

生成系音楽支援システムの Directability 視点からの考察

片寄 晴弘[†] 橋田 光代[†]

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。本稿では、生成系音楽情報処理研究の現状を、自動化処理の可能性と課題の両視点から分析した上で、directability の概念とその基軸の導入の意義について実例を交えつつ紹介する。

Discussion on Directability for Generative Music Systems

HARUHIRO KATAYOSE and MITSUYO HASHIDA

Music systems are roughly divided into ones with the generative goal and ones with the analytical goal. This paper discusses generative music systems, from the views of possibilities and problems of automatic processing. Then this paper addresses the concept of "directability" and its importance, with illustrating some examples.

1. はじめに

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・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。

1950年代のイリアック組曲以来、生成系の研究は脈々と続けられており、歴史的なマイルストーンとなった研究例もいくつか存在する。しかし、最近では、どのような目的でシステム開発が行われているのかがよく判らず議論が噛み合わないというような状況を目にする事も多い。研究領域の中での話に限れば、若干、停滞気味であるようにも思われる。

社会現象の一つとしての iPod に象徴されるように、数年前から、コンテンツがほぼ無尽蔵に流通する時代に突入している。音楽のアクティビティは「作る」「弾く」「聞く」に分けられるが、近年、大衆のニーズ、すなわち、マーケットの大きさに引っ張られる形で、「聞く」に関連するサービスと研究開発がとみに進んでいる。

現在は時代がさらに進み、YouTube や Second Life に代表されるような Web 2.0 と呼ばれる新しいネットワーク社会が認知されるに至っている。今までコンテンツの消費者だった一般大衆が情報の発信者となり、

また、社会インフラの観点からもデザインとその支援に対するニーズはあきらかに拡大している。

以上の背景から、筆者らは、一見停滞気味である生成系音楽情報処理領域の研究に対して、立ち位置を明確化し、新しいネット社会への対応を考慮していくことで、飛躍的に発展していくものと考えている。

本稿では、そのためのキーコンセプトの一つとして、**directability** を取り上げ、その概念とその基軸の導入の意義について実例を交えつつ紹介する。

2. 自動化処理 + α

筆者らは、ここ数年来、コンテンツデザインの支援を基軸とした音楽情報処理研究プロジェクトに取り組んで来た。この章では、現在、取り組んでいる科学技術振興機構 CREST「デジタルメディア領域」CrestMuse プロジェクト*の概略を紹介し、**directability** への導入を行う。

2.1 CrestMuse プロジェクトの概略

あるデザインをそのデザインたらしめている特徴は、一般に言語化されていないことが多い。言語化されていたとしても語彙は乏しく、言葉で説明することも容易ではない。デザインのイメージを伝達するには、言葉よりも事例を用いる方がはるかに容易である。実際、職業的なデザイン分野、特にコンテンツプロダクションにおいては、「ビートルズのあの編曲」、「スタンリー・キューブリック後期作品のシーン展開」などのように、

[†] 関西学院大学/科学技術振興機構 CrestMuse Project
Kwansei Gakuin Univ./CrestMuse Project, JST, Japan

* <http://www.crestmuse.jp/>

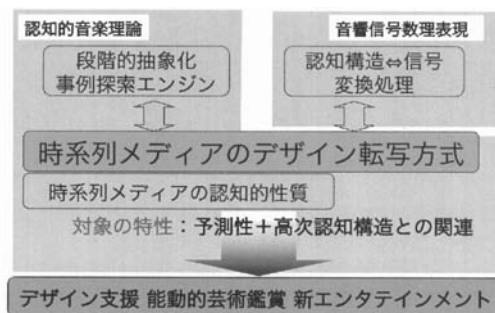


図1 CrestMuse プロジェクトのコンセプト

具体的な目標事例を挙げてデザインイメージの伝達・共有をはかった上で具体的な制作プロセスに入ることが多い。この場合、デザインのイメージは、共通理解¹⁾の元で当事者間の頭の中に構成されており、感性工学という概念空間のような外在化がされているわけではない。このような形のデザイン支援を考えるにあたり、概念空間の操作・形成の部分を一旦留保し、既存デザイン上の意図的な逸脱の操作、転写方式に焦点を当てるといったアプローチが浮かび上がる。

