

タイムスパン木獲得システムの完全自動化

浜中雅俊^{†1}, 平田圭二^{†2} 東条敏^{†3}

^{†1}筑波大学, 科学技術振興機構 さきがけ

^{†2}NTT コミュニケーション科学基礎研究所

^{†3}北陸先端科学技術大学院大学

hamanaka@iit.tsukuba.ac.jp

本稿では、音楽理論 GTTM(Generative Theory of Tonal Music)に基づき、タイムスパン木を完全自動で獲得する手法について述べる。我々はこれまで、GTTM を計算機実装用に拡張した exGTTM を提案し、タイムスパン木を獲得するシステム ATTA (Automatic Time-span Tree Analyzer) の実装を行った。GTTM の計算機上への実装においては、従来、各ルールが競合し分析が困難となる問題が生じていたが、ATTA では、調節可能なパラメータを導入し、ルールに優先度を設定することでその問題を解消していた。しかし、それらのパラメータの設定は手作業で行っていたため、パラメータを曲ごと適切に設定するために多大な労力を必要とした。そこで本研究では、タイムスパン木の安定度および拍節構造の安定度に基づき、ATTA のパラメータの値の最適化を行う FATTA (Full Automatic Time-span Tree Analyzer) を提案する。実験の結果、FATTA はベースラインでの ATTA の性能に比べて高い正解率であることを確認した。

Full Automation of Time-span Tree Analyzer

Masatoshi Hamanaka^{†1} Keiji Hirata^{†2} Satoshi Tojo^{†3}

^{†1} University of Tsukuba / PRESTO, Japan Science and Technology Agency,

^{†2} NTT Communication Science Laboratories,

^{†3} Japan Advanced Institute of Science and Technology

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8573 Japan

We developed a music analysis system called a full automatic time-span tree analyzer (FATTA), which analyzes a piece of music based on the generative theory of tonal music (GTTM). We previously developed an automatic time-span tree analyzer (ATTA), which can acquire a time-span tree of GTTM. Although the ATTA has adjustable parameters for controlling the weight or the priority of each rule, these parameters have to be set manually. This takes a long time because finding the optimal values of the settings themselves takes a long time. The FATTA can automatically estimate the optimal parameters based on the stability of the time-span tree and metrical structure. Experimental results showed that the performance of FATTA outperformed a baseline performance of the ATTA.

1. はじめに

我々は Lerdahl と Jackendoff の提案した GTTM[1]に基づき音楽分析器を構築している[2-3]。音楽理論は、音楽に関する知識、経験、技能を利用して楽曲を分析、解釈する方法論を我々に与えてくれる。音楽は様々な側面から、分析、解釈することができるため、これまで多くの音楽理論[4-6]が提案されており、楽曲の分析や解釈に用いられる様々な音楽的な概念が抽出され、様々

な手順が議論されてきた。

音楽知識を計算機上に形式的に記述する観点から、我々は音楽理論 GTTM が最も有望であると考えている。GTTM の特徴は、音楽が備える多様な側面を包括的に表象しているという点である。GTTM では、旋律の区切りを表現するグルーピング構造とリズムや韻律を表現する拍節構造をもとに、旋律や和声を本質的な部分と装飾的な部分に区別する簡約構造(タイムスパン木、プロロンゲーション木)を抽出する手順が提案されて

いる。GTTMに従えば、旋律、リズム、和声という3つの側面に関して一貫性のある操作の実現が期待できよう。

GTTMの分析の結果得られるタイムスパン木は、曲中の各音の構造的な重要度を階層的に表示したもので、その自動生成は音楽の深層構造の分析[7]を可能とするだけでなく、演奏の表情づけ[8-9]や、音楽情報検索システムの出力結果を提示する際などに有用な、音楽要約[10]を実現する。これらのシステム[7-10]では、楽曲の表層的な構造を直接操作できないユーザでも、タイムスパン木を用いて楽曲を操作することを可能としていた。しかしそこで用いたタイムスパン木は、音楽家が手作業で作成したもので、その自動獲得は実現されていなかった。

