

Using an Expressive Performance Template in a Music Conducting Interface

Haruhiro Katayose

Kwansei Gakuin University
Gakuen, Sanda, 669-1337 JAPAN
<http://ist.ksc.kwansei.ac.jp/~katayose/>

Keita Okudaira

Kwansei Gakuin University
Gakuen, Sanda, 669-1337 JAPAN
keita@ksc.kwansei.ac.jp

ABSTRACT

This paper describes an approach for playing expressive music, as it refers to a pianist's expressiveness, with a tapping-style interface. MIDI-formatted expressive performances played by pianists were first analyzed and transformed into performance templates, in which the deviations from a canonical description was separately described for each event. Using one of the templates as a skill complement, a player can play music expressively over and under the beat level. This paper presents a scheduler that allows a player to mix her/his own intension and the expressiveness in the performance template. The results of a forty-subject user study suggest that using the expression template contributes the subject's joy of playing music with the tapping-style performance interface. This result is also supported by a brain activation study that was done using a near-infrared spectroscopy (NIRS).

Categories and Subject Descriptors

H.5.5 [Information Interfaces and Presentation]: Sound and Music Computing *methodologies and techniques*.

Keywords

Rencon, interfaces for musical expression, visualization

INTRODUCTION

Although it is fun to play a musical instrument, not a few people have experienced embarrassment due to their lack of

skill in playing one. Sometimes this situation may be a reason for quitting playing music and giving up a means of self-expression. Interactive musical instruments are meant to overcome this problem. They are expected to give users a chance to express what they would like to express even if they lack certain musical skills.

The score-follower based on beat tapping and proposed by Mathews [1] is a simple, intuitive music interface to express tempo and dynamics. It is intended especially for amateurs. Mathews' work has been followed by various conducting systems [2,3,4].

If the note descriptions of the score given to the system are nominal (quantized), the players' expression would be limited to the tempi and dynamics. We designed a score-follower called iFP, which utilizes expression templates derived from virtuosi performances. iFP enables the users to enjoy the experience of playing music, as if he or she had the hands of a virtuoso.

The next section outlines the design of iFP. We then describe the procedures to obtain expression templates. After introducing the user-interfaces, we discuss the effectiveness of using the expressive performance template as determined from a subjective evaluation and an observation of the test subject's brain activity.

