

音符表現によらない旋律の木構造表現の検討 (第2報)

Consideration of Non-notewise Tree-based Melody Representation: Part 2

北原 鉄朗

Tetsuro Kitahara

日本大学文理学部

College of Humanities and Sciences, Nihon University

This paper reports our attempts for establishing a methodology for representing hierarchical melody structure without using notewise representations. In our previous paper, we proposed a method for extracting an *melodic contour* using the discrete Wavelet transform with the Haar wavelet. In this paper, we apply this method to other wavelets and report an experiment with “*Ah, vous dirai-je, maman*” (W. A. Mozart).

1. はじめに

本研究の目的は、音楽、特に旋律の階層的な構造を、音符表現を用いずに記述する枠組みを確立することである。

旋律に階層的な構造があることはよく知られており、この階層構造を計算機が扱える形で記述する試みが行われている。たとえば、Generative Theory of Tonal Music (GTTM) [Lerdahl 83] は、ゲシュタルトの原理に基づき設計されたいくつものルールに基づいて、音符間の階層的な結びつきを二分木で表す理論である。GTTM は元々計算機上の実装は行われていなかったが、計算機上で実装する試みや様々な拡張の試みがなされている [平田 02, Hamanaka 07, 松原 14]。GTTM は音符表現を出発点とした理論であるため、これらの拡張も音符表現に基づいたものである。

我々、特に音楽の非専門家が複雑な旋律を聴く際、1つ1つの音符を個別に聴き取っているとは限らない。むしろ、旋律の全体的な流れを聴いている可能性も十分に考えられる。そのような考えから、我々は、個々の音符を用いずに、旋律の全体的な流れを直接記述する方法について検討している [北原 14, 北原 15]。前報 [北原 15] では、旋律概形 (旋律を1つの曲線で表したものを) を離散ウェーブレット変換 (DWT) による多重解像度解析を用いて抽出する方法を定式化した。多重解像度解析は、時系列を低域と高域に分ける処理を再帰的に繰り返す手法で、パラメータを二分木として取得することができる。二分木のある深さ以上のノードの値を0に書き換えてDWTを行うことで、時系列の高域成分が失われ、概形を得ることができる。しかし、結果の分かりやすさを重視するためにマザーウェーブレットとして Haar ウェーブレットを用いていたため、得られた概形は曲線ではなく不連続な直線となっていた。ウェーブレット変換を用いて旋律を分析する研究としては [Velarde 13] があるが、[北原 15] と同様に Haar ウェーブレットのみを用いており、他のマザーウェーブレットは検討していない。

本稿では、[北原 15] で提案した内容を Haar ウェーブレット以外のマザーウェーブレットに対しても試し、一部の変更のみで適用可能であることを示す。また、本手法をモーツァルト「きらきら星変奏曲 八長調 K.265」に適用し、旋律概形同士の距離が主観評価による旋律の類似度とどの程度一致するかを議論する。

2. DWT を用いた旋律概形抽出

2.1 抽出方法の概要 [北原 15]

与えられた旋律に対して音高の時系列を求め、これに対して DWT を適用する。DWT 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。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、これにより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 DWT 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 DWT 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 DWT 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。

2.2 Haar ウェーブレットを用いた場合

モーツァルト「ピアノソナタ K.331」第1楽章の冒頭8小節の主旋律 (図1) を用いて旋律概形抽出の例を示す。この旋律から求めた音高の時系列を求めると図2のようになり、それに対して Haar ウェーブレットを用いて旋律概形を求めると図3のようになる。音高の時系列を 2^n 個の区間に等分割し、区間内で平均を求めたものと等価なものが得られる。

2.3 Daubechies ウェーブレットを用いた場合

Daubechies 4 ウェーブレットを用いて同様のことを行った結果を図4に、Daubechies 8 ウェーブレットを用いて行った結果を図5に示す。どちらも不連続点のない旋律概形が得られたが、Daubechies 4 ウェーブレットの場合は、元のウェーブレットがなめらかな曲線ではないために、得られた旋律概形もなめらかな曲線ではなかったが、Daubechies 8 ウェーブレットではなめらかな旋律概形を得ることができた。



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 .

