

# 生成音楽評価の20年

橋田 光代<sup>1,a)</sup> 鈴木 泰山<sup>2</sup> 奥村 健太<sup>4</sup> 馬場 隆<sup>3</sup> 柴崎 正浩<sup>5</sup>

概要：演奏・学習支援，作編曲，表情付けなど，人間の音楽活動をより豊かにすることを目的とした研究が盛んである．それらの研究評価においては，とくに，生成された音楽に対する評価について，様々な取り組みがなされてきた．本稿は，オーガナイズドセッション「生成音楽の評価学」の導入として，過去の音楽情報科学研究会で取り組まれてきた研究評価について概観する．

## 1. はじめに

音楽は我々の日常生活の中において最も身近な文化活動の一つである．発足から20周年を迎えた音楽情報科学研究会（SIGMUS）においては，かつて「趣味の研究」と言われ[1]，まともに研究対象とみなされてこなかった時代も経験しつつ，一方で，感性工学，あるいはエンターテインメント研究の一分野としてコンピュータテクノロジーの発展と普及に大きく関わってきた[2]．

家庭用コンピュータが普及し始め，またMIDIが正式に規格化された1980年代当初から，Performer, Vision, レコンポーザに代表される音楽アプリケーションが多数開発・発表されてきた．昨今の，スマートフォンに代表される携帯端末用アプリケーションにおいても，単純な楽器演奏ツールから往年のシンセサイザーのエミュレータ，作曲ツール，演奏ゲーム，楽曲検索や音感トレーニングアプリなど多岐にわたってシステムや作品が発表され続けている．

音楽生成系のシステム作りは，「音が出る」という事象自体への興味や憧憬，感動を伴うことから，プログラミング学習の初歩として，また音楽情報科学研究の入口としても取りかかりやすいものである．しかしながら，いざ研究として開発を進めてみると，いずれかの段階で，人間の演奏行為，あるいは音楽（場合によっては，音）そのものを扱うことの奥深さ・難解さと“格闘”することになる．なか

でも研究に直結し，問題になってくるのが，生成された音楽をどのように評価すればよいか，という点であろう．

一般に，研究の有用性や妥当性，先行研究に対する優位性などを示すには，研究による生成物を客観的な手法で評価するプロセスが不可欠である．しかし，こと音楽が題材となると，音楽にまつわる人間の経験則や感性とも照らし合わせることで，すなわち主観的な評価も必要になってくる．この背景には，音楽的感性や経験則の多くが，長い年月をかけて蓄積された一種の暗黙知として流通しており，今なお統一的な言語化や規格化が進んでいないことが挙げられる．この現状を踏まえて，いかに取り組むかが，音楽生成研究をより面白くするか退屈にさせるかの判断の分かれ目となる．

本オーガナイズドセッションでは，音楽生成系研究における生成音楽の評価に着目し，過去の研究者らの取り組みを踏まえて，今後の生成音楽の評価の可能性について議論する．その導入として，本稿では，これまでSIGMUSにて発表されてきた音楽生成系の研究を集積し，生成音楽の評価への取り組み方について概観する．

## 2. 音楽生成を伴う研究の動向

まず，これまでのSIGMUSにおいて発表されてきた研究発表群から，音楽生成系の研究発表を抽出し，その全体像を俯瞰する．

本稿で取り扱う音楽生成系研究とは，楽譜やMIDI，音響信号など，音楽情報を表現するデータを出力するシステムやモデル，技術を提示しているものを指す．代表例としては，作編曲，自動演奏，自動伴奏，演奏表情付け，自動採譜などである．演奏インタフェース（電子楽器）や学習支援システム，音楽構造分析ツールなど，自動生成に限らずユーザの演奏活動を支えることを目的とした研究も含まれる．音楽ジャンル分類や楽曲推薦，歌詞支援，音楽から

