Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC 10). Sapporo, Japan.
Ken'ichi Miyazaki, Yuzuru Hiraga, Mayumi Adachi, Yoshitaka Nakajima, and Minoru Tsuzaki (Editors)

Rencon: Performance Rendering Contest for Automated Music Systems

Mitsuyo Hashida,[*1] Teresa M. Nakra,[#2] Haruhiro Katayose,[*3] Tadahiro Murao,[†4]
Keiji Hirata,[§5] Kenji Suzuki,[¶6] Tetsuro Kitahara[*7]

[*] *Research Center for Human & Media, Kwansei Gakuin University, Japan*
[#] *Department of Music, The College of New Jersey, USA*
[†] *Department of Education, Aichi University of Education, Japan*
[§] *Social Communication Laboratory, NTT Communication Laboratories, Japan*
[¶] *Department of Intelligent Interaction Technologies, Tsukuba University, Japan*

[1] hashida@kwansei.ac.jp, [2] nakra@tcnj.edu, [3] Katayose@kwansei.ac.jp
[4] tmurao@auecc.aichi-edu.ac.jp, [5] hirata@brl.ntt.co.jp
[6] kenji@ieee.org, [7] t.kitahara@kwansei.ac.jp

ABSTRACT

Rencon (Performance Rendering Contest) is an annual international competition at which entrants present the computer systems they have developed for generating expressive musical performances and audience members and organizers judge the performances. Recent advances in performance rendering technology have brought with them the need for a means for researchers in this area to obtain feedback about the abilities of their systems in comparison so those of other researchers. At the ICMPC10, the competition will consist of an autonomous section (for evaluating the ability of the entered systems to generate performances autonomously) and a system-supported section (for evaluating human performances done using computer systems). The autonomous section aims to evaluate performances rendered by autonomous computer systems using e.g. a rule-based or case-based approach, and the system-supported section aims to build common ground for evaluating human performances done using computer systems.

I. INTRODUCTION

Professional musicians estimate the intention of the composer for the piece to be played, plan their performance of it, and finally perform it to the level of their ability. Even though computer programs can now play chess better than professional players, their ability to play musical instruments has not yet reached the level of professional musicians. Many issues, such as how to deal with intention and affection, must be addressed before they can reach this level.

Recent developments in computer science have begun to clarify the thought processes of musicians when they perform, and the latest systems can generate performances that are barely distinguishable from human ones although they still cannot match the level of professional musicians. Recent advances in performance rendering technology have brought with them the need for a means for researchers in this area to obtain feedback about the abilities of their systems in comparison so those of other researchers.

Subjective evaluations of generated performances are needed because sensuousness and emotionality are important in musical performances. Competition is an effective way to obtain such evaluations and should promote further advances. Rencon ("rendering contest"), which was started in 2002, is an annual international competition at which entrants present the computer systems they have developed for generating expressive musical performances and audience members and organizers judge the performances. At the six events held so far, both domestically and internationally, the performances, evaluations and subsequent discussions have proven valuable.

At ICMPC10, the competition will feature a new approach: systems entered in one of the two sections will generate performances of two newly created musical pieces autonomously and on the spot. This is a challenging approach because every performance will be a world première, one that even the system developers will not have previously heard. These tougher conditions should even better promote advances in rendering technology. In the other section, human entrants will render performances using commercial or original music applications.

The rest of this paper is organized as follows: In Section II, we describe the current state of performance rendering and the Rencon Project. Then we present an overview of the competition in Section III. The details of the autonomous and system-supported sections are described in Sections IV and V, respectively. We end with some concluding remarks in Section VI.

