

機関番号：13302

研究種目：基盤研究（B）

研究期間：2008～2010

課題番号：20300035

研究課題名（和文） 生成的音楽理論に基づく自動楽曲分析器の構築

研究課題名（英文） Implementation of Automatic Analyzer Based on Generative Theory of Tonal Music

研究代表者

東条 敏 (TOJO SATOSHI)

北陸先端科学技術大学院大学・情報科学研究科・教授

研究者番号：90272989

研究成果の概要（和文）：本研究課題ではGenerative Theory of Tonal Music (GTTM)を計算機上に実装し、グルーピング解析・拍節構造解析・タイムスパン木生成の自動化を行った。この過程で解析に伴うパラメータのチューニングシステムを構築し、インタフェースを整え、また解析例をデータベースにした。さらに和声解析と結び付けるためのシステム、旋律を計算論的に加工する束上のモーフィングシステムを開発した。

研究成果の概要（英文）：According to the Generative Theory of Tonal Music (GTTM), we have implemented the grouping analysis, the metrical analysis, and the time-span tree analysis. In this process, we have automatized the parameter tuning, designed user interface, and accumulated the resulting analyses. Furthermore, we have developed the chord analysis system and the morphing system on the melody lattice.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	4,800,000	1,440,000	6,240,000
2009年度	3,400,000	1,020,000	4,420,000
2010年度	1,300,000	390,000	1,690,000
年度			
年度			
総計	9,500,000	2,850,000	12,350,000

研究分野：情報科学・人工知能

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：音楽情報処理, 生成論的音楽理論, 計算論的音楽理論

1. 研究開始当初の背景

Generative Theory of Tonal Music (GTTM) は R. Jackendoff と F. Lehrdahl によって 1983 年に提唱された音楽の構造解析の理論である。提唱されて以来既に四半世紀が経過するが、音楽認知・音楽情報処理の多くの研究で今なお言及され続けている。その理論の特徴は以下のようにまとめることができる。

- 楽曲を構成する要素、音はその近隣の音との間で依存関係があり、重要度という観点から比較可能である。より重要な音をヘッド(head)として選び徐々に音の数

を減らしていくと、楽曲は簡易な構造に還元可能(Reduction Hypothesis)である。

- 楽曲を構成する音はその重要度において階層をなし、楽曲全体は木構造(tree)をなす。
- この木構造はその音楽の基本的な構造(basic form; Ursatz)を反映するものであると同時に、人間の音楽認知の過程においてもその音楽の本質(gestalt)を反映するものである。

GTTM は次の4つのサブ理論から構成され、互いに依存性がある。

- (i) グループ構造解析
- (ii) 拍節構造解析
- (iii) タイムスパン木の生成
- (iv) プロロンゲーション木の生成

グループ構造解析とは位置的に近い音群をまとめてグループとみなす処理である。拍節構造解析とは拍(beat)の乗る位置を同定する処理である。タイムスパン木はこの前二つの処理結果に基づき拍節的安定性を表示する木であり、プロロンゲーション木は和声的安定性を表示する木である。

GTTM は以上の解析のためにさまざまな構造構成規則を提示している。この規則群は well-formedness rule (文法規則) と呼ばれる厳格な遵守を要求する規則と preference rule (選好規則) と呼ばれる「好ましい」構造を記述する規則からなる。このうち選好規則は強要しないことがかえって適用の可否を曖昧にしておき、複数の選好規則が両立しない場合にどのように解決するかはもとの理論には記述されていない。よって GTTM は手続き的な規則群で構成される理論でありながら計算機システムへの実装は困難とされてきた。実際これまで多くの試みがなされ、その一部の規則のみが実装例として報告されることはあったが、木構造の創出までを機械化した例はかつてなかった。

2. 研究の目的

われわれは 2004 年以來 GTTM の実装に取り組み、選好規則のコンフリクトに対して各規則に重み付けのパラメータを付与することにより、世界で初めてタイムスパン木と呼ばれる構造を創出できることを示した。このうち、われわれがパラメータ化によって実現したのは前三つ、すなわちグループ構造解析・拍節構造解析とそれに基づくタイムスパン木の創出までである。プロロンゲーション木の創出は未着手であり、かつまた対象となる音楽構造にも制限がある。