SYSTEM OVERVIEW

In this section, we briefly describe the iFP design [5] and

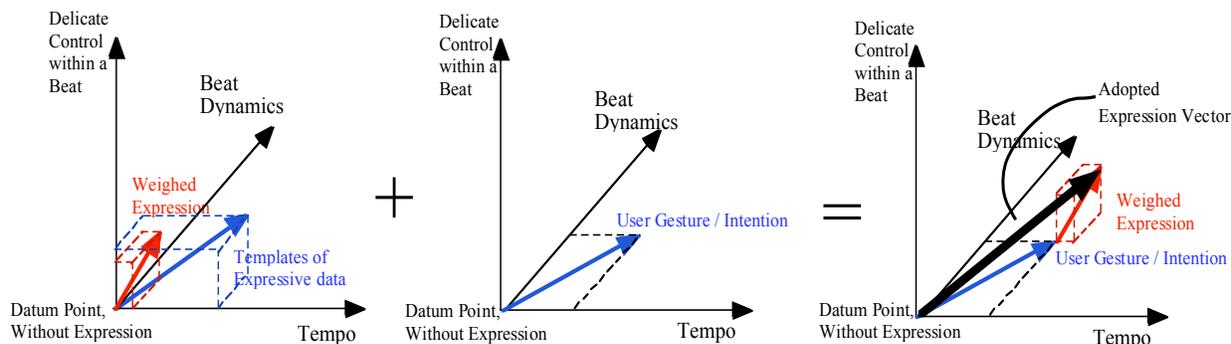


Figure 1. Conceptual overview of performance calculation. The performance data are given by a mixture of the player's intention and expressiveness described in the performance template. In this three dimensional space, the vertical axis denotes the variance of deviations of all notes within the beat.

```

.....
2.00 BPM 126.2 4
2.00 (0.00 E3 78 3.00 -0.11)
=2
1.00 TACTUS 2 4
1.00 BPM 128.1 4
1.00 (0.00 C#4 76 0.75 -0.09) (0.04 E1 60 1.00 -0.13)
1.75 (0.10 D4 77 0.25 -0.14)
2.00 BPM 130.0 4
2.00 (0.00 B3 75 1.00 -0.03) (0.00 G#3 56 1.00 0.03)
3.00 BPM 127.7
3.00 (0.00 B3 72 1.00 0.00) (0.09 G#3 56 1.00 -0.12) (0.14 D3 57 1.00 -0.21)
=3
1.00 TACTUS 1 4
1.00 BPM 127.6 4
1.00 (0.00 B3 77 2.00 -0.05) (0.00 G#3 47 2.00 -0.05) (-0.06 D4 57 2.00 -0.32)
3.00 BPM 129.7 4
3.00 (0.00 F#4 75 1.00 -0.15) (0.00 D4 54 1.00 0.03)
=4
1.00 BPM 127.7 4
1.00 (0.00 D#4 73 0.75 -0.38) (0.02 C4 65 0.75 -0.08)
.....

```

Figure 2. Description of performance template.

some of its functions. We then illustrate how the expressive performance template is utilized in iFP and describe its principal functions.

Utilizing a Template

Beat tapping is an intuitive interface to input tempo and dynamics to a performance system. However, the player cannot express the delicate sub-beat nuance with only beat tapping. The primary goal of utilizing the expressive performance template is to fill in expressions at the sub-beat level. The player's intention and the expression model described in the template are mixed as shown in Figure 1.

The player is allowed to vary the weight parameters dynamically by using sliders, each of which is multiplied with the deviation of tempo, dynamics, and delicate nuance within a beat. If all of these weight parameters are set to 0%, the expression of the template has no effect. On the other hand, if it is set to 120%, for example, the player can emphasize the deviations of the template.

iFP also provides a morphing function to interpolate (extrapolate) two different expressive performance templates of a musical piece.

Outline of Scheduling

Schedulers of interactive music systems have to calculate the timing of notes dynamically. iFP adopts a predictive scheduler, which arranges notes from the current beat to the next beat by using the history of the player's tap. One of the important points of using a predictive scheduler is that tap (beat) detection and scheduling of notes should be independent. This yields the merits of 1) compensation of the delay when using a MIDI-driven *robotic* acoustic instrument, and 2) easy implementation of the automatic performance mode (sequencer of the performance template).

A predictive scheduler might produce undesired gaps between the predicted beat timing and the actual players' tap. Especially if the gap is a delay, it may be perceived as

spoiling the performance. We prepared two countermeasures to improve the response; one is a function to urge the system when player's tap precedes the scheduled time, and the other is for receiving double taps for the given tactus (see the Expressive Performance Template section).

Functions

The features described above and the other characteristic functions are summarized as follows:

- Utilization of expressive performance template
- Real-time control of weight parameters regarding expressions
- Morphing of two different performance templates
- Predictive scheduling which allows the player to tap an arbitrary beat
- Pauses (breath) inserted based on release timing
- Real-time visualization (feedback) of expressiveness.
- Gestural input using a conducting sensor, a MIDI keyboard and a computer keyboard.

EXPRESSIVE PERFORMANCE TEMPLATE

Format

Figure 2 shows a part of a performance template. The left row represents the start timing of the events. Information about each note, except for the start timing, is placed in brackets. Each bracketed term, in order, represents, the deviation regarding the start time, note name, velocity, duration, and the deviation of duration, respectively. In this format, the tempo is described using the descriptor **BPM**. The description is followed by the tempo (in bpm beat per minuts) and the beat name to which the tempo is given.

iFP's predictive scheduler continues the performance, even if the performer does stop tapping. The player does not have to tap every beat. However, there often is the case that the players wish to tap to each note, instead of the beat. We introduced a descriptor TACTUS to explicitly describe how many taps are received for the beat. The following bracketed expression is an example of a TACTUS description. **(1.00 TACTUS 2 4)** This example means that after time 1.00, two taps are received for quarter notes; in other words, the system receives a tap every eighth note, after time 1.00.