GTTMを計算機上に実装する際の問題点は、楽曲の解釈自体が暗黙で主観的なものであるため、GTTMを含めた音楽理論も必然的に曖昧さを持っていることである。それに加え、多くの音楽理論は人間に説明することを目的にしており、計算機上への実装が考慮されていない点も問題である。従来、GTTMの計算機上への実装が試みられたが[11, 12]、それらの手法ではルール競合の問題が解消できていなかった。一方、文献[13]ではタイムスパン木の獲得を可能としていたが、手作業によるルールの適用が必須であった。したがって、これらの手法[11-13]では分析を自動化することはできなかった。

我々はこれまでGTTMを計算機実装用に拡張したexGTTMを提案し、タイムスパン木を獲得するシステムATTA (Automatic Time-span Tree Analyzer)を構築したが、ATTAで正しい分析結果を得るためには調節可能な複数のパラメータの値を曲ごとに手動で適切に設定しなければならず、多大な労力がかかるという問題があった[2-3]。

そこで本研究では、タイムスパン木の安定度および拍節構造の安定度に基づきATTAの各パラメータの値の最適化を可能とするFATTA (Full Automatic Time-span Tree Analyzer)を提案する。FATTAでは、様々なパラメータセットの中からタイムスパン木および拍節構造が最も安定するパラメータセットを自動獲得する。タイムスパン木の安定度は、Lerdahlが音楽理論GTTMの後に提案した音楽理論Tonal Pitch Space[14]に基づき定義する。実験の結果、FATTAはベースラインでのATTAの性能に比べて高い正解率であることを確認した。

2. Generative Theory of Tonal Music

GTTMは、音楽に関して専門知識のある聴取者の直観を形式的に記述するための理論で、グルーピング構造分析、拍節構造分析、タイムスパン簡約、プロロンゲーション簡約という4つのサブ理論から構成されている。このうち本研究ではタイムスパン木の獲得に必要な、グルーピング構造分析、拍節構造分析、タイムスパン簡約を扱う。(図1)。

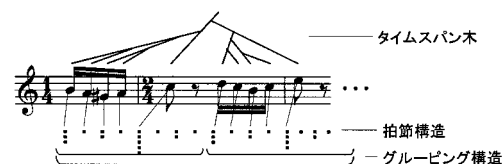


図1: グルーピング構造, 拍節構造, タイムスパン木

グルーピング構造分析は、連続したメロディをフレーズやモチーフなどに階層的に分割するもので、長いメロディを歌うときにどこで息継ぎすべきかを見つけるような分析である。拍節構造分析は、4分音符/2分音符/1小節/2小節/4小節など各拍節レベルにおける強拍と弱拍を同定するもので、聴取者が曲に合わせて手拍子を打つタイミングや指揮者がタクトを振るタイミングを求めるような分析である。タイムスパン簡約は、メロディの重要な部分と装飾的な部分を分離するもので、構造的に重要な音が幹になるような2分木(タイムスパン木)を求める分析である。図2の左図は、メロディとそのタイムスパン木を描いたものであるが、そのタイムスパン(<--->で表された部分)は、右図のようにヘッドと呼ばれる一つの音で代表させることができる(ここではC4の音)。

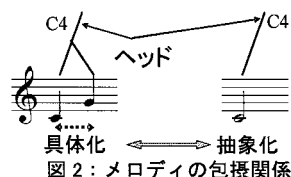


図2: メロディの包摂関係

GTTMの各サブ理論は、構成ルール(WFR: well-formedness rules)と選好ルール(PR: preference rules)の2種類からなる。構成ルールは、構造が成立するために必要な条件や制約を示すルールで、たとえば、「タイムスパンは階層的な構造となっており、各タイムスパンは1つのヘッドを持つ」というような制約がルールとして記述されている。一方、選好ルールは、複数の構造が構成ルールを満たす場合どれが好ましいかを示すルールで、たとえば、「拍節構造が安定するタイムスパン木を優先する」というルールがある。

3. ATTA: タイムスパン木分析器

GTTM の計算機上へ実装する上での問題は (1) 楽曲分析の曖昧さの問題、(2) 楽曲分析に必要な概念の欠如の問題、の 2 つに大別できる。ここで注意すべきは、(1) 楽曲分析の曖昧さ、もさらに 2 つに分けられることである。音楽理論自体に曖昧さがあるため分析結果が曖昧になることと、楽曲の解釈自体に曖昧性が内在していることで分析結果も曖昧になること、は区別すべきと考える。