本研究の目標は GTTM 全体をコンピュータシステムに実装し、音楽情報処理を行うコミュニティに解析ツールとして提供することである。この目的のため、本研究では 2008 年度より 3 年をかけて、次の課題を解決する。

- (a) 自動パラメータチューニング。曲に応じて最適なパラメータセットを獲得するには膨大なパラメータ空間を人手で探索する必要があった。自動化はこの手間を短縮するのが目的である。しかし自動チューニングのためには、数多くの楽曲の解析例からチューニング・アルゴリズムをトレーニングする必要がある。したがってまず解析の正解データを収集する必要があり、これにはプロの音楽家の手になるデータ作りが前提となる。
- (b) 和声的安定性の概念の実装。タイムスパン

木はグルーピング解析と拍節解析の結果を元に、主に拍節的な安定性を反映した音楽構造である。しかし音楽を楽句の連鎖と見るマクロ的な視点からは、メロディや和音の連鎖の安定性 (pitch stability) の概念が必要であり、これをカデンツ (cadence) の処理とプロロンゲーション木生成において実装する。

- (c) ユーザインタフェースの構築。ATTA は GTTM の選好規則間で重み付けを変更するためにパラメータチューニングを行うインタフェースと木構造のビューワーを持つ。この ATTA にパラメータの探索システム、プロロンゲーション木への改編システムを追加し、システムを公開し使用に供することを目的とする。一般ユーザが使用可能であるためにはポータビリティとユーザ・インタフェースを大幅に改善する必要があり、これを最終段階の目標とする。

3. 研究の方法

本申請の研究計画では、最初の 3 つのサブ理論の完成と、残りの和声解析部の実装を 3 年かけて行う。GTTM の 4 つのサブ理論のうち (i) グループ構造解析 (ii) 拍節構造解析 (iii) タイムスパン木の構成は部分的に試作されてきた。この解析が目的とするところは、楽譜から隣どうしの音の重要度を比較して、ボトムアップにヘッドが選択されて木構造が構成されるものである (図 1)。

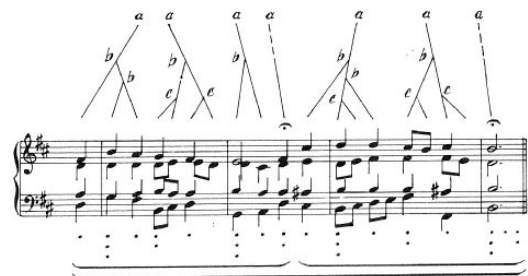


図 1 タイムスパン木の生成

研究開始時点での解析器の構成は、楽曲の XML ファイルを入力として、グルーピング処理・拍節解析処理を施した上にタイムスパン木を出力する。

① 2008 年度

(i) ~ (iii) のサブ理論で未実装部分に着手する。

(a) 自動パラメータチューニング

ATTA は現在ではウィンドウシステムのスライディングスイッチのインタフェースを持つ。これは画面上手動でボタンを左右にシフトすることによって各パラメータの重み付けを変えるしくみであるが、この方法では最適木の探索に膨大な時間を要する。このた

め、20年度では GTTM に経験のあるデータ試作者を募り、楽曲の解析例を蓄積する。

(b) 和声的安定性の実装

和声的安定性(pitch stability)は本研究における最重要課題である。この研究は次の3ステップからなる。

- 和音境界認識
- 和音認識
- カデンツ処理

2008年度はこのうちまず和音認識、すなわちある短いタイムスパンにおける一グループの音群に対し三度と五度の関係を発見し、調と和音を同定する。この和音認識は係留音(suspension)や倚音(appoggiatura)の処理まで含めると困難なタスクである。なお2008年度では(c)ユーザインタフェースは将来の公開を前提とするレベルで準備を始める。

②2009-2010年度、

(a) 自動チューニング

21年度以降はフィードバック規則にホモフォニーに関する処理を加え、FATTAの精度を上げることを目標とする。

(b) 和声的安定性の実装

まずホモフォニーへの拡張処理として単旋律としてのソプラノ部認識に続いて、他の声部をソプラノ部への和声的サポートとしてみなし一つの和音を単音(pitch event)に代わる楽曲の単位として拍節構造規則・タイムスパン木生成規則を実装する。特に各和音のバスが明らかになる場合、選好規則MPR6(バス部の強拍処理)の実装が可能である。また先に述べたとおり、和音認識は次項に述べる和声的安定性の準備とする。和声的安定性の処理の本質部分はカデンツ処理である。人間が楽句(phrase)を認識するしくみとして「曲が終結する感じ」を規則にまとめたものがカデンツである。本研究ではカデンツをサポートするため楽句末尾の和音をまとめる処理を行う。すなわち拍節構造処理のカデンツ処理、係留音処理を実装し、タイムスパン木の構成に関わるカデンツ処理(cadential retention)の選好規則の実装を行う。タイムスパン木における楽句の終止形(cadence)はカデンツ処理をして複数の和音を単一の音としてみなす処理をする必要がある(図2)。

