Ha-Hi-Hun plays Chopin's Etude

Keiji Hirata∗**, Rumi Hiraga**†
∗ NTT Communication Science Research Laboratory, †Bunkyo University
email: hirata@brl.ntt.co.jp

Abstract

A new framework called two-stage performance rendering was proposed in order to make it realize incremental, interactive, and local rendering through direct instructions issued by a user. The first stage translates a user's instruction into the deviations of the onset time, duration, and amplitude of structurally important notes. The second stage spreads the deviations over surrounding notes. Ha-Hi-Hun is a prototype performance rendering system having the framework.

1 Introduction

To achieve a high level of controllability, we assert that a performance rendering (PR) system should be able to (a) properly interpret the user's instructions and (b) synthesize a natural performance that reflects these instructions. For (a), if instructions are given in a natural language, they are usually subjective, equivocal, and even time-varying. The system should be able to be customized or personalized and be context-sensitive. To solve this problem, there can be a method that a user gives some sample performances instead of a natural language. For (b), let us suppose a case in which a user gives an instruction to play note Q louder in a particular part of a piece. If the system naively increases only the amplitude of Q, the generated performance may become unnatural. Considering the role of Q in the piece, the surrounding notes should also be played either louder or softer and even their agogics may have to be adjusted. Thus, to keep a generated performance natural, a PR system must maintain a certain musical consistency, which is represented in the form of the constraints regarding the agogics and dynamics for Q and the surrounding notes. To meet these two requirements, this paper proposes a new framework called *two-stage performance rendering* [1].

2 Two-Stage Performance Rendering

Let us assume that tutor's instructions for expression are issued to salient notes. That is, when a tutor says "play this note carefully", this note means a salient note within a certain time range. The two-stage PR framework is motivated by this assumption.

2.1 Architecture

In Fig. 1, the first stage translates a user's instruction into the agogics and dynamics of structurally important notes in a range and the second stage adjusts the surrounding notes. Here, a structurally important note

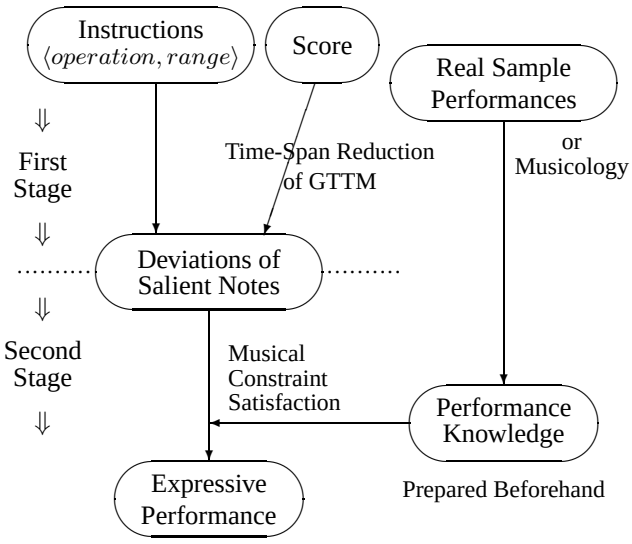

Figure 1: Two-Stage Performance Rendering

means a salient note in the context of the time-span reduction of GTTM.

The inputs for two-stage PR are a score to be performed, instructions given by the user, and sample performances for extracting performance knowledge, which may be substituted with built-in rules or mathematical

expressions derived from musicology. The user's instructions specify an operation and the range of the score to which the operation is applied. The operations include faster, brighter, more passionate, or "imitate this sample performance". The output is an expressive performance of the score with the instructions issued.

