

音楽における自動処理と Directability

片寄 晴弘

音楽プロセスを、計算機を用いて実現しようという取り組みは、他の計算機科学の研究分野と比べても早期から立ち上がり、1957年には自動作曲作品「イリアック組曲」が作曲されている。それ以降、この領域では、自動作曲、演奏の表情付け、自動伴奏など、さまざまな音楽システムの開発がなされてきた。これらのシステムは、計算機科学の可能性を示すものとして大きく注目されてきた。その一方で、「ヒトのための」という視点において、システムはどうあるべきか、どうデザインされるべきかという事項に対する関心も高まりつつある。本稿では、自動音楽処理の代表的な研究例を紹介するとともに、これからの音楽インタフェースの在り方について議論する。

キーワード：音楽システム、自動処理、インタフェース

1. はじめに

音楽の生成とは、どのような心の働きによるものだろうか。感性と科学の関連で音楽を考える取り組みは、古代ギリシャ時代から哲学の範疇で、音律論が議論されるなど、非常に歴史のある研究領域である。

近代科学における取り組みとしては、1950年代のMeyerの情動論[1]、1970年代のHigginsの計算モデル[2]、1980年代初頭のDeutschらの旋律認識に関する研究[3]などがある。1980年代には、音楽理論と認知・心理学的な研究の融合、さらにはその理論化を目指す音楽認知科学という研究領域が成立するに至っている[4]。その後の重要な研究成果としてはLerdahlとJackendoffのGenerative Theory of Tonal Music (GTTM)[5]、NarmourによるImpication-Realization Model (IRM)[6]などがある。

これら「音楽理解のモデリング」に焦点を当てた研究に対し、計算機を用いた生成的なアプローチ、すなわち、音楽に関するアクティビティを計算機上で実現する研究も、1950年代から積極的に取り組まれてきた。最も早期のものとしては、コンピュータによる自動作曲作品「イリアック組曲」(1957年)があげられる[7]。1968年には、積み木を題材とした問題可決に取り組んだWinogradが和声解析を行うシステムの発表を行っている[8]。それ以降、現在までに、自動作曲、演奏の表情付け、伴奏システムなど数多くの音

楽システムの研究開発がなされてきた。

音楽は、他の計算機応用分野と比べても早期から人間に代わる「自動処理」の実現に向けて、研究開発がなされてきた。その背景としては、芸術というある意味最も人間的といえる領域において計算機の可能性を追求していくことの魅力、加えて、楽譜や楽典の存在に見られるように、他の芸術分野と比べて表現 (representation) に関する下地が整っていたことが挙げられる。

本稿では、まず、音楽における自動処理の代表例として、自動作曲、演奏の表情付けの研究概況を述べ、その中の具体例としてCopeによる作曲システムEMI[9]とPachetによる後続音列作曲システムContinuator[10]を紹介する。芸術領域における自動処理は、機械の可能性を示すものとして、驚きをもって迎えられてきた。その一方で、自動処理にはその存在意義に関する哲学的な問題や、ヒトが実施する「デザイン」の支援という文脈で「自動処理」を考えた場合、その追求が必ずしも生産性の向上にはつながらないという問題が存在する。本稿では、この視点を踏まえて、今後の音楽の自動化処理技術、インタフェースの在り方を、筆者が実施している研究プロジェクトの取り組みを紹介しつつ、議論したい。

2. 音楽生成における自動処理

音楽情報処理研究は生成系研究と分析系研究に大別される。生成系研究は、作曲(編曲)、演奏の表情付け、自動伴奏など、音楽情報処理の中でも華やかなイメージを伴った研究対象であり、その可能性に対して音楽以外の情報処理領域からも注目を集めてきた。

かたよせ はるひろ
関西学院大学 理工学部情報科学科
〒669-1337 三田市学園2-1

作曲とは、音の並びに構造を与えることである。広義に考えれば、例えば、星座の星の並びや、オゾンホール面積など某かのデータを音に変換するプログラムも作曲システムとなるが、「ショパンやモーツァルト、あるいは、ポップスなどのように様式感が感じられ、かつ、全く、同じものはない」楽曲を作るというスタンスでの自動システムも作られてきた。そのさきがけとなるのが、冒頭でも述べた「イリアック組曲」で知られる Hiller, Isaacson らの取り組みである。トータルな意味での作曲システム他に、和声付けやハーモナイゼーション、後続音列の作成などの自動作曲の要素技術になるシステムも多数開発されている。Band-in-a-Box¹のように、コード進行を入力し、ジャンル指定を行うだけで、ポップスやロック風の曲が出来上がる商用システムも存在している。