Preparing Expressive Performance Templates

This section describes the procedure to make performance templates. The first step is to identify the canonical time value of each played note of the given expressive performance. It is not easy to obtain quantized notation, because local tempi varies more than twice from the average tempo. Manual quantization is extremely troublesome. One possible approach is to use an automatic

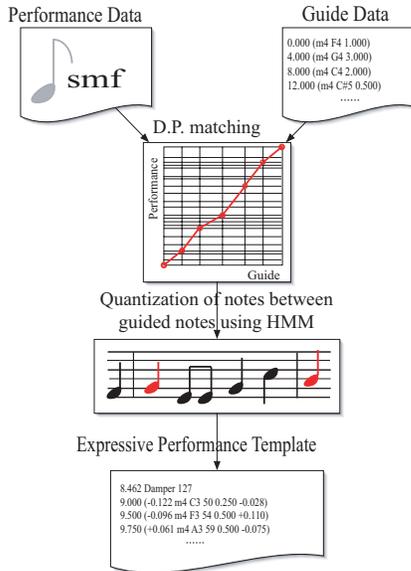


Figure 3. Acquisition of expressive performance template based on D.P. and HMM.

matching procedure that matches the notes in the performance with those in the score. However, the input of score data is itself time-consuming. Therefore, we designed a tool which identifies the notes in the performance given only a sparse score and then assigns a canonical notation value and deviation to all of the notes (see Figure 3) [6]. The DP matching procedure is utilized for the 1st step and a Hidden Markov Model (HMM) is used for assigning the 2nd time value to the notes. This tool enables us to prepare error-free expressive performance templates by giving only 10% of the notes as guides. At present, we have made over 100 expressive performance templates.

SCHEDULER

In this section, we describe the scheduler that realizes a mixture of the player's intention and expressiveness described in the performance template.

Calculation of Tempo

The tempo is calculated using 1) the average tempo obtained from the recent history (tactus-count: α) of the tempo, 2) the tempo to which the player's veer is considered, using the differential of the two most recent tapping, and 3) the prescribed tempo in the performance template [$Tempo_T$]. Let $stdTEMPO$ denote the overall average tempo of the template, and w_H , w_P , and w_T denote the weights for 1), 2) and 3) respectively. The tempo after n_{th} tactus, BPM_n is calculated as:

$$BPM_n = \left\{ w_T \times (Tempo_T - stdTempo) + w_H \left(\frac{1}{\alpha} \sum_{k=n-\alpha}^{n-1} BPM_k - stdTempo \right) \right\} \times \frac{1}{w_T + w_H} \times \left(\frac{BPM_{n-1}}{BPM_{n-2}} \right)^{w_P}$$

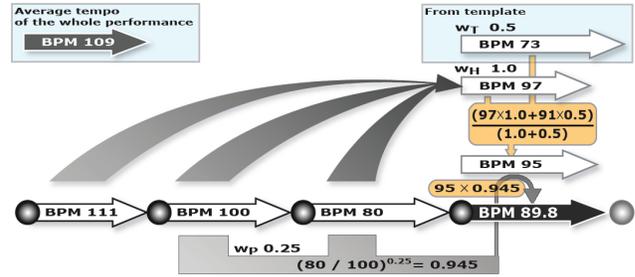


Figure 4. Calculation of tempo

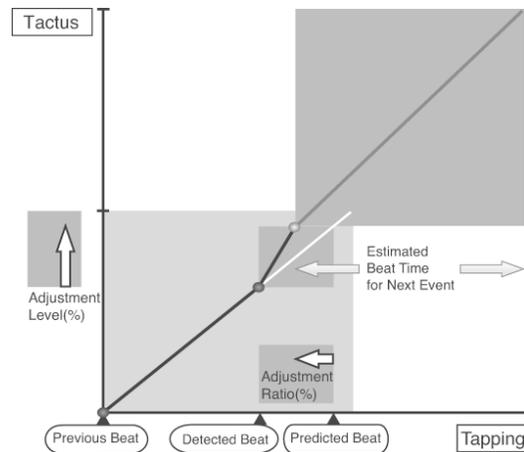


Figure 5. Adjustment ratio and level in a tempo map.

