

計算機による演奏の表情付け聴き比べ —NIME04 Rencon レポート—

A Report of NIME04 Rencon: Listening Contest to Evaluate Performance Rendering Systems

野池賢二*1 Kenzi NOIKE 平賀瑠美*2 Rumi HIRAGA 橋田光代*1*3 Mitsuyo HASHIDA 平田圭二*4 Keiji HIRATA 片寄晴弘*1*5 Haruhiro KATAYOSE

*1 科学技術振興機構 PRESTO/JST *2 文教大学 Bunkyo University *3 和歌山大学 Wakayama University

*4 NTT コミュニケーション科学基礎研究所 NTT Communication Science Laboratories *5 関西学院大学 Kwansai Gakuin University

This paper reports NIME04 Rencon (Performance Rendering Contest) and its paper session held at the International Conference on New Interfaces for Musical Expression. This paper shows the result of NIME04 Rencon with some statistical analyses, and discusses future works.

1. はじめに

非言語メディアを題材とした情報処理研究のひとつに、音楽を題材とした音楽情報処理研究がある。その中で取り組まれている研究テーマのひとつに、あたかも人間が演奏しているかのような速度変化や強弱変化などがついた演奏の生成を目標とする“演奏の表情付け”がある。演奏をはじめとする音楽情報は、サイエンス領域における他の多くの研究題材とは異なり、絶対的な評価尺度を持つものではない。そのため、音楽を扱うシステム、特に生成系のシステムの評価や、それらの生成物である音楽そのものの評価は、定量的に行うことが難しく、また同様の他のシステムとの間で優劣を決めることは困難である。しかし、情報処理研究におけるシステムである以上、なんらかの形で客観的、定量的評価が必要かつ不可欠である。

このような状況を受け、我々は音楽情報処理システムの評価方法のひとつの切り口として、演奏生成システムによる演奏をコンテスト方式で評価するプロジェクト、Rencon を 2000 年から開始した [平賀 02]。プロジェクト名 Rencon は、“Performance Rendering Contest” に由来しており、コンテストそのものを Rencon と呼ぶこともある。期待の意味を込めて描いた Rencon ロードマップを図 1 に示す。

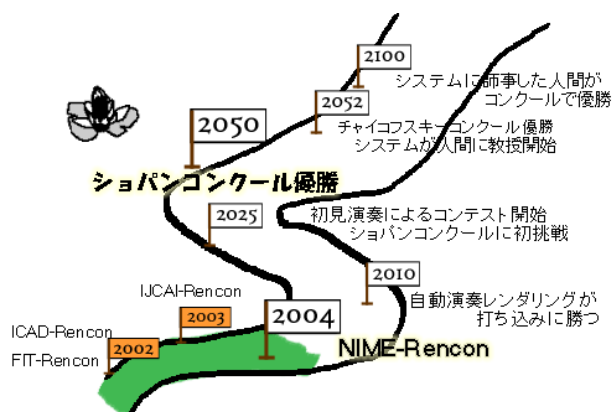


図 1: Rencon ロードマップ 2004

2004 年は、ICAD Rencon, FIT Rencon, IJCAI Rencon に続く通算 4 回目の Rencon を、NIME04 において開催した。我々はこれを、“NIME04 Rencon” と呼んでいる。NIME とは、New Interfaces for Musical Expression の略称であり、「音楽 / 芸術表現のための新インターフェース」に関する国際会議である。2004 年の NIME である NIME04 は、浜松にある静岡文化芸術大学にて、6 月 3 日 (木) から 5 日 (土) の 3 日間開催され、NIME04 Rencon は、それと同期間開催した。

