3G3-OS-15a-4

旋律の微分と簡約の導入 〜 暗意-実現モデルの形式化を目指して

Differentiation of Melody and Its Application to Melodic Reduction - Towards Formalization of Implication-Realization Model

平田 圭二 [*1] Keiji Hirata
大村 英史 [*2] Hidefumi Ohmura
北原 鉄朗 [*3] Tetsuro Kitahara

[*1]公立はこだて未来大学 システム情報科学部 Future University Hakodate
[*2]東京理科大学 理工学部 Tokyo University of Science
[*3]日本大学 文理学部 Nihon University

Towards the formalization of Implication-Realization Model, we propose the representation method for a melody, taking into account the 1st and 2nd differentiations of a melody, and define the reduction operation on the representation. The introduction of a valid reduction operation into a melody means providing a valid metric space of a melody. In Implication-Realization Model, the idiostructure contained by a melody is characterized by the occurrences (places and textures) of basic patterns other than P and D. Hence, let us define the reduction operation in Implication-Realization Model as the operation of transforming the occurences of the basic patterns other than P and D to those only of P and D. Since Implication-Realization Model categorizes the relationships among three adjacent notes and take into account the 1st and 2nd differentiations of a melody, the reduction operation can be implemented by the value attenuation of the 1st and 2nd differentiations. In this paper, we show the case studies of a simple melody and explain the representation method for a melody and the reduction operation based on the differentiation of a melody.

1. はじめに

旋律中の音には，数小節離れた音とも関連性を持つ重要な音から，隣接した音としか関連性を持たない重要でない音まで様々な音が含まれる．旋律中の音どうしの関連性とは，音（あるいは音列）から他の音（あるいは音列）への参照と換言できる．旋律中の音どうしの関連性は，聴取者の心の中に認知的に存在する関連性と，作曲者の心の中にある関連性の 2 種類がある．一般に，前者は聴取した旋律を記憶し連想する処理に基づくので関連付けられる音どうしの距離は楽句（4 小節程度）からせいぜい数十小節（10 秒〜数分程度）であるのに対し，後者は楽章レベルまで離れる場合がある．例えば，繰り返される音列は自分自身を参照しており，聴取者としては次に聴こえる音を正しく予想できるようになる．作曲者としては，短い範囲の繰り返しは強調の効果をもたらすために用いられ，長い範囲の繰り返しは安定をもたらすために用いられる．他の例として，ダイアトニックスケールに沿って上昇する音列に含まれる音はそれまで上昇してきた各音を参照し，音高が上昇する動きに対する期待を高める．

音楽分析の主流である Schenker 理論 [1] や GTTM [4] における簡約 (reduction) は，近接する重要でない音から捨象していく操作である．簡約を繰り返すことで，楽譜上のより離れた音から成るような仮想的な旋律が生み出される．特に，GTTM では，簡約によってどのような仮想的な旋律がどのような順序で得られるかが木構造として表現されている [*1]．

簡約という操作は知識表現においても重要な役割を果たす．ある表現 x を簡約して y が得られたとする．この時，そのような y は x に包摂されると言い，$y \sqsubseteq x$ と書く．この包摂関係は知識表現におけるプリミティブである is_a 関係あるいは $part\_of$ 関係に対応する [*2]．このような半順序が与えられたことは計算のための領域 (domain) が定義されたことを意味する．簡約を定義し計算のための領域 (domain) を定義するアプローチの利点は，第一に対象や概念の世界を抽象度に沿う木構造として体系的に理解できることであり，第二に対象や概念に関する距離空間を構築できることである．Schenker 理論における Ursatz や GTTM におけるタイムスパン木，延長木の basic form は，旋律を抽象度に沿う木構造として体系的に表現している．また我々は，タイムスパン木とその簡約操作の形式化に基づいてタイムスパン木に関する距離空間を構築した [3]．そして，その距離空間は理論から導かれたものであるにもかかわらず，十分な認知的リアリティを持つことを被験者実験によって確認した．