演奏の表情付けは、指定された音の並びに対して、音量やテンポやアーティキュレーション²に変化を与え、音楽を生き活きとしたものとして、実体化する作業である。最近の例では、音楽演奏ロボットが想像されることも多いだろうが、楽譜の構造を把握し、その構造が聴取者にわかるように伝えることが中心課題である。つまり、「弾く（演奏）」も分析と計画に処理の根幹がある。1980年代半ばの演奏ルールの発見的検証[11][12]の取り組みにはじまり、最近では、Rencon³と呼ばれる国際コンテストも開催されている。

演奏に関連したもう一つの技術として、自動伴奏システムがある。自動伴奏システムは、ミスタッチや間違いが入る人間のソリストの演奏に追従して、伴奏を行うシステムである。人間の演奏位置の推定が課題となるが、1980年代半ばに、DannenbergによりDTW (Dynamic Time Warping) を用いる手法[13]、2000年代に入り、武田によりHMMを利用した和音にも対応可能なより一般化された手法[14]が提案されている。

音楽における自動化システムの研究例は枚挙にいとまがない。事例、あるいは、学習セットを用意して、その分解と再合成によりを楽曲や演奏表情を実施するというのが、現在の音楽生成システム系研究の時流である。これら既存事例を用いて音楽生成を行うシステムの中でも、明示的にルールや条件節の候補を用意し

ておき、その関係性やパラメータの学習を行う学習適用型のものと、類似事例の検索と組み合わせによってターゲットの音楽要素の生成を実施する事例ベース型のタイプにほぼ大別される。以下、音楽における自動化システムの具体例として、CopeによるEMIプロジェクトとPachetによる後続音列作曲システムContinuatorについて紹介する。

2.1 EMI

人の知的プロセスの代替という位置付けで開発された最も代表的な自動作曲システムの一つが、D. Copeが1981年に開始した自動作曲に関するプロジェクトEMI (Experiment in Music Intelligence) である[9]。Copeは「作曲とは、今までに作られた作品の事例の解析と再合成によってなされる」との理念のもと、大きくパターンマッチ (モチーフ抽出) プロセスと、ルール解析プロセスで構成される作曲システムを提案・実装した (図1)。

パターンマッチプロセスでは、楽曲からピッチやリズム情報を元に、同じか同型と考えられるモチーフを抽出・蓄積していく。一方、ルール解析プロセスではパート進行や繰り返す音の数、和声概形など次の進行の出現確率を蓄えていく。このようにして作品様式に関する基礎データを取得し、乱数を用いて、モチーフや伴奏を再構成することで作曲が行われる。EMIのホームページ⁴にて、EMIが作曲し、Cope自身が選んだ作品を聞くことができる。試聴できるデータは人手により選ばれ、演奏は人間のピアニストによるものであるが、ショパン、バッハといった作風がよくとら

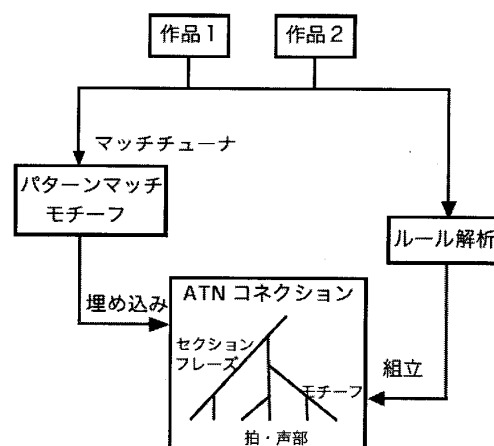


図1 EMIの概要。図中、左サイドがパターンマッチ (モチーフ抽出) プロセス、右サイドがルール解析のプロセスを示している。

¹ <http://content3.e-frontier.co.jp/PG/win/>

² スラー、スタッカート、アクセント等の表現

³ <http://www.renconmusic.org/>

⁴ <http://arts.ucsc.edu/faculty/cope/experiments.htm>

えられている。

2.2 Continuator

Continuator は、2004 年の、Pachet によって提案されたインタラクティブな作曲システムである [10]。

Continuator は、プレイヤーの与えたフレーズに続く、後続フレーズを生成する。演奏パターンは、マルコフ連鎖としてモデル化されている。ユーザが一固まりの音列入力をした後、その入力に対してできるだけ長い連鎖が得られるようユーザ入力とモデルとのパターンマッチングを実施し、選ばれたモデルでの後続音を発音する。その後続音を付け加えた音列に対して、順次、同様の処理を繰り返していくことで、後続フレーズを生成していく。

