

Real-time Initiative Exchange Algorithm for Interactive Music System

Yoichiro Taki Kenji Suzuki Shuji Hashimoto

Department of Applied Physics, Waseda University
3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan
Tel. +81-3-5286-3233 Fax. +81-3-3202-7523
{takism, kenji, shuji}@shalab.phys.waseda.ac.jp

Abstract

In the musical performance played in concert, it is natural that players exchange the initiative of performance with each other, without spoiling the musical harmony. However, in the human-machine interactive musical performance system, the initiative exchange has not been considered explicitly. This paper proposes a novel human-machine interface system which allows the smooth initiative exchange between human and machine during the performance. The system consists of a Sound Analyzer Module, an Initiative Control Module and a Performance Module. The changes in tempo and volume are detected as triggers for the initiative transfer.

1. INTRODUCTION

When humans make a collaborative performance, the initiative exchange often takes place. In some cases the initiative transfer is initiated with gestures such as signing and winking. In other cases the control key of the interactive effects including initiative exchange is embedded in the performed musical sounds. In order to realize an elegant and natural style of performance, it is important for performers to communicate with partners through the musical sounds without spoiling the performance coordination.

There have been reported a variety of interactive music systems some of which enable machines to follow and to synchronize with human performance [Vorcoe, 1984] [Dannenberg, 1984][Katayose et. al, 1993][Grubb et. al, 1997][Inoue et. al, 1993]. Such systems seem to be one-sided, because the machine is not autonomous but a slave performer. If the machine has some degree of independence [Horiuchi et. al, 1993] and an ability to take and give the initiative the performance, it is expected that the human performer will be extremely stimulated to enhance the creativity by the cooperative effect.

In this paper, we propose a human-machine interactive system that makes it possible to exchange the performance initiative elegantly between human and machine. We do not introduce any special switch device for the initiative control. The human intention of initiative transfer is decoded from the performed sound and the machine intention (state) is coded in the generated sound not to spoil the musical harmony. The proposed system is developed in MAX/FTS environment on Macintosh G3 equipped with a microphone and speakers.

2. INITIATIVE EXCHANGE ALGORITHM

2-1. Definition of the Term “Initiative”

We define the term of “Initiative” used through this paper, as the authority to vary the performance tempo and to make another performer follow.

2-2. The Conditions for Initiative Exchange

To realize the smooth initiative exchange between human and machine, the conditions for the occurrence of initiative exchange must be suited for the human intuitive senses. We adopt the following assumptions for initiative exchange.

- (i) Both the human performer and the system change the performance volume or the performance tempo to express the intentions for initiative exchange. The increase and the decrease of volume which exceeds a particular threshold value signify the ‘Claim’ of initiative to another performer and the ‘Renunciation’ of initiative, respectively. On the other hand, if the human performer plays without the conspicuous variation of the performance tempo, we assume that he does not want to keep his superiority but allows the system to take the initiative.
- (ii) The system has several internal patterns in advance, each of which expresses the variation of performance tempo through a piece of music. When the initiative is transferred from the human performer to the system, the system selects the internal pattern with the greatest similarity with the human performance; i.e. the system calculates distances between the human performance tempo variation and the internal tempo patterns during the following phase, to recognize the most similar

performance style S based on the least square error as shown in Equation (2.1).

$$\begin{cases} S = \min s_k \\ s_k = \sum_{n=1}^N (T_{kn} - R_n)^2 \end{cases} \quad (k=1,2,\dots,m)$$

T_{kn} : Ave. tempo of n -th measure in the internal pattern k
 R_n : Ave. tempo of n -th measure, performed actually

(2.1)

Where, m and N represent the number of internal patterns and the number of measures the performer has played until then, respectively