Schenker 理論や GTTM が，簡約という操作によって多くの楽曲が共通に持つ性質の発見を目指しているのに対し，Narmour は簡約によらずに旋律を類型化することで作曲家個人の音楽を他と差別化する特徴 (idiostructure) を取り出すことを目指した [6]．Narmour は，先行する Meyer の考えを発展させ，短い範囲（低レベル）で隣接する音列それ自身が表現する暗意-実現の構造が idiostructure であると考えた．Narmour の暗意-実現モデルでは，人は旋律を聞きながら次に来る音を期待し（暗意），期待がそのまま実現したり裏切られたりして情動が生じる．音が 3 つあれば，最初の 2 音で期待が生じ，3 音めで実現か裏切りに到る．つまり，暗意-実現が生じる最小の音数は 3 音である．これより Narmour は連続する 3 つの音が作る形（基本パターン）を幾つかに分類し，実際の旋律はその基本パターンの組み合わせで構成されるとした．

連絡先: 平田圭二: hirata@fun.ac.jp
大村英史: hidefumi.ohmura@gmail.com
北原鉄朗: kitahara@chs.nihon-u.ac.jp

*1 木構造は階層を表現するが，Lerdahl は “音楽学において reduction は hierarchy を意味する” と述べている [5, p.188]．つまり，GTTM の枠組みでは，簡約，階層，木構造がほぼ同義で用いられている．

*2 属性 (feature) に関する部分集合と考える場合は is_a 関係と，要素の部分集合と考える場合は $part\_of$ 関係と対応する．

本稿では，Narmour の暗意-実現モデルに簡約を導入する試みを述べる．暗意-実現モデルには簡約の概念が含まれないため，大域的な基本パターン（仮想的な旋律）が作る暗意-実現の構造を扱えず，また旋律中の基本パターンの出現に基づく旋律どうしの類似度を計算することもできない．暗意-実現モデルに簡約の概念を導入することで，Schenker 理論や GTTM とは異なる観点 (idiostructure) から旋律を識別したり，操作できるのではないかと期待される [*3]

2. 暗意-実現モデルにおける簡約

簡約が任意回適用可能であることを保証するため，簡約という操作の型は $\mu \rightarrow \mu$ でなければならない（ここで μ は（仮想的な）旋律）．この型を満たすような素朴なものとして，基本パターンを作る 3 つの音のいずれかを head とし，さらに 3 つの隣接する head が 1 つ上のレベルの仮想的な基本パターンを作り，ボトムアップに 3 分木を作る方式が考えられる．しかし，この方式で構成される木構造は，実質的に 2 分木のタイムスパン木と等価であると思われる．なぜなら，一般に，隣接する 2 つの基本パターンでは，先行する基本パターンの 3 音めと後続の基本パターンの 1 音めは同一なので，3 つの音から head を選ぶ処理は，実質的に 2 音から head を選ぶ処理に等しい．また，3 つの音のいずれが head なのかを決める方法が必要となり，Schenker 理論や GTTM と同様の複雑な規則が必要となる．

異なる観点から考える．数値解析では時系列データの 2 点が与えられると 1 階微分が求まり，3 点が与えられると 2 階微分が求まる．これより，暗意-実現モデルにおける 3 音から成る音のパターン類型化は，旋律の 2 階微分までの値の正負や大小によるカテゴリ化として解釈される．よって，暗意-実現モデルにおける簡約を，旋律 p だけでなく，その 1 階微分 $\dot{p}$ と 2 階微分 $\ddot{p}$ を含む三つ組 $\langle p, \dot{p}, \ddot{p} \rangle$（仮想的な旋律）に対する操作として定義することが考えられる．

