Vol.2014-MUS-102 No.2
2014/2/23

クラスタリングと統計的学習に基づく音楽理論σGTTMⅡ：局所的グルーピング境界の検出

金森光平[†1]　　浜中雅俊[†2]

本稿では，楽曲をクラスタリングし統計的学習を行うことで，音楽理論 GTTM の局所的グルーピング構造を自動検出する方法について述べる．音楽理論 GTTM に基づく楽曲の構造分析の際，ルールの競合が生じるため計算機上への実装が困難となる問題が生じていた．この問題に対し，従来の研究では，ルールの優先順位を制御するパラメータの導入や，統計的手法によるルールの優先順位の学習により，計算機上への実装を可能にしていたが，楽曲により適切なパラメータの値や適切な優先順位が異なるという問題があった．そこで本研究では学習曲をクラスタリングし，クラスタごとに統計的学習を行うことで，各楽曲に適した分析を行うことを試みた．

Music Theory based on Clustering and Statistical Learning σGTTMⅡ: Detection of Local Grouping Boundary

KOUHEI KANAMORI[†1]　MASATOSHI HAMANAKA[†2]

This report describes the method of detecting local grouping boundary based on the generative theory of tonal music(GTTM) by clustering and statistical learning. It is difficult to implement GTTM on computer because the rules of GTTM often conflicts and cannot detect grouping boundaries. Previous systems can implement GTTM on computer by introducing adjustable parameter or statistical learning, however, there is a problem which the values of the parameters or orders of priority of rules are different depends on the piece of music. To solve this problem, we attempt to analyze each music appropriately by clustering music and statistical learning by each cluster.

1. はじめに

我々は，楽曲の構造が似ているもの同士をクラスタリングし，クラスタごとに楽曲の構造を音楽理論 GTTM に基づき自動的に検出できるシステムσGTTMⅡを構築している．本稿ではその第一歩として，音楽理論 Generative Theory of Tonal Music (GTTM)[1]の 局所的なグルーピング構造を，最適なクラスタの検出器を用いることで，従来研究より精度良く獲得できることを示す．

音楽理論 GTTM とは，旋律の切れ目(グルーピング構造)やリズムの強弱(拍節構造)など，音楽が持つ多様な構造を包括的に分析できる音楽理論である．GTTM の分析の結果得られるタイムスパン木は，旋律の各音符の重要度を木構造で表したもので，音符の重要度を元に旋律を簡約するなど，楽曲の深層的な構造の操作を可能にする．

これまで GTTM のタイムスパン木を用いて，音楽の深層構造の分析を可能とするシステム[2,3]や，演奏の表情づけを可能とするシステム[4-6]や，音楽要約を可能とするシステム[7,8]があった．しかし GTTM は，楽曲の構造分析のためのルールの中に定義が曖昧なものが存在し，これらのシステム[2-8]はいずれも音楽家の手作業によるタイムスパン木の作成が必要であった．

そこで浜中らは，定義が曖昧なルールのパラメータ化を行って GTTM を拡張した exGTTM を提案し，パラメータを操作するだけでタイムスパン木を獲得できるシステム Automatic Time-span Tree Analyzer (ATTA)[9]を実装した．ATTA は，パラメータを操作することが楽曲構造の多様な解釈に対応し，システムを使うユーザーに楽曲構造の解釈を委ねるようにしていた．一方で ATTA はパラメータの数が多く調節が困難であるという問題があった．

これに対し三浦らは，音楽家が GTTM を用いて分析した 100 曲の楽曲分析データを用いて，統計的学習でルールの優先順位を学習していた．そして局所的グルーピング構造を自動的に検出するシステムσGTTM を提案した[10]．しかしσGTTM による学習の結果できたグルーピング構造検出器は，学習した 100 曲全体で一つの検出器を構築したため，分析する楽曲によっては間違えの多い分析結果となることがあった．

そこで本研究では，楽曲の構造を自動的に検出できることと，システムを使うユーザーが楽曲の解釈を選択できることをシステム構築の方針とし，局所的グルーピング構造の検出に取り組んでいる．我々はまず，楽曲を均等の数に分け，統計的学習を用いて，クラスタごとにルールの優先

†1 筑波大学理工学群工学システム学類
University of Tsukuba, College of Engineering Systems
†2 筑波大学大学院システム情報系
University of Tsukuba, Faculty of Engineering, Information and Systems

ⓒ 2014 Information Processing Society of Japan

順位を発見する．そしてルールの優先順位が似ている楽曲ごとに再びクラスタリングを行い，各楽曲に適したクラスタを構築した．また，システムを使うユーザーが楽曲の構造を複数の選択肢の中から選べるようにした．これは，楽曲の構造に対するユーザーの解釈をなるべく反映させるためである．システムを実装し性能評価を行った結果，適切なクラスタの検出器を用いることにより従来研究の性能を上回ることを確認した．