上記の問題を解決するため、我々は GTTM の完全パラメータ化および外在化により計算機実装用に拡張した exGTTM を提案した。そして、exGTTM を計算機上へ実装したシステム ATTA を構築した (図 3)。計算機実装のための外在化には、大局的な構造に関するルールと局所的な構造に関するルールの両方を適切に組み合わせ階層的な構造を獲得するためのアルゴリズムの導入も含まれる。そのようなアルゴリズムは GTTM[1] では、述べられていなかった。また、完全パラメータ化は、ルールの優先度を決め、ルー

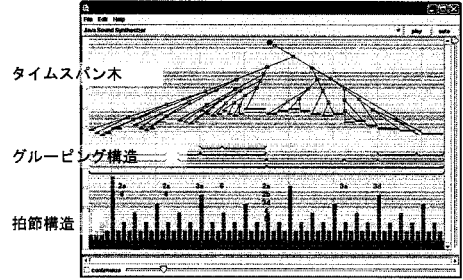


図 3: ATTA の分析結果を表示するビューア

ル間の競合を防ぐだけでなく、階層的なタイムスパン木の構造を変化させることを可能とする。もし exGTTM が、楽曲の正しい分析の結果を出力できない場合、我々は新たなパラメータを exGTTM に追加することで、正しい分析結果が出力できるように試みた。このようにして、外在化およびパラメータ化を繰り返し、最終的に、グルーピング構造分析に関する 15 個のパラメータ、拍節構造分析に関する 18 個のパラメータ、タイムスパン簡約に関する 13 個のパラメータを導入した (表 1)。

表 1: exGTTM の調節可能なパラメータ

	パラメータ	説明
グルーピング構造分析	S_{GPRR} ($0 \leq S_{GPRR} \leq 1$)	グルーピング構造分析の各ルールの強さを表すパラメータ。値が大きいほど、ルールの強さが強くなる。 $R \in \{ 2a, 2b, 3a, 3b, 3c, 3d, 4, 5, \text{ and } 6 \}$
	σ ($0 \leq \sigma \leq 0.1$)	GPR5 で用いる平均をグループの中心とする正規分布の標準偏差。値が大きくなるほど正規分布の裾野が広がる。
	W_m ($0 \leq W_m \leq 1$)	GPR6 で、各音の発音時刻の類似度と音高差の類似度のどちらを重視するか決めるパラメータ。値が大きいほど、音高差の類似度を重視する。
	W_l ($0 \leq W_l \leq 1$)	GPR6 で、並列な区間の長さをどのくらい重視するかを決めるパラメータ。値が大きいほど長い並列的な区間を重視する。
	W_s ($0 \leq W_s \leq 1$)	GPR6 で、音符 i が並列な区間の始端あるいは終端のいずれになる方を重視するかを調整するパラメータ。値が大きいほど終端を重視する。
	T_{GPR4} ($0 \leq T_{GPR4} \leq 1$)	GPR4 で、GPR2,3 の効果が明白であるかどうかを決める閾値。値が小さいほど、GPR4 が成立しやすくなる。
	T^{low} ($0 \leq T^{low} \leq 1$)	低レベルの境界であるかどうかを決める閾値。値が小さいほど、境界と認識されやすくなる。
拍節構造分析	S_{MPRR} ($0 \leq S_{MPRR} \leq 1$)	拍節構造分析の各ルールの強さを表すパラメータ。値が大きいほど、ルールの強さが強くなる。 $R \in \{ 1, 2, 3, 4, 5a, 5b, 5c, 5d, 5e, \text{ and } 10 \}$
	W_m ($0 \leq W_m \leq 1$)	MPR1 で、各音の発音時刻の類似度と音高差の類似度のどちらを重視するか決めるパラメータ。値が大きいほど、音高差の類似度を重視する。
	W_l ($0 \leq W_l \leq 1$)	MPR1 で、並列な区間の長さをどのくらい重視するかを決めるパラメータ。値が大きいほど長い並列的な区間を重視する。
	W_s ($0 \leq W_s \leq 1$)	MPR1 で、拍点 i が並列な区間の始端あるいは終端のいずれになる方を重視するかを調整するパラメータ。値が大きいほど終端を重視する。
	T_{MPRR} ($0 \leq T_{MPRR} \leq 1$)	MPR の各ルールが成立するかどうかを決める閾値。 $R \in \{ 4, 5a, 5b, 5c \}$
	タイムスパン簡約	S_{TSRPRR} ($0 \leq S_{TSRPRR} \leq 1$)
W_m ($0 \leq W_m \leq 1$)		TSRPR4 で、各音の発音時刻の類似度と音高差の類似度のどちらを重視するか決めるパラメータ。値が大きいほど、音高差の類似度を重視する。
W_l ($0 \leq W_l \leq 1$)		TSRPR4 で、並列な区間の長さをどのくらい重視するかを決めるパラメータ。値が大きいほど長い並列的な区間を重視する。
W_s ($0 \leq W_s \leq 1$)		TSRPR4 で、ヘッド i が並列な区間の始端あるいは終端のいずれになる方を重視するかを調整するパラメータ。値が大きいほど終端を重視する。

