

木構造に基づく時系列メディア表現法の提案と その操作系の実現に向けて

平田 圭二^{1,a)} 東条 敏^{2,b)} 浜中 雅俊^{3,c)} 長尾 確^{4,d)} 北原 鉄朗^{5,e)} 松原 正樹^{6,f)}
吉井 和佳^{7,g)}

概要：本稿では、我々が2014年度に開始した研究プロジェクト「木構造に基づく時系列メディアの表現法の提案とその操作系の実現」について述べる。音楽や自然言語のような時系列メディアのもつ意味・意図を表現するには、木構造（階層的順序構造）を用いるアプローチが有望であると考えられる。本プロジェクトの目的は、コンテンツの意図とそれに対する操作を形式化し代数系を構築することである。ここで、コンテンツの意図は木構造として、操作は木構造に対する演算子として形式化される。ここでその木構造と演算子は、人の認知(cognition)を統合的に理解できるようなものでなければならない。そして、その代数系を、音楽や自然言語を含む様々な時系列メディアに水平展開させる。具体的には、まず楽譜データを対象とした音楽構造分析とその演算体系の構築に二つのアプローチで取り組む。一つは、楽譜データを木構造に還元し、その木構造に論理演算を直接適用することで、ユーザにとって直感的な編集やモーフィングを可能にするアプローチである。もう一つは、木構造の背後にある生成規則まで還元し、確率空間において生成規則率を操作することで、より広いクラスの楽譜データを統一的に扱うアプローチである。最終的にはこの二つのアプローチの統合を目指す。これと同時に、音楽以外の時系列メディアとして、ディスカッション議事録や映画を対象とし、言語(発言やセリフ)の持つ意味とその時間軸上での関連性を推論することにより、音楽と同様の木構造の獲得および演算が可能になるかどうか検証を進める。また、これら基礎理論の成果にもとづいて、概念実証のためのユーザアプリケーションの開発にも取り組む。本研究プロジェクトは最終的に、コンテンツ制作の仮想化に寄与すると考えている。

キーワード：時系列メディア、音楽構造解析、ディスカッションマイニング、タイムスパン木、生成文法、確率モデル、Generative Theory of Tonal Music (GTTM)

1. はじめに

自然言語や音楽などの時系列メディアを介したコミュニケーションにおいて、人は自然言語や音楽によるコンテンツをオブジェクトの列として認識する。そのオブジェクトは記号や数学的実体(関数や集合など)として記述される(図1)。人は意図を込めてオブジェクトの列を表出

(presentation)するが、その意図はチョムスキー(Noam Chomsky)を始祖とする階層的順序構造(木構造)として表現(representation)されることが多い。例えば自然言語の単語列に対して、チョムスキー流の句構造規則を適用した結果得られる階層構造は構文木と呼ばれる。同様に、楽曲におけるピッチイベントの並びに対しても、音楽理論GTTM等によって二分木が得られる。

受取り側の人々がオブジェクトの列を理解するとは、表出側と同様な木構造をオブジェクトの列に割り当てることであり、これによって理解が成立したと考える。これまで提案されてきた自然言語や音楽などの理論の多くは受取り側に対応し、与えられたオブジェクトの列から木構造を導く機械的な手続きを教えるものであった。一方、表出側に対応して、人が意図を込めてオブジェクトの列を作り出す理論として、文章の創作法や作曲法などが多数提案されている。しかし、これらは人の思考を直接的にオブジェクトの列として表出するための実践的な方法であり、人の思考を

¹ 公立はこだて未来大学 システム情報科学部
² 北陸先端科学技術大学院大学 情報科学研究科
³ 京都大学 大学院医学研究科
⁴ 名古屋大学 大学院情報科学研究科
⁵ 日本大学 文理学部
⁶ 筑波大学 図書館情報メディア系
⁷ 京都大学 大学院情報科学研究科
a) hirata@fun.ac.jp
b) tojo@jasit.ac.jp
c) masatoshi@kuhp.kyoto-u.ac.jp
d) nagao@nuie.nagoya-u.ac.jp
e) kitahara@chs.nihon-u.ac.jp
f) masaki@slis.tsukuba.ac.jp
g) yoshii@kuis.kyoto-u.ac.jp