Figure 4 shows an example of tempo calculation. If the player sets w_T to a bigger value, more template data will be imported, and the player can feel like conducting a pianist. Setting w_P to a bigger value quickens the response. By setting w_H to a bigger value, the tempo of the music will be stable, affected by the recent average tempo. The user can set the parameters as s/he likes.

Improvement of Response

The problem with using predictive control is the possibility of undesired gaps between the predicted beat time and the actual user input. We introduced the following function in order to fill the gap, when a player's tap for the next beat is prior to the scheduled (predicted) timing.

Figure 5 shows the scheduling status in a tempo map. In this figure, the horizontal axis is the actual time, and the vertical axis is the tactus. The line drawn at a gradient in the map stands for tempo. Here, the adjustment level is a parameter that stands for how much the scheduler has to shrink the gap between the scheduled (predicted) time and the player's real tap, that is, to re-schedule the system beat time, when the player's tap is detected prior to the scheduled time. The adjustment ratio stands for the weight to fix the current beat time, between the player's tap and the scheduler beat time in order that the system can predict the next beat time.

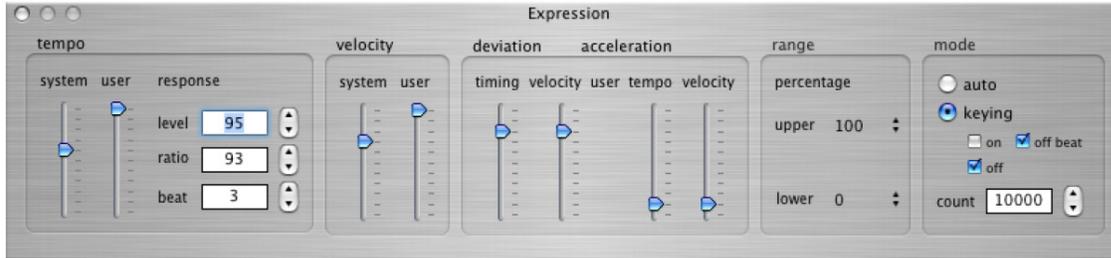


Figure 6. Standard GUI for setting parameters.

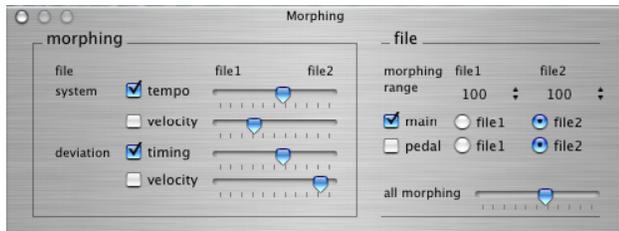


Figure 7. GUI for morphing.

Let denote the $beat_pos_{fix}$ the current beat time to be set, and $beat_pos_{scheduler}$ and $beat_pos_{tap}$ denote the scheduled beat time and the time the player taps for the beat, respectively. If the player's tap is detected prior to the scheduled time and the adjustment level is set to 100%, the system instantaneously issues events that correspond to the note on the beat, before the scheduled time. Then,

$$beat_pos_{fix} (= beat_pos_{scheduler}) = beat_pos_{tap}$$

If the scheduler beat time and the player's tap timing are different,

$$beat_pos_{fix} = A_L / 100 \cdot beat_pos_{tap} + (100 - A_L) / 100 \cdot beat_pos_{scheduler}$$

where, A_L is the adjustment level.