本稿では、NIME04 Rencon の概要とコンテストの結果、そして、次回開催に向けて残された課題について述べる。

2. NIME04 Rencon

2.1 概要

NIME04 Rencon の 3 日間のプログラムを図 2 に示す。

- 3 日 (木)
 - 13:00 - 15:00
ブラインド聴き比べコンテスト
- 4 日 (金)
 - 13:00 - 15:00
ブラインド聴き比べコンテスト
 - 15:20 - 16:20
研究発表
 - * Rumi Hiraga, Roberto Bresin, Keiji Hirata, Haruhiro Katayose:
Rencon 2004: Turing Test for Musical Expression
 - * Haruhiro Katayose, Keita Okudaira:
Using an Expressive Performance Template in a Music Conducting Interface
 - * Hideki Kawahara, Hideki Banno, Masanori Morise:
Acappella synthesis demonstrations using RWC music database
- 5 日 (土)
 - 13:00 - 15:00
オープン聴き比べによる投票。

図 2: NIME04 Rencon 開催プログラム

連絡先: 野池賢二, 〒 669-1337 兵庫県三田市学園 2-1
 関西学院大学理工学部情報科学科 片寄研究室気付,
 e-mail: noike@ksc.kwansei.ac.jp

NIME04 Rencon は、聴き比べコンテストと研究発表セッションからなる。聴き比べコンテストでは、いままでの Rencon がない新しい試みとして、「規定部門でのチューリングテスト

方式による評価」,「自由部門での歌声生成システムの聴き比べコンテスト」,「演奏の機械らしさを評価する逆チューリングテスト」を実施した。詳細を順に述べる。

2.2 聴き比べコンテスト

2.2.1 コンテスト実施要項

聴き比べコンテストは,規定部門,自由部門,逆チューリング部門の3部門からなる。

規定部門 (Compulsory section)

評価判断をなるべく公平にするために,課題曲と聴き比べ時の使用音源に,ある程度の制約を設ける。

課題曲 参加者の選んだショパンのピアノ小曲

使用する音源 Nemesys GIGA Piano

提出形式 SMF 形式

評価方法 「演奏の人間らしさ」,「演奏の好み」の2項目について,聴衆それぞれに5段階評価点を投票してもらう。順位は「演奏の人間らしさ」の評価点の平均点によって決定する。もし,同点の演奏が複数ある場合は「演奏の好み」の評価点によって順位を決める。優勝システムには,Rencon 賞を授与する。

NIME04 Rencon では,システムが生成した演奏のほか人間による演奏も加え,どの演奏が何(誰)による演奏かがわからないチューリングテスト方式で行った。システムによる演奏と,人間による演奏との判断がつかうかどうか,興味ある事柄のひとつである。

自由部門 (Open section)

課題曲や使用音源に制約を設けずに競う。管楽器や弦楽器の演奏,歌声,あるいはNIMEで発表されるような新しいインタフェースを持つ楽器による演奏もエントリーすることができる。

課題曲,使用する音源,提出形式 一切が参加者の自由

評価方法 「演奏の好み」について,聴衆それぞれに5段階評価点を投票してもらい,その評価点の平均点によって順位を決定する。優勝システムには,Rencon 賞を授与する。

ジャンルや楽器が異なる演奏を比べるため,同一の基準で比較評価をしたとはいえないが,各演奏生成システムの性能や機能の現状を知ることができる。また,単純にエンタテインメントとして楽しむこともでき,コンテストを気楽なものとする事ができる。NIME04 Rencon では,特に歌声生成に取り組んでいる研究者に声をかけ,複数の歌声生成システムによる歌声演奏聴き比べを実現した。

逆チューリングテスト部門 (Gnirut section)

規定部門とは逆に,演奏の機械らしさを聴衆の投票によって競う。通常とは逆の基準による評価を行うことによって,演奏生成システムのよりよい評価方法を模索する。我々はこれを,逆チューリングテスト,あるいは,Turing を reverse string して Gnirut と呼んでいる。

課題曲 参加者の選んだパッサのピアノ小曲

使用する音源 Nemesys GIGA Piano

提出形式 SMF 形式

評価方法 「演奏の機械らしさ」,「演奏の好み」の2項目について,聴衆それぞれに5段階評価点を投票してもらう。順位は「演奏の機械らしさ」の評価点の平均点によって決定する。もし,同点の演奏が複数ある場合は「演奏の好み」の評価点によって順位を決める。