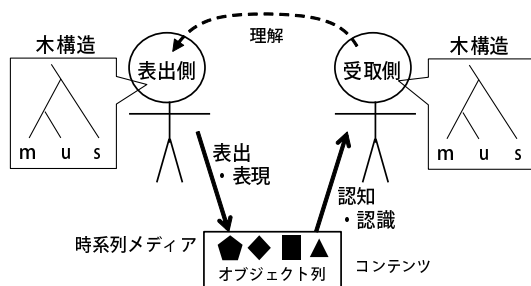


図 1 チョムスキー流コミュニケーションと木構造の役割

計算機上での意図表現に変換する形式的な理論としては不十分である。現在、人が意図を込めるような処理を機械で実現しようとすると、多数のヒューリスティクスを導入したり、generate & test 法や遺伝的アルゴリズム等（直接それを生成するシステムの実現は難しいが評価したり識別する関数や手続きを与えることは比較的容易であるという一般的な性質を利用した手法）を用いることが多い。

本プロジェクトの目的は、自然言語や音楽などの時系列メディアに関して、意図を込めた表出を計算論的な立場から定式化することである。計算論的な立場とは、入力された木構造を加工して出力するような演算が定義できることを指す。意図を込めたオブジェクトの列つまりコンテンツ表出のための技術やツールには前掲の generate & test 法や遺伝的アルゴリズムの他に、エディタ、事例に基づく推論、シミュレーション、機械学習などがある。これらはコンテンツを直接表出する方法であるのに対し、我々は、性質が良く知られている基本演算あるいは使い慣れた基本機能をユーザが自由に組み合わせるプログラミング的なアプローチを採用する*1。このアプローチの利点は、コンテンツ作成中（試行錯誤的に基本機能を組み合わせている時）にコンテンツ最終形での意図や状態が予想できることである。これは、作成しているコンテンツの性質を基本機能の性質から推論し構成できるからである。プログラミングの最中のプログラマが自分の書いているコードの実行時動作を予想しながらプログラムを書くのと同様である。

2. プロジェクト全体構成

本プロジェクトは、主に以下の三つの研究課題からなる。

- (1) 音楽における理論的枠組みの構築（図 2 左上）：楽譜（楽曲）から意図を抽出し木構造を得る方法（分析法）を検討する。計算機上での木構造表現と木構造に適用する演算体系を構築する。
- (2) 音楽における理論的枠組みの応用（図 2 左下）：プログラミング的アプローチに則った音楽応用を構築し、基本機能の組合せに関する実践的方法論を確立する。
- (3) 他の時系列メディアへの展開（図 2 右）：上の (1) で

*1 画像処理分野に数理形態学 (mathematical morphology) と呼ばれる手法がある [1]。画像（形状）を集合として扱い、基本演算をその集合に適応し変形していくことで画像の意味理解を行う。

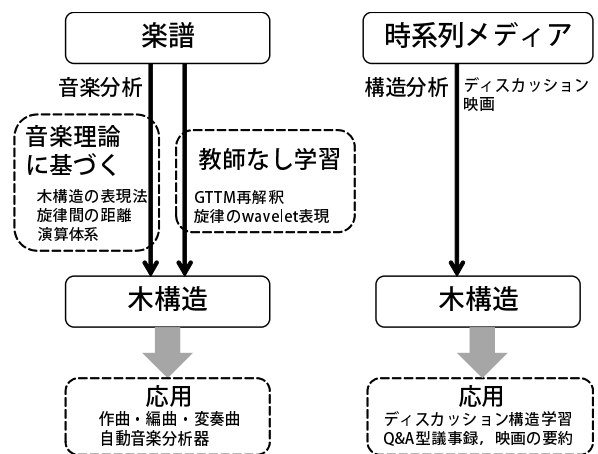


図 2 各研究課題の位置付け

定式化された枠組みを拡張し、様々な時系列メディア（ディスカッション、映画、物語、音楽など）に適用する。同じく応用を構築することで、理論的枠組みの応用範囲、有効性を検証する