exGTTM に導入したパラメータは以下の 3 つの
 カテゴリーに分類できる。まず第 1 のカテゴリーは、
 GTTM でその存在は明らかであったものの具体的
 な値が与えられていなかったパラメータである。例
 えば、あるルールが成立するかどうかを、 D_{rule} とい
 うパラメータであらわし、ルールが成立すれば 1 に、
 不成立ならば 0 に対応付けた。第 2 のカテゴリーは、
 GTTM ではその存在自体が暗黙であったものを明
 示化したパラメータである。この例として、GTTM
 の各選好ルールの強さを重みづけするパラメータ
 が挙げられる。GTTM ではルールの適用に際して
 競合が起きることは認識されており、何らかの方
 法で解決しなければならなかった。このパラメー
 タを導入することによって、ルールの強さの制御
 が可能となり、競合を解消することが可能となっ
 た。第 3 のカテゴリーとして、GTTM ではその存
 在自体が議論されていなかったパラメータがある。
 この例として、並列性に関するルールにおいてフ
 レーズ間で並列性がどの程度で成立するかを表す
 パラメータが挙げられる。GTTM では並列性につ
 いて詳細な定義が与えられていなかった。

4. タイムスパン木分析器の完全自動化

ATTA は GTTM のタイムスパン木の獲得を可
 能としていたが、正しい分析結果を得るためには
 調節可能なパラメータを曲ごとに手動で適切に
 設定しなければならず、多大な労力がかかるとい
 う問題があった。本節では、ATTA の各パラメー
 タの値を最適化するための手法について述べる。

GTTM のルールのうち、ATTA に未実装のルール
 として、タイムスパン木・プロロンゲーション木の
 安定度に関する選好ルール GPR7 および拍節構造の
 安定度に関する選好ルール TSRPR5 がある。

GPR7 (time-span and prolongational stability):
 タイムスパン木、プロロンゲーション木が
 安定するグルーピング構造を優先する。

TSRPR5 (metrical stability): タイムスパン木
 のヘッドを選ぶ際、拍節構造が安定となる
 ものを優先する。

これらのルールは、タイムスパン簡約やプロロンゲ
 ーション簡約の結果など GTTM の分析過程の後方
 で得られる情報と前方にある分析過程とのインタ
 クションを必要とする。しかし、その具体的な
 方法やタイムスパン木の安定度や拍節構造の
 安定度がどのようなものであるかについては
 GTTM に記載がないため、実装が困難であった。

そこで本研究では、分析の結果得られたタイム
 スパン木を用いて、タイムスパン木の安定度およ
 び拍節構造の安定度を評価するアルゴリズムを与
 えることで、ATTA の各パラメータの値を自動的

に最適化する方法を提案する。その際用いるタイ
 ムスパン木の安定度を定義するにはいろいろな尺
 度が可能であるが、ここでは調とコード進行の安
 定性という観点に立ち、音楽理論 Tonal Pitch
 Space[14]に基づき定義する。一方、拍節構造の安
 定度は、タイムスパン木と整合性の高い拍節構造
 ほど安定であると考え、タイムスパン木と拍節構
 造の整合性の高さを安定度として定義した。図 4
 は、ATTA および ATTA の分析結果を用いてパラ
 メータの最適化を行うパラメータ最適化部からな
 る FATTA の処理の流れを示したものである。