The first stage maps the user's subjective instructions in a natural language or with a sample performance itself to the deviations of onset time, duration and amplitude (velocity) for every prominent note. (Fig. 2). In

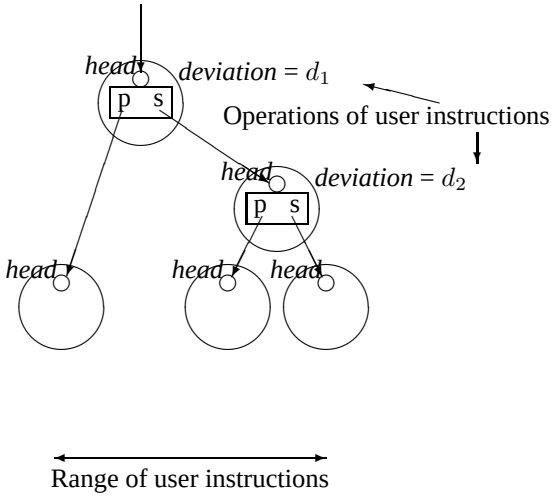

Figure 2: Setting of deviations to salient notes

the figure, the small circle stands for the *head* attribute, the big circle the *ts* object, p within the box the *primary* attribute, and s the *secondary* attribute. The *deviation* attribute is written outside of the big circle.

At the first stage, there are several ways for calculating the values of deviation d_1 and d_2 from user's instructions. Since the values are derived, for instance, from the heuristic values provided a priori, case-based reasoning, and/or learning on a user behavior and an operation environment, they highly depend on user's subjectivity and context in general.

On the other hand, notes included in a score are grouped hierarchically according to the time-span reduction. Then, the reduction identifies salient notes in groups at every level. The range of an instruction is mapped to the time span of a group.

The second stage propagates the deviations set up to salient notes at the first stage to their surrounding notes. (Fig. 3). At that time, the deviations of salient notes are unchanged, and only agogics and dynamics of the surrounding notes are adjusted. This stage is introduced to bring about a musically natural performance. The performance knowledge for the propagation should be obtained

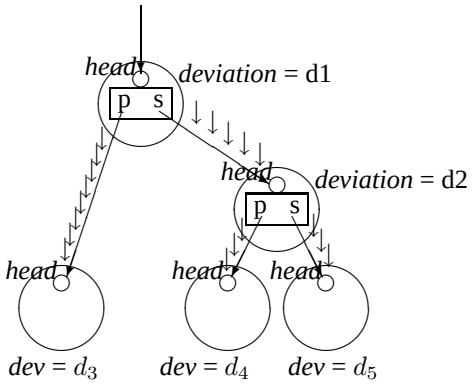

Figure 3: Musical Constraint Satisfaction

in advance, and may be acquired from the analysis of sample performances with some musical theories, such as GTTM and I-R model [3].

The arrows $\downarrow$ in the figure show the propagation of the deviations of d_1 and d_2 to the lower *ts* objects. While the values of d_1 and d_2 are unchanged, those of d_3, d_4 and d_5 are changed. The performance knowledge used here can be considered a constraint regarding the agogics and dynamics for salient notes and their surrounding ones.

3 Performance Rendering by Ha-Hi-Hun

At the listening comparison, Ha-Hi-Hun plays Etude Op.10, No. 3 by F. Chopin with Paderevski Edition using Chopin's Nocturn Op. 32, No. 2 that is rendered by M. Hashida as a case performance. The rendered performance and the case performance will be the same as those of O-SLuR and Kagurame-II. At the second Rencon last year, Ha-Hi-Hun rendered the same Etude with several case performances. The rendered music gave listeners the artificial impression and was not appreciated well. This time Ha-Hi-Hun uses a single case performance in order to give the consistency to a rendered performance.

The descriptions in Section 1 and Section 2 are the excerpt from the paper [2].

References

[1] Hirata, K., Hiraga, R. and Aoyagi, T.: Next Generation Performance Rendering – Exploiting Controllability, *Proc. of ICMC*, pp. 360–363, 2000.

[2] Hirata, K. and Hiraga, R.: Ha-Hi-Hun: Performance Rendering System of High Controllability, *ICAD 2002 Rencon Workshop Proceedings*, 2002.

[3] Lerdahl, F. and Jackendoff, R.: Generative Theory of Tonal Music, The MIT Press, 1983.