

作品系コンテンツのための自動作曲システムに向けて(1)

長嶋洋一 (SUAC/ASL)

概要： 作品系のショートムービーやフラッシュなど、作品系マルチメディアコンテンツを制作する際に活用する背景音楽(BGM)パートを、お手軽に「使える音楽データ」として自動生成するシステムの開発に向けて、作品系コンテンツの分類整理検討を行い、音楽的な特徴について考察した。

Towards the automatic composition system for artistic contents (1)

Yoichi Nagashima (nagasm@computer.org)

Absreacts : This report is a part of the research of the system that composes the music part to produce the artistic contents automatically. I collected a lot of FLASH works, classified those features, and discussed the tendency.

1. はじめに

筆者はこれまでComputer Musicの作曲家・研究者として、主としてメディア・アート領域での作品創作発表・研究を行ってきた[1]。最近ではネットワーク利用の音楽セッションシステム[2][3]、コンテンツにおける映像と音楽のビートの同期[4][5]など、エンタテインメントやメディア心理学のテーマに取り組んでいる。本研究はこの関連として、Webコンテンツの一つであるFLASHと音楽にまたがる領域を対象とし、「作品系FLASHコンテンツの[使える]音楽パートを自動生成するシステム」(2005年前期IPA「未踏ソフトウェア創造事業」採択)の開発とも関係している。

2. 研究の背景 - FLASHをめぐる状況

マルチメディア系コンテンツの開発環境という視点から、筆者は著作の中でいくつかの具体例を検討紹介してきた[6][7][8]。

図1はSun社が提唱し世界中に普及したJavaの開発中の風景である。Javaは仮想マシンの採用によりplatform非依存の初めての成功例として、インターネット時代のキラコンテンツ第1号(全てのブラウザが標準プラグインに対応)となった。セキュリティと提供サービスとのトレードオフ、エミュレーションによる遅さなどの課題を抱えながらも、亜流でない「100% pure Java」は、サーバーとともにも今後も発展すると思われる。

図2はMacromedia社がJava以前から提供しているDirector(Webコンテンツとしては「Shockwave」)の開発中の風景である。Directorはもともと、複雑・大規模・高度なコンテンツを制作するためのオーサリングツールであり、C言語などのプログラミングスキルの無いデザイナーやアーティストの開発を支援するが、製品はプロ用ソフトとしてかなり高価である(Shockwave Playerプラグインはフリー)。広告業界などで有名な映像制作ソフトの「Motion

Dive」自体もDirectorで開発されるなど、現在でもアプリケーション開発とコンテンツ制作の主流ツールである。

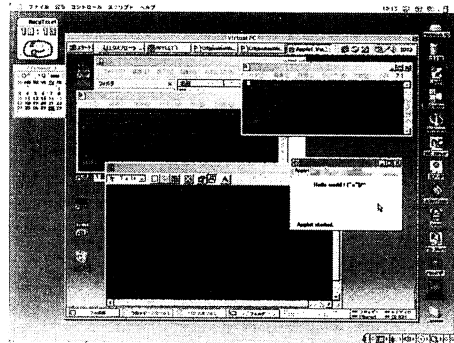


図1 Javaによる開発風景

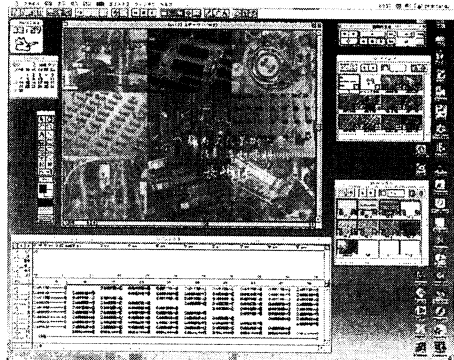


図2 Director (Shockwave)による開発風景

図3はFLASHによる開発中の風景である。歴史的経緯としては、元々インターネット向けに「軽いアニメーション」コンテンツを制作する「Future Splash」というソフトがあった。Directorはグラフィクスがbitmap起源で、パン