以上のような視点から、CrestMuse プロジェクトでは、デザインの転写技術を基盤として、音楽を中心としたコンテンツデザイン支援、新たな表現手段の提供することを目標に、研究開発を進めている。図1にプロジェクトの概要を示す。

2.2 知性・感性の自動化技術の流れ

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、システムインテリジェンス、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、その部分に固執しすぎると、研究の目標がわかりにくくなったり、アプリケーションの利用状況にそぐわないという問題が生じる。以下、コンピュータチェス・将棋、音楽分野の状況、CREST 森島チーム（デジタルアニメーションラボ）* の取り組みを紹介し、課題を論じる。

2.2.1 コンピュータチェス・将棋

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たのを、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言われている。

チェス・将棋の世界では、計算機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、ソフト同士との対戦にしても、勝敗の予想が予めつくような状態になってしまうと対戦そのものに対する面白味が半減してしまうだろうが、プロ棋士の羽生は、詰みのチェックや類似した戦局の検索に計算機技術が有効であると冷静に指摘している²⁾。加えて、コンピュータチェス・将棋はアマチュアの練習相手や娯楽を目的として現在なお一般に普及しており、今後とも、この状況は変わらないと予想される。

2.2.2 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たのに対し、生成系音楽システムがそうなるのは、まだまだ先、あるいは、将来的にも不可能という見方もある。しかし、質の問題を深く議論しないのであれば、音楽デザインの代替手段として利用できるレベルに達していると言える。

Copeの一連の自動作曲の研究³⁾やRencon**での取り組みに見られるように、自動作曲、表情付けにおける機械学習、事例ベースアプローチが定着しつつある。自動化研究が事例活用の方角に進んでいる一方で、生成結果をどのように修正するか、そのための事例活用のインタフェースがどのようなものがよいか、という課題については後回しにされてきた。その結果、生成結果に対し修正を行いたいという場合、別のアプリケーションを立ち上げて、音符レベルの操作を行うことを余儀なくされることが少なくない。

芸術や音楽の存在の意義に照らせば、人間の支援ということをしかりと見据えた自動化技術の開発が求められる。

2.2.3 CREST 森島チームの取り組み

森島らは、筆者らと同じくCREST「デジタルメディア領域」の助成を受け、アニメーション制作支援技術の研究開発を進めている。

このプロジェクトでは、物理モデルの有効活用を研究の一つの柱としているが、複数の現場のアニメーションデザイナーから、「長年培ってきた演出技法を使えない、すなわち、デザイナー自身の書き加えができないようなツールは使えない」という厳しい指摘を受けた

* <http://www.cavie-x.net/top.html>

** <http://www.renconmusic.org/>

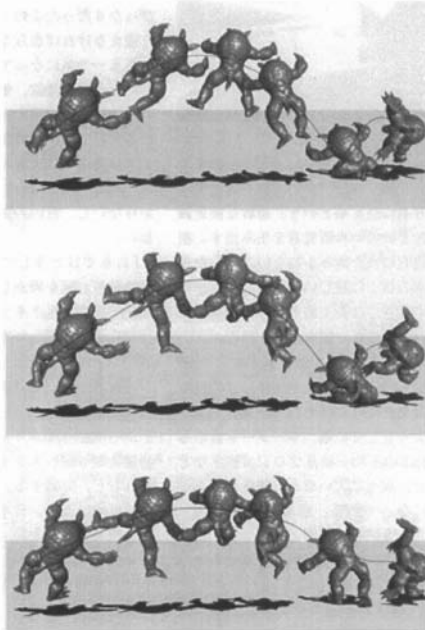


図2 アニメーションのためのツール（科学技術振興機構 CREST 「デジタルメディア」領域資料集より）（上）：物理モデルで計算された動き、（中）：デザイナーがキャラクタの体勢を修正、（下）デザイナーが動きモデルを修正

という。

この指摘に対し、森島らは、物理モデルで作成したモーションに対し、後に、シームレスなインタフェースで手を加えられるようにする（図2）ことで、デザイナーにとっての *directable* インタフェースを提供した。