Calculation of Note Event Timing

The timing of each note event (note-on, note-off) is calculated using IOI_n given by the inverse of BPM_n (see the Calculation of Tempo section), as follows,

$$Time_{each_issue} = IOI_n \cdot (pos_{T_each_note} + dev_{T_each_note} \cdot w_{T_dev})$$

where, $Time_{each_issue}$ [s] is the time after the identified current beat, pos_{each_note} is the scheduled time of the note without deviation, $dev_{T_each_note}$ is the value of the deviation term of the note, and the w_{T_dev} is the weighting factor for the template. When $w_{T_dev} = 0$ is given, the temporal deviation under beat level will be mechanical.

Calculation of Velocity (Note Intensity)

The notes are classified into the control notes (note on the beat) and the remainder. First, the system decides the beat velocity V_{beat} for the beat. It is calculated, considering how loud/quiet the player and the machine (performance template) intend to play the note of the beat.

$$V_{beat} = V_{std} + \{ (V_T - V_{std}) \cdot w_{T_v} + (V_U - V_{std}) \cdot w_{U_v} \} / (w_{T_v} + w_{U_v})$$



Figure 8. With a conducting gestural interface using capacity transducers.

where, V_{std} is the standard (average) velocity of the all notes of the template, V_T is the average of the note-on velocity within the beat, V_U is the velocity that the player gives, and w_{T_v} and w_{U_v} are the weights for V_T and V_U , respectively. When w_{T_v} and w_{U_v} are 0, the intensity deviation under beat level will be mechanical.

The velocity of the each note V_{each_issue} is calculated as:

$$V_{each_issue} = V_{beat} \cdot (1 + V_{T_each_dev} + V_{U_dev})$$

where, $V_{T_each_dev}$ stands for deviation in the template, and V_{U_dev} stands for the player's intensity veer.

$$V_{T_each_dev} = (V_{T_each_note} - V_T) / V_T \cdot w_{T_dev}$$

$$V_{U_dev} = (V_{U_current} - V_{U_prior}) / V_{U_prior} \cdot$$

$$(pos_{T_each_note} + dev_{T_each_note} \cdot w_{T_dev}) \cdot w_{U_dev}$$

where, $V_{T_each_note}$ is each note-on velocity within the beat, and V_{U_n} denotes the velocity given by the n th player's tap and w_{U_dev} denotes the weight for the player's intensity veer.

USER INTERFACE

GUI and Gestural Interface

Figure 6 shows the standard GUI to characterize the performance. The users are given sliders so they can set the weight parameters regarding tempo (template/user), dynamics (velocity: template/user), and deviation (delicate control within a beat: template). Figure 7 shows the GUI for morphing two performance templates. The player can interpolate and extrapolate the performance using each of the morphing parameters. The player is also allowed to use peripherals of MIDI instruments instead of the software sliders. If the radio button "keying" is selected, the system accepts the player's beat tap. If "auto" is selected, the

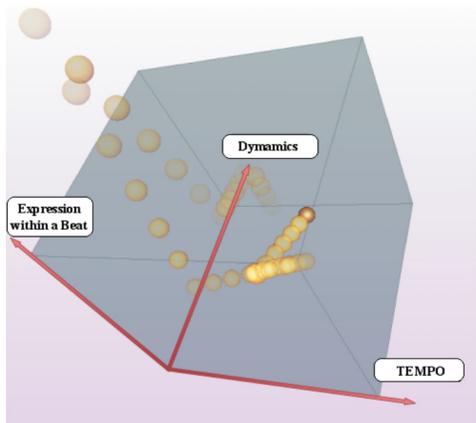


Figure 9. Example of visualization of a performance. K.331 played by S. Bunin.

system does not accept the beat tap. The expressive performance template is played without the control of the player's beat tap (automatic mode).

The gestural conducting sensor is based on capacity transducers (see Figure 8). We used hardware of DigitalTheremin manufactured by Yume-system (<http://www.moosys.co.jp/>), for the prototype. The beat signal is issued when the hand is located at the lowest position. The range of hand movement is assigned to the dynamics for the next beat. When the hand is lower than a certain threshold, the system holds the performance, i.e. gives a rest.