II. CURRENT STATE OF PERFORMANCE RENDERING AND REMCON PROJECT

A. Performance Rendering by Automated Systems

Research that ushered in performance rendering systems dates back to the 1980s (Frydèn and Sundberg 1984, Clynes 1984). Since the 1990s, approaches involving music recognition theories such as the generative theory of tonal music (Lerdahl and Jackendoff 1983) and the implication-realization model (Narmour 1990), learning systems (Arcos et al., 1998, Bresin et al., 2000), and example-based reasoning (Ishikawa et al. 2002, Widmer 1996) have been proposed. In addition, a competition for system-rendered performances has been held since 2002 (Hiraga 2002a/b). Moreover, a lot of commercial software for desktop music and digital audio workstations has been published.

Figure 1 diagrams a typical performance rendering system. Generally, automated performance rendering systems take a score as the input, generate a performance of it using an original rendering process, and output the rendered performance in MIDI file format. These systems are often categorized into rule-based and case-based systems. In the rule-based approach, which is used by many commercial music software

ISBN: 978-4-9904208-0-2 © 2008 ICMPC10

systems, the performance of a score is generated using musical knowledge (rules). In the case-based approach, the system finds a melody (or other sequence) similar to the target melody (or sequence) and transfers its expression directly. This approach enables the user to produce musical expressions even if the user does not know the rules of expression for the target melody (or sequence).

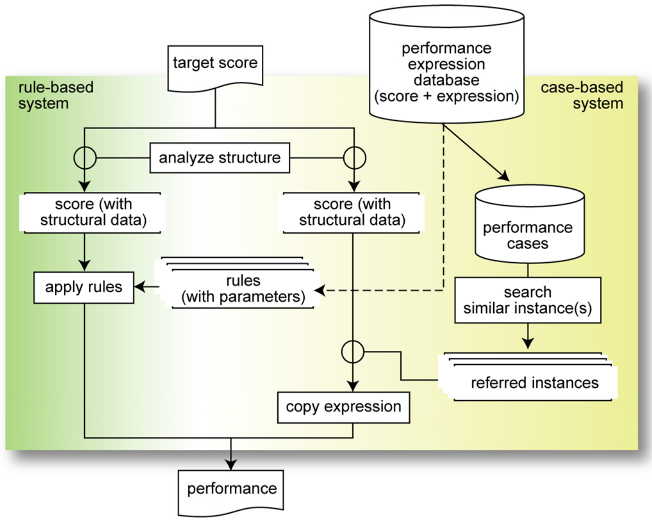

Figure 1: Illustration of typical rendering system

Some rule-based systems have been done on extracting rules with parameters from human performances as case-based systems do. Use of these rules as examples for reference is one trend in performance rendering studies. Recently, structural information contained in the score it is used both in rule-based and case-based systems, to emulate the way a musician renders a musical performance.

B. Rencon Project

At previous Rencons, the participants were asked to send beforehand a computer file containing the performance, rendered offline, of a set piece designated in advance. Sound files generated from these performance files were evaluated by human judges. This approach, introduced to lower the technical hurdles, gave the participants a chance to customize their system to best fit the set piece.

To make the evaluation fairer, we now require that each system render an unknown piece and that the rendering process be transparent. This would have been difficult even a couple of years ago, but progress in autonomous processing has made it realistic.

Another important issue regarding performance rendering is user generated content (UGC). The number of commercial performance rendering systems has been increasing along with the popularity of Web 2.0 services. The human performance design process realized with computer software should be discussed and the autonomous function of performance rendering systems should be evaluated.

III. ICMPC-Rencon

A. Overview

ICMPC-Rencon will have two sessions, the competition and a topical session, as shown in Table 1.

Table 1: ICMPC-Rencon Schedule

Aug. 25	**Competition Session (2 hours)**	
	Anonymous section	**System-supported section**
	1. Score input & pre-processing 2. Performance rendering 3. System introduction by entrant (1 min) 3. Performance playing 4. Audience evaluation 5. Free discussion	1. Performance rendering 2. Free discussion
Aug. 27	**Topical Session (2 hours)**	
	1. Invited talk Teresa M. Nakra (The College of New Jersey) "Conductor's Gestures and Performance Rendering" 2. Panel discussion 3. Awards ceremony	

In the competition, there will be two sections: an autonomous section and a system-supported section. The autonomous section is for autonomous computer systems, such as systems using the rule-based or case-based approach, to render a performance. Entrants are not allowed to hand-edit the performances during the rendering process. The aim of this section is to evaluate performances rendered by autonomous computer systems using e.g. a rule-based or case-based approach. The system-supported section is for entrants using commercial music software or an original application to render a performance. The aim of this section is to build common ground for evaluating human performances done using computer systems as well as to make Rencon more widely appealing.