2. 音楽理論 GTTM とその問題点

F.Lerdahl と R.Jackendoff によって提唱された音楽理論 Generative Theory of Tonal Music (GTTM)は，他の音楽理論に比べルールが比較的厳密に定義されているが，計算機へ実装する上では，曖昧さが残されている点が問題となっていた．

本節では，2.1 節で GTTM の楽曲分析方法について説明した上で，2.2 節で GTTM の問題点について述べる．

2.1 音楽理論 GTTM の楽曲分析方法

GTTM は以下の 4 つの分析を順に行うことで楽曲の構造を分析する． グルーピング構造分析は，旋律をいくつかのグループに分割する分析である．拍節構造分析は，強迫や弱拍などの，楽曲のリズム構造を同定する分析である．タイムスパン簡約は，旋律の各音符に対し重要度をつけ，重要度の関係を木構造で表す分析である．プロロンゲーション簡約は，旋律の緊張，弛緩構造を木構造で表す分析である．図 1 に GTTM によるタイムスパン簡約までの楽曲構造の分析結果の例を示す．

この 4 つの構造分析はそれぞれ構成ルールと選好ルールから成っている．構成ルールとは構造が成り立つために必要な条件を示したルールであり，選好ルールとは複数の構造が構成ルールを満たした場合，どれがより好ましい構造かを求めるためのルールである．またそれぞれの構造は階層的になっていて，構成ルールと選好ルールに従い，それぞれの構造を順番に分析することで階層的な楽曲の構造を分析する．これら 4 つの分析のうち，我々が本稿で扱うのは(1)のグルーピング構造分析である．

グルーピング構造の構成ルールと選好ルールはそれぞれ GWFR(Grouping Well-Formedness Rules)，GPR(Grouping Preference Rules, GPR)と呼ばれる．GPR はさらに階層的な構造のうち最も低次な(局所的な)グルーピングに対応するルール(GPR 1,2,3)と，高次な(大局的な)グルーピングに対応するルールの 2 種類に分けられる．局所的な GPR が適応された音符間は，旋律の切れ目となる可能性があることを示している．図 2 に局所的グルーピング構造の分析例を示す．

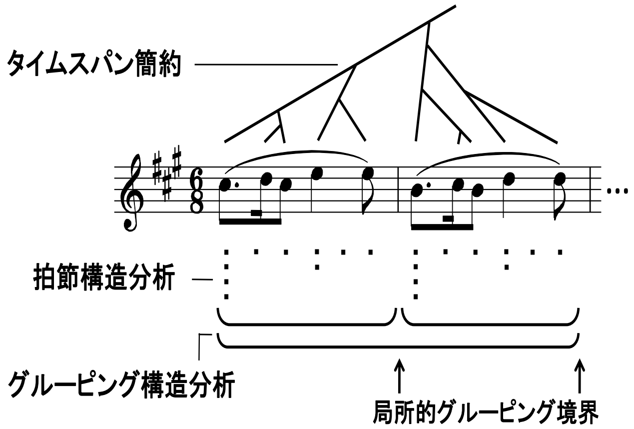


図 1　GTTM による楽曲構造分析の全体図

2.2 音楽理論 GTTM の問題点

GTTM を用いて分析を行う際に問題となるのは，より好ましい構造を決定する選好ルールに優先順位が存在しないために，選好ルール間で競合が起こってしまうことである．この選好ルール間の競合は，計算機上に実装するに当たりいくつかの問題を生ずる．本稿で扱う局所的なグルーピング構造の分析では，以下の 2 つの問題が考えられる．

第一の問題は，グルーピング境界を検出する際に，ルールの競合が生じることである．図 2 の分析例では，17 番目の音符と 18 番目の音符の間(以下, 17-18 音符間などとする)に GPR2a(休符やスラーの切れ目に適応されるルール)と GPR2b(比較的音符の音価が大きい音符に適応されるルール)が適応され，18-19 音符間に GPR3a(比較的音高の差が大きい音符に適応されるルール)が適応されているが，この両方を境界とすることは，GPR1(単音のみとなるグルーピングは避けるべきとするルール)により禁止されているため，どちらか一方を境界として選択しなくてはならない．分析例では 17-18 音符間が境界となっているが，どちらの音符間を境界とするかを選ぶ方法は GTTM では規定されていない．

