

非言語メディアのデザイン支援の課題と可能性

Toward Computer-supported Design for Non-verbal Media

片寄晴弘^{*1} 平田圭二^{*2} 原田利宣^{*3} 平賀瑠美^{*4} 笠尾敦司^{*5} 宮田 一乗^{*6}
 Haruhiro Katayose Keiji Hirata Toshinobu Harada Rumi Hiraga Atsushi Kasao Kazunori Miyata

^{*1}関西学院大理工 Kwansai Gakuin University ^{*1}さきがけ研究 21 PRESTO, JST ^{*2}NTT CS 研 NTT Communication Science Laboratories

^{*3}和歌山大システム工 Wakayama University ^{*4}文教大情報 Bunkyo University ^{*5}東京工芸大芸術 Tokyo Institute of Polytechnics

^{*6}北陸先端科学技術大学院大学 JAIST

Demands for multimedia contents are increasing. Computer-supported design for non-verbal media is supposed to be one of the most crucial information technology services. The goal of this paper is to illustrate a guide for productive design defined as the process to elaborate the artifacts toward the expression goal. This paper summarizes common procedures in the design processes, based on observation of research and systems explored in music, painting, plastic modeling and motion design. We think hierarchical formalization of design-artifacts and utilizing examples are the crucial approach for computer-supported design for non-verbal media. The assertion is examined from the views of formalization regarding art and design, and taxonomy of design models.

1. はじめに

非言語メディアを論理的に記述し、その内容を他者へ伝達することは一般に困難である。例えば、音楽、映像、造形、舞踊を、いったん自然言語あるいは形式言語による記述に置き換えてしまうと決して元の音楽、映像、造形、舞踊を復元することはできない。非言語メディアには本質的に互いに曖昧かつ主観的な記述でしか表現できない要素が含まれており逆にその部分が非言語メディアの存在意義でもあった。このような背景のもと、これまで、音楽、絵画、造形、モーションといった非言語メディアのデザイン支援の研究は、人工知能ともあまり関連付けることなく、また互いにも関連付けることなく、個別に進展してきたが、実際に非言語メディアのデザイン支援の研究事例を俯瞰してみると意外にも共通点が多いことがわかる。

本論文では、非言語メディアとして音楽、絵画、造形、モーションを取り上げ、各領域で個別に研究されてきたデザイン支援の研究事例について、共通点と相違点を整理し議論することにより、非言語メディアのデザイン支援システム構築にむけての課題と可能性を探る。

2. 具体的な取り組み

本章では、特に、コンテンツ制作における生産性の向上という点に着目し、音楽、造形、絵画、モーションの4領域において、実際に構築されたデザイン支援システムを紹介する。

2.1 音楽領域での取り組み

EMIはCopeによって1981年から開始された自動作曲に関するプロジェクトである[Cope 91]。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」という考え方に基づき、自動作曲・編曲システムを構築した。EMIの処理概要を図1に示す。EMIは大きくわけてパターン

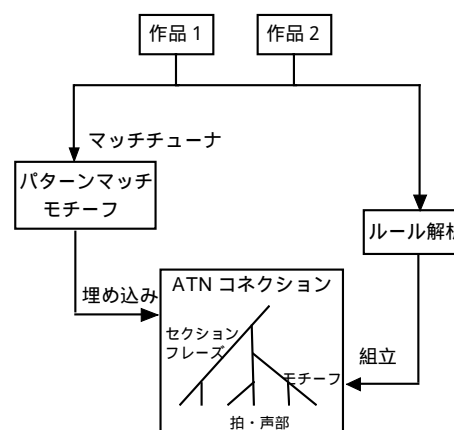


図1: EMIの概要。図中、左サイドがパターンマッチ(モチーフ抽出)プロセス、右サイドがルール解析のプロセスを示している。

マッチ(モチーフ抽出)プロセスとルール解析のプロセスから構成されている。パターンマッチでは、ピッチのみ、リズムのみ、ピッチとリズムを合わせたものの3つの基準から、同一または同型と考えられるモチーフの発見を行う。一方、ルール解析では、パートの進行方向、繰り返される音の数、和声の概形などからモチーフの出現確率を計算し、これら作品の様式に関する基礎データから、乱数を用いて、モチーフを再構成することで作曲が行われる。ユーザが与えるデータは楽譜(音のシーケンス)である。使用者が把握できる内部変数は、モチーフ、和音の推移確率などである。