パターンマッチングにおいては、データスパースネス問題に対処する処理を用意しておく必要がある。Continuator では、探索の条件を、ピッチと音長と音量、その条件に合致するものがなければ、ピッチの代わりに音域と音量というように、徐々に探索の条件を緩和していくことによりこの問題に対処している。

モデル（マルコフ連鎖）は、あらかじめシステムに与えたデータベースを解析することによって得られるが、固定のモデルのみに基づいて処理を実施すると、生成フレーズが固定化されてしまうという問題がある。この問題に対処するものとして、Continuator では、プレイヤーの近々の演奏で演奏された音の頻度（確率）を蓄えておき、よく使われた音の発音に重みをつけて発現させる処理が実装されている。

3. 自動処理と Directability

2 節で述べてきたように、音楽における自動処理には、さまざまなアプローチでの研究が行われており、コンテンツデザインにおける生産性の向上や新しい音楽エンタテインメントの創成に至っている。

筆者らも既存音楽事例の活用、デザイン転写技術によって音楽デザインを支援する研究に取り組んでおり、2005 年から国家プロジェクト（＝CrestMuse プロジェクト）として、研究を実施する機会を得て活動を続けている [15]。デザイン転写技術とは自動処理に他ならない。デザイン支援という目的において、転写技術の開発だけではいけないことが、これまでの活動を通じて判ってきた。現在では、デザイン転写技術の開発に加えて、directability をプロジェクトのキーコンセプトに位置づけて研究に取り組んでいる。以下、CrestMuse プロジェクトの概要と、directability 基

軸の導入に至った過程を紹介する。

3.1 CrestMuse プロジェクト

デザイン支援に関する代表的なアプローチとしては、デザイン素材の印象の多変量解析とメディア提示技術を組み合わせたものが知られており、感性工学として、さまざまな応用システムが作られてきた。この手法は、商品群の中からイメージに合ったものを探すといった検索型のタスクに対しては有効である反面、イメージ空間設定の煩雑さ、印象空間上での分解能の問題から、生成系のデザインプロセスを支援する目的には必ずしも実用されてこなかった。

自分の所望のデザインを言葉で伝えるのは容易なことではない。形容詞を用いて意図を伝えるより、具体的な事例を示した方が、はるかにイメージを伝えやすい。実際、職業的なデザイン分野、コンテンツプロダクションにおいては、「Beatles のあの編曲」、「Stanley Kubrick 後期作品のシーン展開」などのように、具体的な目標事例を掲げてデザインイメージの伝達・共有をはかり、その上で、具体的な制作プロセスに入ることが少なくない。このデザインの実施・支援の形を、音楽を対象に実現しようということで CrestMuse プロジェクトが起案された。

図 2 に CrestMuse プロジェクトの起案時のキーコンセプトを示す。ここで、エラボレーション (elaboration) とは、デザイン上の意図的な逸脱のことをさす。エラボレーションはヒトの音楽認知構造と因果関係がある。未来のエラボレーションは観測されたエラボレーションから予測されるという性質のもと、これらの性質に着目した時系列メディアの階層的なデータ記述とデータの相互変換、段階的探索によるエラボレーションの転写機構の実現を目指したわけである。

現在までの CrestMuse の具体的な研究成果については、ビデオとして、それぞれ、「音楽の愉しみ (2006 年)」、「デザイン転写技術」、「コミュニティ・人」、「能動的音楽鑑賞」、「directability」(以上 2008 年)として一般公開している。プロジェクトのホームページ⁵ から閲覧可能であるので、興味のある方はご覧いただきたい。

3.2 Directability

歌詞を入力に基づく作曲システム Orpheus [16]、旋律予測に基づくセッションシステム BayesianBand [17]、フレーズ表現に着目した表情付けシステム