2.3 知性・感性の自動化技術の応用展開

コンピュータチェス・将棋の事例は、使用者自身のトレーニングや支援、楽しみといったニーズに応えるという目的で、自律型の知能システムの存在価値があることを示唆している。表情付けの研究であれば、音楽教育や自動伴奏システムに展開された場合、その存在価値が確保されるということになる。

コンピュータチェス・将棋の場合、強いアルゴリズムを作るという単純明快な目標を用意することが出来る。芸術やデザインシステムの場合、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。その是非について論じるのはここでは避けるが、自動化技術を人間の支援のための基礎技術として位置づけた時に、改めて考慮していかなければならない事項が存在する。その一つが、森島らも指摘している *directability* というキーワードであろう。以下の章では、*directability* に対する我々の定義を与え、特

に、事例を利用したデザイン支援インタフェースにおける *directability* の構成について論じていく。

3. Directability

既存事例を利用したデザイン支援形態において、個々の具体的なデザイン対象を直接操作するのではなく、人の知覚・認知に適合した、より抽象度の高い操作子を用いて対象の操作が可能な状態にすることは、最も根幹的な目標の一つに位置づけられる。ここでは、この状況を *directability* と呼ぶことにする*。

以下、事例利用型インタフェースにおける *directability* の実現の方策について述べる。

3.1 加算型ハンドラ

デザイン転写において、*directability* を実現する第一の手段は、デザイン上の意味のある偏倚 (*elaboration*) と基準となるデータ（無表情データあるいは平均データ）に分離するデータ記述を用意し、デザインをその記述を用いて表現することである。こうすることで、偏倚の部分を加算する形でデザインの特徴を転写したり、特徴を強調したり、モーフィングの操作を行ったりすることが可能となる。

対象を、適当な階層レベルで記述できれば、さらに、操作手段は広がる。音楽には、楽譜や MIDI に基づくデータ記述が存在し、さらに、*elaboration* と相関を持つと考えられる認知構造が存在する（図3⁴）。これらは *directability* の実現に向けての有利な材料である。

この視点により構成された音楽インタフェースの一例としては、iFP⁵ があげられる。iFP では、名演奏における *elaboration* を、拍単位のテンポ、ダイナミクス、拍内の微細な表現に分解したデータを演奏表情テンプレートとして利用する。演奏者はそのニュアンスを保持しつつ、指揮動作や拍打によって自身の意図でテンポとダイナミクスを制御する（図4）。二つの演奏事例を用い、テンポ、音量、拍内の微細表情のそれぞれのパラメータを内挿（外挿）することによって作り出した新たな演奏表情テンプレートを利用することもできる。

3.2 積算型ハンドラ

加算型ハンドラが、どちらかというと対象の物理的な差異に注目した *directability* の実現アプローチであるのに対し、知覚領域においてできるだけ直交

* 図2の例では、物理モデル計算の処理の後にデザイナーの修正（=*directability*）が実施される。これに対し、我々の定義では、事例を利用した計算過程の段階でユーザが関与できる形態、すなわち、密なインタラクションをより意識したものとなっている。