第二の問題は，ある音符間に GPR の局所的なルールが適応された際に，その音符間が必ずしもグルーピング境界となるわけではないことである．図 2 を例にとると，5-6 音符間と 23-24 音符間はともに GPR3a が適応されて局所的グルーピング境界となっているが，11-12 音符間，18-19 音符間，29-30 音符間，32-33 音符間はいずれも GPR3a が適応されているのにも関わらず局所的グルーピング境界となっていない．この問題を解決する方法も GTTM では規定されていない．

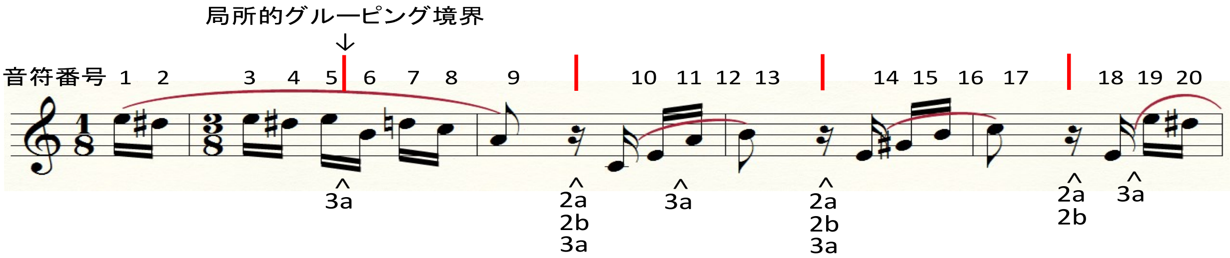


図 2　局所的グルーピング構造の分析例

3.　σ GTTM II の構築

我々は, 局所的な GPR の優先度は楽曲によっていくつかの傾向があるという前提のもと，楽曲を傾向ごとにクラスタリングし，楽曲の傾向に沿った分析を行うことを方針として σ GTTM II を構築している.

本節では，3.1 節で音楽理論 σ GTTM II の概要，3.2 節で学習データ，3.3 節で三浦らによる σ GTTM による局所的 GPR の優先度決定方法についてそれぞれ述べる.

3.1　音楽理論 σ GTTM II

本稿で構築したシステム σ GGTTM II は，学習フェーズと構造決定フェーズからなる(図 3). 学習フェーズでは，局所的な GPR の競合の問題を解決するために， σ GTTM で用いられた決定木学習の手法により，GPR に優先度を付ける．次に，決定木によって得られた局所的な GPR の優先度に基づき，楽曲を複数のクラスタに分割する．この決定木による学習とクラスタ分割を繰り返すことで，より楽曲に適した検出器を構築する． 構造決定フェーズでは，学習フェーズの出力である複数の検出器のうちどの検出器を使って構造決定するのかを, ユーザーが選択できるようにする. これにより，楽曲の構造に対するユーザーの解釈を反映させることを可能にした.

3.2　学習データ

学習データには，exGTTM を評価する際に用いた正解データを採用する．この正解データは，GTTM をよく理解している 1 人の音楽家が，クラシックの曲から８小節の長さ(８小節で区切りの悪い場合は８小節以上で区切りの良いところまでとした)で切り出した１００個のメロディの楽譜データを，GTTM に基づき手作業でグルーピング構造分析したものである．さらに，この正解データが明らかに間違った解釈でないことを３人の GTTM の専門家がクロスチェックしている.

3.3　σ GTTM による局所的 GPR の優先度決定方法

本研究では, 局所的な GPR の優先度決定の方法に三浦らによる σ GTTM で用いられた決定木を用いた手法を採用する．決定木とは，データのある注目する属性(目的属性)に関する重要な知識を，木構造によるルール(条件属性)の組み合わせで表現したものである．決定木では，条件属性の中から適切なルールを一つ選び，そのルールが 1 か 0 かによって，データを２つに分割する．この分割を分割テストと呼び，これを繰り返すことで決定木を構築していく.