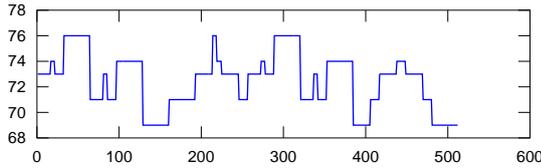


図 2: 図 1 の旋律から求めた音高の時系列

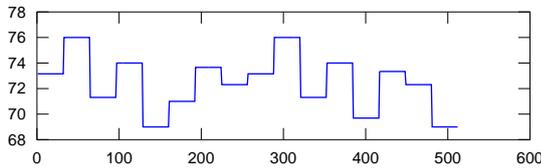


図 3: Haar ウェーブレットを用いて求めた旋律概形 (深さ 4 までを残した場合)

3. 旋律概形に対する距離の定義

3.1 距離の定義の概要 [北原 15]

DWT を用いることで旋律概形をパラメータの二分木として得ることができたが、この二分木の 2 つの部分木に対して距離尺度を定義する。ただし、深さの等しいノードを两部分木の根にするものとする。いま、二分木:

$$\{d_{-1,1}, \dots, d_{-1,N/2}; d_{-2,1}, \dots, d_{-2,N/4}; \dots; d_{-n,1}\}$$

に対して、 $d_{-I,j'}$ を根とする部分木を

$$T' = \{d'_{-1,1}, \dots, d'_{-1,N(1)}; \dots; d'_{-I,1}\},$$

$d_{-I,j''}$ を根とする部分木を

$$T'' = \{d''_{-1,1}, \dots, d''_{-1,N(1)}; \dots; d''_{-I,1}\}$$

とする。ここで、 $N(i)$ は高さ i のノードの個数である ($N(1) = 1$)。元の二分木が完全二分木であり、部分木の根として取り出すノードは同じ深さのものを取り出すので、 T' と T'' は要素数が等しくなる。 T' と T'' の間の二乗距離 $\text{dist}^2(T', T'')$ を次のように定義する:

$$\text{dist}^2(T', T'') = \sum_{i=1}^I \left\{ \frac{1}{N(i)} \sum_{j=1}^{N(i)} (d'_{-i,j} - d''_{-i,j})^2 \right\}.$$

3.2 Haar 以外のウェーブレットを使うための改良

Haar ウェーブレット ($\psi(t)$ で表す) は $t \geq 1$ のときに $\psi(t) = 0$ であるのに対し、それ以外のウェーブレットではこれを満た

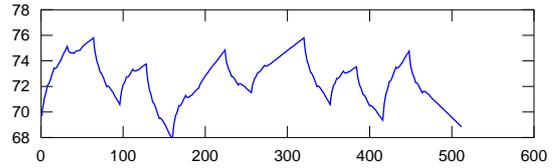


図 4: Daubechies 4 ウェーブレットを用いて求めた旋律概形 (深さ 4 までを残した場合)

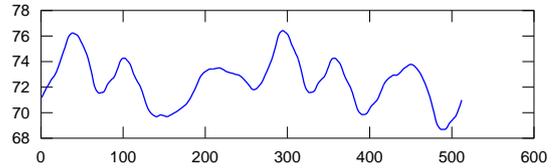


図 5: Daubechies 8 ウェーブレットを用いて求めた旋律概形 (深さ 4 までを残した場合)

