

計算機による演奏はどのように評価され得るか

How to evaluate performance rendering systems?

平賀 瑠美*¹ 片寄 晴弘*² 平田 圭二*³
Rumi Hiraga Haruhiro Katayose Keiji Hirata

*¹文教大学 *²関西学院大学・さきがけ研究 21
Bunkyo University Kwansai Gakuin University・PRESTO/JST

*³NTT コミュニケーション科学基礎研究所
NTT CS Research Laboratory

Performance rendering is a design process of generating expressive performance using a musical sheet by appending a performer's musical plan. A software system of realizing such a process has the the difficulty in having its goal, because the output has been subjectively appreciated. Rencon, CONtest for performance RENdering systems to pursue finding an evaluation method for performance rendering system and to provide a forum for related researchers, started in 2002. We describe the performance rendering research, Rencon, and the analysis of the result of the second Rencon.

1. はじめに

演奏生成システムは、楽譜という共通情報に対して演奏者の意図を付加するというデザインプロセスを実現する。このようなシステムは、その出力が主観的に受け入れられるものであったため、一定の評価基準というものが存在しておらず、システムの価値を共通理解することが困難であった。

従来、演奏生成システムの評価は、システム独自に閉じた環境下で生成演奏を被験者に聴かせる主観テストや特定の事例演奏との比較で行われることが多かった。しかしこれでは、お互いのシステムを共通の土上で判断することができない。そこで、一定の条件に基づいて演奏を生成、聴取することで複数のシステムを比較することによる評価、具体的には音楽コンクールが提案された。これは一定の条件という枠組みの上で各システムが行うデザインプロセスの違いのみを見分けるのに有効な方法となり得る。

2002年より、演奏生成システムを技術と演奏によりコンクール形式で総合的に評価するワークショップ Rencon (performance RENdering CONtest) で、演奏生成システムの評価方法を確立するための試みが始まった [Rencon HP]。以下、AIシステムとしての演奏生成システムと Rencon の概要を述べ、2002年に行った2回の Rencon のうち、FIT (Forum on Information Technology) の特別企画として行われた第2回 Rencon (以後 FIT-Rencon と呼ぶ) でのコンテストに参加した演奏を用いての聴き比べ結果についての考察を行う。

2. 演奏生成システム

演奏生成システムの出力は人間が演奏するような音楽であるが、入力として楽譜情報のみを用いたのではそれは実現できない。必要な情報には人間の演奏家が演奏プランを立てるときに用いるような情報があり、これは暗黙的なものも多いが、楽曲の構造についてのいくつかの音楽分析研究や、それらと演奏の関係の研究、また、演奏を理解する研究などの成果を用いることも有効である。

演奏生成システムは、人間の恣意的な介入の程度により、支援型と自立型に分けることができる。いずれのタイプも、表情付けの要因、例えば楽曲の構造を演奏生成のために用いる。これらのシステムは、より豊富な音楽的意味をもつ入力を可能にするインタフェースを提供するところがシーケンスソフトウェアを用いた従来の打ち込みによる演奏生成とは異なるが、支援型の場合、要因を入力する際のパラメータ値については、人間が判断している。自立型は、そのパラメータ値を自動的に模範演奏から取り出し利用するものである。

エージェントシステムとして考えれば、演奏生成システムは楽譜を認知し (出力が MIDI 形式の場合は) 固定された音源という環境の上で演奏を生成するという行動をとる。システムの大きな問題点であり、評価のためのワークショップ Rencon が必要とされる理由の一つにゴールが曖昧であるということが挙げられる。また、現在は、限られたジャンルの限られた楽器の楽曲を限られた演奏要因を用いて生成しているものがほとんどである。ここからは、容易に構築できるシステムであるという見方ができるかもしれないが、逆に、それでは何故今まで既に演奏生成システムが完成されなかったのかということに対して誰も答えられない。これは、デザインプロセスとしての演奏生成に未解明の部分が多いということが原因の一つである。

しかしながら、演奏の自動生成ということは、単に人工知能の一手法を演奏に応用すればできるわけではなく、行動において適当な演奏プランの採用と環境に聴取者の心理までを入れて考えていかなくてはならないであろう。こういったことは、計算機から演奏生成にアプローチするだけでは成就することはできず、音楽学、認知科学、心理学などの研究の総合的な成果としてなし得るものである。

3. Rencon 概要

既に述べたように、演奏生成システムは評価が困難であることと、複数の研究領域の協調作業が必要となる。そこで、評価方法を見出すということと、関連研究領域の研究者のためのフォーラムとして我々は Rencon を 2002 年より開始した。

2002年には2回の Rencon を行った [Hiraga 03]。第1回は、International Conference on Auditory Display 付随の終日のワークショップで、コンピュータ、認知科学、音楽学の各分野からの研究発表と一般聴衆による聴き比べを行った。第2

システム名	順位						
	FIT Rencon*	1	2	3	4	5*	6*
Muse	1	1	1	1	1	1	1
(Machine Learning)	2	3	2	2	3	2	3
MIS	3	2	5	5	2	3	2
CiP2	4	5	3	6	5	4	5
Yutaka	5	4	4	3	4	6	8
DM	6	7	7	4	7	5	4
(Manual)	6	8	6	7	6	8	7
CiP1	8	6	8	8	8	7	6
Kagurame	10	9	9	10	10	9	9
HHH	9	10	10	9	9	10	10