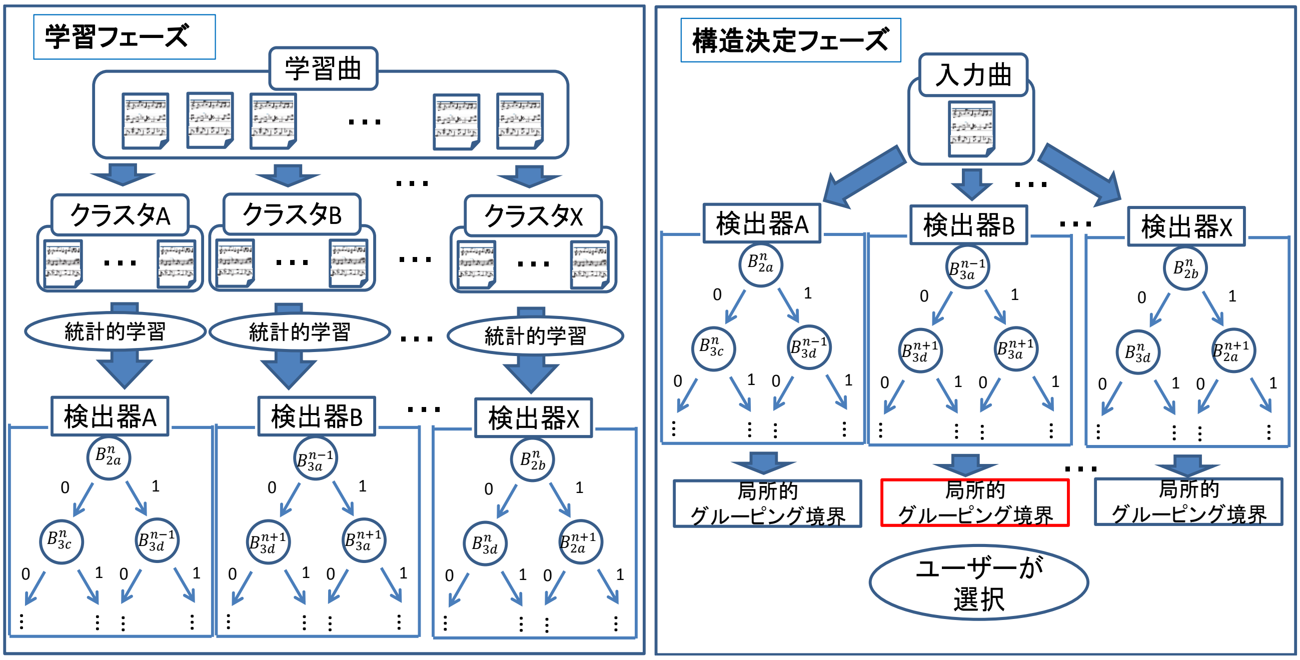


図 3　システムの全体像

ⓒ 2014 Information Processing Society of Japan

3.3.1 データの抽象化方法

学習データから目的の決定木を得るためにはデータの抽象化が必要である. 局所的 GPR の優先度を求めるという目的の下では，ある音符間に注目したとき，その音符間に局所的な GPR(1,2,3)が適応されているかが重要であり，また GPR 1 を考慮すると，注目音符間の 1 つ前の音符間と 1 つ後ろの音符間のルールの適応状況も重要となる.

そこでデータの種類をB^{n}_{GPR}という形で定義する．上付き文字によって一つ前の音符間を n-1，一つ後ろの音符間を n+1,真ん中の音符間を n と表し，下付き文字によって GPR の種類を表すとする. すべての種類のデータを書きだすと，

B^{n-1}_{2a}，B^{n-1}_{2b}，B^{n-1}_{3a}，B^{n-1}_{3b}，B^{n-1}_{3c}，B^{n-1}_{3d}，B^{n}_{2a}，B^{n}_{2b}，B^{n}_{3a}，B^{n}_{3b}，B^{n}_{3c}，B^{n}_{3d}，B^{n+1}_{2a}，B^{n+1}_{2b}，B^{n+1}_{3a}，B^{n+1}_{3b}，B^{n+1}_{3c}，B^{n+1}_{3d}

となる．この計 18 種類のルールの有無(ルールが適応されていれば 1，そうでなければ 0 の値を持つとする)を目的属性とするデータを用意し，学習を行った.

3.3.2 決定木の分割テスト

決定木の分割テストの際の評価関数には判別問題に良く利用される相互情報量を利用する．相互情報量は目的属性(bとする)に関する条件属性の重要度を比較するのに有用であり，相互情報量が最も大きい条件属性のルールを用いて分割テストをすることで，より性能の高い決定木を構築することが出来る.

相互情報量を求める手法を具体的に説明する．まず，あるデータ集合 S の事象 X に関する「あいまいさ」を，エントロピー関数を用いて次式のように定義する.

$$H(S) = H(X) = -\sum_{j=1}^{k} p_j \log_k p_j$$

ここで，p_jはXのk種類ある事象$a_1, \ldots, a_k$のうち，事象$a_j (1 \leq j \leq k)$の出現率とする. ただし, $a_1 \cap a_k = \emptyset\ (i \neq j)$であるとする．今回の学習データを例として考えると，目的属性bをX, 条件属性(ここではB^{n}_{2a}とする)を Y1 として, bとB^{n}_{2a}の出現率の関係（図 4）を用いると，

$$H(X) = -0.239 \log_2 0.239 - 0.761 \log_2 0.761 = 0.794$$

となる．同様にして H(X|Y1=1)と H(X|Y1=0)が求まる.