さない場合が多い。一方、ウェーブレットのシフトは 1 ずつ行われることから、時系列上の各点はシフトされた複数のウェーブレットの重ね合わせで表されることとなる。たとえば、レベル $-i$ における DWT の結果が $\{d_{-i,1}, d_{-i,2}, d_{-i,3}, d_{-i,4}\}$ だった場合、Haar ウェーブレットであれば $0 \leq t < 1$ 以外は常に $\psi(t) = 0$ なので、 $d_{-1,1}$ は与えられた時系列を 4 等分した 1 つ目の区間の特徴を表し、 $d_{-1,2}$ は 2 つ目の区間の特徴を表すが、それ以外のウェーブレットではそのようにならない。そのため、部分木間の距離を計算する際に上の式をそのまま用いると、部分木が表す区間外の特徴に左右されてしまう。これは特に部分木の根に近い側で顕著である。そこで、根に近い方から何段か (k とする) を距離計算から除外することで、この問題を解決する。すなわち、次のように距離尺度を再定義する:

$$\text{dist}^2(T', T'') = \sum_{i=1}^{I-k} \left\{ \frac{1}{N(i)} \sum_{j=1}^{N(i)} (d'_{-i,j} - d''_{-i,j})^2 \right\}.$$

k をいくつにすればよいかは、用いるマザーウェーブレット (のタップ数) によって異なる。

3.3 Daubechies ウェーブレットを用いた実行例

2.3 節で旋律概形を抽出する際に得られたパラメータの二分木に対して、部分木間の距離計算を行った。マザーウェーブレットには Daubechies 4 ウェーブレットを用いた。パラメータの二分木を図 6 に示す。この二分木に対して、図中の (*1)~(*6) および (*7), (*8) を根とする部分木を取り出し、これらの間の二乗距離を計算した。 k は 1 とした。計算結果を表 1 に示す。この旋律は 1~2 小節目と 5~6 小節目が繰り返しのようになっており、それに対応する (*1) を根とする部分木と (*3) を根とする部分木の二乗距離が、他に比べて十分に小さい結果となった。また、1 小節目と 2 小節目、5 小節目と 6 小節目も (全体的な音高は異なるものの) 繰り返しのようになっており、これらに対応する部分木間の二乗距離 ((*5) を根とする部分木と (*6) を根とする部分木の二乗距離, (*7) を根とする部分木と (*8) を根とする部分木の二乗距離) も十分に小さい値となった。

Daubechies 8 ウェーブレットを用いて同様のことを行った。 k は 2 とした。計算の結果 (表 2), Daubechies 4 ウェーブレットを用いた場合 (表 1) と同じ傾向が得られた。

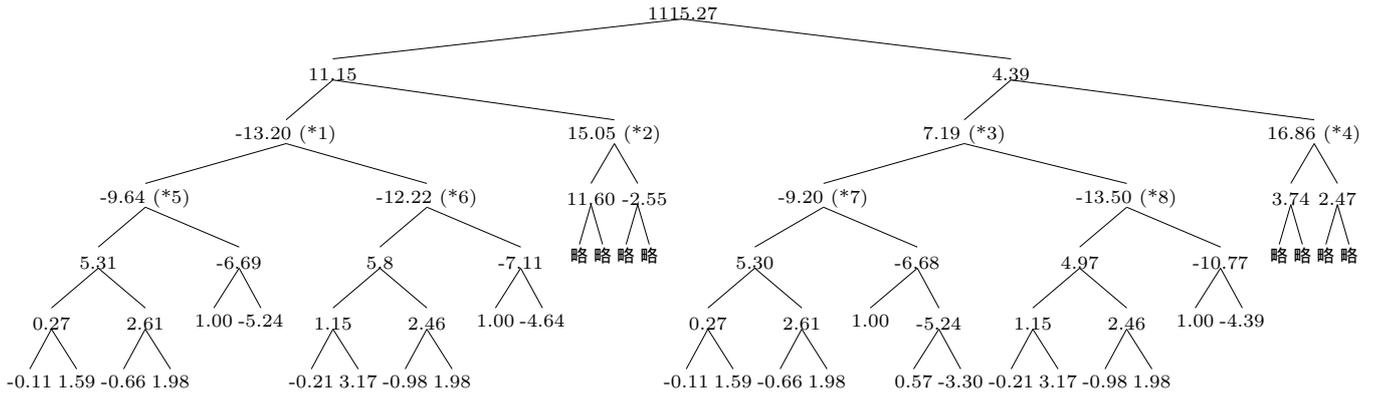


図 6: 旋律概形を表す二分木．本来完全二分木であるが，左右両方の子ノードの値が 0 のものは省略した．

表 1: 部分木間の二乗距離の計算例 (Daubechies 4). $T(\text{root} = (*n))$ は，図 6 中のノード $(*n)$ を根とする部分木を表す．