研究課題 (1) に関して、一般に、意図表現の正確さと操作の簡便さとはトレードオフの関係にある。例えば、意図通りのコンテンツを作ろうと思うと、精密で大量な操作・制御を必要とする場合が多い。ここで、認知的リアリティ*2を持つ定式化が実現できれば、ユーザには直観的に理解できる基本機能を直観的に理解できる方法で組み合わせる枠組が提供される、つまり、基本機能の組み合わせは compositional かつ predictable であるという仮説を我々は立てている。その結果、大規模なコンテンツ作成における意図の表現が見通し良く実現できることを期待している。

3. 音楽における理論的枠組みの構築

本章では、第一の課題である音楽における理論的枠組みの構築について述べる。具体的には、まず楽譜データを対象とした音楽構造分析とその演算体系の構築に二つのアプローチで取り組む。一つは、楽譜データを木構造に還元し、その木構造に論理演算を直接適用することで、ユーザにとって直感的な編集やモーフィングを可能にするアプローチである。もう一つは、木構造の背後にある生成規則まで還元し、確率空間において生成規則率を操作することで、より広いクラスの楽譜データを統一的に扱うアプローチである。最終的にはこの二つのアプローチの統合を目指す。

3.1 意図を表す木構造に求められる情報

我々は、意図を表す木構造として音楽理論 Generative Theory of Tonal Music (GTTM) [12] のタイムスパン木を採用する [7]。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

*2 木構造を仮定すると人の認知 (cognition) を整合的に理解できるという意味で、その木構造は認知的リアリティを持つと呼ぶ。

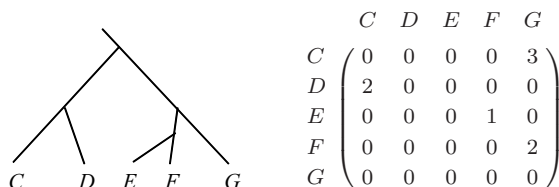


図 3 タイムスパン木の例とその行列表現

グ構造解析, 拍節構造解析, タイムスパン簡約, プロロンゲーション簡約と呼ばれる4つのサブ理論から構成されている。具体的には, 旋律の区切りを階層的に表現するグルーピング構造とリズムやアクセントを表現する拍節構造をもとに, ある旋律や和声を本質的な部分と装飾的な部分とに階層的に簡約化することで, ボトムアップに木構造を獲得する手順(タイムスパン簡約)が定義されている。さらに, タイムスパン簡約の結果を用いて, 楽曲全体の大局的な構造を分析した二分木を求める手順(プロロンゲーション簡約)も定義されている。

タイムスパン木のヘッドは支配下の部分木をすべて支配するローカルキーであると考えられ, 我々はその支配域を最大タイムスパンとして定式化してきた [6, 19]。しかし, GTTM では枝の接合点の高さ(簡約のレベル)情報が暗黙的なため, 簡約可能な枝の決定に曖昧性が残っていた。もし簡約におけるレベルの概念を明示化し, 枝の接合点の高さ情報を適切に決定できると, GTTM のタイムスパン簡約結果をシェンカー理論による簡約結果と一致させることができるかも知れない。このため, 我々は枝の接合点の高さ情報も含む木構造の表現法として, どのピッチイベントがどのピッチイベントにどの高さで接合しているかの高さ情報を成分として持つ行列表現を提案する(図3)。

また, プログラミング的なアプローチを実現するためには, 全てのタイムスパン木を含むようなドメインを定義する必要がある。ここでタイムスパン木のドメインの大きさを考えてみる。一般に n 個のシンボル列に対する二分木の種類は $n-1$ のカタラン数であることが知られている。GTTM による解析結果では, 各分岐点においてヘッドの概念, すなわち二つの枝の間で優位・下位(primary/secondary)の概念が生ずる。したがって n 個のピッチイベントに対するヘッド付き二分木のトポロジーの数は, $n-1$ 個ある各接点の primary/secondary を考えると, カタラン数を 2^{n-1} 倍したもので, すなわち $T_n = C_{n-1} \times 2^{n-1}$ 個である。