4.1 Tonal Pitch Space に基づく GPR7 の実装

GPR7 は、タイムスパン簡約およびプロロンゲ
 ーション簡約からグルーピング構造分析へのフ
 ィードバックに関するルールであり、タイムス
 パン木やプロロンゲーション木が安定するよう
 なグルーピング構造を優先する。タイムスパン木
 の安定度について、GTTM ではプロロンゲーション
 簡約を行って評価できるようになっている。しか
 し、コード進行の安定性を評価する手法について
 記載がなかった。これを補う手法として、同著者
 (Lerdahl) による Tonal Pitch Space[14]という
 理論が有用で、この理論に基づく音間の距離を利
 用する方針とする。Tonal Pitch Space では、2 つ
 の調間の距離、2 つのコード間の距離、2 音間の距
 離をそれぞれ定義し、その距離が最も短くなるよ
 うに調、コード進行を決定する。そして、Tonal Pitch
 Space では、タイムスパン木の安定度に関して以下
 のような定義がある。

Time-span Stability: あるタイムスパンのヘッド
 i は、その直上のタイムスパンのヘッドを j とす
 ると、 i と j の距離が最小になるよう決定する。

たとえば、 i が直上のタイムスパンの *primary* のヘッ
 ドと等しいなら距離はゼロとなり最も安定となる。

D_{GPR7} は GPR7 が成立する度合いで、0 から 1
 までの連続値をとる。本研究ではタイムスパン木
 の各ヘッドの音符間の距離 (Tonal Pitch Space で
 定義される距離) が近いほど大きな値となるよう
 D_{GPR7} を次式のように定義する。

$$D_{GPR7} = \frac{\sum_i \text{distance}(p(i), s(i)) \times \text{size}(i)^2}{\sum_i \text{size}(i)^2} \quad (1)$$

ただし、 i は、タイムスパン木のヘッドを表す。
 タイムスパン木は 2 分木であるから、各ヘッドは、
 重要な枝 *primary* と重要でない枝 *secondary* を含ん
 でいる。そして、 $p(i)$ は、ヘッド i の *primary*
 の音符、 $s(i)$ は、ヘッド i の *secondary* の音符を
 表す。distance(x,y) は音符 x と音符 y の距離を
 Tonal Pitch Space[14]に基づき求めた値である。
 値は、0 から 1 の値となるよう正規化している。

size(i)は、タイムスパンの長さで、2 乗したものを重みとして掛けて正規化している。size(i)や log size(i)など様々なものを試した結果、一番性能が良かったため、2 乗したものを採用した。

4.2 TSRPR5 の実装

TSRPR5 (Metrical Stability) は、タイムスパン簡約から拍節構造分析へのフィードバックに関するルールであり、拍節構造が安定するタイムスパン木を優先する。タイムスパン選好ルールには、TSRRPR5 以外にも TSRPR1 (Metrical Position) という拍節構造に関するルールがあるが、TSRPR1 が、拍点の多い音をヘッドとして優先するという局所的なルールであるのに対し、TSRPR5 は、簡約の結果が安定な拍節構造となるタイムスパン木を優先するという大局的なルールであるという違いがある。

TSRPR5 を実装するためには、タイムスパン簡約の結果がどのぐらい拍節的に安定しているかを評価する必要がある。本研究では、タイムスパン

木と整合性の高い拍節構造が安定な拍節構造であると考え、次式でタイムスパン簡約の結果がどのぐらい拍節的に安定しているか、すなわち TSRPR5 が成立する度合いを定義する。

$$dot(x) \text{は拍節構造分析の結果として求まる各音符の}$$

$$D_{TSRPR5} = \frac{1}{\sum_i size(i)^2} \times \sum_i \begin{cases} size(i)^2 & dot(b(i)) \geq dot(l(i)) \\ 0 & dot(b(i)) < dot(l(i)) \end{cases} \quad (2)$$

拍点の数である。D_{TSRPR5} は 0 から 1 の値を出力し、タイムスパン簡約の結果重要と判断とされた音符が拍節構造分析でも重要と判断されているほど大きな値となる。

4.3 適切なパラメータセットの獲得

タイムスパン木の安定度 (4.1 節) や、拍節構造の安定度 (4.2 節) を求めるためには、GTTM の分析過程の後方で得られるタイムスパン木の情報を必要とする。したがって、これらは循環的な定義であると言うことができる。この循環的な定義を