<sup>1</sup> 相愛大学音楽学部  
Soai University

<sup>2</sup> 株式会社ピコラボ  
Picolab Co. Ltd

<sup>3</sup> 関西学院大学理工学研究科  
Kwansei Gakuin University

<sup>4</sup> 名古屋工業大学  
Nagoya Institute of Technology

<sup>5</sup> 芝浦工業大学  
Shibaura Institute of Technology

a) hashida@soai.ac.jp

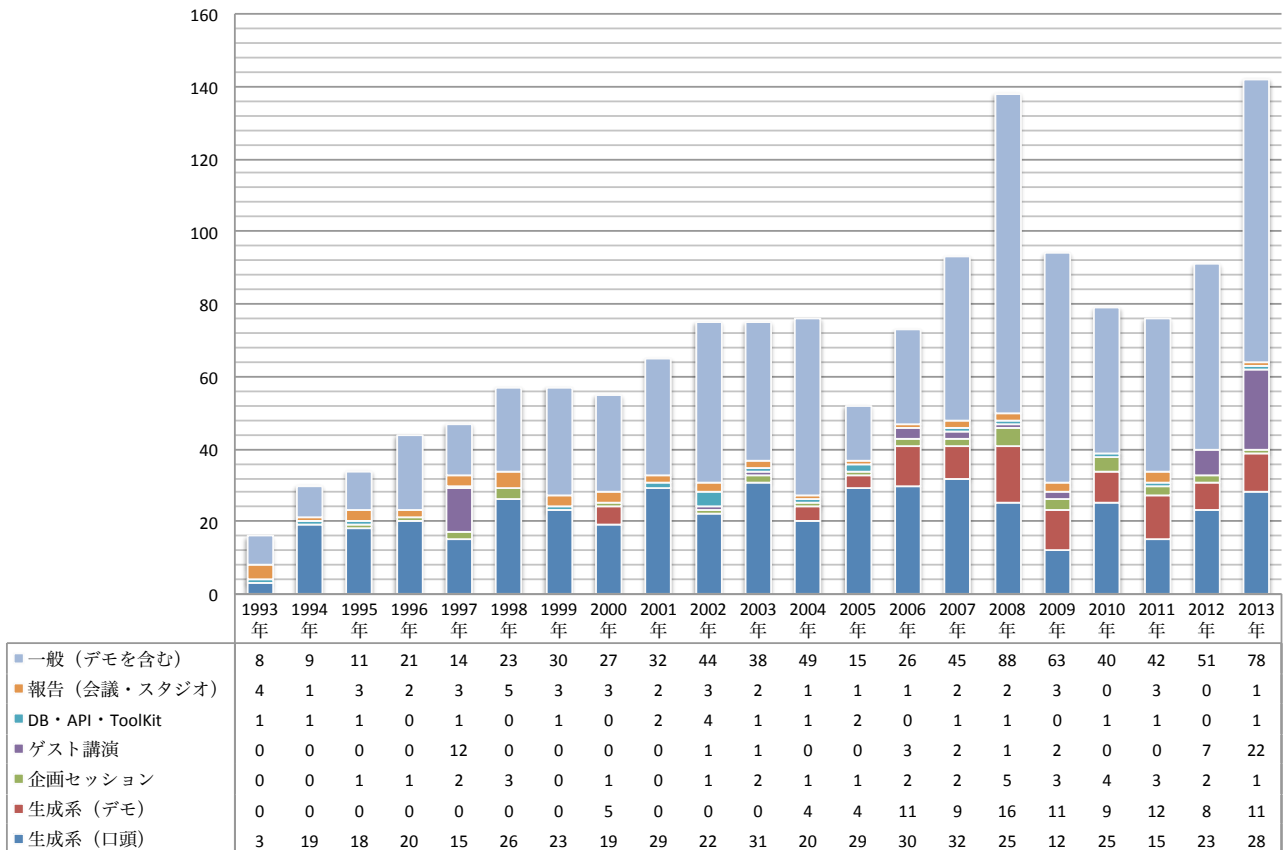


図 1 SIGMUS での年別のおおざっぱな発表分類。

の映像生成など、音や音符の出力と関連の薄いものについては対象外とする。

第 1 回 SIGMUS (1993 年) から前回の第 101 回 (2013 年) まで、発表件数は総数 1,568 件を数える。この数字は、情報処理学会の電子図書館「情報学広場」の研究報告目次に掲載された全タイトル 1,418 件と、8 月例会 (夏シンポ) 等で開催されるデモセッションへのエントリー 170 件を合算したものである。これらについて、年度別に分けた音楽生成系研究の発表数の推移を図 1 に示す。

### 2.1 音楽生成系研究の口頭発表数の推移

研究会全体を通して、発表件数は年々増えており、音楽情報科学研究が着実に成長発展を遂げているのはこれまでに報告されてきた。その中で音楽生成系研究は、1993 年当時から、概ね例年 20 件前後で発表されてきている (図 1 「生成系 (口頭)」を参照)。増減はあるが、長年の間、発表数がほぼ安定しているというのは、時代の変遷によって研究で利用される PC 環境やデバイスが移り変わって行くなか、常に音楽生成に対する関心が保たれていることを伺わせる。全期間を通じて、その総数は 464 件 (全体の 32.7%) となった。