$$H(X|Y1=1) = \frac{0.101}{0.136}\log_2\frac{0.101}{0.136} - \frac{0.035}{0.136}\log_2\frac{0.035}{0.136} = 0.817$$

$$H(X|Y1=0) = \frac{0.138}{0.864}\log_2\frac{0.138}{0.864} - \frac{0.726}{0.864}\log_2\frac{0.726}{0.864} = 0.634$$

よって，H(X|Y1)が以下のようにして求まる.

$$H(X|Y1) = 0.136 \times H(X|Y1=1) + 0.864 \times H(X|Y1=0) = 0.659$$

このエントロピー関数値の減少量,

$$H(X) - H(X|Y1) = 0.135$$

が$Y1$のXに関する相互情報量である．また，bをX，B^{n}_{2b}を$Y2$とし，bとB^{n}_{2b}の出現率の関係（図 5）を用いて，同様に$Y2$のXに関する相互情報量を求めると,

$$H(X) - H(X|Y2) = 0.125$$

となる.

このようにして全ての条件属性に対して相互情報量を求め，相互情報量が最も大きい条件属性のルールを見つけ

Proportion	$B^{n}_{2a}=1$	$B^{n}_{2a}=0$	Total
$b=1$	0.101	0.138	0.239
$b=0$	0.035	0.726	0.761
Total	0.136	0.864	1.000

Proportion	$B^{n}_{2b}=1$	$B^{n}_{2b}=0$	Total
$b=1$	0.109	0.131	0.240
$b=0$	0.051	0.709	0.760
Total	0.160	0.840	1.000

図 4　b とB^{n}_{2a}の出現率の関係

3.3.3 統計的学習による選好ルールの優先度判定

決定木の構築には, J.R.Quinlan が開発した C4.5 という決定木生成アルゴリズムを用いた[11]．C4.5 はエントロピーの概念を用いて，正解データから決定木を生成することができる.

図 5 は完成した決定木の一例を示している．この決定木の各終端ノードは，そのノードに分類されるデータの局所的グルーピング境界の有無 b の頻度が大きい値でラベル付される．これから各条件属性の組み合わせに対する境界の有無の確率を求めることができる. この条件付き確率が 0.5 以上の場合には，局所的グルーピング境界が存在する(b=1)と判定し，0.5 未満の場合とその条件属性の組み合わせのデータが無い場合には，局所的グルーピング境界が存在しない(b=0)と判定した．例えば図 5 ではB^{n}_{2a}=1 かつB^{n}_{2b}=1 の時，b=1 となるので，GPR2a と GPR2b が同時に成り立つ音符間には, 全て局所的グルーピング境界があると判定する.