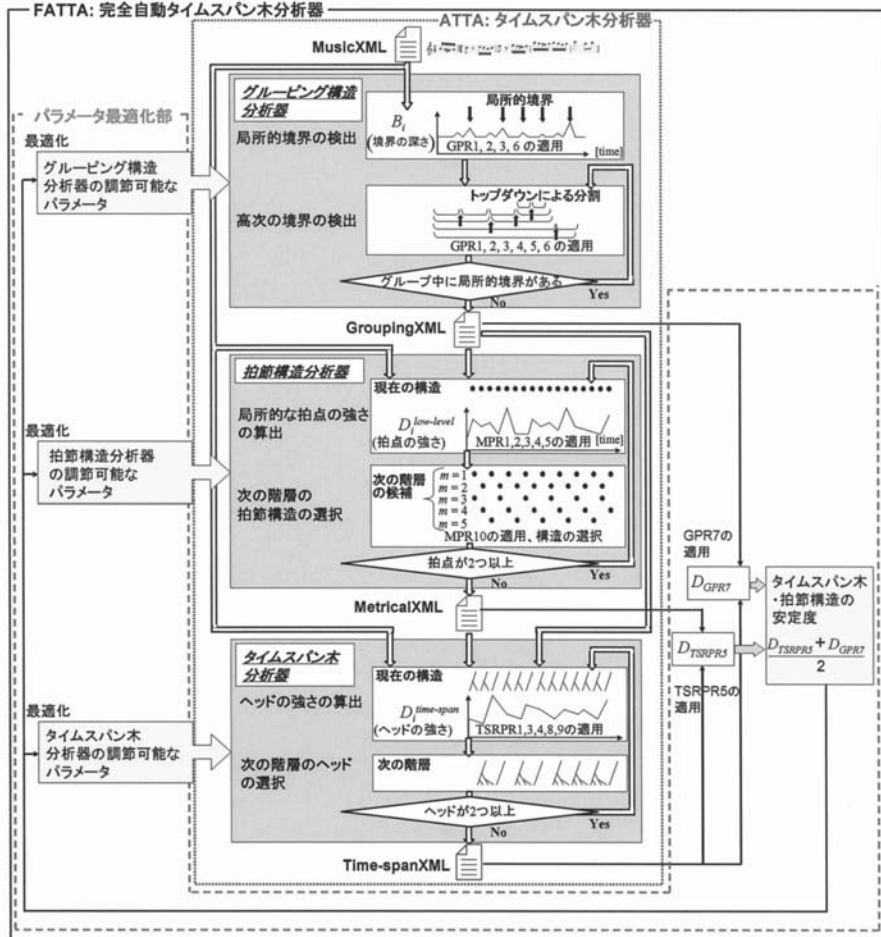


図 4: FATTA の処理の流れ

素朴に実装すると解が収束しない可能性がある。そこで本研究では、相互に依存するモジュール全体を対象として、その全パラメータを最適化する問題として定式化した。具体的には、タイムスパン木の安定度 D_{GPR7} および拍節構造の安定度 D_{TSRPR5} の両方が大きくなるような ATTA の各パラメータの値を適切なパラメータセットとして獲得する。ここでは、 D_{GPR7} と D_{TSRPR5} の相加平均が最大となるようなパラメータの値を探索する。ATTAには46個の調節可能なパラメータがあるため、探索の際、全探索を行うと効率が悪い。そこで以下のようなアルゴリズムを用いることにした。

- (1) あるひとつのパラメータを値域の最小から最大まで動かし、 D_{GPR7} と D_{TSRPR5} の相加平均が最大となるような値に設定する。ただし、パラメータの値は値域の10分の1ずつ変化させるものとする。相加平均が最大となるような値がいくつもある場合には、値域の中央値に近いほうの値とする。
- (2) (1) をすべてのパラメータについて行う。
- (3) D_{GPR7} と D_{TSRPR5} の相加平均が前回の繰り返しよりも大きくなるまで (2) を繰り返す。

5. 実験結果

FATTA の性能の評価を適合率 P (正解データと同じグループ/拍点/ヘッドが、システムの出力に含まれている割合) と再現率 R (システムの出力と同じグループ/拍点/ヘッドが、正解データに含まれている割合) を組み合わせた F 値を用いて評価する。このような評価を行うためには、グルーピング構造、拍節構造分析、タイムスパン簡約の正解データが必要である。正解データは、GTTM をよく理解している一人の音楽家が、クラシック曲から切り出した8小節の長さの100個のメロディについて分析を行ったものを、3人のGTTMの専門家がクロスチェックしたものである。