コン用CDROMコンテンツはともかく、Web用データ(dcrファイル)サイズが大きい欠点があった。Future Splashはvectorグラフィクスでサイズが小さいが、専用プラグインが敬遠された。そこでMacromedia社はFuture Splashを会社ごと買収しFuture Splash Animator(Flash 1)として発表、以後、FLASHとして両者のプラグインを統合したShockwave(フリー)を提供した。現在ではほぼ全てのブラウザがShockwave(swfファイル)標準対応となり、FLASHは一気にWebコンテンツ制作ツールの主役に躍り出てきた。2005年には、静止画/ドキュメントベースの世界標準WebコンテンツのAcrobat(PDF)を提供するAdobe社が約4000億円でMacromedia社を買収すると報じられた。Macromedia社はこれまでもDirectorとFlashの機能を相互に高めて(互いの長所を重複して共有)いたが、これによりますますダイナミックなWebコンテンツの制作環境が進展することを期待したい。FLASHコンテンツの多彩な表現と軽さはパソコンだけでなく、携帯電話などインターネット対応の電子機器への組み込みにも好適で、Sonyは携帯電話用ポストペットを全てFLASHで提供すると発表した。Java(iアプリ)のエミュレーションに対し、専用プレーヤにより「小さく・速い」FLASHは、これからさらに爆発的に普及する可能性を感じさせる。

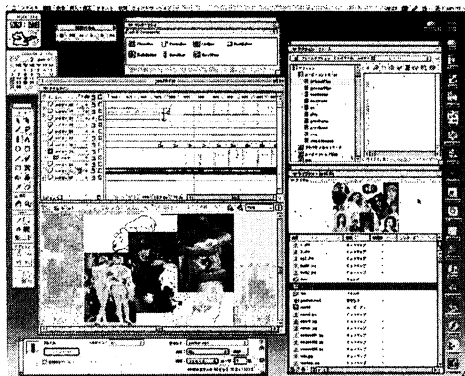


図3 FLASH MXによる開発風景

FLASHのオープン性も重要である。筆者はまだ日本語Java資料の無い時代に、sun社がWebで公開するドキュメントだけを頼りに、Javaバイトコードを逆コンパイルするツールを容易に開発できた[9]。FLASHについてもMacromedia社はこのオープン性を重視し、詳細なドキュメントをフリーで公開している[10][11]。この190ページのPDF資料により、他社ソフトはデータをFLASHフォーマットで書き出すことができ、研究者はFLASHコンテンツを生データのまま詳細に解析できるのである。

演奏された音楽音響を対象とした分析には「音源分離」「ビートトラッキング」「フレージング抽出」などの難問がある[12]。そこでMIDI音源のMIDI情報受信から発音までの遅れのばらつきを考慮した上で、MIDI化音楽演奏情報

の解析が有効である[13][14][15]。コンテンツの視覚的情報と音響情報との関係を研究する場合、映像トラックとサウンドトラックとが別々のブロックに圧縮されているようなムービー/ストリームデータでは、DVテープを記録した際に起きる「音ずれ」[16]のような同期性の問題も起きる。FLASHはタイムラインをベースとしたデータ体系であり、SMF同様に仕様が詳細に公開されている。そこで生データをそのまま分析することで、実験環境の時間的遅れや同期のずれを心配することなく、マルチメディア間の関係を正確に解析できる。メディア心理学実験において、提示データについて明確に定義された実験をデザインでき、被験者データの解析部分に安心して集中できるのは、研究上、非常に大きなメリットであろう。

3. 研究の目標 - 「使える音楽」の提供

本研究と関係するプロジェクトとして、中期的テーマ「誰でも手軽に作品系FLASHコンテンツの音楽パートを自動生成するシステム」の開発が、IPA「未踏ソフトウェア創造事業」に採択された[17]。この開発の目標は、IT時代の色々なコンテンツ、特にShort MovieやFLASH等の作品系マルチメディアコンテンツを制作する際に必須となる「背景音楽(BGM)パート」として、お手軽に「使える音楽データ」を自動生成するシステムを開発することである。最終成果のシステムとアルゴリズムはWeb上でフリーに公開・発表し、FLASH等の作品系コンテンツ制作支援環境とも自由に統合できるように提供する計画である。