3.2 木構造に対する演算体系

プログラミング的なアプローチでは, プログラムを組むように, 人は意図を表す木構造に対し基本演算を適用していく。認知的リアリティを持つ表現法と基本演算という制約のもとで, 基本演算として, これまで簡約操作(reduction), join/meet を導入してきた [19, 20] (図4)。直感的には, reduction は部分集合の計算に, join は和集合の

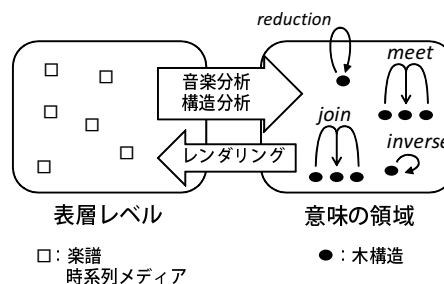


図 4 代数的枠組を持つ音楽理論

計算に, meet は積集合の計算に対応付けられる。ここまででは半群でしかない。現在, さらに演算体系の記述力を上げるために, 補元(正確には相対擬補元 [5])を計算する演算子 inverse を検討している。もちろん, 認知的リアリティに関する制約を満たさなければならない。inverse の導入によって群となり, 反対例から遠ざかる修飾, 類推に基づく編曲(正比例)などのアルゴリズムが実現できる。

3.3 ピッチを考慮したタイムスパン木による旋律間類似度

我々は GTTM タイムスパン木生成において曖昧であった音高と和声を考慮した選好ルールを拡張し, ピッチを考慮したタイムスパン木(pitch sensitive time-span tree, 以降 pTS 木)を提案した [13]。まず pTS 木において音高間の距離を計算するために, GTTM の後継である音楽理論 Tonal Pitch Space [11]におけるピッチクラス階層性の枠組みを用いた。次に和音間の距離を計算するために, 3.1節で触れたローカルキーの概念を用いた。pTS 木は従来のタイムスパン木に対する演算体系(reduction, join/meet 演算子)をそのまま適用できるため, 従来のタイムスパン木同様に, 音高と和声を考慮した pTS 木間の距離が定義され, 旋律間類似度が計算できるようになる。聴取実験により心理的な距離と比較したところ, より高い認知的リアリティを持つ表現であることが確認された [14]。

現状では, pTS 木間の距離定義を与える際には, ローカルキーの支配域は最大タイムスパンに対応すると仮定している(3.1節)。しかし, 音価が最も短い簡約レベル(例えば16分音符)までローカルキーが存在すると考えるのは人間の直観にあっていない。ローカルキーの支配域については今後の検討課題である。

3.4 確率的生成モデルとしての GTTM の再解釈

GTTM は音楽の三大要素である旋律・リズム・和声の規則に関する包括的な理論体系である。音楽構造を分析するために与えられている選好規則は, 「~と分割しても良い」, 「~がグループを作ることが望ましい」というような成立しない場合を許容するタイプの言明である。実際に音楽構造を分析(簡約)してみると複数の選好規則を同時に適用できない場所や状況が発生する。選好規則はこのような競合

を解消する特性を本来に備えており、照合の程度が低い
選好規則は適用されない。人が選好規則の適用/不適用の
判断をする時は専門家の暗黙的な知識を用いる。しかし、
計算機による自動分析を実現するためには、規則間の競合
を決定的に解決するため暗黙的な知識を明示的なパラメ
ータとして導入し、手動で適切に設定する必要があった。そ
のパラメータを自動獲得する試みも行われているが、分析
精度の面で依然として改善の余地がある [2,9]。

この問題を解決するため、我々は統計的機械学習の見地
に立ち、確率的生成モデルとして GTTM を再解釈するこ
とが有望であると考えている。特に、GTTM の核をなす
タイムスパン簡約の確率モデル化と楽譜データに基づくモ
デルパラメータの教師なし学習に取り組みたい。具体的
には、ある開始記号が階層的に分割されることで、時間・音
高空間内の音符配置が確率的に生成される過程を考える。
このトップダウンな生成過程を順問題と捉えると、タイム
スパン簡約は音符配置からボトムアップに木構造を獲得す
る逆問題として定式化できる。すなわち、簡約規則は逆方
向の生成規則として再定式化され、簡約規則の競合を決
定的に解決するため用いられた選好規則は、生成規則の確
率（使われやすさ）として音楽の確率的生成文法に自然に
取り込むことができる。