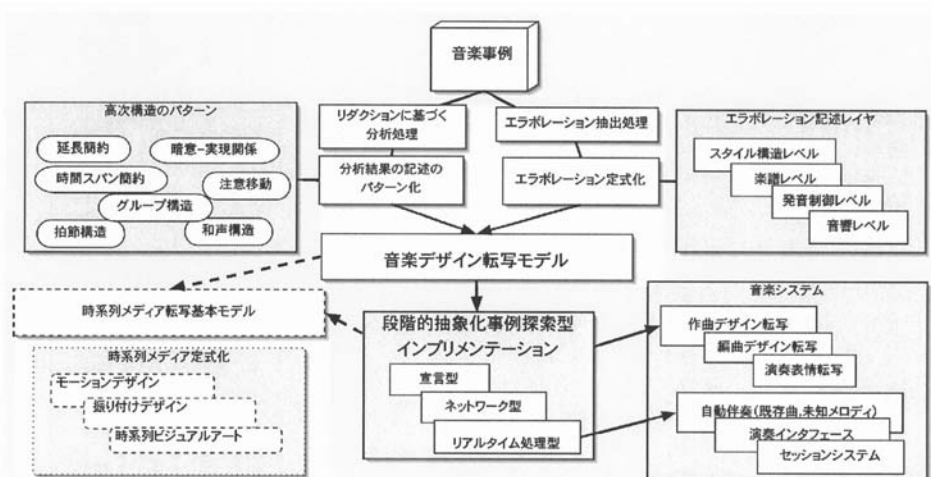


図3 音楽におけるデザイン転写の全体像

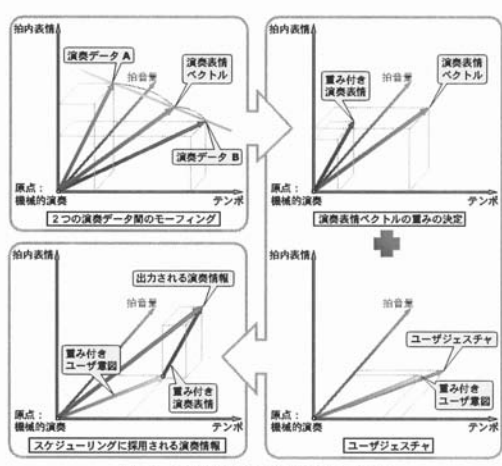


図4 iFPにおける演奏表情の合成



図5 iFP (指揮インタフェース)を用いて演奏している様子

するような積算型ハンドラを用意して、ユーザインタフェースを構成するという手法がある。積算型ハンドラの発見は必ずしも容易ではないが、うまくハンドラが設定できれば、計算機技術ならではの新しいデザイン手段が提供できるようになる。

「美空ひばりが宇多田ヒカルの曲を歌うとどうなるか？」そのような夢の第一歩として我々が構築を進めている歌唱モーフィング⁶⁾は、知覚的ハンドラの提供を目指したデザイン支援インタフェースと言える。

このインタフェースでは、音声分析変換合成システム STRAIGHT⁷⁾を利用して歌唱を声質と歌い回しの二つの時系列パラメータによって記述し、二つの歌唱の時間軸での対応点を求めた上で、図6のようなイン

タフェースを用いて歌唱のデザインを実施する。現時点での知覚実験では、ポピュラー音楽では声質が歌手の同定に大きく貢献していることが確認されている。

3.3 利用事例のアクセス手段の抜本的な改変

加算型ハンドラ、積算型ハンドラの提供によって、デザイン対象のスムーズな操作が可能となる。これに加えて、利用事例のアクセス手段の抜本的な改変というアプローチが考えられる。

2.2節において、自動化処理の使いづらさを指摘したが、1) 利用されうる有効な事例が十分に集積されており、2) それらからの効率的な事例の検索手段もしくは複数の検索手段が提供されており、3) 即時的なデザイン転写(ユーザへのフィードバック)が実現されていれば、それは、directableなインタフェースとしてユーザに解される。

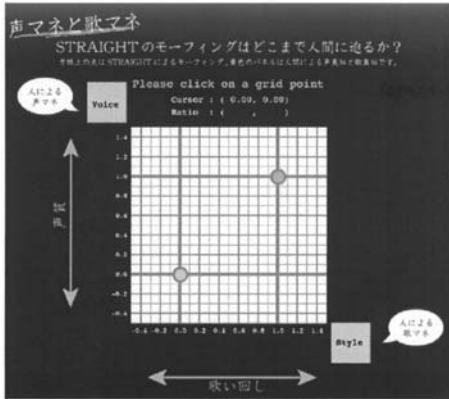


図 6 歌声モーフィングのインタフェース

デザインの計算過程において密なインタラクションが実現されていない状態であっても、即応性と、今までに無かった機能レベルのオプションが提供されていれば、ユーザは新鮮な操作感を覚える。この例として、後藤らの一連の能動的音楽鑑賞インタフェース⁸⁾、吉井らの Drumix⁹⁾ や伊藤らの表情付けシステム Itopul¹⁰⁾ をあげることができる。

4. 議 論

本稿では、事例活用のためのインタフェースにむけての初期的な検討を実施してきた。その内容は、インタラクションデザイン*に密接に関係しているが、現在、他の具体的研究と比較できる段階には無い。以下では、このアプローチでの研究展開の質的な可能性に関連して、メタ認知形成の支援、楽しみの提供、Web2.0 時代における展開、評価研究のあり方について論じる。