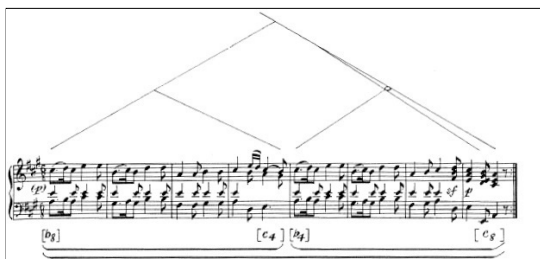


図2 モーツァルト K331 のカデンツ処理

(c) ユーザインタフェースの構築

本研究の終了時には、XML 入力された楽曲に対して、その楽譜・適用されたグループ解析/拍節構造解析の規則を構築する。このタイムスパン木を得る過程においては、いくつか既得のパラメータセットから特定のセットを選択した上、さらに手動でパラメータの再セットができるものとする。タイムスパン木の表示部分は簡単なマウス操作でプロローゲーション木に再構成されるものとする。本申請の最終的な課題はこのような解析ツールを音楽情報処理の共通なプラットフォームとして提供することであり、ポータブルであり、かつあらゆるユーザにとって使いやすいインタフェースシステムの設計・開発を目標とする。

4. 研究成果

(1) 言語と音楽の構造認知

小鳥の発声は紛れもなく彼らのことばであり、それを歌というのはわれわれの勝手なレトリックである。彼らは鳴管(咽喉に相当する器官)を使って発声し、それによってコミュニケーションを行う事実には変わりがない。ジュウシマツの歌はチョムスキー階層における正規文法であり、有限状態オートマトンで表現されることが知られている。人間の言語が他の生物の言語と決定的に異なるのは、それが階層的な構文を持つということである。この階層的という概念をもとに本稿では言語の構文認識と同じしくみが音楽の構造認識にも用いられる可能性を講演した。

(2) σ GTTM

本研究では、音楽理論 GTTM と、統計的学習を組み合わせた音楽理論 σ xGTTM の構築を目指して、まずグルーピング境界の検出に取り組んだ。我々はこれまで、GTTM を計算機実装用に拡張した exGTTM を提案した。しかし exGTTM で正しい分析結果を得るためには、適切なパラメータを曲ごとに手動で設定しなければいけない困難な作業が必要であった。本研究では、局所的なグルーピング境界の自動検出をする際、データの集合を法則化する数学的手法である決定木によって、ルールの優先度を求めることでルールの競合が起こらないようにした。また不確実性を含む事象の予測が可能な Bayesian Network によって、EM アルゴリズムによる学習をすることでグルーピング境界の判定基準の曖昧さを扱うことを試みた。実験の結果、本手法を実装した検出器はベースラインでの exGTTM の性能を上回ることを確認した。

(3) FATTA に基づくメロディ予測

本研究では、自動楽曲分析器 FATTA

(Full-Automatic Time-span Tree Analyzer)に基づき、現在演奏中のメロディ（単旋律）の後続音を予測する手法を提案した。従来多くのメロディ予測手法では、統計的学習手法に基づいていたために、学習データに含まれていないメロディに対しては適切でない予測となる可能性があった。これに対し本研究では、音楽理論に基づきメロディの適切さを評価し予測を行った。具体的には、音楽理論 GTTM および TPS (Tonal Pitch Space) に基づき構築した自動楽曲分析器 FATTA を用いて、複数存在する後続音の候補に対してメロディの安定度を評価した。本手法の特長は、メロディの表層構造だけでなく GTTM および TPS に基づき得られる音楽の深層構造を利用してメロディの安定度を評価している点である。実験の結果、提案手法がメロディの後続音を適切に予測していることを確認した。