⁵ http://crestmuse.jp/crestmuse_research_2008_j.html

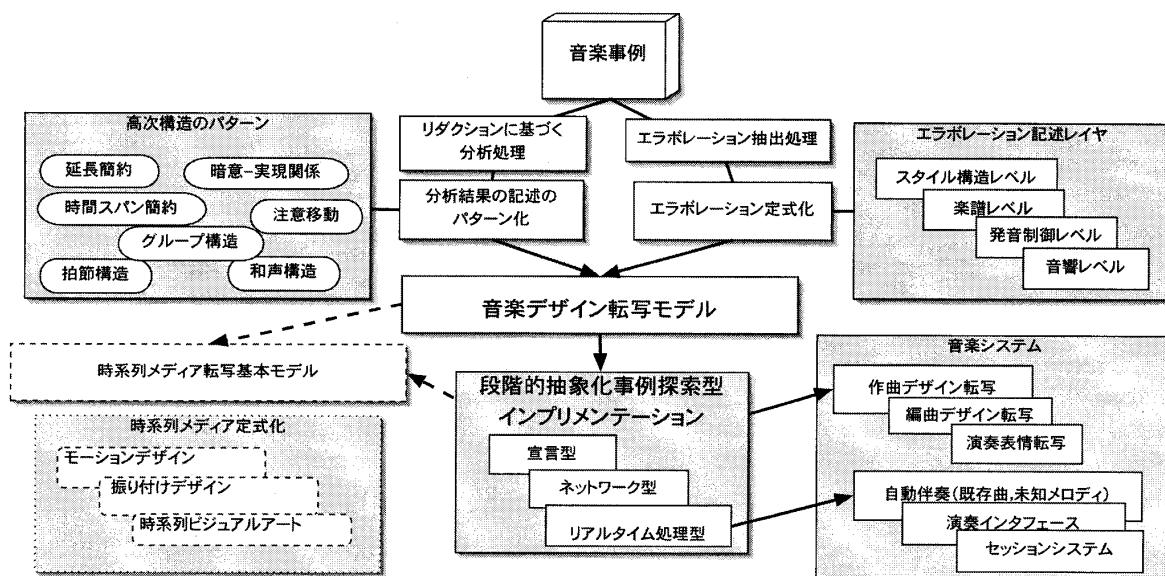


図2 CrestMuseプロジェクトの研究実施上のキーコンセプト (起案時)

Itopul[18]など、デザイン転写技術に関する具体的なアプリケーションの開発が進む一方で、システムの実利用視点での課題がクローズアップされるようになってきた。参照事例データベースの切替えとサイコロの振り直しでしか、音楽生成結果に関与できないというシステムを積極的に使いたいと考えるユーザはほとんどいない。「機械に完全御任せ」というユースケースにおいて自動処理は機能する[19]が、デザインタスクにおいてそのようなケースは稀である。ユーザが主役、それをどう補助すれば良いかという視点でシステム開発が求められる。

事例参照によるデザイン支援は、ユーザが対象に対しての呼び名がわからない、しかし、その特徴を利用してデザインを実施したいという場合に有効に機能する。このことと、上述のデザイン支援システムにおける要求事項を満たすものとして取りあげたキーコンセプトが directability である。参照事例を絵の具のように混ぜ合わせたり、特徴の部分をこね回したりといった操作が直接結果となって反映されるようなインタフェース、その操作性を directability と呼んでいる。現在、デザイン転写技術の開発とならぶ重要研究課題として、研究開発を実施している[15]。以下、directability を考慮して実装された音楽デザインシステムを2事例紹介する。

(a) v. morish

v. morish は、二名の歌唱を、実時間でモーフィングすることによって、歌唱のデザイン機能の提供を目指した歌唱合成インタフェースである[20]。TAN-

DEMSTRAIGHT[21]を用いて、歌唱を節回し(横軸)、声質(縦軸)の2成分によって分解・記述し、左下、右上のそれぞれに、歌手A、歌手Bの歌唱を配置した平面上で、歌唱のモーフィングを実施する(図3)。このインタフェースを利用することで、例えば、自身の歌唱の歌い回しや声質をプロの歌唱のそれと置き換えたり、任意の比率で合成したり、といった歌唱のデザインが可能となる。

(b) Mixtract

Mixtract は、フレーズの組立・分析・編集に焦点を当てた統合型のインタラクティブ演奏デザイン支援システムである[22]。指定した演奏データ、もしくは、データベースからフレーズにおける各階層における演奏表現(テンポ、ダイナミクスカーブ)を抽出し、演奏デザインを実施する対象曲のフレーズ表現の基本テンプレートを作成し、その基本テンプレートを、GUIを使って適宜修正していくことで演奏デザインを実施する。

4. 議論

CrestMuseプロジェクトでの取り組みに限らず、情報系科学技術の研究においては、マシンインテリジェンス、処理の自動化の部分に力が注がれる場合が多い。これらの取り組みがデザイン支援や他の応用領域の技術基盤となることに疑いはないが、3節でもふれたように、処理が高度化するにつれて、デザイン支援の在り方に加えて、哲学的な問題も生じつつある。4節では、この点について議論を進める。