確率的な枠組みを用いれば、大量の楽譜データに対する
尤度最大化という明確な基準のもとで、旋律・リズム・和
声の生成規則の確率を自動的に学習できる。さらに、音楽
情報処理分野でも最近注目されているノンパラメトリック
ベイズモデルに基づく確率的生成文法を定式化すれば、生
成規則自体を自動的に獲得することも可能であると考えて
いる。実際、いくつかの研究 [17,21] では、音楽データか
ら音長・音高空間に二次元拡張された確率的文脈自由文法
(PCFG) を学習しようとする試みがなされている。本研究
では、GTTM の枠組みを生かし、旋律・リズム・和声の包
括的な確率モデルを構築することに取り組む。この結果、
既存の旋律間モーフィング [3] では扱うことが難しかった、
一見それほど類似していない旋律間のモーフィング、さら
には和音を含む複旋律間のモーフィングも可能になると考
えられる。

3.5 Wavelet による旋律構造の階層表現

我々は、音楽の非専門家は旋律を音符レベルではなく
もっと大まかな形で理解しているという仮説を立て、旋律
の音符表現によらない認知表現を探索している。我々はす
でに、音符レベルの情報を明示せずに 1 本のなめらかな曲
線（旋律概形）で旋律を近似する方法を提案している [22]。
本手法では、音符列として与えられる旋律から音高の時系
列を求め、それに対してフーリエ変換を適用して低次の
フーリエ係数のみ残して逆フーリエ変換を行うことで、旋
律概形を抽出する。

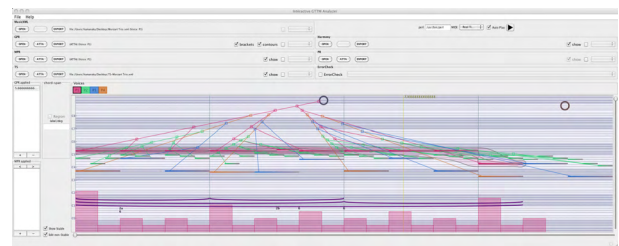


図 5 複旋律楽曲分析器のプロトタイプ

我々は、旋律同士を概形レベルで演算する方法の確立を
目指している。旋律には階層的な拍節構造があり、拍節構
造は二分木で自然に表されるので、旋律表現も二分木で表
すことが自然である。そこで、旋律概形を多重解像度解析
（離散ウェーブレット変換を再帰的に繰り返す方法）に基
づいて再定義することを試みている [10]。この処理によっ
て、旋律概形のパラメータを簡約概念を内包する二分木と
して表現できると、3.2 節で述べたような木構造に対する
種々の演算子が適用できるようになる。

4. 音楽における理論的枠組みの応用

本章では、第二の課題である音楽における理論的枠組み
の応用について述べる。構築した理論体系を一般ユーザが
自在に利用できるアプリケーションとして実装するには、
適切な可視化インタフェースと直感的な操作体系の開発が
必要である。

4.1 自動変奏曲生成システム

先の 3.3 節で述べた pTS 木の演算体系を用いて、音高と
和声を考慮した編曲システム構築を考えている。既存の最
大タイムスパンによる GTTM タイムスパン木編曲アルゴ
リズム [3,8] に音高や和声を考慮する改良を加えることで、
和声構造を維持した旋律の変奏などのより音楽的な編曲が
実現可能となる。また階層的半順序構造であるタイムス
パン木の特性を活かして、簡約 (reduction) とその逆演算
(elaboration) を用いた編曲も検討中である。

4.2 自動音楽分析器

音楽理論の計算機上への実装が難しいのは、一般に、音
楽理論に基づく楽曲の解釈が一意に定まらないためであ
る。その原因には次の 2 つが考えられる：音楽家が楽曲を
分析することを目的として書かれた音楽理論それ自体に
曖昧性が含まれていることと、1 つの楽曲でも複数の解釈
ができるという解釈の多義性による曖昧性である。我々は
音楽理論 GTTM に基づく楽曲分析器 ATTA (Automatic
Time-span Tree Analyzer) を実現するにあたって、後者
の曖昧性は積極的に認めつつ、前者の音楽理論の曖昧性を
なるべく解消することを目指した。具体的には、暗黙的な
知識を多くの調節可能なパラメータとして外在化した [2]。