(4) Tonal Pitch Spaceの実装

楽曲の和声は、各和音がそれぞれもっている調や機能といった属性や、和声の規則的な繋がり（カデンツ）などの要素より成り立っている。和音は単に時間的な随意に選択された音高の集まりではなく、一般的には音楽理論から導かれるいくつかの制約のもとで構成音が選択されている。中でも重要な制約として調が存在し、構成音は音階内より選択される。この調についての情報は基本的には楽譜中の調号によって示される。しかし、すべての転調が調号によって記述されているわけではなく、実際の楽曲中には暗黙的な短い転調が多く含まれている（図3）。



図3 短い転調の例

このような和声を構成する要素の解析は音楽情報処理の基礎であり、計算機による編曲、採譜、曲検索などに応用が期待できる。これら和声的要素の解析を行うための手法は多く提案されているが、いずれも解析精度において改良の余地が残っている。これらの問題の根源的な原因は音楽理論自体の不整備である。特に和声学は実践理論として作曲者に向けて構成、発達したものであり、既存の楽曲を解析するという観点から省みられることが少なかつた。その結果、和声学自体も完全なものではなく、曖昧性を含むと同時に、作曲の過程において敢えて規則が無視されることもあり、計算機による解析において前述のような問題を生み出すこととなったといえる。そこで本研究では、従来の和声学と近似していな

がら、より全域的且つ厳格に音高間、和音間、調間の遠近を論じることのできる音楽理論である Tonal Pitch Space (TPS) の和音間距離（図4）の概念を導入することで、以上の問題を解決しつつ、和音列の調性、機能、カデンツ構造を総合的に認識する和声進行の解析手法を提案した。

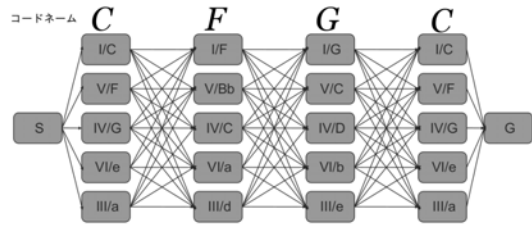


図4 和音間距離のネットワーク

(5) 旋律のインタラクティブな生成

本研究では、類似した楽曲構造を持つメロディの生成手法として、元のメロディと同じ局所的なグルーピング構造を持つメロディの生成方法を提案した。類似した楽曲構造を持つメロディの生成は、雰囲気や特徴が似ているメロディを作るという直感的な作曲に役立つ。本手法の特長は、グルーピング構造の変更を、本手法により提示された修正候補を選択することでインタラクティブに行う点である。本手法を用いて、元のメロディと同じグルーピング構造を持つ新しいメロディが生成できることを確認した。

(6) モーフィングアルゴリズムの検証

GTTM/ATTA によって生成されるタイムスパン木は編曲やモーフィングにおいて有用である。簡約 (reduction) というのは本来一つの曲に対して定義されている構造の簡素化の過程であるが、本研究ではこの概念を敷衍し、複数の木の間にも簡約と同様な順序関係を仮定した。次に木の間で meet と join が定義されよって束 (lattice) を構成するように定式化を行う（図5）。これにより複数の木から新たな楽曲が創成できるための条件を考察した。

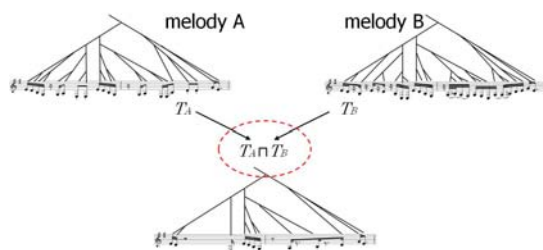


図5 ラティス上の演算

(7) タイムスパン木の単一化可能性

GTTM/ATTA によって生成されるタイムスパン木は編曲やモーフィングにおいて有用であ

る。簡約(reduction)というのは本来一つの曲に対して定義されている構造の簡素化の過程であるが、本研究ではこの概念を敷衍し、複数の木の間にも簡約と同様な順序関係を仮定した。次に木の間でmeetとjoinが定義されよって束(lattice)を構成するように定式化を行った。これにより複数の木から新たな楽曲が創成できるための条件を考察した。