- (iii) The system enters the state to receive the signal for initiative exchange when the accumulated square difference between machine tempo and human tempo exceeds a certain threshold after taking the initiative. We call this state "Critical State" in the sense that the initiative transfer can happen very easily compared to the ordinary state. Then the system performs with a slightly low or high volume to inform the performer that the system is in the "Critical State". and the consequence which takes the initiative later depend on the assumption (i).
- (iv) The opportunity that the performer can exchange the initiative is not restricted the case of (iii). The human performer can also exchange the initiative without entering the critical state if the sound volume exceeds another threshold value that is larger than the threshold value required to exchange the initiative in the critical state.
- (v) When the human performer desires to obtain the initiative from the system, the performer has to synchronize his performance with the tempo of the system to some extent. To do so, the performer can generate the critical state. But once the critical state is archived, it is no necessary to perform along with the system. On the contrary, the off-synchronization is regarded as one of the positive attitude for taking initiative after the critical state.

2-3. The Renewal of the Internal Pattern

The tempo variation of the rendering performed is stored as one of the internal patterns which is a template of the machine performance. In this case, the most dissimilar pattern which is the internal pattern with the maximum squared errors, is replaced with the one performed. Because the performance pattern which is actually played once by the human performer is considered to contain his musical idea and to be possibly played again.

2-4. The Initiative Control Flow

In this section, we describe how the initiative is

transferred between the human performer and the system in the interactive performance.

At the beginning of the performance, the human performer takes the initiative; i.e. the system accompanies the human performance. When the accumulated squared errors S exceeds a particular threshold value, the system enters the Critical State regarded as a special stage in which the initiative transfer from the human performer to the system is likely to occur. The human can distinguish the state change, because the system turns the performance volume high to ask for the initiative exchange.

In the stage, when the human performer takes a concessive or approval attitude, that is, playing with lower volume or stop performing, the initiative transfer from the human performer to the system will take place. If the performer takes a defiant or negative attitude, that is, playing loudly, the initiative transfer will not occur because it is regarded as the expression of denial.

On the other hand, the initiative transfer from the system to the performer will take place when the performer indicates his intention, that is, increasing the volume. In this case, the system plays with lower volume to express that it is prepared to give the initiative to the performer. Therefore if the performer does not wish to obtain the initiative, he ought to perform with the lower volume.

3. SYSTEM OVERVIEW

All of the system is developed in MAX/FTS environment on Macintosh G3 without any other special equipment. In this section, we describe the details of three modules; Sound Analyzer Module, Initiative Control Module and Performance Module.

3-1. Sound Analyzer Module

This module is used to analyze various types of sounds; the singing voice, instrumental sound and clapping sound. Firstly, it extracts the volume information of input sounds at the rate of 100Hz to recognize the tempo and volume. Then, the value of volume is calculated with Equation (3.1).

$$Volume = 10 \log_{10} \left\{ \frac{1}{N} \sum_{n=0}^{N-1} x^2(t + n\Delta t) \right\} \left/ \frac{A^2}{2} \right\} \quad [dB]$$

(3.1)

Where, x , A and N indicate the input amplitude, the maximum amplitude and the number of samples per a frame, respectively.

In addition to the information of the volume, this module detects the pitch from the singing voice and instrumental sound. To detect the pitch, the fundamental frequency is calculated with the cepstrum method assuming the harmonizing structure of the sound. For the purpose of evading the sampling errors, the fundamental frequency is