In the anonymous section, all of the entered systems will first render the performances, and then the rendered performances will be played by an automated grand piano and evaluated by the audience. In the system-supported section, the entrants will render performances using commercial or original music applications.

In the topical session, there will be an invited talk, a panel discussion on performance rendering, and an awards ceremony.

B. Set Pieces and Rendering Procedure

All systems are required to render set pieces of music, which were newly composed and provided by Professor Tadahiro Murao (Aichi University of Education).

<u>Autonomous section:</u>

1. Piano piece (Chopin-style, 1 min.)
2. Piano piece (Mozart-style, 1 min.)

<u>System-supported section:</u>

1. Piano piece (Chopin-style, 1 min.)

The styles of the compositions apply either the ideo-structure of Chopin or that of Mozart. There are no measure changes during the pieces. The playing time (without expressiveness) of each piece is around 1 minute.

The participants will be given the compositions at the beginning of the competition. The participants in the autonomous section will be required to render both pieces. The participants in both sections will have 60 min to render the performance. The rendering process will be open for all participants to observe.

C. Entered Systems

As of April 15, five systems have been entered in the autonomous section and two have been entered in the system-supported section. Sections IV and V introduce the systems that have been entered. Since systems can be entered up until the day of the competition, there may be more systems competing.

D. Evaluation and Awards

After the 60-min rendering period, the rendered performances will be played by an automated grand piano in turn.

The performances for the autonomous section will be evaluated by the attendees and Professor Murao, taking into account the degree of humanness and expressiveness. The attendees can vote on-site or later on the Internet. The results will be announced at the topical session. There will be three awards.

- The Rencon Award will be presented to the entrant whose rendered performance gets the highest score based on the attendee voting.
- The Rencon Technical Award will be presented to the entrant whose system is judged to be the best from a technical point of view by the Rencon Organizing Committee.
- The Rencon Murao Award will be presented to the entrant whose rendered performance affects the composer the most.

The performances in the system-supported section will be evaluated by Professor Murao, taking into account how they affect him. There will be one award.

E. Topical Session

In the topical session, there will be an invited talk, a panel discussion, and an awards ceremony. In the invited talk, Dr. Teresa M. Nakra (The College of New Jersey), who has brought a high-tech interaction paradigm to classical and traditional music as the Founder and Artistic Director of Immersion Music, will talk about "Conductor's Gestures and Performance Rendering." The panel will discuss topics related to the competition, such as automated methods of generating musical expression and interfaces for performance rendering systems.

IV. AUTONOMOUS SECTION

As explained in Section III, the autonomous section is for performances rendered by autonomous computer systems using, for example, a rule-based or -based approach. The systems in this section should be able to

- read score data in MusicXML or standard MIDI format,
- render an expressive performance (using, for example, a rule- or case-based approach), and
- output the generated data in standard MIDI format.

Figure 2 illustrates what the entrants are allowed and not allowed to do. For example, they are not allowed to hand-edit the rendered performances.