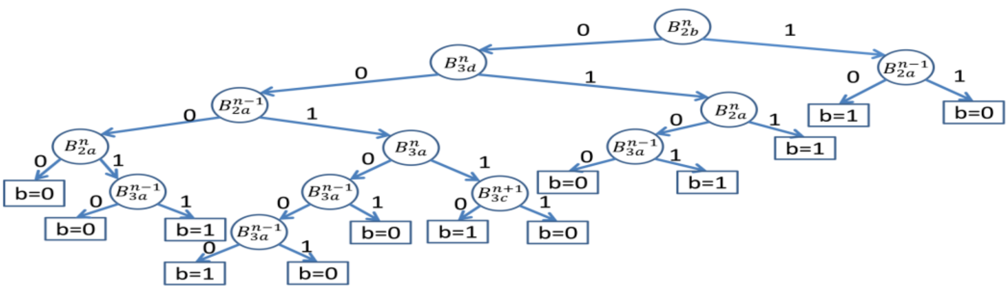


図 5　生成された決定木

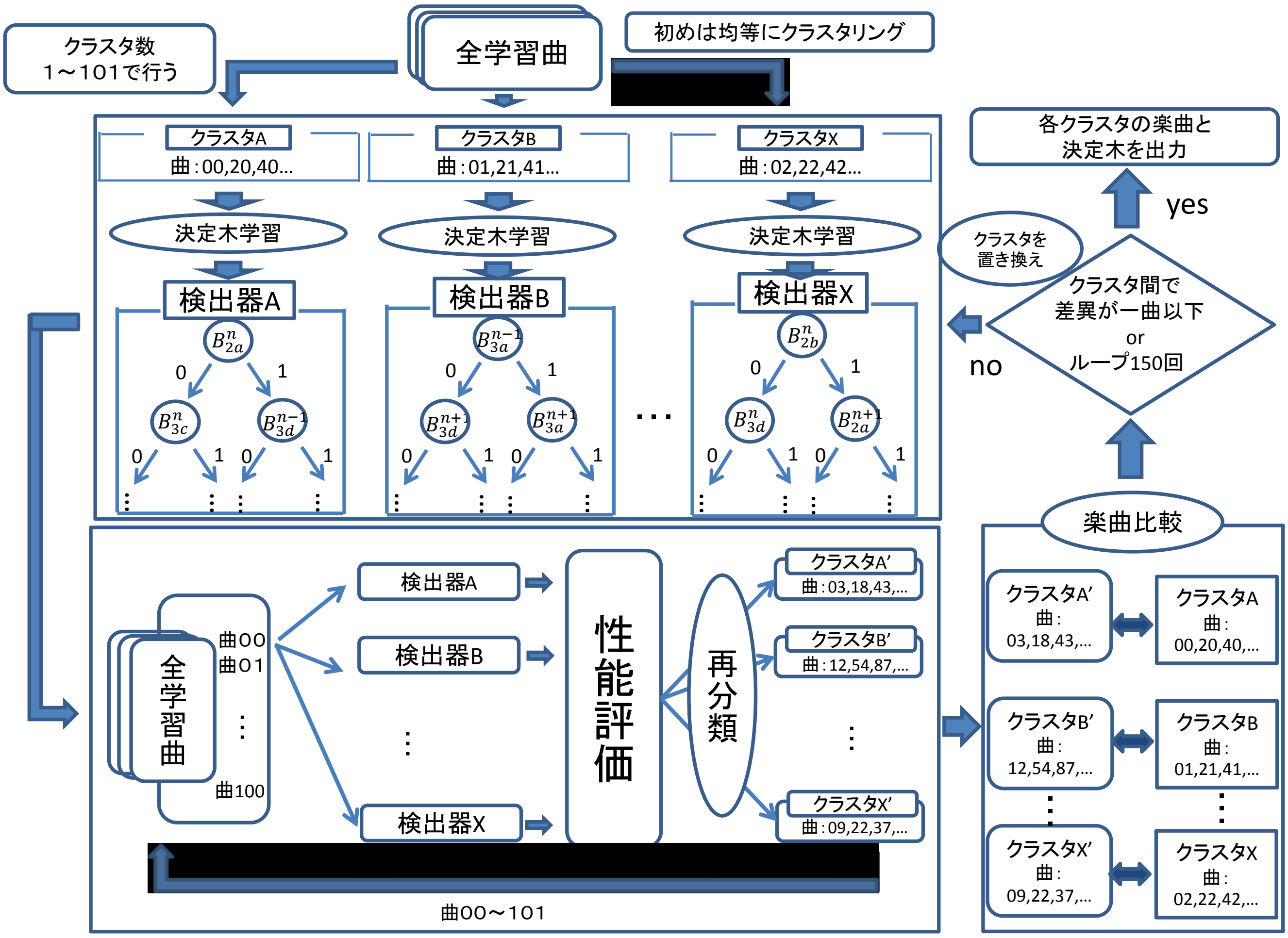


図 6　システムの詳細

4. 楽曲のクラスタリング方法

本システムでは，楽曲をクラスタリングし統計的学習を行うことで，楽曲の局所的 GPR の優先度が似ている楽曲ごとに検出器を作り分ける．その際，より GPR の優先度が似ている楽曲ごとにクラスタリングするために，クラスタリングと統計的学習を再帰的に繰り返すことで，クラスタ内の楽曲の傾向を徐々に近づけるシステムを考えた．図 6 に本システムの詳細を示す．

本節では，4.1 節でシステムの性能評価方法について，4.2 節でシステムの詳細について述べる．

4.1 システムの性能評価方法

本研究では，σGTTMⅡの局所的なグルーピング境界検出器の性能評価を，適合率 P(precision)と再現率(recall)を組み合わせた F 値で評価する．適合率 P とは，正解データと同じグルーピング境界がシステムの出力に含まれている割合のことで，再現率 R はシステムの出力と同じグルーピング境界が正解データに含まれている割合のことである．F 値の式は，

$$F_{値} = 2 \times \frac{P \times R}{P + R}$$

で表され，適合率 P と再現率 R の値が高いほど F 値の値も高くなる．F 値が高いほどシステムの性能は高いと言える．

4.2 システムの詳細

まず，学習曲を均等に各クラスタに割り振る．これは，初期の段階でクラスタごとに…. 次にクラスタごとに決定木学習を行う．(以後，学習フェーズと呼ぶ)この学習の結果，クラスタごとに GPR の優先度が木構造で表される．しかしこの方法ではクラスタ内の楽曲の全体の傾向を表してしまうので，このままではクラスタ内に的外れな学習結果となる楽曲が存在する可能性がある．

そこで，学習フェーズでできた木構造で全学習データを一曲ずつ性能評価して，最も GPR の優先度が近い木構造を作ったクラスタに再分類する．そして再分類前と再分類後でクラスタ内の楽曲がどの程度一致しているかを比較する．クラスタ内の楽曲の種類に変化が無ければ無いほど，学習フェーズでできた検出器がクラスタ内の個々の楽曲の GPR の優先度の傾向を全て網羅した検出器であると言え

ⓒ 2014 Information Processing Society of Japan

る.