* のついた聴き比べはシステム解説後に行った。

表 1: 聴き比べ結果

回は、FIT の特別企画として半日のオープンセッションを行い、やはり聴き比べを行った。

従来主観的に受け入れられていた演奏を成果物とするシステムの評価には、まず共通の土台が必要となる。第 1 回は、生成楽曲のジャンルなどについては一切制限を設けなかったが、演奏はどれもヤマハのピアノプレーヤー（自動演奏ピアノ）上で行った。これは各システムが普段用いている環境とは全く異なるものである。第 2 回の詳しい結果については 4.1 で述べる。

1. で述べた一定の条件というのは、現在のところ入力情報の一部である楽曲の楽譜と出力環境である音源が挙げられる。また、コンテストに評価するポイント、例えばアーティキュレーションを特に考慮するなどを決め、これも事前に公表すべきである。このような評価ポイントを積み重ねていくことと、回答例を示すことにより、今後ベンチマークテスト的なものを構築していくことができるであろう。また、各システムがデザインプロセスに専念できるように、主催者側から楽譜データと演奏データを提供し、演奏エントリをしやすくすることも近いうちに実現する予定である。

4. 考察

4.1 コンテスト結果の解析

表 1 に FIT-Rencon およびその後同じデータを用いて異なる聴衆に対して行った聴き比べの結果を記す。聴き比べでは、(1) 演奏はどの程度好きか (2) 演奏はどの程度自然に聞こえるかの 2 点について 1 (最低) から 5 (最高) までの 5 段階で点数をつけた。表は、聴衆の音楽経験の背景については問わず、あくまでも二つの問いに対して主観的に回答したものをまとめたものである。参加システムに対しては、“モーツァルトまたはショパンによる曲” という限定のみを行った。当日のリハーサルを通して、音源は各システムで個別に最適と考えるものを選んで演奏を行った。演奏は、2. で述べた 2 種類のシステムによるものと、シーケンスソフトウェアを用いた打ち込みによるものである。それぞれのシステムの簡単な説明や HP は、文献 [Hiraga 02] を参照されたい。表中 FIT-Rencon および、5 と 6 の聴き比べでは、システムの解説を行ってから投票を行った。

この結果から以下のことが明らかになった。

- 聴衆の構成によらず、聞き比べは一定の傾向を持つ。

- FIT-Rencon での演奏は事前に音楽の専門家が聴き、音楽の構造と演奏という点から評価をした。表にあるように演奏は三つのグループに分けることができるが、音楽の専門家も上位グループの構成はほぼ同じであり、音楽経験によらない聴取判断の傾向がうかがえる。

- 生成システムの解説の有無は主観判断に影響を与えないようである。

- 上位グループのうちの二つは自律システムであり、演奏生成システムの将来の可能性を示唆する。

4.2 深く？ 広く？

Rencon では、2050 年にショパンコンテストで演奏生成システムが優勝するという目標を掲げている。これは、演奏システムがどのようなものであるかを分かりやすく一言で説明している。しかし、これは演奏生成システムの将来の姿の一つの可能性であり、演奏生成システム自体は取り組む研究者がそれぞれに研究目標を持っているのは確かだが、2. で述べたように、共通のゴールは明示できないというのが現在の事実である。

この目標の Rencon 実施に対する影響は、楽器と課題曲の制限に現れる。FIT-Rencon での聴き比べでは、モーツァルトも容認したが、ショパンコンクールをめざすならば、聴き比べはショパンに限定する方が自然であり、聴衆にとっても、聴き比べることが容易になる。しかしこのことにより、コンテストへの新規参入システムが増えないという問題も抱えることになる。つまり、音楽はピアノやショパンに限ったものではなく、演奏を扱うならばジャンル、楽器は制限すべきでないという考えもある。実際、管楽器を扱い出力を波形データとするシステムも存在し、それらが Rencon に参加することで、フェスティバルという形で盛り上げていくという方向も一つ考えられる。また、比較のために課題曲を決め、好き嫌いを判断するだけという一種のゲーム形式の余興として Rencon の音楽コンテストの部分を広めていくということも考えられる。

5. おわりに

Rencon は、演奏生成システムを対象としてコンクール形式で評価しようとする試みである。本稿ではその背景と考察を述べた。タイトルに対する答えは、現在与えられなくても、少なくとも「計算機による演奏は評価され得る」という段階に入りつつあると断言することができる状況である。

元々人間の音楽コンクール自体に対して様々な意見が存在するところへ、システムが生成した演奏を参加させようという Rencon に対しては、種々の意見が存在するが、評価方法の不在という現状を変えていく新しい試みとして、国際的イベントという形で継続的に進めていく予定である。

参考文献

[Hiraga 02] Hiraga, R. et al.: The Second Rencon- Performance Contest, Panel Discussion, and the Future, *Proc. of FIT2002* (2002).

[Hiraga 03] Hiraga, R. et al.: Rencon in 2002, *Working Notes of IJCAI 2003 Workshop on Methods for automatic music performance and their applications in a public rendering contest* (2003).

[Rencon HP] <http://shouchan.ei.tuat.ac.jp/rencon/>