2.2 造形領域での取り組み

造形領域においては、デザインプロセスの定式化を行い、その知見をデザイン支援に応用する研究が行われてきた。その端緒は、G. Kepesによって説かれた視覚言語である[Kepes 44]。しかし、今までの研究の多くは、造形研究家やデザイナーの直

連絡先: 片寄晴弘, 関西学院大, 〒669-1337 三田市学園 2-1,
 tel.&fax. 0795-65-7861,
 http://ist.ksc.kwansei.ac.jp/~katayose/


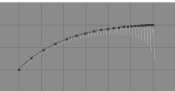

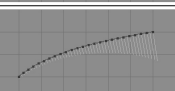

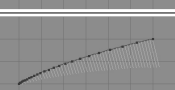

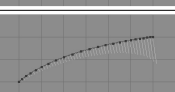

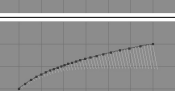
	curve C の傾き	曲線例	印象
単調リズム曲線	発散型 傾きが「負」 		「切れのある」 「勢いのある」
	定速型 傾きが「0」 		「安定した」 「静的な」
	収束型 傾きが「正」 		「線にたまりのある」 「求心的な」
複合リズム曲線	山型 傾きが「正」+「負」 		発散していた曲線がある境から収束していく印象
	谷型 傾きが「負」+「正」 		収束していた曲線がある境から発散していく印象

図 2: 視覚言語化 (定式化) された 5 つの曲線タイプ

感を頼りにしており、形状の物理量などに基づいたものではなかったため、成果の一般的な有用性や適用可能性、説明の科学的な客観性は乏しかった。

原田らは視覚言語を扱う対象の物理量から抽出 (定式化) することを目指し、自動車の曲線 (面) デザインを題材として、その定量化や認知科学的アプローチを試みた (図 2) [原田 98]。ここで体系的に定式化された視覚言語を再構成することで、創成 (デザイン) が可能となる曲線の解空間全体を同定することに成功した。さらに、デザイン支援として、任意のデザイン概念 (イメージ) を具現化する視覚言語とそれらを組み合わせる統語法 (様式) との関係性を形式化し、自動車の曲線デザインにおいて検証した。しかし、自動車全体の曲面までを網羅した様式の実装までには至っておらず、このレベルの様式の形式化、表現や操作の探究が今後の課題である。

2.3 絵画領域での取り組み

芸術家の Harold Cohen が Lisp で構築したコンピュータ画家 “Aaron” は、Cohen の芸術活動における過程 (精神の働き) をマシン上に実装したものである。“Aaron” においては、Cohen 自身の描画に対する知識が形式化され、プログラムとして具現化されている。“Aaron” は Cohen の表現に関する知識の一部を持ってはいるが、それはプログラム中に埋め込まれたものであり、一般の人は利用できない。

これに対し、笠尾・宮田らは「多くの人と描画知識を共有しながら、新たな表現を生み出す」描画サーバーシステムの構築に着手している。この描画サーバーシステムの基本エンジンには、描画ソフトウェア SIC [笠尾 01] が利用されている。SIC はまず、元となる写真から画像の構造を抽出し、次にその構造をもとに、数段の表現ステップを踏むことで、表現したい画風を作り出す (図 3)。

SIC のプログラムには、各ステップにおける描画知識が埋め込まれているが、それらを統合し一つの表現に結びつける知識については、プログラムから切り離すようにしている。つまり、各ステップにおける多くの“処理”の組み合わせ方やパラメータ設定を、スクリプトとして外部記述するようにしている。SIC を用いて生成された作品は、スクリプトの記述内容によって、多様に変化する。その対応関係を集積していくことにより、描画知識の整理、共有、再利用ができるようになると