これは短期的にはクリエイター(とその卵)のためのシステムであるが、3年後には到来する「クリエイション(作品創造)の大衆化」という時代的要請を視野に入れている。ITが普及する中、「デジカメ写真を自動スライドショー化」「ホームビデオを半自動編集してmovie作品化」「FLASH作品のお手軽制作ツール」など、一般大衆がマルチメディア作品系コンテンツを「自分で作る」というエンタテインメントの時代が到来する。そこで問題となる音楽パートの創作を、著作権の問題をクリアし、音楽的な専門知識を要求することなく、容易にコンテンツに適した「使える」音楽を自動生成するためのシステムの開発を目指す。

4. 作品系FLASHの収集と分類

本研究においては「作品系FLASH」と対象を明確に限定した。そこでまず第一に「Webサイトのナビゲーション機能としてのFLASH」を対象から除外した。これはFLASHの目的がまったく違うからである。ナビゲーションFLASHの例としては、資生堂・ソニー・日産自動車など、日本の多くの企業サイトに見ることができる。

基本的にはこれ以外のWebページ(FLASHプラグイン)で見ることができるものは、一応FLASHの「作品」(作家が制作した成果物という意味)で

ある。ただしそれに加えて、本研究では「どつきり系FLASH」「エロサイト系FLASH」「グロサイト系FLASH」「ゲームFLASH」「外国語能力の必要なFLASH」なども除外した。この理由・詳細については、文献[18]を参照されたい。

以上のような基準を整理した上で、2005年1月から6月にかけて「ネット上からFLASHをとにかく集める」作業を行った。Webには所謂「FLASH集サイト」が多数あるが、この中には上記で除外した「どつきり系/エロFLASH/グロFLASH」等も多数、直リンクされており、収集作業を行った筆者は望まなくても全てからその直撃を食らった。著作権や肖像権を侵害したり名誉毀損(誹謗中傷)ネタのために、ペンネームあるいはクレジットのまったく無いFLASH作品もとても多く、これらはリンク先から消え“Not Found”であるものが多かった。そこで、実験のためのFLASHライブラリとしてリンクを使用することを断念し、ソースのswfファイルを全て手元にダウンロードした。FLASHを作る作家/グループはアマチュアもプロも世界中に多数おり、またFLASH作品は制作まで1年かけた熟練の名作から「FLASHを始めて3日です」的な稚拙なものまで千差万別である。収集作業において閲覧したFLASHは7000本以上、その中で上記「対象外FLASH」「駄フラ」「糞フラ」等を捨てて、とりあえずswfファイルを手元に保存したFLASHは6月25日時点で約4000本となった。

本研究で対象とした作品系FLASHの大部分は、Wikipediaの分類で言えば「アニメーション」と「音楽再生」ということになる。しかし作家の創造性と表現方法は無限であり、機械的に類型化することにはもともと限界がある。分類項目の一つとして「その他」は最後まで必要であると思われる。以下、作業の中で注目したいいくつかの代表的な「作品系FLASH」の例を列記するが、紙面の関係でこの詳細な解説は省略する。

音楽系FLASH

MTV/プロモ系 (cool)
Motion Graphics (VJ)系 (cool)
Motion Typography系 (cool)
ジュークボックス系 (多ch活用)
J-POP等のカラオケPV (歌詞付)
空耳音楽系・外国語/方言系 (笑系)
ラップ系・会話音声の音楽化 (笑系)

物語系FLASH

恋愛もの・冒険もの・歴史もの
生命もの(生・死・運命)(感動系)
テーマもの(ホラー・和風・ご当地)
ドキュメンタリー・Project Xもの
不条理もの・不思議系・抽象芸術
キャラもの(ゲーム・アニメ・2ch)