このクラスタ内の楽曲比較の結果，クラスタ間の楽曲の差異の合計が二曲以上の場合は，再分類後のクラスタで再び学習フェーズへと戻り，クラスタ間の楽曲の差異の合計が一曲以下または，ループが 150 回に達した時点で各クラスタ内の楽曲と検出器を出力して終了させた．このシステムにより, 局所的 GPR の優先度がクラスタ内の楽曲全てにより良く当てはまる検出器を作ることを試みた.

5. 実験結果

提案システムを評価するため，システムの最終的な出力であるクラスタ内の楽曲を，そのクラスタでできた検出器にかけて性能を評価した. 本システムを計算機上に実装し, クラスタ数を 1 から 100 まで変化させ評価した結果を図 7 に示す.

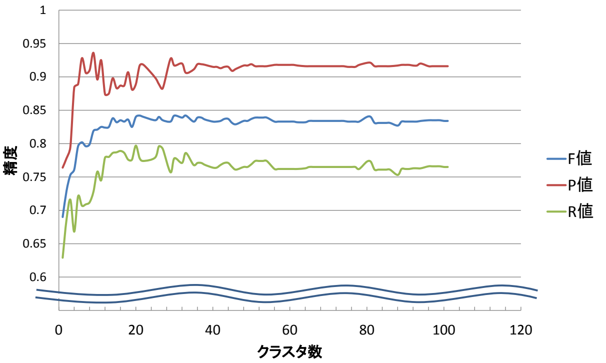


図 7 クラスタ数による精度の推移

このとき，初期のクラスタ数が大きくなるにつれて，システムの出力にクラスタ内の楽曲の数が 0 となるものが現れ，初期のクラスタ数とシステムの出力で楽曲の数が一曲以上となるクラスタ数が異なることがあった．図 8 に初期のクラスタ数とシステムの出力で楽曲の数が一曲以上あるクラスタの数との関係を示す.

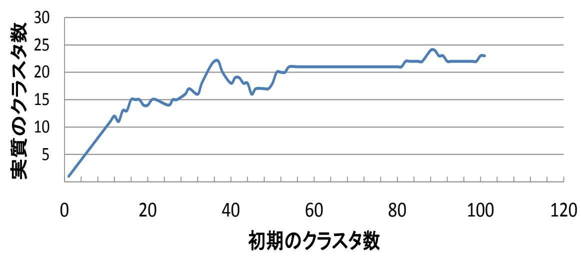


図 8 初期のクラスタ数に対する実質のクラスタ数の推移

次に，本システムと従来手法の性能を比較するために，同じ楽曲で従来手法の ATTA と σ GTTM でそれぞれ局所的なグルーピング構造を検出した．また，本システムである σ GTTM Ⅱ の性能はクラスタ数 10 のときのものとした．実験の結果，複数ある検出器のうち最適な検出器を選ぶことにより， σ GTTM Ⅱ の性能は従来手法の性能を上回ることを確認した．図 9 にこの評価実験の結果を示す．なお，※の個所は文献 9)で記述されていなかったため，記載できなかった.

	適合率P	再現率R	F値
ATTA (パラメータ調節後)	※	※	0.77
σGTTM	0.764	0.630	0.690
σGTTM Ⅱ (クラスタ数10)	0.905	0.734	0.811

図 9 評価実験の結果(closed)

学習していないデータに対する本システムの性能を評価するために, 3.2 節の学習データと同様に，音楽家が GTTM を用いて分析し，3 人の GTTM の専門家がクロスチェックした 130 曲の楽曲を用いて，性能を評価した．また参考までに先行研究である σ GTTM も同条件で評価した.
実験の結果，最適な検出器を選ぶことにより，学習していないデータに対しても高い性能であることを確認した．図 10 にこの評価実験の結果を示す.

	適合率P	再現率R	F値
σGTTM	0.467	0.736	0.571
σGTTM Ⅱ (クラスタ数10)	0.684	0.916	0.783

図 10 学習していないデータに対する性能評価(open)

今回の実験のうち, クラスタ数 10 の時に生成されたクラスタと検出器の中から，2 つ分の検出器と，そのクラスタに属する楽曲のうち 2 曲分の局所的グルーピング境界を検出した例を付録に示す.