ATTA ではタイムスパン木までの分析を実現したが、殆どのパラメータは手動で調節していたため分析に多大な時間がかかったこと、楽曲によっては分析精度が低かったことが課題として残った。この点については、確率的な手法の導入による分析の自動化を試みている [9]。また、GTTM では分析対象を単旋律 (monophony, homophony) に限定し複旋律 (polyphony) を対象外としている。ATTA も複旋律を対象外としたが、現在は応用範囲の拡大を目指して、複旋律にも対応した分析器の構築を進めている [4] (図 5)。

4.3 旋律概形を利用した旋律生成

先の 3.5 節で述べた旋律概形は、音楽の専門知識に依存しない旋律表現であるため、非専門家でも容易に操作できる。現在開発中のシステムでは、ユーザがマウスなどを使って旋律概形を描き直すと、それが即座にフーリエ変換またはウェーブレット変換のパラメータに変換され、それに基づいて新たな旋律が生成されることを目指している。

本プロジェクトによって旋律概形間の演算体系を構築できれば、旋律のモーフィングなどを実現することができよう。既存の旋律間モーフィング [3] では、タイムスパン木に共通部分がないと (ある程度類似した旋律間でないと) 適切にモーフィングを実現できなかった。しかし旋律概形をベースにすることで、それほど類似していない旋律間でも柔軟にモーフィングを実現できる可能性がある。

5. 他の時系列メディアへの展開

本章では、第三の課題である音楽における他の時系列メディアへの展開について述べる。音楽と言語はいずれも木構造で表現できると考えられるが、基本とする単位 (音符と単語、フレーズと文) など異なる点が多々存在するため、その差異を解消あるいは吸収する技法が必要となる。

5.1 ディスカッション構造の学習

ディスカッションは木構造で表すことができる。名古屋大学長尾研究室では、約 8 年間に渡る研究室でのディスカッションを記録し分析している [16]。ディスカッションでは、スライドを用いたプレゼンテーションに対して、議論参加者が意見を言い、発表者がそれに応え、他の参加者も関連した意見を言う。この結果、スライド (のあるページ) をルートとし各発言がルートや先行する発言に結び付くような木構造を形成する (図 6)。ここで、新しく話題を提示する発言を導入発言 (start-up)、以前の発言と同じ話題について意見を言う発言を継続発言 (follow-up) と呼ぶ。1 個の導入発言と 0 個以上の継続発言が話題ごとのまとめ (それを議論チャンクと呼ぶ) を形成する。

また、各発言をクラスタリングし、より抽象度の高いシンボルとして扱うために、word2vec を応用した *statement2vec* の開発を進めている。発言に含まれる単語列を

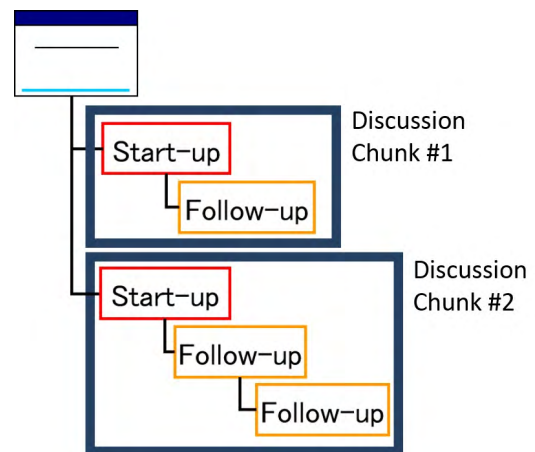


図 6 議論の木構造

ベクトル空間にマッピングすることにより、ベクトル空間上で k-means アルゴリズムに基づくクラスタリングを適用できる。さらに、クラスラベル付きノードを持つ木構造の集合から、ディスカッションに関する確率文脈自由文法 (PCFG) を学習することを目指している。