(8) 木構造を用いたメロディ生成

本研究では、曲の雰囲気や特徴を変えずに、楽曲構造に基づきメロディ中に新たな音を追加して、メロディを生成する手法について考察した。本手法の特長は楽曲構造を利用したことであり、その結果、生成されたメロディは元のメロディの雰囲気を忠実に反映することができた。提案手法では、まず、メロディを音楽理論 GTTM に基づき分析し、タイムスパン木を獲得する。次に、得られたタイムスパン木に、既存の曲のタイムスパン木を参考にして、新たな枝を追加することで、音数の増加した新たなメロディを生成する。実験の結果、簡約したメロディから楽曲構造に沿って具体的なメロディを生成できることを確認した。

(9) メロディの内挿による作曲支援

本研究の目的は、音符・休符だけでなく、音楽構造を考慮したメロディ操作により音楽初心者でも直感的に作曲ができる作曲支援システムの構築をすることである。ここで音楽構造とは、音楽理論 GTTM に基づく楽曲分析の結果得られる木構造で、メロディを階層的な木構造で表し、本質的な部分(構造的に重要な部分)と装飾的な部分に分けることを可能としている。木構造は音楽構造であり、根を上側としており各枝の音の構造的な重要度を表している。従来の作曲用ソフトウェアでは、楽譜上の音符や休符などの表層的な部分の操作に限定されており、その適切な操作には専門知識が必要であった。

音楽知識のないユーザであっても作曲が可能な手法として、あるメロディに他のメロディの特徴を加え新たなメロディの生成を可能とするメロディのモーフィング手法(以下、内挿手法)がある。ここで類似したメロディの検出を可能とするため、我々は音楽構造に沿ってメロディを分割し、メロディの部分的な類似部分を検索することを考える。従来、音楽構造に基づくメロディの部分的類似度の算出手法が提案されていたが、検索する長さの範囲を手動により設定しており、検出数が設定の仕方に依存してしまう問題があった。また、この手法では類似度の算出に、各音の重要度を計算していないため、類似度の高いメロディが必ず本質的な部分で類似したメロディにはならない。さらに、この手法

では2つのメロディで類似とする条件が厳しく、類似したメロディの検出は困難である。

本研究では、より多くのメロディを検索するため、検索するメロディの長さの範囲を設定せず、検索対象のメロディの長さに対し、比較するメロディの分割が正しいかを判断した。具体的には、あるメロディを分割していき、検索対象のメロディの長さより短いメロディになった時、比較を行う。この時、そのメロディの長さが分割前の長さと比較して近い方のメロディを類似度判定の対象とする。内挿にはメロディの本質的な部分が類似したメロディが必要となるため、類似度には音楽構造を考慮した重要度から各音の重み付けを導入した。さらに、システムを携帯アプリとして利用し、操作の抽象度を上げつつ、音符1つずつの操作も行えるユーザインタフェースの実装を試みた。実験の結果、内挿可能なメロディの検出数の向上を確認した。

(10) タイムスパン木の meet/join

単一化不可能な2つの旋律から1つの旋律を合成する問題を考える。そのような2つの旋律(タイムスパン木) S_1, S_2 に対して $S_1 \subseteq \text{join}(S_1, S_2) \wedge S_2 \subseteq \text{join}(S_1, S_2)$ (*) という条件が成立するような join 演算を提案した。当該問題への対処法として、われわれはこれまで単一化に幾つかの拡張を加えることを試みてきた。しかしこのアプローチでは上の条件(*)を成立させることが難しい。そこで単一化アルゴリズムを拡張するのではなく、タイムスパン木の表現法を工夫するアプローチを検討し、タイムスパンを表現する構造化音長という型を導入した。これより、join 演算の適用範囲を広げその有用性を高めることができた。一方 meet 演算に関しても同様の議論を展開することができた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

浜中雅俊. 音楽メディアの情報処理と感性的な利用. 電子情報通信学会誌 92(11) 2, pp. 958-9, 2009. 査読有

[学会発表] (計19件)

(1) K. Hirata, S. Tojo, M. Hamanaka. Melodic Morphing Algorithm in Formalism. Mathematics and Computation in Music. 2011.6.17. IRCAM, Paris, France.