2.2.2 コンテスト実施とその結果

聴き比べコンテストのエントリー一覧を図3に示す。それぞれ順に,システム名,エントリ代表者名(所属機関),エントリ曲を記している。全エントリを合わせた総演奏時間は30分弱あり,1日に2時間あるセッションの中で,4回の聴き比べができる。コンテスト当日は,NIME04参加者の多くが少人数ずつ途切れることなく訪れ,3日間を通して多くの聴衆に参加してもらえた。聴衆は,熱心に聴いて投票してくれる人が多く,有効投票を51票,集めることができた。

各部門の上位3位を図4に示す。全順位については,Rencon Web サイト中の http://shouchan.ei.tuat.ac.jp/~rencon/NIME04/NIME04_result.html をご参照いただきたい。

まず,規定部門の結果をみると,人間による演奏が上位を占めている。これは,現在の演奏表情付けシステムの演奏が,人間の演奏ほどには上手く表情付けできていないことを示している。ただし,本稿には載せられなかったが「演奏の好み」の評価点で,橋田らのシステム WALT (後の PopE) の演奏のほうが人間の演奏のひとつよりも高い平均点を得ており,必ずしも演奏表情付けシステムが劣っているわけではないことも示唆された。また,規定部門では,Etude Op.10, Nr.3 “Chanson de l'adieu (別れの曲)” によるエントリが4エントリあり,楽曲と演奏音源がまったく同一である条件下での聴き比べができた。その結果は,上位から順に,(1)人間の演奏,(2)COPER,(3)Kagurame Phase-II,(4)Ha-Hi-Hun,となり,やはり,人間の演奏が最上位であった。

自由部門では,歌声生成システムが4エントリ^{*1},ピアノ曲,協奏曲が,それぞれ1エントリあった。その中で最上位を獲得したのは,黒澤らのエントリ,SuperConductorTMによる協奏曲の演奏であった。SuperConductorTMは,Manfred Clynes氏によって手がけられたシステムであり,製品にもなっている。4エントリある歌声生成システムだけの順位を見ると,上位から,(1)STRAIGHT,(2)WONDER HORN,(3)HMM-based singing voice synthesis system,(4)Vocaloid,であった。

自由部門では特に制約を設けずにエントリを募集したため,技術的な良し悪しよりも,演奏生成時の作り込み度合いや,選曲,聴衆の好みに順位が左右されたような印象を受けた。

逆チューリング部門では,Bresin氏のシステム The Director Musices Program による演奏が1エントリ,人間の演奏が2エントリあった。その結果は,若干の差ではあるが,やはり Bresin 氏のシステムによる演奏が最も「機械らしい」と判断された。この結果から,人間が機械らしく演奏しても,システムの演奏とは区別がついてしまうことが明らかになり,人間の演奏には意図していない「機械らしくない」特徴が含まれているであろうことが示唆された。