5.2 Q&A 型議事録システム

人は、会議の各発言の間にある賛成/反対や質問/回答などの関係に基づき、隣接する発言間だけでなく離れた発言間の関係も考慮して、会議の内容つまり議論の構造を理解する。発言や発言列にはある観点から見て重要なものやそうでないものがあるので、我々はその重要度に基づく発言列に関する簡約の概念を定義し、会議の発言列を分析して簡約の概念を内包する議論の木構造 (議論タイムスパン木) を導く方法を提案した [15]。議論タイムスパン木が会議の意図を表現していると考えれば、プログラミング的なアプローチ (第 1 章) を適用して、発言列に対して意図を込めた変形を施すことが可能となる (図 2 右)。応用として、結論に大きく寄与した発言の抽出、質疑応答が続いて盛り上がった箇所の同定、ある提案に対する反対意見の列挙、あるいは、発言力のある出席者の同定、前回と同じ議論を繰り返してしまった部分の抽出などが考えられる。

我々は、システムとユーザがインタラクティブに Q&A を繰り返しながら会議に関する情報を抽出するような、対象を発言録に特化した検索エンジンの開発を目指している (Q&A 型議事録システムと呼ぶ)。会議の発言録に対して問い合わせ (Q) を行い、その結果 (A) を受け、さらに新たな問い合わせを繰り返すことで、その会議をパーソナルな視点から振り返ることができる。現在は、関係者全員に同一の議事録を配布するだけの場合が一般的であるが、Q&A 型議事録システムはその人のためだけのパーソナルな議事録を作成することが特徴である。パーソナルな議事録には、その人が興味を持った議論の記録が含まれていたり、その人が押さえておくべき要点が含まれている。

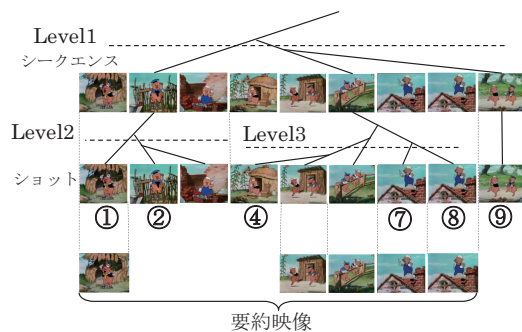


図 7 タイムスパン木を用いた映画の要約

Q&A 型議事録システムを実現するための研究課題としては、発言や発言列の重要度を決定する観点の指定法・制御法、正確な議論の木構造を生成する方法、議論の木構造を変形するための基本演算子の定義、ユザビリティ向上のため議論の木構造の可視化法などが挙げられる。

5.3 映画の要約

時系列メディアの1つである映画を対象として、それを要約するための構造化に取り組んでいる。個人や状況によって目的の異なる要約が必要とされている。例えば、大体的な内容が理解できればよい、自分の定めた制約時間長にしたい、自分の関心の強い部分だけをまとめて視聴したいなどである。我々は、映画のストーリーにおける重要な場面と、各場面の従属関係をタイムスパン木で表現し、そのタイムスパン木を編集することで個人の目的や状況に合わせた柔軟な要約の生成を目指している [18] (図 7)。

6. 今後の展望

計算機技術の歴史は、実世界を仮想化する歴史といえる。まず、CPU・記憶装置・ネットワークなどのハードウェアが仮想化され、次に、UI やコンテンツなどのソフトウェアが仮想化された。これにより、機能やインターフェースが標準化され、分割・コピー・統合・変換が容易になってきた。昨今のクラウドはその仮想化技術に根差したサービスであり、FabLab は仮想化の恩恵を受けたもの作りである。我々が興味を抱くのは、計算機科学の新たな歴史として、コンテンツ制作の仮想化、あるいは音楽制作の仮想化である。意図を込めて二つのコンテンツを混ぜる。ある意図をもってコンテンツの一部を取り出したり、一部を修正する。そのような処理において、人はどんな「計算」をしているのだろうか。我々は、その本質を見極めることに取り組みたい。いつか、多くのユーザがメディアの種類や細かいテクニックなどを気にすることなく、自分の意図をコンテンツに込めることができる時代がやってくると信じている。

謝辞: 貴重な議論をしていただいた寺澤洋子氏(筑波大学)に感謝する。本研究の一部はJSPS 科研費 26280089, 25330434, 26700020 の助成を受けた。