4.1 メタ認知形成の支援

メタ認知とは、自分の認知過程そのものを認知することである。諏訪の定義¹¹⁾によれば、メタ認知における「認知過程」とは、

- 言語的思考 (言語的に考えている過程)
- 環境からの知覚 (環境を五感で知覚する過程)
- 身体動作の知覚 (自分の身体部位がどう動き、その結果、どうな感触を得ているかを知覚する過程)

である。3.1 節で紹介した iFP では、名演奏をテンポ、ダイナミクス、拍内表情に展開し、自らの表現行為を伴った形でなぞることができる。iFP の演奏感覚は、懼れずにナイーブな発言をするなら、演奏者の意図やフレーズがわかったような気になるというも

* 人間-システムの系を通じた作業全体における効率化をどのくらい進めることができるかを視点においたデザイン、あるいは、その研究領域。

のである。最近の脳研究においては、運動野の脳活動が記憶や物事の理解に関与していることが明らかになりつつある。この他にも、複数のフレーズ手法に対するルールを用意し、そのパラメータの制御をユーザにゆだねることによって、フレーズの理解を促すというタイプの切り口も存在する¹²⁾。directability を備えたインタフェースは、メタ認知形成の支援に利用されうる。

4.2 楽しみの提供

音楽行為は本来的に人間の愉しみと直結していると考えられる¹³⁾。音楽プロセスの自動化を進めることは、人間の楽しみをうばうということに直結するが、directability を備えた音楽インタフェースとして、ユーザ側に再提示することで、音楽の新たな愉しみを提供できるようになる。3.3 節で紹介した能動的音楽鑑賞インタフェースはその代表例である。また、生成系の音楽インタフェースは、音楽鑑賞インタフェースとは異なった能動的な没入感 (= flow)¹⁴⁾ を誘発する環境として機能する可能性が高い。flow 状態への誘導と評価法の確立は学問領域としても魅力的である¹⁵⁾、directability と関連した形で研究が活性化することが期待される。

4.3 Web 2.0 時代におけるデザイン支援

現在、インターネット上では無数のコンテンツが流通している。また、本来は、個人が持っていたデータをネットワーク上のサーバに保管し、アクセスするというサービスの提供が進んでいる。事例参照型デザイン支援の課題の一つは参照事例の確保であるが、ネットワーク上での資産を用いることでこの問題は解決できるものと期待される。

blog、YouTube、ニコニコ動画等のネットワークサービスへの愛好者数が急増している。投稿されるコンテンツには、既存コンテンツに付加的情報を加えたというものが少なくない。Web 2.0 サービス事業が急速に普及している背景としては、サービス提供者側が、ユーザにとって簡便なコンテンツ投稿手段を提供したことにより、人間の本質的な自己表現に関する要求に火がついたと考えられる。本稿で議論してきたデザイン支援は、このような潜在的な要求に応えるものであり、技術を普及させることによって、生産されるコンテンツの質・量の底上げに対する貢献が期待できる。

4.4 評価研究のあり方

ユーザインタフェース系の研究においては、どのような支援を目的とするのかを明らかにした上で、それに適した、実装と評価を進めることが求められる。

この課題に対して、中小路は、支援システムを、(a)

ダンベル型, (b) ランニングシューズ型, (c) スキー板型に分類し, それぞれのタイプにあった評価を行っていくべきであると指摘している¹⁶⁾. ダンベル型は, それを使用することによる使用者の筋肉の強化, すなわち, ユーザの能力向上や教育を目的とする. ランニングシューズ型は, より早く, 快適に走ることが目標である. スキー板型の場合, それを使用しないと「スキー」というアクティビティ自体が行えない. スキーという体験そのものを可能にするツールである. なお, スキー板型については, 「スキー」体験をデザインするのを目指すのか, 「良い」スキー体験をするために今までのものとは違ったタイプのスキー板をデザインするのかが視点の差も重要である.

本稿の議論の起点は, 事例を活用したデザイン支援において, いかにして, デザインの効率性を高めるか, という点にあった. その意味においては, (b) ランニングシューズ型の評価・検討を進めていくことが求められる. ところが, **directability** の意義, 構成に関する議論を進めていくうちに, (c) スキー板型のタイプとして考慮・展開していくべき事項が少なからずあることが判ってきた. この場合, 評価研究に対しての意識が強すぎると発想の自由度を狭める危険性がある. 帰納的研究方法の実践も不可欠である.