図 3: SIC の利用に基づく描画

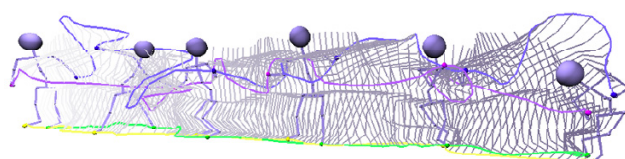


図 4: 概略記述からの連続動作の生成。動作生成の結果を軌道表示したところ。[星野 03] より

期待している。描画サーバーシステムでは、スクリプト、写真画像、写真画像から表現された作品の 3 つを一組として、共有スペース上で管理している。

2.4 モーション領域での取り組み

エンタテインメントや教育などの分野では、CG による人物アニメーションに対するニーズが高まっている。この制作には複雑な技術や専門知識に加え、膨大な制作時間が必要で、個人、とりわけ初心者にとっては敷居が高い。CG で人物動作を生成する場合、一般的にキーフレーム補間法が使用されるが、自然な動作を生成するためには多数のキーフレームを設定して動きを与えてやる必要がある。この作業を代替するものとして、少数のキーフレームのみを指定し、計算モデルによってモーションを補間する研究が行われている [星野 03]。

星野らは、アニメーションのデザインプロセスを、概略的な図表現による基本構造と運動レベルの詳細化に分離することで、複雑な人物アニメーションを容易に生成する手法を提案している。ここでの目的は、ストーリーボードによって記述されたシナリオやシーンデザインなどの様々なレベルで、モーション事例を再利用することである。まず、ビデオ映像からの 3 次元的人物動作を、輝度値と関節駆動力の最小化問題に帰着させて推定する。次にキーフレーム間のモーション補間を行うため、ストーリーボードに書かれた概略的なキャラクタ動作の記述に、動作データベースに蓄積された動作セグメント (事例) を適用する (図 4)。

3. 非言語メディアを対象としたデザイン

3.1 デザインと elaboration

一般にデザインとは、表現すべき抽象的な課題が与えられ、それを適切な実装技術やインタフェース技術等によって効率良く達成する表現行為 (その過程あるいは結果) を指すことが多い。コンテンツ制作においては、そのコンテンツをより多くの受容者に受け入れて貰うという目的をもってデザインがなされ

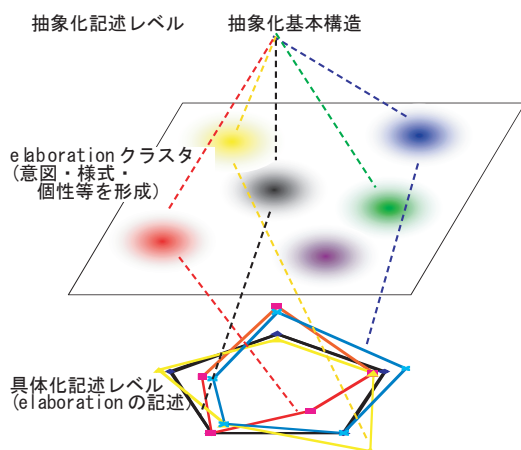


図 5: elaboration と様式・意図の関係

る。非言語メディアコンテンツに対して各受容者の持つ印象や感想自体は曖昧かつ主観的であるが、我々は生物学的及び社会的に共通の背景や概念を持っているため、ある程度の一般性は仮定できると考えている。

本稿では、非言語メディアデザインの表現行為において、受容者の印象や感想を制作者の意図に沿ったものにするために意図的に配された逸脱を *elaboration* と呼ぶ (図 5)。例えば、演奏者はある意図を実現するため、譜面通りの機械的な演奏からテンポ、音長、音量などを逸脱させる。これは演奏表現レベルでの *elaboration* である。また舞踊においては、特定のポーズに対する体の各部分のちょっとした配置や角度の違いが *elaboration* である。一般に、*elaboration* には個人差が認められる。受容者は、そこに、個性や様式を感じとる。