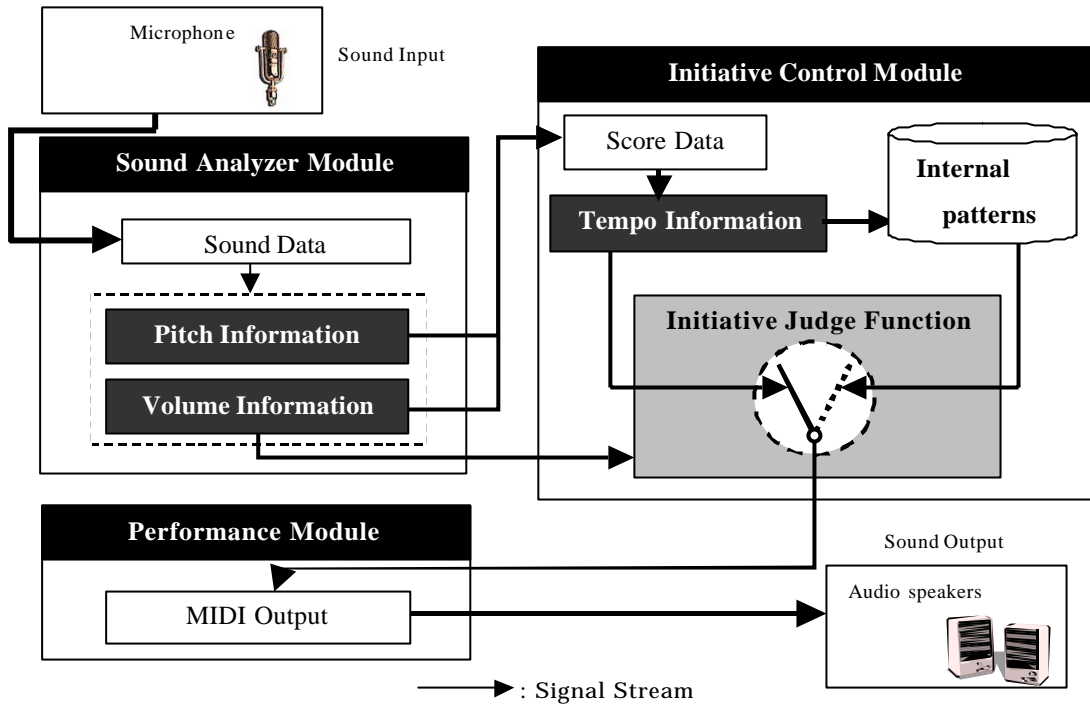


Figure 1. System Overview

calculated as follows.

$$f_{fund.} \approx \frac{\sum_{k=1}^p f_k}{\sum_{k=1}^p k} \quad (3.2)$$

f_k : k -th peak of frequency

The information of pitch, which is extracted from input sound, is used to infer the current location on the musical score by comparing with the note information of the score which is given in advance.

When the input sound is clapping, this module can not detect the pitch but the volume of the sound. In this case, this module estimates the performance tempo by measuring the time interval between claps. These results are used for the system to accompany with the human performer.

3-2. Initiative Control Module

This module receives volume and tempo information from Sound Analyzer module. It recognizes the style of the performance by calculating the squared errors between the tempo variation pattern of the performer's sound and the internal patterns and control the initiative transfer based on the conditions described in the section 2-2.

3-3. Performance Module

Performance Module generates MIDI signal and generates an accompaniment sounds. It follows the decision that the Initiative Control Module did. So the Performance Module not only follows the human performer but also leads the performer's playing on occasion.

4. EXPERIMENT

In this section, we show a typical example of the interactive performance generated by using the proposed system.

We adopt the sounds of MIDI data as a system accompaniment for the human performer. The system stores the information of musical score to recognize the current position of the performance. The five internal patterns which are some performance models created by the preliminary performances are stored in the system, too. In order to accompany with the human performer, the system estimates the performance tempo by the linear prediction method.

Figure 2 shows the result of the interactive performance between the human singer and the system. In this figure, the flat line at the bottom of this figure indicates the possessor of the initiative. The higher level of this line indicates that the initiative belongs to the human performer. To the contrary, the lower level represents that the system takes the initiative of performance. The middle height of these two levels signifies critical state. Hence this line has no relation to the scale of both Y-axes.

In this figure, the initiative transfer during performance is illustrated. The sudden dropping of 'Performer Volume' indicates that the human performer stopped performing and the initiative exchange occurred. After that, the system did not follow the human performer. And we can see the initiative exchange in the critical state also took place.