2.3 研究発表

研究発表のセッションでは,3件の研究発表があった。

*1 4エントリのうちのひとつ,Vocaloidによる演奏は,今回は参考出品として Rencon 運営側で用意した。

規定部門	
1.	The Director Musices Program: Roberto Bresin (KTH), Prelude Op.28 No.7
2.	Ha-Hi-Hun: 平田圭二 (NTT コミュニケーション科学基礎研究所), Etude Op.10, Nr.3 “Chanson de l’adieu”
3.	(人間の演奏): はるか (ピアノ教師), Etude Op.10, Nr.3 “Chanson de l’adieu”
4.	COPER: 野池賢二 (PRESTO/JST), Etude Op.10, Nr.3 “Chanson de l’adieu”
5.	Kagurame Phase-II: 鈴木泰山 (株式会社デューオシステムズ), Etude Op.10, Nr.3 “Chanson de l’adieu”
6.	(人間の演奏): はるか (ピアノ教師), Valse Op.64, Nr.1
7.	WALTS (表情なし): 橋田光代 (PRESTO/JST, 和歌山大学), Impromptu, Op.66 “Fantasie-Impromptu”
8.	WALTS (テンプレート適用): 橋田光代 (PRESTO/JST, 和歌山大学), Impromptu, Op.66 “Fantasie-Impromptu”
9.	WALTS (テンプレート適用 + 2 小節目第 4 拍目手動調節): 橋田光代 (PRESTO/JST, 和歌山大学), Impromptu, Op.66 “Fantasie-Impromptu”
自由部門	
1.	SuperConductor TM : 黒澤 隆 (フロンティアKインターナショナル株式会社), J. S. Bach, Brandenburg Concerto No.5, D Major, 3rd Movement
2.	RUBATO: Guerino Mazzola, J. S. Bach, The Art of Fugue, Contrapunctus III
3.	WONDER HORN: 青野裕司 (NTT サイバースペース研究所) 森山直太郎, さくら (独唱)
4.	HMM-based singing voice synthesis system: 酒向 慎司 (東京大学大学院) しあわせなら手をたたこう
5.	STRAIGHT: 河原 英紀 (和歌山大学), 武満 徹, Small Sky
6.	(参考出品) Vocaloid: (YAMAHA, Zero-G), John Bettis / Richard Carpenter, “Yesterday Once More”
逆チューリングテスト	
1.	The Director Musices Program: Roberto Bresin (KTH), J. S. Bach, The Well-Tempered Clavier, Prelude I, BWV 846
2.	(人間の演奏): はるか (ピアノ教師), J. S. Bach, Invention I, C Major
3.	(人間の演奏): はるか (ピアノ教師), J. S. Bach, The Well-Tempered Clavier, Prelude I, BWV 846

図 3: NIME04 Rencon 聴き比べコンテストエントリー一覧

平賀らは、これまでの Rencon の歩みと、演奏生成システム評価へのチューリングテスト導入について発表した。

奥平らは、演奏表情テンプレートをを用いた拍打型演奏インタフェース iFP と、演奏表情の視覚化機能について発表した。

河原らは、高品質音声分析変換合成システム STRAIGHT による高音質音声モーフィングと、多声アカペラ合成について

規定部門	
1 位	(人間の演奏: Valse Op.64, Nr.1), 平均得点 3.75
2 位	(人間の演奏: Etude Op.10, Nr.3), 平均得点 3.49
3 位	WALTS (テンプレート適用), 平均得点 3.47
自由部門	
1 位	SuperConductor TM , 平均得点 3.06
2 位	STRAIGHT, 平均得点 2.90
3 位	WONDER HORN, 平均得点 2.80
逆チューリング部門	
1 位	The Director Musices Program, 平均得点 3.41
2 位	(人間の演奏: Invention I, C Major), 平均得点 3.37
3 位	(人間の演奏: Prelude I, BWV 846), 平均得点 2.63

図 4: NIME04 Rencon 聴き比べコンテスト投票結果上位 3 位

発表した。

研究発表セッションの最後には、聴き比べコンテストにエントリーしたシステムそれぞれについての簡単な紹介をエントリー代表者を交えて行い、その後、コンテストの投票結果順位を発表した。各部門の最上位獲得システムの代表者には、Rencon 運営代表の平賀から賞状が授与された。

3. 検討

3.1 順位の有意性検定

聴き比べコンテスト結果順位の有意性をみるために、獲得評価点の母集団平均値区間推定による検定を、有意水準 5%にて行った。紙面の都合により、本稿では規定部門の「演奏の人間らしさ」の評価点についてだけを図 5 に示す。図 5 の縦軸は「演奏の人間らしさ」評価点であり、9 エントリを平均評価点の高いものから順に左から並べてある。平均評価点の上下には、上下方の信頼限界も示した。順位に有意性があるエントリ間には、縦破線を記した。