3.2 感性工学と事例に基づくデザイン支援

非言語メディアを扱う応用指向技術の一部は感性工学と呼ばれており [長町 93, 大澤 00]、次のような方法論を採用している: 多数の形容詞によって評定したデザイン素材に対する印象を多変量解析によって低次元の特徴量空間に縮約する、その特徴量空間内の距離が感性的な近さに対応することを利用してメディアの提示と選択等を行う、GUI 等の技術で特徴量空間内の配置されたメディアを操作する。応用実績としては「車」「眼鏡」「商標」等の検索インタフェースがある。

感性工学においてシステムを操作するためのインタフェースは「美しい」「激しい」「都会的」などの形容詞 (情緒表現) に基づいている*1。情緒表現の中には上述の逸脱、すなわち、*ideosyncratic* (そのコンテンツをそのコンテンツたらしめているコンテンツ中の特徴) なデザインの機微や個人差に対応するものもある。しかし、特異な情緒表現は次元を縮約する際にうもれてしまう。それゆえ、感性工学が有効なのは、一般性がある程度まで仮定でき、その特徴量が良く知られている領域における識別、検索、照合的なタスクに限定される。したがって、感性工学的アプローチは、コンテンツ制作を目的とするデザイン支援システムには本質的に適していないと思われる。

一方、現実のメディア制作現場では、考えている様子やイメージを指し示すのに、具体的な事例を利用し、情緒表現は補足的に用いることが多い [片寄 03]。例えば「ビートルズの

*1 情緒表現は、さらに知覚情報の言い換えに相当する知覚表象表現、文化的な基盤において共有されているクラスの表象に相当する文化的表象表現、嗜好に関連し、一般性の仮定が困難な嗜好関連表現に分類できる。

あの編曲のレゲエ風テイスト使おう」「スタンリー・キューブリックのあのシーン展開での激しさが欲しい」などの会話が飛び交う場面は少なくない。事例からその内容 (の一部) を転写する利点は、*ideosyncratic* なデザインの機微や個人差を情緒表現で記述することなく伝達できる点にある。事例をデザイン支援に利用するという方法論の欠点としては、事例に基づく推論 (CBR) でも言われているように、事例の適切なインデキシング、開いた状況 (文脈) への対応が難しいということが挙げられよう。この問題は、事例の記述 (コンテンツ表現) の問題に帰着される。特に、デザイン支援においては逸脱も含めて適切に表現されなければならない。

4. 非言語メディアの記述

4.1 音楽における記述の階層構造

各非言語メディアにおける記述の階層構造を理解するには、それらの階層構造どうしを対応付けてみるのが良いだろう。音楽には楽譜が存在する。楽譜*2 には、音符のシーケンス、楽器種、演奏表現の手がかりとなる発想記号等が含まれる。楽譜における音の長さ、音の大きさ、音の高さは量子化されたものであり、実際の演奏にはそれに逸脱が加わる。楽譜だけでは元の演奏を正確に再現することはできないが、楽譜の記述力、記述のコスト、再現性のトレードオフとして現在の楽譜のような記法が定着したと考えられる。一方、楽譜という抽象化された記述階層が設定されたことで、楽曲そのものの記述 (楽譜) と音楽演奏の表現に関する部分 (逸脱) を分離することが可能となったとも見なせる。現在の DTM においても、楽譜情報と逸脱情報を分離して記述する方式が一般的となっている。

次に演奏の記述を考える。演奏者は与えられた楽譜に対して発音を制御してニュアンスを表現するが、そのために付加された逸脱を、ここでは演奏記述レベルの情報と呼ぼう。音響信号レベルでは、立ち上がり、立ち下がり、周波数スペクトル、及びそれらの時間変化等を意味する。例えば、ピアノやオルガン等の打鍵系楽器の場合、演奏記述レベルの情報は、打鍵時刻、離鍵時刻、音量で記述することができる。バイオリン、トランペット等の楽器については、さらに弓の圧力、息の制御などの情報も記述する必要がある。