暗意-実現モデルの 8 個の基本パターンは，旋律の 1 階微分の値を正負と大きさ（ゼロ/小/大）でカテゴリ化し，2 階微分の値を正/ゼロ/負でカテゴリ化した内の部分集合に当たる．この意味において，本手法は，暗意-実現モデルの妥当な一般化であると考えられる．ただし，8 個の基本パターンの内 ID (Intervallic Duplicate) は，さらに 1 階微分の値が時間軸との間に作る面積（3.2 節）に関する制約を満たす必要がある．

旋律内での暗意-実現構造の出現パターンが前章で述べた idiostructure を表現していると考えられる．したがって，簡約操作が満たすべき条件は，簡約によって暗意-実現構造が削除されると，削除されたあとの音列は P（Process, 継続の暗意）または D（Duplicate, 反復の暗意，Process の特殊形）となることである．Schenker 理論や GTTM での簡約は重要な音を残すこと，あるいは重要でない音そのものを減らすことを意味するが，暗意-実現モデルにおける簡約は音の変化量（動き）を減らすことに対応する．つまり，Schenker 理論や GTTM の簡約は旋律内で音高が変わる回数を減らし，最終的に Ursatz や延長木の basic form を導くのに対し，暗意-実現モデルの簡約は同じ音数を保ったまま音高の変化量を減らし，最終的に P または D のみの旋律を導く [*4]．

*3 実は，Narmour は上述した Schenker 理論や GTTM における仮想的な旋律に関して，“現在の音楽分析の多くは，木構造に注目してレベルの高低にかかわらず簡約を行うが，これは非常に悪い習慣である” と述べている [6, p. xi].

*4 P と D 以外の暗意-実現構造の出現パターンが idiostructure とも換言できる

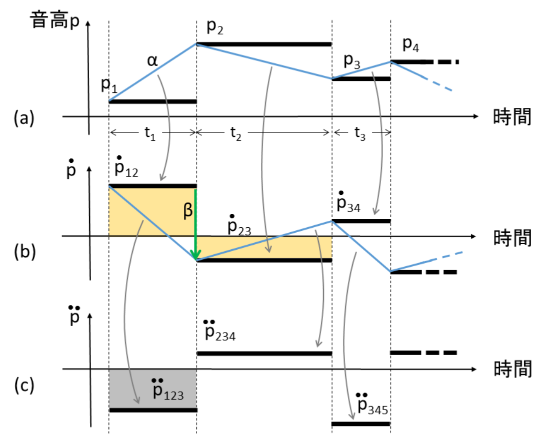

図 1: ピアノロール形式 (a) とその 1 階微分 (b) と 2 階微分 (c)

本簡約手法は音高方向の動きだけを簡約化している．音価の変化を減じる意味での簡約も考えられるが，本稿の範囲を超えるので扱わない．音符の個数を減じる意味での簡約は実質的に GTTM の簡約と等価である．

3. 旋律の表現と簡約操作

3.1 ピアノロール形式の微分と積分

まず，p を音高とする．時間的に連続する 3 音 p_i，p_{i+1}，p_{i+2} が存在すると，p，$\dot{p}$，$\ddot{p}$ が計算できる（$\dot{p}$ は p の 1 階微分を，$\ddot{p}$ は p の 2 階微分を意味する）．図 1 に，ピアノロール形式で与えられた旋律 (a) を微分する様子を示す．図中 (a) の p_i と t_i (i = 1..4) は，ピアノロール形式で表現された旋律中の各音のそれぞれ音高と音価を表す．図中 (b) の $\dot{p}_{i\ i+1}$ (i = 1..3) は，p_i と p_{i+1} の 2 音から得られた微分の値 $(p_{i+1} - p_i)/t_i$ を表す [*5]．$\dot{p}_{i\ i+1}$ の時間幅は p_i のそれを引き継ぐ．例えば，p_1 と p_2 の音の開始時点を青い線 α で結んでいるが，その線の傾き $(p_2 - p_1)/t_1$ が図中 (b) の $\dot{p}_{12}$ の値となる．同様に，図中 (c) の $\ddot{p}_{i\ i+1\ i+2}$ (i = 1..3) は，(b) に表示された $\dot{p}_{i\ i+1}$ と $\dot{p}_{i+1\ i+2}$ の 2 つの微分値からさらに得られた微分の値を表す（2 階微分）．