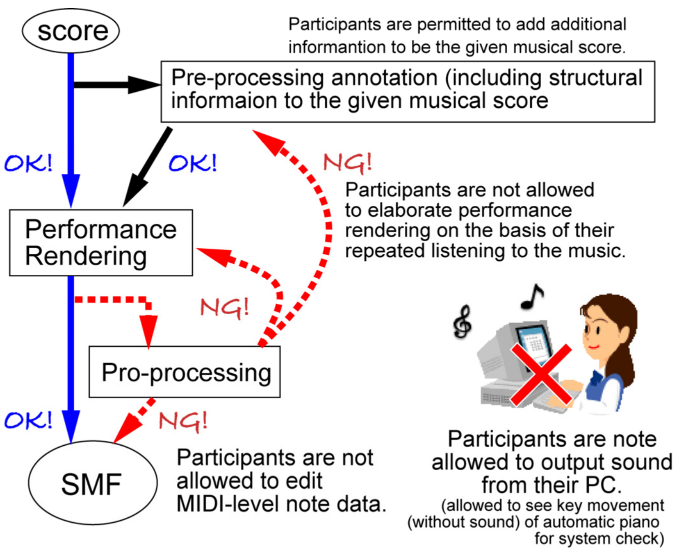

Figure 2: Performance rendering in autonomous section

A. Input Data Format

The input score data is in MusicXML format (http://www.recordare.com/xml.html), standard MIDI file format, and printed score format. The data in all three formats will be provided to each entrant at the beginning of the competition.

For the MusicXML format, both data and expression marks (f, p, crescendo, andante, slur, staccato, and so on) will be included. For the standard MIDI file format, both Formats 1 and 0 will be provided. The note velocity is set to 64. The program change (instrument) value is set to 0 (acoustic grand piano). The tempo (BPM) is set to 120 a quarter note. The data will not include any control messages or pedal information. The data will not describe the complete music structure, including the phrase structure, harmony, and chord progress, too.

B. Entered Systems

1) NAIST Music Rendering Model

Contributors: Keiko Teramura et al.

Affiliation: Graduate School of Information Science, Nara Institute of Science and Technology,

The NAIST Music Rendering Model uses a Gaussian statistical machine learning method for generating an expressive music performance and requires that only a few parameters be set manually. The model learns the musical performance so as to predict a musician's performance, which includes conscious and unconscious variations on the musical score. The musical score information is given as input, and the corresponding actual performance—dynamics (note strength), attack time, and release time—is given as output. The model can directly approximate input-output relationships that might include rules that are hard to describe explicitly. It should be able to imitate a master performer if it is trained using a large enough dataset.

2) The Grachten System

Contributors: Maarten Grachten et al.
Affiliation: Department of Computational Perception, Johannes Kepler University

The rendering algorithm in this system consists of a Bayesian network/HMM trained on a large set of performance data comprising performances of Mozart Piano Sonatas by Roland Batik on a Bösendorfer SE290 grand piano. Expressive timing and dynamics are estimated on the basis of observed score features. Among the score features are not only typical candidates such as pitch interval and inter onset interval, but also metrical position, current chord function w.r.t the key (Kostka & Payne, 2004), phrase information, and implication-realization information (Narmour, 1990). These more abstract score features are computed from the score using Temperley's Melisma software (Temperley, 2001) and a parser for implication-realization analysis (Grachten, 2006). Rather than rendering the polyphonic score by projecting it along a predicted tempo and dynamics curve, predictions are made per note, where notes are treated as belonging to voices, using voice decomposition algorithms (Temperley, 2001, Madsen & Widmer, 2006).

3) COPER

Contributor: Kenji Noike
(No Affiliation)

"COPER on the Web" is a Web-based performance rendering system using a corpus of expressive music performances. This system generates expression for a given musical piece by applying expressions in the musical segment extracted from a corpus. The segments are extracted using queries related to the "musical group boundary" or "pitch, duration and transition of the notes" in a melody sequence. The corpus consists of a set of cases (actual examples) of human musical performances, described using descriptors of nominal (quantized) note information and musical deviation terms. The deviation terms consist of whole tempo transition, damper pedal on/off action with depth degree value, and delicate temporal and dynamics tactus control.

The version for ICMPC10 uses a new consolidated corpus: the corpora prepared for the previous version of *COPER* and "CrestMusePEDB," a music performance expression database consisting of 35 pieces (about 10,000 musical segments).