お笑い系FLASH

Motion Typographyもの(セリフ+字幕)
パロディー・コラージュ
事件・政治家・被告等をおちゃらける
4コマ漫画的な連載/テーマもの

メッセージ/風刺/告発系FLASH
政治テーマ・国際テーマ
市民運動テーマ・環境テーマ
イベントのCM/プロモ
パソコンもの(Windows批難系)

ネットマンガFLASH

紙芝居系(静止画をクリックで進める)
TVアニメのWeb版(出版社が運営)

インタラクティブ系FLASH(純ゲームを除く)

浮遊感覚もの
シミュレーションもの

その他

本研究の場合には、(1)「紙芝居」的に静止画をクリックして送るタイプ、(2)セリフが字幕だけで出る(読む)サイレントムービー、(3)サウンドがインタラクティブ性に附随した効果音響だけである作品、(4)空耳系やパロディー系で対象となる音楽がテーマとして限定される作品群、(5)演出効果として「無音」を意図したもの、など、制作においてオリジナル音楽を必要としないFLASHについては、最終的には対象から外れるという性格がある。ただし、研究の途中ではこれらも考察・検討の対象として注目した。

5. 「使えない」問題点の検討

本研究で対象とする「作品系FLASH」の制作において、3年後の一般大衆、現在のクリエイター(とその卵)が熱烈に希望するのは、「使える」音楽データをお手軽に欲しい、という事である。もちろん、DTM(打込み音楽)ホビーストや音楽の専門知識のある人であれば、MIDI機器やDTMソフトウェアによって、オリジナルの音楽を作って楽しむ環境は、ITの進展とともにとても充実している。しかし、いわば音楽の専門家でない一般大衆・ビジュアルデザイナー・クリエイターにとっても、マルチメディアコンテンツを創造するために、「使える音楽」が求められていて、これは現状では解決されていない。以下は、現状でどう「使えない」のか、のうち主な3点である。

(1) 入手する - 著作権のため使えない

いかに自分がお気に入りの楽曲であっても、プロアマ問わず他人の作った音楽は、作品をWebで公開する等の用途には無断では使えない。その楽曲を自分で「耳コピー」したり市販の楽譜を打込んでMIDIデータを作る、すなわちオリジナルアレンジであっても駄目である。現在、クリエイターの卵(コンテンツデザインの学生など)が行っている対策は、(a)知り合いに作曲・提供してもらう、(b)著作権フリー音楽集の利用、(c)音楽シェアウェア(有料で使用許諾)の利用、などである。その結果、「あの作品と音楽は同じじゃん」等のフラッシュ作品があちこちに生まれ、せっかくの創作が淋しいものになる。

(2) 市販ソフト - いくつかの問題点

これまでに発表・市販されてきた多くのMIDIベースの「自動作曲」「作曲支援」ソフトの中にも、音楽的知識により個々の音符を打ち込むことなく「お手軽作曲」する、という音楽生成ソフトウェア製品は存在する。しかし「ポップス風」「カントリー風」等の別売ライブラリを買い揃えないと実用にならず、ミニマル自動生成系のソフトでは開発元のマニアックな思い入れからか機能/操作が複雑で、面白い楽曲出力を得るためには、グラフィックソフト以上の投資額と専門的知識が要求される。お手軽に音楽を欲しい立場にとっては、これはかなり敷居が高い。

多数のサウンドフレーズ断片を選択して並べて繋げる、というMac付属の“GarageBand”はシンプルで画期的な製品である。しかし実際にフラッシュ作品のサウンドトラックを制作してみると、多くのユーザが好むフレーズは膨大なライブラリの一部に集中して、バリエーション不足(同じような音楽の重複。発売されているオプションフレーズ集の購入が必要とより高つく)、さらに後述の時間的編集性の点で不満が残る。