図 1 の (a) → (b) → (c) が微分の向きだとすると，(c) → (b) → (a) は積分の向きに対応する．例えば，図中 (c) の $\ddot{p}_{123}$ の値を積分することを考える．$\ddot{p}_{123}$ が時間軸との間に作る面積 $\ddot{p}_{123} \times t_1$（灰色の部分，$\ddot{p}_{123}$ は負なので面積も負）が積分値となり，それは (b) の β の緑線の部分に対応する．ただし，一般に積分では定数項の分だけ積分全体の値は不定になるが，この場合の定数項は $\dot{p}_{12}$ に対応する．

3.2 ケーススタディ

旋律の簡約を旋律に含まれる音の変化を減らす操作だと考えると，それは 1 階微分あるいは 2 階微分の値をゼロに近づけることに相当する（いったん微分の値をゼロに近づけてから積分し旋律を生成する）．ここではまず簡単な例として，2 階微分の値をゼロにした場合，つまり (c) の $\ddot{p}_{123}$ をゼロにした場合を考える（図 2）．$\ddot{p}_{123}$ を積分して $\dot{p}$ の値を得ると，定数

*5 ピアノロールの形を関数の値だと考えると，その微分は音高が変化する時点にデルタ関数が置かれたような形になる．本稿で扱うピアノロールは，ある意味で質点のように理想化されていて，発音開始タイミング (onset) のみが存在するように理想化されている．よって，微分の値もピアノロールのように表現される．

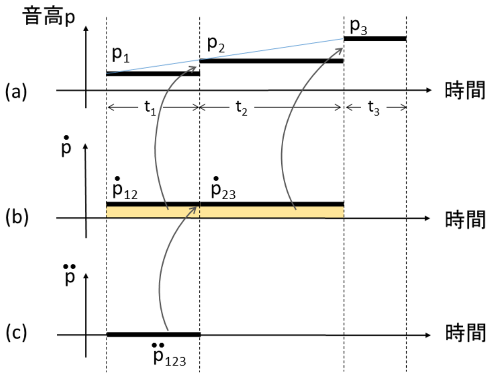

図 2: 2 階微分の値をゼロに減じることによる簡約

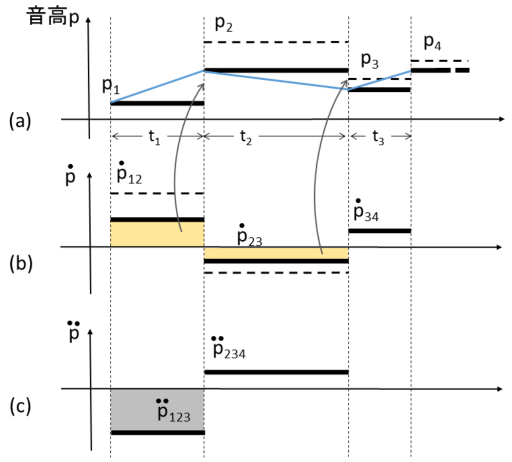

図 3: 1 階微分の値を 50%に減じることによる簡約

項 C のみとなる．つまり $\dot{p}_{23} = \dot{p}_{12} + C$ である．ここで，1 階微分の積分値に着目する（図 1(b) の $\dot{p}_{12}$ と $\dot{p}_{23}$ が時間軸との間で作る部分（図中黄色の部分）; 図 2(b) でも同様に黄色の部分）．この 2 つの面積の値が等しくなるように，図 2 において $\ddot{p}_{123}$ を積分すると，$\dot{p}_{12} = \dot{p}_{23} = (p_3 - p_1)/(t_1 + t_2)$ を得る．この例では $C = 0$ とした．また，図中黄色部分の面積の値は $p_3 - p_1$ である．