4) Kagurame-Phase II

Contributors: Tatsuya Hino et al.
Affiliation: Shibaura Institute of Technology

Kagurame Phase-II is a performance rendering system that can generate various types of performances for a single piece depending on the parameter settings. This is made possible by using a case-based approach rather than a rule-based one for generating musical expressions. A large amount of human performance data is used as a knowledge base for musical expression generation instead of rules or methods for musical expression. The parameter settings are called the "performance condition." It can include various types of parameters such as style, mood, and performer.

5) Itopul

Contributors: Shunji Tanaka et al.
Affiliation: Kwansei Gakuin University

Itopul is a case-based performance rendering system that runs in a Java environment. It is characterized by 1) the combination of a phrasing model and a pulse model, 2) the use of a hierarchical music structure to avoid the data sparseness problem, 3) visualization of the rendering process as it proceeds, and 4) music structures directly modifiable by the user.

V. SYSTEM-SUPPORTED SECTION

As mentioned, the system-supported section aims to build common ground for evaluating human performances done using computer systems. Entrants will perform a musical piece using commercial music software or an original application. As shown in Figure 3, the entrants are allowed to elaborate the performance expression at any step while listening to the playback. They are also allowed to generate expressions by using the mouse, keyboard, or an abstracted body movement like hand conducting. They are not allowed to play their musical instrument directly.

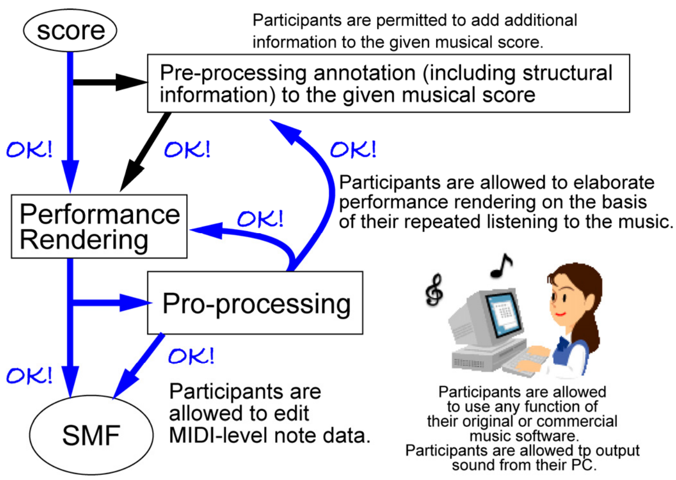

Figure 3: Performance rendering in system-supported section

A. Input Data Format

The participants will be given a dead-pan MIDI file (SMF, Format 1) at the start of the competition. The note velocity is set to 64. The program change (instrument) value is set to 0 (acoustic grand piano.) The tempo (BPM) is set to 120 a quarter note. The data will not include any control messages or pedal information. A printed score will also be provided for reference.

B. Entered Systems

6) ESP System

Contributors: Elaine Chew et al.
Affiliation: University of Southern California

The Expression Synthesis Project, ESP, explores the use of the driving metaphor for expressive performance. The ESP system enables a user to interactively create a performance while navigating a road that suggests an interpretation based on musical structure. The "driver" controls a "car" while road bends encourage slow-downs and straight sections encourage speedups. A dashboard shows a roadmap and speedometer (in bpm); buttons control the articulation and (de)coupling of dynamics with acceleration. The system enables users to experiment with performances prior to mastering an instrument or composition and enables composers to affect performance decisions by, for example, crafting a roadmap.