(2) 平田圭二, 東条敏, 浜中雅俊. タイムスパン木の join と meet について. 第88回

- 情報処理学会 音楽情報科学研究会.
2010.12.4. 昭和音楽大学 (神奈川県川崎市)
- (3) 矢澤櫻子, 寺澤洋子, 平田圭二, 東条敏, 浜中雅俊. 暗意実現モデルにおける基本類型を用いたメロディ構造分析. 第87回 情報処理学会 音楽情報科学研究会. 2010.10.14. KDDI 研究所 (埼玉県ふじみ野市)
- (4) 西田智, 浜中雅俊, 平田圭二, 東条敏. 音楽理論 GTTM に基づく木構造を用いたメロディ生成手法. 第86回 情報処理学会 音楽情報科学研究会. 2010.7.30. つくばグランドホテル (茨城県つくば市)
- (5) 東条敏, 平田圭二, 浜中雅俊. タイムスパン木の単一化可能性. 第86回 情報処理学会 音楽情報科学研究会. 2010.7.30. つくばグランドホテル (茨城県つくば市)
- (6) S. Onuma, M. Hamanaka. Piano Arrangement System Based on Composers' Arrangement Processes. International Computer Music Conference. 2010.6.2. Stony Brook, New York, USA.
- (7) 平田圭二, 東条敏, 浜中雅俊. 旋律モーフィングアルゴリズムの形式的検証. 第85回 情報処理学会 音楽情報科学研究会. 2010.5.27. 東北大学 青葉山キャンパス (宮城県仙台市)
- (8) 村上雄一, 浜中雅俊. メロディの内挿手法に基づく作曲支援システム. 情報処理学会 第72回全国大会. 2010.3.9. 東京大学
- (9) 西田智, 浜中雅俊, 平田圭二, 東条敏. 類似した楽曲構造を持った旋律のインタラクティブな生成方式. 情報処理学会 第84回音楽情報科学研究会. 2010.2.16. 関西学院大学
- (10) M. Hamanaka, S. Tojo. Interactive Gttm Analyzer. 10th International Conference on Music Information Retrieval conference (ISMIR2009). 2009.10.27. 神戸
- (11) Y. Miura, M. Hamanaka, K. Hirata, S. Tojo. Use of Decision Tree to Detect GTTM Group Boundaries. International Computer Music conference (ICMC2009). 2009.8.17. McGill University, Montreal, Canada
- (12) M. Hamanaka, K. Hirata, S. Tojo. Melody Extrapolation in GTTM Approach. International Computer Music conference (ICMC2009). 2009.8.17. McGill University, Montreal, Canada
- (13) 浜中雅俊. 掛けモーフ, デモンストラーション. 音楽情報処理の研究紹介VII. 情報処理学会 第81回音楽情報処理研究会. 2009.7.21. 飯坂温泉
- (14) 坂本鐘期, 東条敏. Tonal Pitch Space を用いた楽曲の和声解析. 情報処理学会 第80回音楽情報処理研究会. 2009.5.21. 筑波大学
- (15) 東条敏. 音楽と言語の構造認知. 電子情報通信学会 思考と言語研究会, 2008.11.15, 東京
- (16) M. Hamanaka, K. Hirata, S. Tojo. Melody Expectation Method based on GTTM and TPS. International Society for Music Information Retrieval, 2008.9.16, Philadelphia, USA.
- (17) M. Hamanaka, K. Hirata, S. Tojo. Melody Morphing Method based on GTTM. International Computer Music Conference, 2008.8.29, Belfast, UK.
- (18) 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 σ GTTM: 局所的グルーピング境界の検出. 情報処理学会 第76回音楽情報処理研究会. 2008.8.7, 名古屋大学, 名古屋
- (19) 浜中雅俊, 平田圭二, 東条敏. FATTA に基づくメロディ予測システム (アプリケーション). 情報処理学会 第76回音楽情報処理研究会. 2008.8.7, 名古屋大学, 名古屋
- [その他]
研究成果ホームページ
<http://music.iit.tsukuba.ac.jp/hamanaka/gttm.htm>
6. 研究組織
- (1) 研究代表者
東条 敏 (TOJO SATOSHI)
北陸先端科学技術大学院大学・情報科学研究科・教授
研究者番号: 90272989
- (2) 研究分担者
平田 圭二 (HIRATA KEIJI)
NTT コミュニケーション科学基礎研究所・研究員
研究者番号: 30396121
- 浜中 雅俊 (HAMANAKA MASATOSHI)
筑波大学・システム情報工学研究科・講師
研究者番号: 30451686