6. おわりに

本稿では，統計的学習を用いて適切にクラスタを選択することで，GTTM の局所的なグルーピング境界を自動で検出する方法について述べた．本手法では，決定木を用いて局所的な GPR の優先度を求めることで, 選好ルールの競合が起こらないようにした．さらに，学習曲を局所的な GPR の優先度が似ている楽曲ごとにクラスタリングを行い，クラスタごとに最適な検出器を生成した．本手法を実装した検出器を用いて実験した結果，適切なクラスタを選ぶことにより，先行研究の性能を上回ることを確認した.

今後， σ GTTM Ⅱ を高次のグルーピング境界の検出や，拍節構造，タイムスパン簡約に拡張することで，システムが自動的に出力した複数の楽曲の構造の中から，ユーザー

が自身の解釈に合った構造を選択できるシステムを構築していく.

参考文献

1) F. Lerdahl and R. Jackendoff, A Generative Theory of Tonal Music. Cambridge: The MIT Press, 1983.

2) 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.

3) K. Hirata, and T. Aoyagi. “Computational Music Representation on the Generative Theory of Tonal Music and the Deductive Object-Oriented Database.” Computer Music Journal 27(3), 73–89, 2003.

4) N. Todd. A Model of Expressive Timing in Tonal Music. Musical Perception, 3:1, 33-58, 1985.

5) Widmer, G. "Understanding and Learning Musical Expression", Proceedings of International Computer Music Conference, pp. 268-275, 1993.

6) Hirata, K., and Hiraga, R. "Ha-Hi-Hun plays Chopin’s Etude", Working Notes of IJCAI-03 Workshop on Methods for Automatic Music Performance and their Applications in a Public Rendering Contest, pp. 72-73, 2003.

7) 平田圭二, 松田周, パピプーーン: GTTM に基づく音楽要約システム, 情報処理学会研究報告 2002-MUS-46, pp.29-36, 2002.

8) 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.

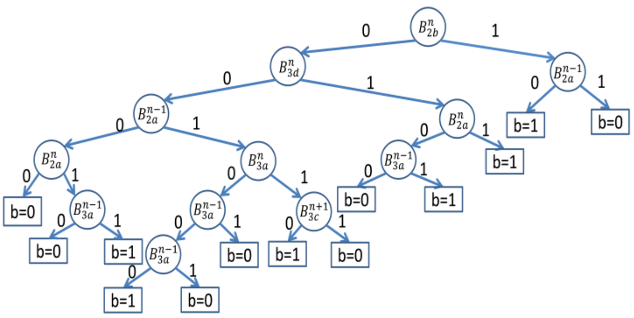
9) 浜中雅俊, 平田圭二, 東条敏: "ATTA: exGTTM に基づく自動タイムスパン木獲得システム", 情報処理学会 音楽情報科学研究会 研究報告 2005-MUS-61-4, Vol.2005, No.82, pp.19-26, August 2005.

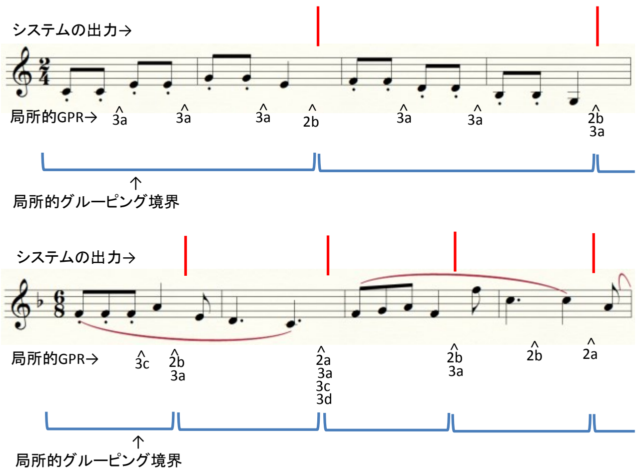
10) 三浦右士, 浜中雅俊, 平田圭二, 東条敏: "統計的学習に基づく音楽理論 σGTTM : 局所的グルーピング境界の検出", 情報処理学会 音楽情報科学研究会 研究報告 2008-MUS-76-14, Vol.2008, No.78, pp.75-82, August 2008.

11) J.R.Quinlan, C4.5:Programs for Machine Learning. Morgan Kaufman Publishers,1993.

付録 σGTTMⅡによる分析結果(クラスタ数 10)

付録 A.1 クラスタ１の決定木と，クラスタ内の楽曲の局所的グルーピング構造分析例





付録 A.2 クラスタ２の決定木と，クラスタ内の楽曲の局所的グルーピング構造分析例