Visualization

iFP provides real-time visualization of the performance trajectory in three-dimensional space, as shown in Figure 9. The axes are the tempo, the dynamics, and the summed variance of the expression deviation within a beat. The user can observe the trajectory from various viewpoints.

If the player uses iFP with automatic (sequencer) mode, this visualization function should be the view of the expressiveness of the performance template.

EVALUATION

We conducted two experiments to verify the effectiveness of using expressive performance templates. One was an evaluation regarding the players' introspection, and the other was a brain activity measurement using near-infrared spectroscopy (NIRS).

Introspection Evaluation

We focused on “controllable” and “expressive” aspects for the introspection study. “Controllable” stood for difficulty in playing music. “Expressive” stood for how well the player could express the music. For the experiment, we used a conducting interface and an expressive template for “When You Wish Upon A Star” for piano. The system parameters for this experiment were those of a music

teacher who is also a conductor so that the performance taste would be close to conducting, as shown in Figure 6.

We interviewed forty subjects, whose music experience was 0~33 years. We asked them “Which performance (with / without the performance template) is more “controllable” or more “expressive”? We limited the time to practice to 10 minutes in this experiment. The results are shown in Table 1.

Table 1. Introspection Regarding Expression Template Use: The value in each column is the number of subjects who preferred the condition.

		Controllable		sum
		Better: with Template	Better: without Template	
Expressive	Better: with template	13	15	28
	Better: without Template	0	12	12
sum		13	27	40

We investigated the response of the 27 subjects who answered, “controllability is better without template”, by changing the parameters affecting controllability. All of the subjects answered that the controllability was improved when the parameters of both adjustment level and ratio were 100%. This meant dis-coincidence of the player's taps and heard beats makes the performance difficult. However, some of the experienced subjects commented that this dis-coincidence was indispensable to gain expressiveness.

Next, we investigated learning effects, for five of the 15 people who answered “expressive performance template contributes to expressiveness, but it does not contribute to controllability”. Four people among five subjects changed their opinion to “prefer to use a template also for controllability” after learning.

These results seem to indicate that the expressive performance template contributes to both *expressiveness* and *controllability*, after one has learned how to play the music using iFP.

Evaluation using NIRS

Physiological measurements are good for verifying subjective introspection results. Brain activity is a most promising measure for what a subject is thinking and feeling. Recently, a relatively new technique, near-infrared spectroscopy (NIRS), has been used to measure changes in cerebral oxygenation in human subjects [7]. Changes in oxyhemoglobin (HbO) and deoxyhemoglobin (Hb) detected by NIRS reflect changes in neurophysiological activity, and as a result, may be used as an index of brain activity. It is reported that the brain in the Fz area is deactivated (HbO decrease), when a subject is relaxed, in meditation, or in immersion in playing games. We measured brain activities around the Fz area when the subjects played with the iFP, or did other musical tasks (control tasks).

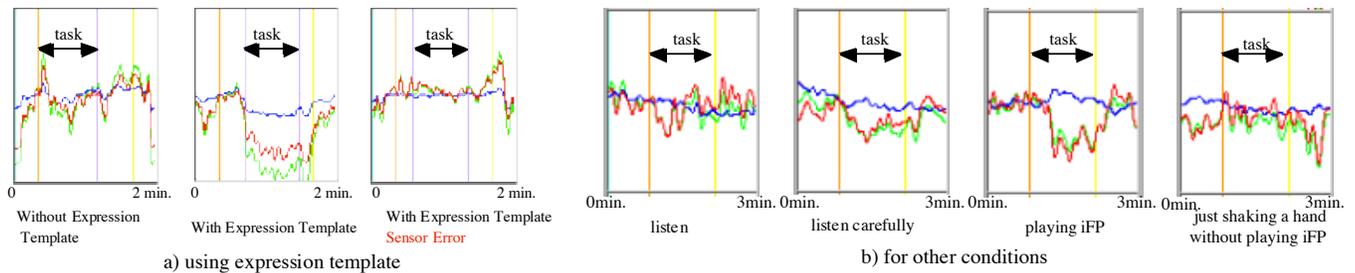


Figure 10. Brain activity measured with NIRS. Measured during listening and playing “When You Wish Upon A Star” using iFP. Arrows shows the duration of the performance. A single emitting source fiber was positioned at Fz. The red line shows the sum of Oxi-Hb.