さらに $\dot{p}$ を積分して p を得る時，図 1(b) の黄色い部分の面積を保存するように定数項を決めるとすると，定数項は p_1 となり，$p_2 = \dot{p}_{12} \times t_1 + p_1$ を得る．t_1 秒後は p_2 という値になり．$t_1 + t_2$ 秒後は p_3 という簡約前と同じ値になる．図 1(b) の面積の値を図 2(b) でも同値に保つことで，簡約する領域の両端の音高を簡約前と同じに保てる．

次に，1 階微分の値を 50%に減じて積分する場合を考える（図 3）．図中 (b) の黄色の面積部分が $\dot{p}$ の積分に相当し，(a) では p_1 と p_2 の音高差として現れる．比較のため図 1 の p と $\dot{p}$ の値を破線で示している．図 1 の (a) より図 3 の (a) の方が，変化が少なくなっている．初期値としての p_1 は同じだが，p_3 の値や p_4 の値は変化している．

最後に，2 階微分の値を 50%に減じて積分する場合を考える（図 4）．図中 (c) の灰色の面積部分が $\ddot{p}$ の積分に相当し，(b) では $\dot{p}_{12}$ と $\dot{p}_{23}$ の差として現れる．ここで，$\dot{p}$ レベルの面積保存という制限を課す；つまり図 4(b) $\dot{p}_{12}$，$\dot{p}_{23}$，$\dot{p}_{34}$ が時間軸との間に作る面積（黄色の部分）が図 1(b) の相当する部分に等しいという意味である．この条件を満たすため，$\dot{p}_{12}$ の

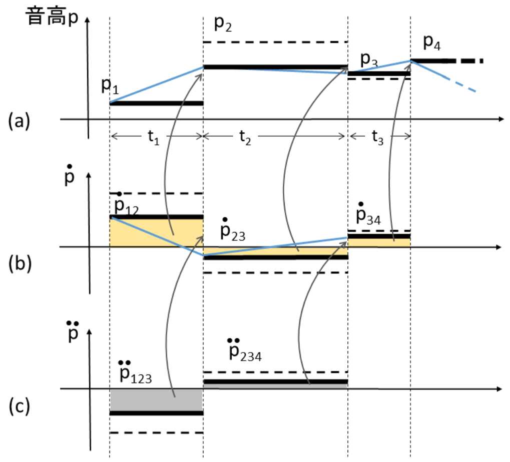

図 4: 2 階微分の値を 50%に減じることによる簡約

値を図 1(b) の $\dot{p}_{12}$ の値より負の方向にシフトする操作を行った．ここで，時間軸より下に現れた部分は負の面積と見なす．さらに積分して p の系列を得ると，p_2 と p_3 の値は異なっているものの，p_3 の値は簡約前後で不変である．

上の例では，1 階微分や 2 階微分を表すピアノロールの各セグメントを時間軸に近づけることで面積（積分値）を減じて簡約を実現している．一方，複数のセグメントから構成されるピアノロールの形を保ったまま，その全体を上下に移動させて面積を減じる方法もある．

3.3 簡約に関する制約と包摂関係の定義

前節で議論したように，簡約するということは 1 階微分あるいは 2 階微分の値を減ずることである．ただし 2 階微分の値を減ずる際，1 階微分の積分値（各々の線分が時間軸との間に作る面積）は同値という条件を付した（図 2，図 4 の例）．これは，簡約した前後で，始音と終音の音高を等しく保つためである．

また，旋律を簡約する際に，現実的な要求として，この音だけは同じタイミングかつ同じ音高に固定しておきたいというものがある．これは，1 階微分の積分値を保存するという制約で実現できるだろう．