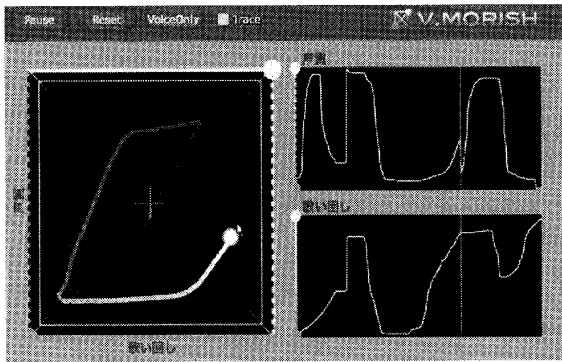


図3 v. morishのGUI. 左側の2次元座標入力インタフェースを用いたモーフィングパラメータの実時間制御, および, 右側のインタフェースを用いたオフライン入力が可能である.

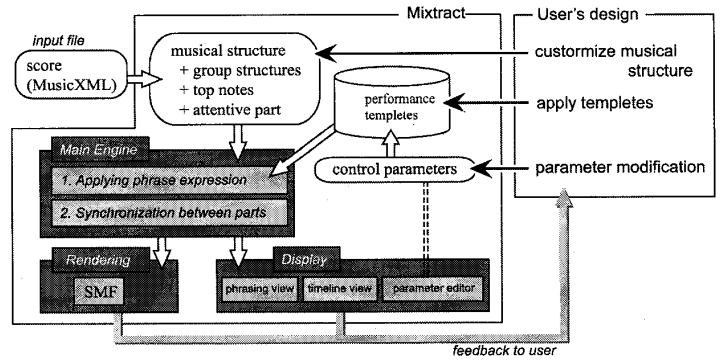


図4 Mixtractの概念図

4.1 コンピュータチェス・将棋の場合

コンピュータチェス・将棋は, 代表的な人工知能応用研究である. 1950年に Shannon, Turing がコンピュータチェスの可能性を示す論文を示し, 将棋については1970年頃開発が始められたのを皮切りに研究開発が続けられてきた.

1997年には Deep Blue がチェスチャンピオンのカスパロフに勝ち, 2005年にはパソコンのソフトウェアがチャンピオンに事実上勝った. 探索の場合の数が多いコンピュータ将棋については, 現段階の棋力はアマチュア五段レベルであるが, 今後約10年で人間の名人に匹敵するようになるともいわれている.

チェス・将棋の世界では, 計算機の能力が人間を凌駕する状況が現実のものとなりつつある. 人間との対戦にしても, ソフト同士の対戦にしても, 勝敗の予想がつくような状況になれば, 対戦そのものに対する面白味は半減してしまうだろう. これに対して, プロ棋士羽生は, 棋士の支援として, 計算機技術が, 詰みのチェックや類似した戦局の検索に利用できると冷静に指摘している [23]. コンピュータチェス・将棋には, 娯楽, アマチュアの練習相手としてのニーズがあり, この状況は, 今後とも, 変わらないと予想される.

4.2 音楽分野の状況

コンピュータチェス・将棋が, 人間のトッププレイヤーに匹敵するようになってきているのに対し, 生成系音楽システムはどうか? 大量消費レベルの音楽を想定, 生成結果の選び出しは人間が責任を負う, という条件であれば, 現在の自動音楽生成システムは, ほぼ, 人間の能力に匹敵する. しかし, 人間の活動に倣い, 生成されたものを評価して取捨選択するかどうかを判断するという部分まで含めて考えるなら, 将来

的にも, 自動音楽生成システムの実現は, かなり難しいということになる. 美しさや完成度の客観的評価は困難であり, そのモデリングとなるとさらに難しい.

芸術という目的に対し, そもそも機械の自律性を追求すること自体に哲学的な課題が存在する. Cope は, 最近, EMI に関する一切のシステム, コードを破棄してしまった. 理由は明らかにされていないが, 創作, デザインにおける人間の関わり方に関する葛藤が背景にあったものと考えられる.

人間と同様の能力を発揮できるような機械の開発は技術者にとっての大きな夢の一つである. 一方で, 音楽を始め, 芸術は, ヒトの人生・生活・精神を色彩豊かなものにするに存在理由がある. 自動処理技術の開発もさることながら, ヒトの活動をサポートするには何が必要かをしっかりと議論した上で, 技術開発が行われなければならない.