$\text{dist}^2(T(\text{root} = (*1)), T(\text{root} = (*2)))$	326.63
$\text{dist}^2(T(\text{root} = (*1)), T(\text{root} = (*3)))$	4.27
$\text{dist}^2(T(\text{root} = (*1)), T(\text{root} = (*4)))$	305.39
$\text{dist}^2(T(\text{root} = (*2)), T(\text{root} = (*3)))$	347.72
$\text{dist}^2(T(\text{root} = (*2)), T(\text{root} = (*4)))$	68.41
$\text{dist}^2(T(\text{root} = (*3)), T(\text{root} = (*4)))$	343.33
$\text{dist}^2(T(\text{root} = (*5)), T(\text{root} = (*6)))$	0.78
$\text{dist}^2(T(\text{root} = (*7)), T(\text{root} = (*8)))$	9.13

表 2: 部分木間の二乗距離の計算例 (Daubechies 8). $T(\text{root} = (*n))$ は，図 6 中のノード $(*n)$ を根とする部分木を表す．

$\text{dist}^2(T(\text{root} = (*1)), T(\text{root} = (*2)))$	92.34
$\text{dist}^2(T(\text{root} = (*1)), T(\text{root} = (*3)))$	3.00
$\text{dist}^2(T(\text{root} = (*1)), T(\text{root} = (*4)))$	73.38
$\text{dist}^2(T(\text{root} = (*2)), T(\text{root} = (*3)))$	85.77
$\text{dist}^2(T(\text{root} = (*2)), T(\text{root} = (*4)))$	6.82
$\text{dist}^2(T(\text{root} = (*3)), T(\text{root} = (*4)))$	67.78
$\text{dist}^2(T(\text{root} = (*5)), T(\text{root} = (*6)))$	2.60
$\text{dist}^2(T(\text{root} = (*7)), T(\text{root} = (*8)))$	7.47

表 3: 提案手法，タイムスパン木による Theme と各 Var. の距離および主観評価による類似度

	Proposed	TS-based	Subjective
Theme-Var. 1	74.24	13.31	-0.73
Theme-Var. 2	775.12	33.00	-0.91
Theme-Var. 3	364.99	18.42	-1.09
Theme-Var. 4	473.87	34.92	-0.82
Theme-Var. 5	1079.42	8.88	1.18
Theme-Var. 6	460.03	32.94	-1.00
Theme-Var. 7	224.63	13.44	-1.45
Theme-Var. 8	36.03	19.25	-0.64
Theme-Var. 9	38.37	11.25	1.36
Theme-Var. 10	629.47	47.06	0.64
Theme-Var. 11	38.27	26.50	0.73
Theme-Var. 12	28.36	51.63	1.00
Rank Corr.	0.231	0.119	—

Proposed: 提案手法 (DWT による旋律概形) による二乗距離
 TS-based: タイムスパン木による距離 [Matsubara 14]
 Subjective: 主観評価による類似度 ([Hirata 13] より)

よる距離尺度と主観評価の両方で Theme に近いという結果になったにも関わらず，3.2 節の距離尺度では極めて大きい距離が出力された．これは上述の通り，音脈分凝の機構がないために，原曲と同じ主旋律を奏でる音脈をとらえられなかったからと考えられる．

4. 「きらきら星変奏曲」への適用

DWT による旋律概形抽出をモーツァルト「きらきら星変奏曲 八長調 K.265」に適用した．この曲は童謡の「きらきら星」の旋律を様々に変形させた旋律からなる曲で，変奏された多くの旋律では原曲の概形を維持されていると期待される．実験には，[Hirata 13] に掲載されている単旋律に編曲された主題と 12 種の変奏を用いた．