以上の考察から，包摂関係（簡約操作）の定義は次のようになるであろう．まず，p, q をピアノロール形式で表現された旋律とする．モデル化に必要な構造は $\langle p, \dot{p}, \ddot{p} \rangle$ という三つ組で記述される（q についても同様）*6．この時

$$\langle p, \dot{p}, \ddot{p} \rangle \sqsubseteq \langle q, \dot{q}, \ddot{q} \rangle \Leftrightarrow \mathcal{A}(\dot{p}) \leq \mathcal{A}(\dot{q}) \wedge \mathcal{A}(\ddot{p}) \leq \mathcal{A}(\ddot{q})$$

と定義する．ここで，$\mathcal{A}(\cdot)$ は，引数として与えられる $\dot{p}$ や $\ddot{p}$ が時間軸との間に作る面積（積分値）を表す．$\dot{p}$ や $\ddot{p}$ が時間軸より下にある場合は負の値をとる．

さらに，簡約化の逆の精緻化の操作に関して，例えば，2 階微分の成分を M 倍 (M > 1.0) すると，旋律をある意味で強調できるかも知れない．このように簡約化と精緻化を一元的に扱える点も本方式の特長である．

4. おわりに

本稿では，暗意-実現モデルの一般化として，旋律の 2 階微分までを考慮に入れた旋律の表現法を提案し，その表現上の簡

*6 構造 $\langle p, \dot{p}, \ddot{p} \rangle$ は簡約によって同じ構造に写像される．つまり $\mu \rightarrow \mu$ である．

約を定義した．旋律が持つ idiostructure は，P や D 以外の基本パターンが旋律中にどのように出現するかということで特徴づけられる．これより，暗意-実現モデルにおける簡約を P や D 以外の基本パターンを P や D に変形するあるいは近づける操作と定義し，旋律の 1 階微分の値あるいは 2 階微分の値を減じる操作として実現した．本稿では単純な旋律のケーススタディのみを示し，旋律の微分による表現法と簡約操作を説明した．

本方式によるピアノロール譜の簡約は音階と和声を考慮していないので，簡約後の音高はダイアトニックスケールに従うとは限らない（むしろ殆どの場合はダイアトニックスケールから外れると思われる）．簡約後の仮想的な旋律から，実際の旋律を生成するには，北原らによる旋律概形の抽出手法 [2] を用いることができる．北原らの手法では，フーリエ変換と逆フーリエ変換による旋律概形の平準化によってやはりダイアトニックスケールに乗らない音高を得る．そして，それらをダイアトニックスケールに量子化するため隠れマルコフモデルを使っている．我々も同様に隠れマルコフモデルを使って簡約後のピアノロール譜をダイアトニックスケールに量子化することができよう．

今後の課題として，3.1 節で導入した旋律の表現と包摂関係の認知的リアリティを確認する必要がある．被験者実験を行い，簡約された旋律の類似度を計測する予定である．

謝辞

本研究は JSPS 科研費 26280089，25330434 の助成を受けたものです．

参考文献

[1] Cadwallader, A., Gagné, D.: *Analysis of Tonal Music: A Schenkerian Approach, Third Edition*. Oxford University Press (2010)

[2] 土屋裕一, 北原鉄朗: 音符を単位としない旋編集のための旋律概形抽出手法, 情報処理学会論文誌 (テクニカルノート), Vol.54, No.4, pp.1302-1307 (April 2013).

[3] Hirata, K., Tojo, S., Hamanaka, M.: An Algebraic Approach to Time-Span Reduction. *Computational Music Analysis*, David Meredith (Ed), Chapter 10, pp.251-270, Springer (2016)

[4] Lerdahl, F., Jackendoff. R.: A Generative Theory of Tonal Music. The MIT Press (1983)

[5] Lerdahl, F.,: Genesis and Architecture of the GTTM Project, Music Perception Vol.26, Issue 3, pp.187-194 (2009)

[6] Narmour, E.: The Analysis and CogniTion of Basic Melodic Structures: The Implication-Realization Model, The University of Chicago Press (1990).