Figure 10 shows the results of some of these experiments. These are data of a subject, who answered, “The expressive performance template contributes to both *expressiveness* and *controllability*.” The subject is educated in music, and received her Master of Music degree from a music university. Figure 10.a) is a comparison of using and not using the expressive performance template. We can see the decrease of HbO, when the subject played with iFP using the expressive performance template. The right data of Figure 10.a) were obtained by chance. It is interesting to see the response of the subject when something unexpected happened. Figure 10.b) is a comparison with other music activities. HbO was lower when the subject listened to the music carefully imagining that the subject was playing the piano, and played with the iFP. The decrease was more salient at with the iFP. These results correspond to the reports of the subjects’ introspection regarding pleasantness very well. Although the interpretation of deactivation at Fz itself is still controversial [8], we may say that introspection of using iFP is supported by the NIRS observation of brain activity.

CONCLUSION

This paper introduced a performance interface called iFP for playing expressive music with a conducting interface. MIDI-formatted expressive performances played by pianists were analyzed and transformed into performance templates, in which the deviations from the printed notation values are separately described. Using the template as a skill-complement, a player can play music expressively over and under beat level. The scheduler of iFP allows the player to mix her/his own intension and the expressiveness in the performance template. The results of a forty subject user study suggested that using the expression template contributes to a player’s joy of expressing music. This conclusion is also supported by the results of brain activity measurements.

We are just beginning our experiments using NIRS. We would like to trace the changes of the subjects’ introspection and brain activity, as they learn to play with the iFP. We are also interested in investigating interactions between brain regions when the subjects are playing music.

Another important task is to provide more data to be used in iFP. So far, a tool to produce a performance template from MIDI-formatted data has been completed. We would like to improve the tool, so it can convert acoustic music into the expressive performance template.

ACKNOWLEDGEMENT

The authors would like to thank Mr. Kenzi Noike, Ms. Mitsuyo Hashida, and Mr. Ken’ichi Toyoda for their contributions to the study. Prof. Hiroshi Hoshina and Mr. Yoshihiro Takeuchi made valuable comments, and they were very helpful. This research was supported by PRESTO, JST, JAPAN.

REFERENCES

- Mathews, M. The Conductor Program and Mechanical Baton, *Proc. Intl. Computer Music Conf.*, (1989) 58-70.
- Nakra, T. M. Synthesizing Expressive Music Through the Language of Conducting. *J. of New Music Research*, 31, 1,(2002), 11-26.
- Morita, H., Hashimoto, S. and Ohteru, S. A Computer Music System that Follows a Human Conductor. *IEEE Computer*, (1991), 44-53.
- Usa, S. and Mochida, Y. A Multi-modal Conducting Simulator,” *Proc. Int. Computer Music Conf.* (1998), pp.25-32.
- Katayose, H., Okudaira, K. sfp/punin: Performance Rendering Interfaces using Expression Model, *Proc. IJCAI03-workshop, Methods for Automatic Music Performance and their Applications in a Public Rendering Contest*, (2003), 11-16.
- Toyoda, K., Katayose, H., Noike, K. Utility System for Constructing Database of Performance Deviation, *SIGMUS-51, IPSJ*, (2003 in Japanese) 65-70.
- Eda, H., Oda, I., Ito, Y. et al. Multi-channel time-resolved optical tomographic imaging system. *Rev. Sci. Instrum.* 70, (1999) 3595-3602.
- Matsuda, G., and Hiraki, K. Frontal deactivation in video game players, *Proc. Conf. of Intl. Simulation And Gaming Assoc.(ISAGA)*, 110 (2003).