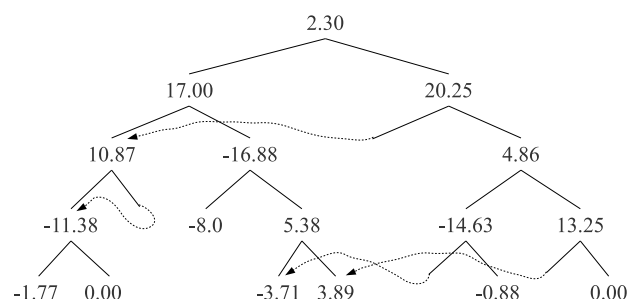


図 6: 距離がしきい値以下の部分木を共有した例. 部分木が削除され別の部分木につながれた箇所は, 破線の矢印で示している.

検出を行う方法を検討した. しかし, Haar ウェーブレットを用いたため, 旋律概形が本来の形であるためならぬ曲線になっておらず, 単に音高の時系列を等分割した各区内で平均を取ったものと等価なものになってしまっている. 今後は, Haar ウェーブレット以外のマザーウェーブレットでも試し, 適切な距離計算や繰り返し検出ができるかどうか確かめる必要がある. また, GTTM などの他の旋律簡約手法との関連, 得られる二分木の共通点や相違点なども考察していく必要がある.

謝辞 本研究に関して有益なご意見をくださった平田圭二氏, 東条 敏氏, 松原正樹氏に感謝する. 本研究の一部は, 科研費 26280089, 26240025 の支援を受けて行われた.

参考文献

- [Lerdahl 83] Lerdahl, F. and Jackendoff, R.: *A Generative Theory of Tonal Music*, MIT Press (1983)
- [土屋 13] 土屋 裕一, 北原 鉄朗: 音符を単位としない旋律編集のための旋律概形抽出手法, 情報処理学会論文誌, Vol. 54, No. 4, pp. 1302-1307 (2013)
- [北原 14] 北原 鉄朗, 土屋 裕一: 旋律を簡約・操作する一手法, 人工知能学会第 28 回全国大会, 1K4-OS-07a-3 (2014)