5. おわりに

機械で, 音楽に関するプロセスを実現しようという歴史は, 計算機の歴史にほぼ相当する. この領域では, 自動作曲, 演奏の表情付け, 自動伴奏など, さまざまな音楽システムの開発がなされ, 計算機科学の可能性を示すものとして大きく注目されてきた. 処理が高度化するにつれて, 「ヒトのための」という視点において, システムがどうあるべきか, どうデザインされるべきかという事項に対してしっかりとした哲学を持つことが求められている.

デザインの支援という目的においては, 自動処理もさることながら, 人間が介在するインタラクションプロセスを起点としてシステムデザインを実施する必要がある. 本稿では, 既存事例の活用と, directability の実現がその鍵になると論じた.

音楽は, ヒトの人生・生活・精神を色彩豊かなもの

にしてきた。この事実を加速するものとして、技術開発が進むことに期待したい。

参考文献

- [1] L. B. Meyer, "Emotion and Meaning of Music," University of Chicago Press, 1956.
- [2] H. Longuet-Higgins, "Perception of Melodies, Nature," Vol. 263, pp. 646-653, 1976.
- [3] D. Deutsch and J. Feroe, "The Internal Representation of Pitch Sequence in Tonal Music," Psychological Review, Vol. 8, pp. 503-522, 1981.
- [4] 波多野誼余夫編, 「音楽と認知」, 東京大学出版会, 1987.
- [5] F. Lerdahl and R. Jackendoff, "A Generative Theory of Tonal Music," MIT Press, 1983.
- [6] E. Narmour, "The analysis and Cognition of Basic Melodic Structure," University of Chicago Press, 1990.
- [7] L. Hiller and L. Isaacson, "Experimental Music," McGraw-Hill, 1959.
- [8] T. Winograd, "Linguistics and the computer analysis of tonal harmony," Journal of Music Theory, Vol. 12, No. 1, pp. 2-49, 1968.
- [9] D. Cope, "Computers and Music Style," Computer Music and Digital Audio Series, 1991.
- [10] F. Pachet, "The continuator: Musical interaction with style," Journal of New Music Research, Vol. 31, No. 1, 2002.
- [11] M. Clynes, "Secretes of Life in Music," Proc. ICMC, pp. 225-232, 1984.
- [12] L. Fryd'en and J. Sundberg, "Performance Rules for Melodies. Origin, Functions, Purposes," Proc. ICMC, pp. 221-225, 1984.
- [13] R. B. Dannenberg, "An On-Line Algorithm for Real-Time Accompaniment," Proc. ICMC, pp. 93-198, 1984.
- [14] 武田晴登, "音楽演奏の確率モデルに基づく自動採譜と自動伴奏に関する研究," 東京大学博士論文, 2007.
- [15] 片寄晴弘, "デザイン転写による音楽制作支援," 情報処理学会誌, Vol. 48, No. 12, pp. 1359-1364, 2007.
- [16] 深山覚, 米林裕一郎, 山本遼, 和泉洋介, 西本卓也, 嵯峨山茂樹, "Orpheus 歌詞の韻律に基づく自動作曲システム," インタラクティブ東京, p. 15, 2008.
- [17] 北原鉄朗, 徳網亮輔, 戸谷直之, 片寄晴弘, "Bayesian-Band: 旋律の予測に基づいた自動伴奏システム," インタラクション 2009, pp. 31-32, 2009.
- [18] M. Hashida and H. Katayose, "A Directable Performance Rendering System: Itopul," Proc. New Instruments on Music Expression (NIME), pp. 277-280, 2008.
- [19] ドナルドノーマン, 安村他訳, "未来のモノのデザイン," 新曜社, 2008.
- [20] 森勢将雅, 大西壮登, 河原英紀, 片寄晴弘, "v. morish: 歌唱モーフィングのインタフェースデザイン," インタラクション 2009 予稿集, pp. 25-26, 2009.
- [21] 森勢将雅, 高橋徹, 河原英紀, 入野俊夫, "分析時刻に依存しない周期信号のパワースペクトル推定法を用いた音声分析," 電子情報通信学会論文誌, 和文 A 分冊, Vol. J92-A, No. 3, pp. 163-171, 2009.
- [22] M. Hashida, S. Tanaka and H. Katayose, "Mixtract: A Directable Musical Expression System," Proc. ACII, 2009.
- [23] 羽生善治, 伊藤毅志, 松原仁, "先を読む頭脳," 新潮社, 2006.