これをみると、順位に有意性があるのは、“WALTS (テンプレート適用+2 小節目第 4 拍目手動調節) - COPER 間”、“The Director Musices Program - WALTS(表情なし) 間”、“WALTS(表情なし) - Ha-Hi-Hun 間”の 3 箇所であり、他の隣合う順位には有意性があるとはいえないことがわかる。したがって、上位 4 位に含まれる人間による演奏 2 つと WALTS による演奏 2 つの 4 エントリの順位は有意であるとはいえず、WALTS による演奏は、人間による演奏と区別ができなくなるほどに迫っている可能性があると言える。

3.2 実施上の検討項目

3.2.1 聴き比べ用音源の仕様公開

NIME04 Rencon では、公平性を確保するために、自由部門以外は SMF による提出を義務づけ、聴き比べ時の音源は Rencon 運営側で用意した。その結果、ダンパーペダルデータを含んだ演奏の発音の不自然さが目立つエントリがあり、エントリ者の意図とは異なる印象を持つ演奏となった可能性がある。この現象は、用意した音源の発音特性として、ダンパーペダルが踏み込まれたときの音の減衰時間が比較的長いということに起因する。聴感上の印象は順位への影響が大きいため、これは早急に解決する必要がある。この問題については、(1) 聴き比べ時の音源のコントロールデータなどの仕様を、できるだけ詳細に公開する、(2) エントリ者が聴き比べ時の音響デー

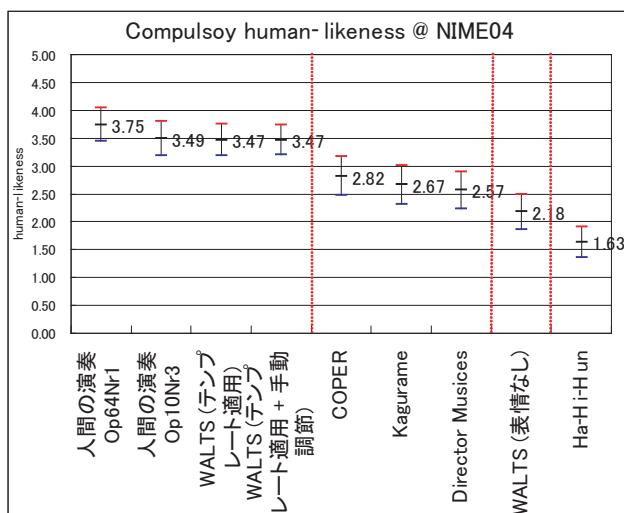


図 5: 規定部門の「演奏の人間らしさ」順位の有意性検定

タを実際に確認できる「試し聴きシステム」を提供する、という方法による解決を計画している。「試し聴きシステム」は、Web サーバやメールサーバを利用した人手を介さない自動返信システムとしての実現を検討中である。

3.2.2 投票者の負担への対処

NIME04 Rencon では、きちんと投票するために聴き比べを一通り行くと 30 分弱の時間がかかる。これは、投票者にとっては無視できない負担であろう。この負担を軽減するためには、被験者自身へのフィードバックを考慮していく必要があると考えている。具体的には、(1) チューリングテスト方式の場合は、答え合わせができるようにする、(2) 自分の投票が結果に反映されることをその場で見せる、(3) 過去の聴取実験結果と照らし合わせ、被験者の聴き方タイプ（傾向）を示す、などのサービスを提供することを考えている。また、時間拘束負担を軽減する他の手段の一つとして、インターネットを利用した投票実施に向けての準備を進めていきたいと考えている。

3.2.3 共通学習データセット、評価データの準備

規定部門において共通の評価基準の下での聴き比べを実現するためには、共通の学習データセットと評価データを用いることが望ましい。これについては、MusicXML をベースとした MusicXML(4R) 形式 [Hirata 03] での学習・評価データセットの準備を進めているところである。