以上から、ここでは、音楽を基準となる非言語メディアとし、楽譜を中心に、縦軸に記述レベルの抽象化度をとり音楽の表現法を整理する (図 6)。図中、楽譜記述レベルの上に高次認知構造としての楽曲構造の層がある。例えば、楽曲構造を楽譜記述レベルから解析的に抽出するための理論としては、GTTM [Lerdahl 83] などがある。また、音楽の各抽象化度 (音響信号、楽音、楽譜記述、楽曲構造) において、それぞれ対応する演奏記述レベルが存在している (図 6 中右側)。演奏表現のためには *elaboration* の付加が必要であり、この *elaboration* の分布を解析することで様式、デザインに込められた意図、デザインの個性等との関連情報を得ることが期待される。

4.2 非言語メディアの記述階層の対応付け

各非言語メディアの記述の階層構造を理解するために、音楽の階層構造と他の非言語メディアの階層構造を対応付ける (図 7)。図中最左カラムに参考として自然言語の階層構造を示した。各領域の各階層にはすでに名称が付けられ概念として区別されているが、ある程度の対応関係を付けることができる。

前述した原田らの曲線の視覚言語 (2.2 節) は、図 7 中の造形領域の自己アフィン性を持つ曲線の抽象化度に対応する。例

*2 楽譜には様々なものがあるが、ここではいわゆる五線譜を想定して議論を行っている。音楽の全体像を考える際には不十分である。

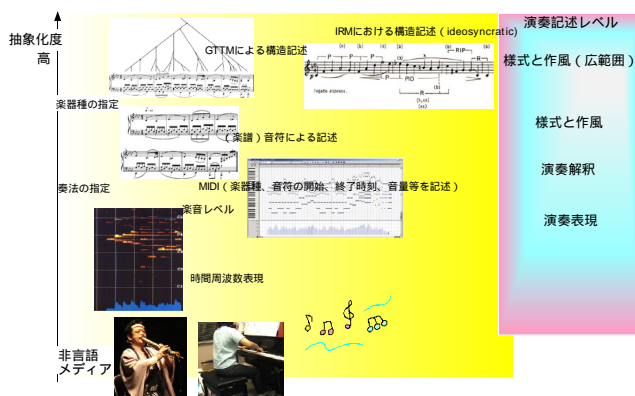


図 6: 音楽における表現法

自然言語	音楽	絵画	造形	モーション
文章	楽曲 曲全体	構図 絵画全体の構成	グラフィックを描き込んだ曲面 曲面にラップやウィンドウを施した最終形	ストーリーボード ストーリーパターン シーンデザイン
段落	楽節 4~16小節程度の楽曲の一部	モチーフ 構図を構成しているパーツ	曲面/曲面群 キーラインに基づいて形成された曲面、およびその組み合わせ	動作ベクトル (キーフレーム) 動作 (時系列) 中の代表点
文	旋律/和音進行 区別感や終止感をもたらす部分	筆触 一筆で描く範囲	キーライン 曲面の主要断面やアクセントとなる曲率早期曲線の組み合わせ	動作ベクトル (運動) 各フレームに対応する動作
句	短いフレーズ/和音/リズム 音程方向あるいは時間方向の最小のグループ		自己アフィン性を有する曲線 複写言化された曲率早期曲線群 (図3参照)	
音素	楽音 五線譜に記述された音階と音値を持つ音符	画素 (ピクセル) 色を有した微小面積	曲率 曲率早期曲線を構成する各構成点での曲率	
音声信号	音響信号 サンプル信号列		曲線の構成点 曲率早期曲線を構成する各構成点	動作映像 ビデオ情報

図 7: 各非言語メディアにおける記述 (定式化) の一例. 音楽は西洋調性音楽, 造形については自動車における曲面デザイン, モーションについては CG における人物等の動作を想定している.

例えば, 3.2 節で述べたように, 逸脱も含めた適切なコンテンツ表現が定義できればユーザにとって直感的で理解しやすい操作インタフェースが提供できよう. 事例の導入で情緒表現を経由せずに済むようになれば, 制作者間の意志疎通が円滑になり, 使い心地 (desirability) の確保につながる事が期待される.