5. おわりに

Web2.0 と呼ばれる新しいネットワーク社会状況が広く認知されるに至り, コンテンツデザインとその支援に対するニーズが拡大しつつある.

今後のコンテンツデザインの形として, 既存事例の活用と, このアプローチに即したインタラクティブデザインの重要性がさらに高まると予想される. 本稿では, そのためのキーコンセプトの一つとして, **directability** を取り上げ, その概念とその基軸の導入の意義について事例を交えつつ紹介した.

現在のアカデミック系のコンテンツ技術開発は, とすれば, 特定の作品や用途に特化した機能実装に力が注がれがちであるが, 波及効果を意識して研究・開発に取り組むことが求められる.

謝辞 謝辞本研究は, 科学技術振興機構 CREST の支援を受けて実施されています.

参考文献

- 1) Gibson, J. J.: *The ecological approach to visual perception*, Lawrence Erlbaum Associates (1986).
- 2) 羽生善治, 伊藤毅志, 松原仁: 先を読む頭脳, 新潮

- 社 (2006).
- 3) Cope, D.: *Computers and Music Style*, A-R EDITIONS (1991).
- 4) 片寄晴弘, 平田圭二, 野池賢二, 原田利宣, 笠尾敦司, 宮田一乗, 平賀瑠美: 非言語メディアのデザイン支援に向けて, 人工知能学会論文誌, Vol. 20, No. 2, pp. 129-138 (2005).
- 5) Katayose, H. and Okudaira, K.: iFP A Music Interface Using an Expressive Performance Template, *ICEC 2004, Lecture Notes in Computer Science*, Vol. 3166, pp. 529-540 (2004).
- 6) 河原英紀, 生駒太一, 森勢将雅, 高橋徹, 豊田健一, 片寄晴弘: 歌唱音声モーフィングに基づく声質と歌い直し転写の知覚的検討, *インタラクシオン 2007*, pp. 113-120 (2007).
- 7) Kawahara, H.: Exploration of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds, *Acoustic Science and Technology*, Vol. 27, No. 6, pp. 349-353 (2006).
- 8) 後藤真孝: 音楽音響信号理解に基づく能動的音楽鑑賞インタフェース, 情報処理学会音楽情報科学研究会 MUS70, pp. 59-66 (2007).
- 9) Yoshii, K., Goto, M., Komatani, K., Ogata, T. and Okuno, H. G.: Drumix: An Audio Player with Real-time Drum-part Rearrangement Functions for Active Music Listening, 情報処理学会論文誌, Vol. 48, No. 3, pp. 1229-1239 (2007).
- 10) 伊藤洋介, 橋田光代, 片寄晴弘: 楽曲大局構造とパルスの表現を併用した事例参照型演奏生成システム, 情報処理学会第 69 回全国大会, 6N-5 (2007).
- 11) 諏訪正樹: 身体知獲得のツールとしてのメタ認知的言語化, 人工知能学会誌, Vol. 20, No. 5, pp. 525-532 (2005).
- 12) Mitsuyo, H., Nagata, N. and Karayose, H.: Pop-E: A performance rendering system for the ensemble music that considered group expression, *Proc. ICMPC*, pp. 526-534 (2006).
- 13) 片寄晴弘: 音楽とエンタテインメント, 日本バーチャルリアリティ学会誌, Vol. 9, No. 1, pp. 20-24 (2004).
- 14) Csikszentmihalyi, M.: *Beyond boredom and anxiety: Experiencing flow in work and play*, Jossey Bass Incorporated publisher (2000). (邦訳: 今井浩明 訳, 楽しみの社会学, 新思索社).
- 15) Katayose, H., Nagata, N. and Kazai, K.: Investigation of brain activation while listening to and playing music using fNIRS, *Proc. ICMPC*, pp. 107-114 (2006).
- 16) Nakakoji, K.: Seven Issues for Creativity Support Tool Researchers, <ftp://ftp.cs.umd.edu/pub/hcil/CST-NSF/Papers/sevenissues.pdf> (2005).