分析結果は、パラメータの値によって変化する。そこでまず、ATTAのベースラインの性能として、パラメータ調節前の性能の評価した。パラメータの初期値は値域の中央値とした。そして、FATTAを用いて、適切なパラメータを獲得した結果、F値がATTAのベースラインの性能より向上していることが確認できた。

表2: FATTAとベースラインでのATTAのF値

メロディ	グルーピング構造分析		拍節構造分析		タイムスパン簡約	
	ATTA	FATTA	ATTA	FATTA	ATTA	FATTA
1. ソナチネ (ハイドゥン)	0.21	0.32	0.88	0.88	0.37	0.41
2. 楽興の時	0.24	0.60	0.95	1.00	0.58	0.74
3. トルコ行進曲	0.67	0.67	0.91	0.96	0.68	0.80
4. アニトラの踊り	0.29	0.71	0.82	0.82	0.55	0.52
5. 子犬のワルツ	0.04	0.28	0.87	0.95	0.17	0.57
	:	:	:	:	:	:
平均(100のメロディ)	0.46	0.40	0.84	0.88	0.44	0.35

6. まとめ

本研究では、手作業によるパラメータの調節なしに GTTM のタイムスパン木の自動獲得を可能とする FATTA について述べた。FATTA では、タイムスパン木の安定度および拍節構造の安定度に基づきパラメータの最適化を実現した。具体的には、式 (1) でタイムスパン構造とグルーピング構造との整合性を自動的に計算し、式 (2) でタイムスパン構造と拍節構造との整合性を自動的に計算することで、全パラメータの値を自動的に最適化できるようになった。実験の結果、FATTA は ATTA のベースラインの性能に比べて高い正解率であることが確認できた。

参考文献

- [1] F. Lerdahl, and R. Jackendoff. A Generative Theory of Tonal Music. Cambridge, Massachusetts: MIT Press, 1983.
- [2] 浜中雅俊, 平田圭二, 東条敏, 音楽理論 GTTM に基づくグルーピング構造獲得システム, 情報処理学会論文誌, Vol. 48, No. 1, pp. 284-299, 2007.
- [3] M. Hamanaka, K. Hirata, and S. Tojo, ATTA: Automatic Time-span Tree Analyzer based on Extended GTTM, Proceeding of the 6th International Conference on Music Information Retrieval (ISMIR2005), pp. 358-365, 2005.
- [4] G. Cooper, and L. B. Meyer. The Rhythmic Structure of Music. The University of Chicago Press, 1960.
- [5] E. Narmour. The Analysis and Cognition of Basic Melodic Structure. The University of Chicago Press, 1990.
- [6] D. Temperley. The Cognition of Basic Musical Structures. MIT press, Cambridge, 2001.
- [7] 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, 情報処理学会論文誌 Vol.43, No.2, 2002.
- [8] N. Todd. A Model of Expressive Timing in Tonal Music. Musical Perception, 3:1, 33-58, 1985.
- [9] Widmer, G. "Understanding and Learning Musical Expression", Proceedings of International Computer Music Conference, pp. 268-275, 1993.
- [10] K. Hirata, and S. Matsuda. "Interactive Music Summarization based on Generative Theory of Tonal Music." Journal of New Music Research, 32:2, 165-177, 2003.
- [11] 井田健太郎, 平田圭二, 東条敏, GTTM に基づくグルーピング構造および拍節構造の自動分析の試み, 情報処理学会研究報告 2001-MUS-42, pp.49-54, 2001.
- [12] 東洋武士, 平田圭二, 東条敏, 佐藤健, グルーピング規則適用を改良した GTTM の実装, 情報処理学会研究報告 2002-MUS-47, pp.121-126, 2002.
- [13] T. Nord. A. Toward Theoretical Verification: Developing a Computer Model of Lerdahl and Jackendoff's Generative Theory of Tonal Music. Ph.D. Thesis, The University of Wisconsin, Madison, 1992.
- [14] Lerdahl, F. Tonal Pitch Space, Oxford University Press, 2001.