具体性の高いデータからより抽象度高いデータ記述への変換過程 (認識処理) についてはいくつかの課題がある. 例えば, 自然言語処理において, 一部の例外を除きほぼ自明として扱われていた句の分離も, 実際に人間が使用している音声を対象とする場合には, 別途, 対応が必要となる [Fujisaki 04]. これらの問題は, 非言語メディアにおいては, より顕著である. 特に, 群化形成のモデリングは, これからの重要な研究対象となっている. また, 領域毎に固有の問題が存在するため, 方法論の共有化をはかる際には, その対処も必要である.

この記述の階層構造は静的な構造であり, 対象メディアの側面を表現しているに過ぎない. 他にも, 例えば, 制作と観賞プロセスの構造を対応付けることもできる. 音楽や舞踊といった時系列メディアの鑑賞者は, 作家, パフォーマーと同じ時間軸上を進まなくてはならないのに対し, 絵や写真といった非時系列メディアでは, 鑑賞者の時間の制御はもっぱら観賞者に任されている. ただし時系列メディアでも, 現在を見 (聴き) なが

ら過去を想起したり未来を予測したりしている点で, 鑑賞者は, 制作者 (作家, パフォーマ) の時間からある意味で独立と言える. 他方, 絵画においても, 鑑賞者の視線の移動を促すよう意図されたデザインがある.

5. おわりに

本論文では, 音楽, 造形, 絵画, モーションの各領域におけるデザイン支援に関する研究事例を紹介し, 非言語メディアのデザイン支援における課題と可能性を探ってきた.

非言語メディアのデザイン支援に向けての我々の主張のポイントは以下のように整理される. 1) elaboration を記述する為のデータ記述を, 可能なら, 認知構造との階層化を考慮した形で用意すべきである. 2) ユーザ支援の観点からは, 「個性」や「様式」の代表として, 事例を参照する形をとることが望ましい. 3) 具体的なデータ記述からより抽象度のデータ記述へ変換する過程として, 群化形成のモデリングが重要な課題となる. この領域の研究を進めるメリットとしては, 文字通り, 4) 職業的なデザインにおける生産性の向上, がある. その他に, 5) 知育メディアへの応用・展開も視野に入ってくる.

デザイン支援における事例に基づく方法論の研究は, 「真似る」すなわち「学ぶ」ことに関する科学である. 我々は, この新しい研究分野が, 実应用到に深いつながりを持つと同時に, 非常に基礎的な人工知能に関する研究テーマであると筆者らは感じている.

参考文献

- [Cope 91] Cope, D.: *Computers and Music Style*, A-R EDITIONS (1991)
- [Fujisaki 04] Fujisaki, H.: Retrospects and Prospects of Speech Communication Research, in *Proc. of the 18th International Congress on Acoustics*, pp. V3763-V3768 (2004)
- [原田 98] 原田 利宣, 森山 真光, 吉本 富士市: 視覚言語を用いた曲線創成システム, *デザイン学研究*, Vol. 45, No. 3, pp. 63-70 (1998)
- [星野 03] 星野 准一, 中野 敦, 森 博志: モーションデザインの再利用によるアニメーション作成支援, *人工知能学会全国大会論文集*, 1B1-07 (2003)
- [笠尾 01] 笠尾 敦司: Linux GIMP のプラグイン: SIC で目指す芸術と科学の融合-数行のプログラムの変更がもたらす個人的な CG 表現の世界-, in *Linux Conference 2001* (2001)
- [片寄 03] 片寄 晴弘, 平田 圭二, 原田 利宣, 平賀 瑠美, 笠尾敦司: 事例に基づくデザイン支援と評価基盤の構築, *人工知能学会論文誌*, Vol. 18, No. 1, pp. 24-28 (2003)
- [Kepes 44] Kepes, G.: *The Language of Vision*, Paul Theobald (1944)
- [Lerdahl 83] Lerdahl, F. and Jackendoff, R.: *A Generative Theory of Tonal Music*, MIT Press (1983)
- [長町 93] 長町 三生: 感性商品学: 感性工学の基礎と応用, 海文堂出版 (1993)
- [大澤 00] 大澤 光編: 印象の工学とは何か, 丸善 (2000)