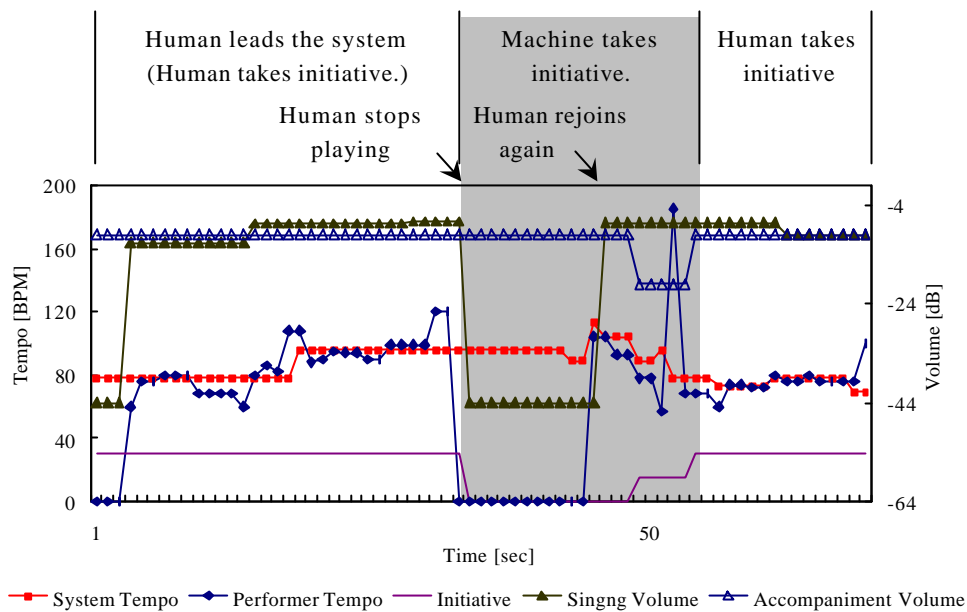


Figure 2. The result of the Interaction

5. CONCLUSION AND FUTURE WORK

We have proposed a real-time initiative exchange algorithm for the interactive music system. As one of the applied example of this technique, we have practically constructed a new accompaniment system. The notable point of the proposed method is that the system not only follows the human performer but also collaborates with by exchanging the initiative in a natural way to play music together.

From the experimental results, we have shown that the initiative exchange algorithm is available to bring the interactive effects between the human performer and the system and to realize the deep collaboration between them.

Another advantage of the system from the artistic point of view is that the initiative transfer between the human and the system can realize the performance with an unexpected style to stimulate the human creativity bringing novel ideas for the music performance and composition.

We are now considering to implement the proposed method into the humanoid robot system which create the dance performance in cooperation with human dancer [Suzuki et.al, 1999].

REFERENCES

- [Vorceo , 1984] B.Vorceo. 1984. "The synthetic performer in the context of live performance" Proc. of ICMC '84, pp.199-200.
- [Dannenberg, 1984] R.B.Dannenberg. 1984. "An On-Line Algorithm for Real-Time Accompaniment" Proc. of ICMC '84, pp.193-198.
- [Katayose et.al, 1993] H.katayose, T.Kanamori, K.Kamei, Y.Nagashima, K.Sato, S.Inokuchi and S.Shimura. 1993. "Virtual Performer" Proc. of ICMC '93, pp.138-145.
- [Grubb et. al, 1997] L.Grubb and R.B. Dannenberg. 1997. "A Stochastic Method of Tracking a Vocal Performer" Proc. of ICMC '97, pp.301-308.
- [Inoue, et. al, 1993] W.Inoue, S.Hashimoto and S.Ohteru. 1993. "A Computer Music System for Human Singing" Proc. of ICMC '93, pp.150-153.
- [Horiuchi et.al, 1993]Y.Horiuchi and H.Tanaka. 1993. "A Computer Accompaniment System with Independence" Proc. of ICMC '93, pp.418-420.
- [Suzuki et. al, 1999] K.Suzuki, T.Ohashi, and S.Hashimoto. 1999. "Interactive Multimodal Mobile Robot for Musical Performance" Proc. of ICMC '99, pp.407-410.