(3) 研究者の自動作曲システム - 面白くない

複数(有限個)の素材を用意し選択し接続していく、という手法には上記のような問題があり、自動作曲システムの本質的な指導原理としては「新たに音楽要素を生成する」というアルゴリズム作曲の手法(既におよそ50年ほどの歴史)が本命である。音楽情報科学の領域では、このようなアルゴリズム作曲による自動作曲の研究は、現在でも多くの大学等で細々と繰返されているが、その成果は惨澹たるもの(ポップでもなく現代音楽風でもなく芸術的でもなくただひたすら稚拙な音楽)で、使える成果は得られていない。その最大の理由は、「音楽的基礎/素養/情熱/勉強」の不足した工学者が、安易にIT応用のため皮相的に音楽テーマを選択している事であり、音楽を分かっている工学者の開発する自動作曲システムには、今後もあまり期待できない(少なくとも筆者の個人的な期待は20年近く裏切られ続けている)。

6. 作品系FLASHの音楽の特徴

本研究は「あらゆる音楽をなんでも自動作曲する」などという無謀な野望を抱いていない。市販の高機能シーケンスソフトやマニアックな自動作曲ソフトでは、あれも出来るこれも出来る、と「どのような音楽スタイルでも作曲できる」万能性・網羅性を指向することで、ある目的を持つ専門的でないユーザの希望から乖離している。

ここで対象と考えている、ショートムービーやフラッシュ作品のためのサウンドトラックとしては、

- ・全体/部分の時間的長さ(「尺」を固定)
- ・イントロ、エンディング
- ・リピート、ループ、ブレイク
- ・ブリッジ、フェード

などの音楽的要素・音楽構造に独特の傾向と要請がある。また、より高度なマルチメディアコンテンツ作品の創作において、グラフィックスのパートと音楽のパートとの同期は非常に重要な要素である。ここでは、

- ・サウンドのある部分だけを時間的に伸縮
- ・没入感/躍動感/幻惑感/浮遊感などの演出
 - 映像と音楽のビートを微調整
 - 「ノリ」「タメ」「グルーブ」

などの操作を駆使して演出効果を盛り込むことを指向するが、これは“GarageBand”のようなサウンドファイル形式のシステムでは困難である。

7. システム開発研究の基本戦略

以上の分析検討を受け、本研究では、「使えない」部分の課題に対して、以下のような戦略を検討・構想・実験して、「使える音楽データ」としての解決を目指している。

著作権については、自動生成アルゴリズムを採用し、本質的に既存の楽曲情報を利用しない。これと同時に、生成された楽曲の特徴情報を圧縮してインターネット検索できる手法を、システムの一部として構想している。この手法を活用すると、既存の楽曲の特徴情報を同様にデータベース化することにより、本システムで生成された楽曲が既存の楽曲と天文学的確率で偶然に一致する、という偶発事象的な著作権侵害を未然にチェックできる可能性を提供でき、オリジナル作品としての著作権の懸念は霧散し、安心してコンテンツの音楽パート生成に「使える」ようになる。

音楽データのアルゴリズム作曲自動生成は、基本的にMIDI情報ベースで行う。これは上記の「作品系コンテンツ創作における特徴」に対応するためである。その出力は、ユーザが好みの音源システムで音響データに変換し、さらにFLASH化のためMP3データ化するシステムと統合する。アルゴリズムに乱数(不確定)要素を持たせ度合いについては、自動作曲の確率統計アルゴリズム中でユーザが選択できるようにする。

対象を「作品系マルチメディアコンテンツの音楽パート」と限定した自動作曲アルゴリズム・エンジンのための新しい概念とモデルの概要は、現時点では以下のようなものであり、今後、多くの専門家との議論により内容を向上させていく。対象となるコンテンツにおいて、本質的に音楽パートは「唯一の主役」ではない。これは、ダンス系のDJ/VJの音楽や演劇/ミュー