### 2.2 デモ発表数の推移

2004 年から、SIGMUS では、主に夏シンポにおいてデモセッションが設けられるようになった。

音楽に限らず、コンテンツ制作・生成を行う研究の場合、論文という形式では伝えにくい部分を「百聞は一見にしかず」で提示する重要性が近年認識されるようになってきている。研究の意図や面白さを伝えるために、デモや動画を通じてシステムの挙動を実際に見せる機会も増えてきている。音楽生成系研究における生成物は、すなわちその大半が音そのものであり、聴取によって初めて生成の正否が正しく理解できる場合が多い。しかし、発表件数が多い場合などでは、1 件あたりの発表時間が短縮されることもしばしばで、実際にはせいぜい 1 ~ 2 分程度で断片的にしか見せられないのが実情である。

デモセッションは、会場でシステムの動作をじかに確認しながら、研究者と充実した議論を行えるよう (1) 予稿の執筆や口頭発表の負担を減らし、(2) 既発表、開発中、あるいは議論そのものを目的とした内容で気軽に参加できるよう設計されているのが大きな特徴である。セッション全体では、例年 10 ~ 20 件程度の発表がエントリーされている。そのうち、音楽生成系のデモ発表 (図 1 「生成系 (デモ)」を参照) はほぼ毎回半数以上を占めている。

ここで、音楽生成系研究の口頭発表ならびにデモ発表の

表 1 音楽生成系研究の評価方法：口頭発表 464 件のうち、2007 年以前については、期間中の全音楽生成系研究のうち年度別に約 3~4 割ずつをサンプルとして任意抽出した。

口頭発表	評価方法【件数(割合)】 複数実施を含む							
	定性的評価					定量的評価		記述なし
	研究者による 分析・考察	専門家による 批評・考察	被験者による主観・体験			手法比較	データ比較	
10 人未満			10 人以上	人数不明				
1993~2007 年 (184 件)	48 (26.1%)	2 (1.1%)	11 (6.0%)	8 (4.3%)	13 (7.1%)	16 (8.7%)	26 (14.1%)	60 (32.6%)
2008~2013 年 (138 件)	22 (15.9%)	2 (1.4%)	22 (15.9%)	27 (19.6%)	5 (3.6%)	12 (8.7%)	18 (11.6%)	32 (23.2%)
合計 (322 件)	70 (21.7%)	4 (1.2%)	33 (10.2%)	35 (10.9%)	18 (5.6%)	28 (8.7%)	42 (13.0%)	92 (28.6%)

件数を合算してみると、2006 年から 2008 年にかけて、また本年 2013 年は、それぞれ 40 件に達していることがわかる。前述のとおり、デモ発表は、研究の完成度合いよりも、システムを動作させることのほうが重要である。半ページという予稿の紙面の都合もあり、評価についてまでは触れずに済むことも多い。その気軽さが、発表件数を増加させる一因を担っていると考えられる。

### 3. 音楽生成系研究の評価手法の分類

次に、音楽生成系研究の各事例で取り上げられた評価手法について分類したものを表 1 に示す。ここでは、2.1 節で抽出した音楽生成系研究の口頭発表 464 件のうち、(1) 直近 6 年間となる 2008 年以降の 138 件すべてと、(2) 2007 年以前については年度別にそれぞれ年度件数の約 3~4 割ずつをサンプルとして任意抽出した 184 件、の合計 322 件を対象とする。調査方法としては、「情報学広場」記載の概要や予稿本文を資料として、評価に関する記述をリストアップし、著者らで手作業による分類を行った。なお、1 件の発表の中で、複数の評価手法を行っているものについては、それぞれをひとつの評価手法として数えている。

全体として、評価手法は大きく定性的評価と定量的評価に分けられる。以下、表 1 に沿って述べる。

#### 3.1 定性的評価

##### (1) 研究者による分析・考察

第 1 章で述べたように、音楽生成系研究は、「音が出る」という事象自体への興味や憧憬、感動を伴うことから、まずは研究者自身が納得のいく生成を目指すのは自然なことであると思われる。多くの場合、研究者自身がその音楽生成に関連する何らかの音楽経験を有していることが考えられる。その場合、なんとなく好きという漠然としたものではなく、その分野に関しては周囲の人間より深い知識やこだわりを保っているため、まずは本人が評価するのが妥当という考え方である。

そこで評価の第一段階として、自身の経験知に基づいて、生成物の聴取評価やシステムの使用感などを分析している発表がこの項目である。2007 年以前の研究事例はこの自己分析による考察が最も多く、期間全体の 26.1%を占めている。

##### (2) 専門家による批評・考察

研究者自身がそれなりに音楽的知識を備えている場合においても、本人の主観的な評価だけではやはり客観性に欠ける側面は十分にある。そこで、第三者が生成物を評価しても納得できるレベルであることを確かめる方法のひとつとして、対象分野における職業専門家に生成物を評価してもらうことが挙げられる。例えばプロとして活動しているピアニストであったり、指揮者、作曲家、ジャズ奏者などが考えられる。研究遂行にあたり、そのような職業専門家と交流する機会を得ることが実際には難しいところであり、発表数も数年に 1 度と頻度は少ない。だがこれが実現すれば、研究の質をより高めることも可能となる。