結果を図 7 に示す．一部を除き概ね Theme と似た形の旋律概形が得られた．Var. 5 と Var. 10 については Theme に似ていない旋律概形が得られた．これらの変奏は単旋律であるものの主旋律と伴奏の 2 つの音脈に分凝して聞こえるような演奏になっている．現在の旋律概形抽出手法では音脈分凝を考慮した抽出はできないため，このような結果になったと思われる．

3.2 節で定義した二分木間の距離尺度を用いて，Theme と各 Var. の二乗距離を求めた．結果を表 3 に示す．同表には，音高を考慮したタイムスパン木による距離計算結果 [Matsubara 14] および主観評価による類似度 ([Hirata 13] より抜粋) も掲載した．3.2 節の距離尺度とタイムスパン木による距離尺度の双方に対して主観評価との順位相関を求めたところ，前者が 0.231，後者が 0.119 となった．Var. 5 についてはタイムスパン木に

5. おわりに

本稿では，[北原 15] で提案した DTW による旋律概形抽出手法および旋律概形の距離尺度が，Daubechies ウェーブレットを用いても有効に働くことを示した．また，モーツァルト「きらきら星変奏曲 八長調 K.265」に適用し，旋律概形の距離尺度が人間による主観評価と弱いながらも相関があることが分かった．今後は，認知科学や音楽心理学における旋律聴取の知見との整合性について調べていくとともに，音脈分凝の仕組みを取り入れ，より人間の旋律認知メカニズムに合った旋律の木構造表現を探索していく．

謝辞 本研究は，筑波大学 松原正樹氏の研究室に短期滞在した際に行った研究である．研究環境を用意していただき，議論にお付き合いいただいた同氏に深く感謝する．また，日頃よりご議論いただいている平田圭二氏，東条 敏氏，大村英史氏にも感謝する．本研究は，科研費 26280089，26240025 の支援を受けて行われた．