7) iFP

Contributors: Takashi Baba et al.
Affiliation: Kurashiki Sakuyo University

iFP is an interface for playing expressive music. It promotes a pianist's expressiveness with its tapping-style input. MIDI-formatted expressive piano performances were first analyzed and transformed into performance templates in which deviations from the canonical description were separately described for each event. Using one of the templates as a skill complement, a musician can expressively play a musical piece over and under the beat level. The iFP scheduler enables players to mix their own intentions and expressiveness into the performance template. An iFP performer can experience the thrill of virtuosity by varying weight parameters dynamically to control the deviations in tempo, dynamics, and delicate nuances within the beat. That is, iFP provides functions for controlling the intention level of the player and the model performance. It also provides a morphing function that interpolates and extrapolates two different expressive performance templates of a musical piece. The peripheral interface can be a PC keyboard, a musical keyboard, or a conducting interface using capacity transducers.

VI. CONCLUDING REMARKS

This paper introduced the current state of performance rendering and summarized the performance rendering competition to be held at ICMPC10.

The organizers will provide our original performance interface (iFP) and some commercial applications to those interested. We expect this competition to trigger a discussion of interaction design for performance interfaces, music interpretation models, and their application to music education.

We aim to contribute to the development of modeling techniques for human mental activities, the formulation of musical performance expression, its application to education, and the creation of novel music enjoyment. We intend to make many people aware of performance generation systems.

ACKNOWLEDGMENTS

The Rencon Committee sincerely thanks Mayumi Adachi and all the ICMPC10 organizers for co-hosting this event with us. We are grateful to the Yamaha Corporation for providing the C-III Grand Piano, the Roland Corporation for providing the Rencon prizes, Tadahiro Murao for composing the set pieces and providing useful advice, and Toshie Matsui for serving as the human pianist. We also thank all the participants and attendees at ICMPC10.

REFERENCES

Arcos, J., de Mantaras, R. & Serra, X. (1998). SaxEx: A Case-Based Reasoning System for Generating Expressive Musical Performances. *Journal of New Music Research*, Vol. 27, No. 3. 194–210.

Bresin, R. & Battel, G. (2000). Articulation strategies in expressive piano performance: Analysis of legato, staccato, and repeated notes in performances of the andante movement of Mozart's Sonata in G major (k 545). *Journal of New Music Research*, Vol. 29, No. 3. 211–224.

Clynes, M. (1984). A Composing program incorporating Microstructure. *Proceedings of International Computer Music Conference (ICMC)*. 225–232.

Frydén, L. & Sundberg, J. (1984). Performance Rules for Melodies. Origin, Functions, Purposes. *Proceedings Of International Computer Music Conference (ICMC).* 221–225.

Grachten, M. (2006). Expressivity-aware Tempo Transformations of Music Performances Using Case Based Reasoning. *Ph. D Thesis of University Pompeu Fabra*.

Hiraga, R., Hashida, M., Hirata, K., Katayose, K. & Noike, K. (2002). RENCON: Toward a New Evaluation Method for Performance Rendering System. *Proceedings of International Computer Music Conference.* 357–360.

Hiraga, R., Hashida, M., Hirata, K., Katayose, H. & Noike, K. (2002). Rencon: toward a common basis for performance rendering concours, *Proceedings of 2002 Rencon Workshop, International Conference on Auditory Display*. 51–53.

Ishikawa, O., Katayose, H. & Inokuchi, S. (2002). Identification of Music Performance Rules Based on Iterated Multiple Regression Analysis. *IPSJ Journal*, Vol. 43, No. 2. 268–276 (in Japanese).

Kostka, S. & Payne, D. (2004). *Tonal Harmony: With An Introduction To Twentieth-Century Music* (5th ed.). McGraw-Hill.

Lerdahl, F. & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Madsen, S. T. & Widmer, G. (2006). Exploring Pianist Performance Styles with Evolutionary String Matching. *International Journal on Artificial Intelligence Tools*, Vol. 15. 495-514

Narmour, E (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*, University of Chicago Press.

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. Cambridge, MA: MIT Press.

Widmer, G. (1996). Learning Expressive Performance: The Structure-Level Approach. *Journal of New Music Research*, Vol. 25, No. 2. 179–205.