##### (3) 被験者による主観・体験

音楽生成系研究において最も典型的な評価手法のひとつに数えられる。被験者を集めて、主観評価を集積、分析するものである。おもに、被験者には、(1) 生成物を視聴する、(2) 実際にシステムを使う、などをしてもらい、その感想を、自由記述や n 段階評定などの方式で評価してもらう。中には実験計画を工夫し、主観評価に頼らず、被験者の行動分析を計測可能なデータとして取得し、それを統計的手法によって分析するという定量評価に応用する方法もある。

何人集めるか、どのような音楽経験を持つ者が望ましいかなど、被験者の選定においては、研究の目的によってさまざまな制約条件を必要とする。一般論として、統計的に優位な評価知見を得るには、数百~数千人という大規模な被験者群を集める必要があるが、残念ながらこれを実現した研究事例は見当たらなかった。今回取り上げた発表の中で、あえて被験者数による区別を行ったところ、概ね「10 人未満」「10 人以上」で同程度の件数数となった。

#### 3.2 定量的評価

##### (1) 手法比較

手法比較は、既存の関連研究で用いられている手法と提案手法とでそれぞれに生成処理を行い、出力結果を数値的に比較して提案手法の優位性を検討するものを指す。全く同一の目的やフレームワークで研究される音楽生成系システムは、実際にはそう多いものではない。たいていの場合、同じ研究者ないし研究室における先行研究との比較を行う

ために用いられている。中長期的に継続している研究においてこの手法が良く採られている。

## (2) データ比較

データ比較は、対象楽曲を設定し、教師データや人間の演奏データを正解（学習）事例とみなして、提案手法を用いて生成した場合にどれだけ正解事例を再現できるかを比較検討するものを指す。正解データがあるという点で、定量的に比較しやすい手法である。ただ、作曲システムなど、新規の音楽データを生成するタイプの研究においては、「正解」の定義が難しいという難点がある。

## 3.3 記述なし

取り上げている研究事例は、いずれも研究会報告であり、査読付き論文ではない。研究によっては、中間報告のような形で、システムやモデル、要素技術の提案のみを扱っている場合もある。そのような発表の場合、評価については多くの場合「記述なし」で済ませていることがある。ただ予稿提出から研究会当日までは約1ヶ月程度の期間があるため、研究会当日にスライドを通じて口頭で評価について説明することもあるが、それらについては記録資料としては残らないため、本稿ではこの項目ですべてまとめることとした。

## 4. まとめにかえて

本稿では、SIGMUSの過去発表を対象として、音楽生成系研究における件数の推移と、扱われて来た評価商法の分類について簡単に述べた。3章および表1で扱わなかった他の生成評価手法として、演奏表情付け研究を対象とした複数のシステムによる合同聴取評価（Rencon）が挙げられる。複数のシステムが寄り集まって、ある程度の同じ生成条件のもと、同一の課題曲、生成時間、聴取者群をもって、演奏表現の違いを一時に聴き比べるというものである。これについては、文献[3]で詳述する。

研究の目的や立場によって、定性的評価と定量的評価は適切に使い分けられるべきであろう、本オーガナイズドセッションでは、演奏支援システムを題材にして、それぞれの評価手法に対する取り組みの違いを議論する[4]。

最後に、生成音楽の評価として、聴取によってその良し悪しをいかに判断すればよいか、という課題について、オーガナイズドセッションのなかで、いくつかの研究事例において生成された演奏データを用いて公開聴取実験を行う予定である。「一聴は百聞にしかず」とまでは言わないが、あらためて、音楽を耳で聴いて評価するということが自体について、会場で活発な議論ができることを期待している。

## 参考文献

- [1] 平田圭二：誰も聴いちゃいねえ，情報処理学会研究報告 2003-MUS-50， Vol. 2003, No. 48, pp. 51-54 (2003).
- [2] 鈴木孝：音楽情報処理の研究目標，情報処理学会研究報告， Vol. 93, No. 109, pp. 17-20 (1993).
- [3] 馬場 隆，奥村健太，柴崎正浩，鈴木泰山：演奏表情付けにおける合同聴取評価：出展を通じた評価の再考，情報処理学会研究報告 [音楽情報科学]， Vol. 2014, No. 19, pp. 1-8 (2014).
- [4] 奥村健太，竹川佳成，堀内靖雄，橋田光代：評価のための問題設定：演奏支援システムの事例から，情報処理学会研究報告 [音楽情報科学]， Vol. 2014, No. 1820, pp. 1-6 (2014).