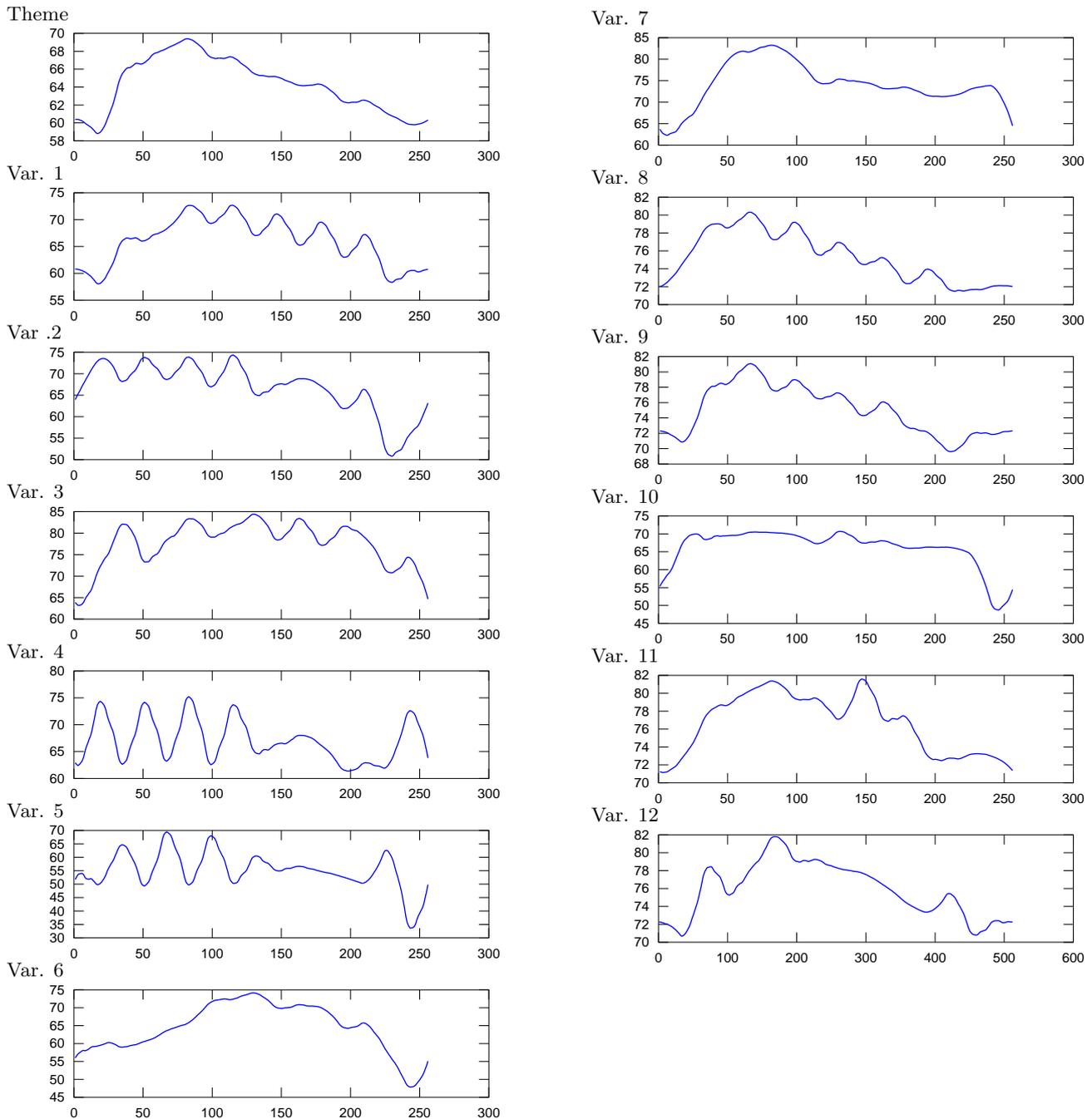


図 7: 「きらきら星変奏曲 八長調 K.265」に対して旋律概形抽出を適用した結果

参考文献

- [Hamanaka 07] Hamanaka, M., Hirata, K., and Tojo, S.: Implementing A Generating Theory of Tonal Music, *Journal of New Music Research*, Vol. 35, No. 4, pp. 249–277 (2007)
- [Hirata 13] Hirata, K., Tojo, S., and Hamanaka, M.: Cognitive Similarity Grounded by Tree Distance from the Analysis of K.265/300e, in *Proc. CMMR 2013*, pp. 415–430 (2013)
- [Lerdahl 83] Lerdahl, F. and Jackendoff, R.: *A Generative Theory of Tonal Music*, MIT Press (1983)
- [Matsubara 14] Matsubara, M., Hirata, K., and Tojo, S.: Distance in Pitch Sensitive Time-span Tree, in *Proc. Joint Conference of ICMC 2014 and SMC 2014* (2014)
- [Velarde 13] Velarde, G., Weyde, T., and Meredith, D.: An Approach to Melodic Segmentation and Classification based on

- Filtering with the Haar-wavelet, *Journal of New Music Research*, Vol. 42, No. 4, pp. 325–345 (2013)
- [松原 14] 松原 正樹, 東条 敏, 平田 圭二: 音楽理論 GTTM に基づく木構造を用いた旋律の認知的類似度の導出, 人工知能学会 2014 年全国大会, 1K4-OS-07a-2 (2014)
- [平田 02] 平田 圭二, 青柳 龍也: 音楽理論 GTTM に基づく多声音楽の表現手法と基本演算, *情処学論*, Vol. 43, No. 2, pp. 277–286 (2002)
- [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)
- [北原 15] 北原 鉄朗: 音符表現によらない旋律の木構造表現の予備検討, 人工知能学会 2015 年全国大会, 2C4-OS-21a-2 (2015)