参考文献

- [1] Goutsias, H., Heijmans, H. J.A.M. (Eds): *Mathematical Morphology*, IOS Press (2000).
- [2] Hamanaka, M., Hirata, K., Tojo, S.: Implementing "A Generating Theory of Tonal Music", *Journal of New Music Research*, Vol.35, No.4, pp.249-277 (2006).
- [3] Hamanaka, M., Hirata, K., Tojo, S.: Melody Morphing Method Based on GTTM, *ICMC*, pp.155-158 (2008).
- [4] Hamanaka, M., Hirata, K., Tojo, S.: Toward Developing a Polyphonic Music Time-Span Tree Analyzer, *MCM* (2013).
- [5] 平田圭二, 東条敏: 楽曲構造束とその上の演算系, 人工知能学会全国大会(第 20 回) 論文集, 1D2-4 (2006).
- [6] Hirata, K., Tojo, S., Hamanaka, M.: Cognitive Similarity Grounded by Tree Distance from the Analysis of K.265/300e, *CMMR*, pp.415-430 (2013).
- [7] 平田圭二, 東条敏: パースタインの「答えのない質問」再考: 計算論的音楽の理論の枠組みについて, 人工知能学会全国大会(第 28 回) 論文集, 1K4-OS-07a-1 (2014).
- [8] Hirata, K., Tojo, S., Hamanaka, M.: Algebraic Mozart by Tree Synthesis, *ICMC/SMC*, pp.991-997 (2014).
- [9] Kanamori, K., Hamanaka, M.: Method to Detect GTTM Local Grouping Boundaries based on Clustering and Statistical Learning, *ICMC* (2014).
- [10] 北原鉄朗, 土屋裕一: 旋律を簡約・操作する一手法, 人工知能学会全国大会(第 28 回) 論文集, 1K4-OS-07a-3 (2014).
- [11] Lerdahl, F.: *Tonal Pitch Space*, Oxford University Press (2001).
- [12] Lerdahl, F., Jackendoff, R.: *A Generative Theory of Tonal Music*, The MIT Press (1983).
- [13] 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 ~ パッハ BWV582 パッサカリアとフーガの分析を例に ~ , 人工知能学会全国大会(第 28 回) 論文集, 1K4-OS7a-2 (2014).
- [14] Matsubara, M., Tojo, S., Hirata, K.: Distance in Pitch Sensitive Time-span Tree, *ICMC/SMC*, pp.1166-1170 (2014).
- [15] 三浦寛也, 長尾確, 平田圭二, 音楽理論 GTTM に基づく議論タイムスパン木の生成方式とその評価, 情報処理学会論文誌, Vol.56, No.3 (2015).
- [16] Nagao, K., Kaji, K., Yamamoto, D., Tomobe, H.: Discussion Mining: Annotation-Based Knowledge Discovery from Real World Activities, *Advances in Multimedia Info. Proc.*, Aizawa, K., Nakamura, Y., Satoh, S. (Eds), Springer, LNCS, Vol.3331, pp.522-531 (2005). (Also presented in PCM 2004)
- [17] Nakano, M., Ohishi, Y., Kameoka, H., Mukai, R., Kashino, K.: Bayesian Nonparametric Music Parser, *ICASSP*, pp.461-464 (2012).
- [18] 竹内星子, 浜中雅俊: 音楽理論に基づく映画の構造化, 人工知能学会全国大会(第 28 回) 論文集, 1K5-OS7b-4 (2014).
- [19] Tojo, S. and Hirata, K.: Structural Similarity Based on Time-span Tree, *CMMR*, pp.645-660 (2012).
- [20] Tojo, S. and Hirata, K.: Distance and Similarity of Time-span Trees, *Journal of Information Processing*, Vol.21, No.2, pp.256-263 (2013).
- [21] 土屋政人, 落合和樹, 亀岡弘和, 嵯峨山茂樹: 楽譜の文脈自由 2 次元木構造表現に基づく多重音スペクトrogram 生成モデルによる音響信号からの自動採譜, 情報処理学会音楽情報科学研究会, 2012-MUS-96, (1) (2012)
- [22] 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).