ジカルの音楽と共通する特性であり、イメージを一言で例示すれば「Drum'n Bass」である。この音楽は、ビジュアルあるいはダンスパブルな主役のメディアとともに、コンテンツの聴取者に快適な音楽を提供して没入感その他の演出効果を上げるが、骨格部分において、古典的な音楽の要素から、メロディー・和声・(リズム)を捨象している。

これまでの音楽情報科学研究者の自動作曲のアプローチでは、大部分がこのメロディー・和声からスタートしているから「使えない」とも言えそうである。本システムの自動作曲アルゴリズムでは、自動生成の基幹はドラムとベースであり、ビートとスケール、という最小限の音楽的枠組みを豊富な自由度とともに生成して、その後他の音楽的要素を付加してユーザが「使える」楽曲に成長させる、という戦略を取る。「Drum'n Bass」というのはあくまで入口の例示であり、この発想はスローバラード的な音楽でも、クラシカルな音楽でも、レゲエでも演歌でも、気付いてみれば共通の重要性を持っていると考える。

8. おわりに

本稿では「誰でも手軽に作品系FLASHコンテンツの音楽パートを自動生成するシステム」の開発研究における、音楽的な検討としては第1報となる報告を行った。研究の背景と目的について整理・紹介するとともに、既存の多数のFLASH作品を収集し、特徴を検討して分類し、作品系コンテンツ、という対象の特性について検討した。今後、具体的な研究をさらに進めて、目標とするシステムの開発・応用へと進めていきたい。

参考文献

- [1] <http://nagasm.org/>
- [2] Y. Nagashima : "IMPROVISESSION-II" : A Perfprming/Composing System for Improvisational Sessions with Networks, Proceedings of International Workshop on Entertainment Computing, 2002
- [3] Y. Nagashima : "GDS (Global Delayed Session) Music - new improvisational music with network latency, Proceedings of 2003 International Computer Music Conference, ICMA, 2003
- [4] 長嶋洋一 : 音楽的ビートが映像的ビートの知覚に及ぼす引き込み効果, 芸術科学会論文誌 Vol.3 No.1, 芸術科学会, 2003
- [5] Y. Nagashima : "Drawing-in effect on perception/cognition of musical beats and visual beats, Proceedings of International Symposium on Musical Acoustics, ISMA, 2004
- [6] 長嶋洋一 : 「Java & AKI-80」, CQ出版, 1997
- [7] 長嶋洋一 : 「コンピュータサウンドの世界」, CQ出版, 1999
- [8] 長嶋洋一 : 「よくわかる組み込みシステムのできるまで」, 日刊工業新聞社, 2005
- [9] Java逆コンパイラ "rejava.c"
<http://nagasm.suac.net/ASL/indy/>
- [10] Macromedia Flash SWF File Format Version 7,
<http://www.digitalpreservation.gov/formats/fdd/fdd000130.shtml>
- [11] The Macromedia Flash SWF File Format SpecificationVersion7,
http://download.macromedia.com/pub/flash/flash_file_format_specification.pdf
- [12] 長嶋・橋本・平賀・平田 : 「コンピュータと音楽の世界」, 共立出版, 1998
- [13] 長嶋洋一 : MIDI音源の発音遅延と音源アルゴリズムに関する検討, 情報処理学会研究報告Vol. 99, No. 68 (99-MUS-31), 情報処理学会, 1999
- [14] 長嶋洋一 : MIDI音源の発音遅延と音楽心理学実験への影響, 日本音響学会音楽音響研究会資料 Vol.18, No. 5, 日本音響学会, 1999
- [15] Y. Nagashima : "Measurement of Latency in Interactive Multimedia Art, Proceedings of International Conference on New Interfaces for Musical Expression, 2004
- [16] DVから記録したDVDの「音ずれ」を考える
<http://nagasm.suac.net/ASL/otozure/>
- [17]<http://www.ipa.go.jp/jinzai/esp/2005mit01/koubokekka.html>
- [18] 長嶋洋一 : 作品系FLASHコンテンツの分類と傾向について, 情報処理学会研究報告 Vol. 2005, No. 59 (2005-EC-1), 情報処理学会, 2005