3.2.4 評価方式の考慮

公平な評価、評価結果の信頼性の向上、そして生成に関する恣意性の排除のためには、生成にかかった「人手の介入度合い」の考慮が必要である。今回は試みとして、介入度合いに関する自己申告式のアンケートをエントリー時に Web 上で実施した。このアンケート結果の評価への反映は、現段階ではなされていないが、これの評価への組み込み方を、今後は検討していく必要がある。また、投票時の評価項目としては、単に「演奏の人間らしさ」や「演奏の好み」だけではなく、フィギュアスケートの採点などにみられる「技術点評価」も加えた複合評価方式も今後は考えていく必要がある。

3.2.5 エントリー楽曲の著作権

聴き比べコンテストは、厳密には公の場での公開演奏であり、エントリー楽曲の著作権の管理については十分な注意が必要である。

Rencon 運営側では、著作権の切れた自由に使える楽曲での

エントリー、または、エントリー者自身で作曲した楽曲でのエントリーを推奨し、そうでない場合は、エントリー者自身で解決するように求めてきた。NIME04 Rencon では、森山直太朗の「さくら（独唱）」でのエントリーを希望する WONDER HORN のエントリー代表者に、自身で著作物使用許諾を得る手続きを行っていただいた。他の使用許諾を得ていないエントリー曲については、演奏データを発表論文 CD-ROM には収録しないことで対処した^{*2}。今後は、他者に権利のある著作物でエントリーしなければならない場合の手引書を Rencon 運営側で用意し、著作物使用に関するより一層の注意を促す予定である。また、Rencon 運営側としては、全エントリーデータが著作権上で問題がないことの確認を徹底する必要がある。

4. まとめ

本稿では、演奏生成システムをコンテスト方式で評価するプロジェクト Rencon の活動紹介として、NIME04 Rencon 開催報告と、次回の開催に向けた検討を行った。NIME04 Rencon では、「チューリングテスト方式による評価」、「歌声生成システムの聴き比べコンテスト」、「演奏の機械らしさを評価する逆チューリングテストの実施」、などの新しい試みを行った。その中でも、歌声生成システムの聴き比べコンテストは、聴衆に好評であった。歌声をはじめとする、ピアノ演奏以外の演奏生成システムの聴き比べ評価は、今後さらに充実させていく予定である。

世界的な最近の動向に目を向けると、ISMIR [ISMIR 05] や RoboCup [RoboCup 05]、KabuRobo [KabuRobo 05] に代表されるように、システムの能力を問題解決技術をコンテスト形式で評価する試みが盛んに行われるようになってきた。Rencon も、共通学習データの配布や、試し聴きシステムの提供などを実施し、「日本発の音楽システム評価プロジェクト」として、より一層の充実を図っていきたい。なお、2005 年の Rencon は、ICMC2005 において開催する予定であり、現在、それに向けての準備が進行中である。

参考文献

- [Hirata 03] Hirata, K., Noike, K., and Katayose, H.: Proposal for a Performance Data Format, in *In Working Notes of IJCAI-03 Workshop on methods for automatic music performance and their applications in a public rendering contest*, pp. pp. 65 – 69 (2003)
- [ISMIR 05] ISMIR 2005 Web site:
<http://ismir2005.ismir.net/> (2005)
- [KabuRobo 05] KabuRobo.jp Web site:
<http://kaburobo.jp/> (2005)
- [Rencon 05] Rencon Web site:
<http://shouchan.ei.tuat.ac.jp/~rencon/>
- [RoboCup 05] RoboCup 日本委員会公式ホームページ:
<http://www.robocup.or.jp/index.html> (2005)
- [平賀 02] 平賀 瑠美, 平田 圭二, 片寄 晴弘: 蓮根, 目指せ世界一のピアニスト, 情報処理, Vol. 43, No. 2, pp. 136–141 (2002)

*2 STRAIGHT によるエントリーの代表者からは、自身の研究グループで作曲した楽曲への差し替えが可能であるとの連絡を受けていたが、日程の都合などの兼ね合いで差し替えることをお断りさせていただいた。