Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Estimating musical score and proficiency at playing drums

Yuki KONISHI (1) and Masanobu MIURA (2)

(1) Graduate School of Science and Technology, Ryukoku University, Japan
(2) Faculty of Science and Technology, Ryukoku University, Japan

PACS: 43.75.-z Music and musical instruments

ABSTRACT

We designed a system that estimates the musical score of a player's performance based on the Bayesian method. We also used a method of evaluating performance proficiency outlined in a previous study. The proposed system calculates an adaptation probability of the player's performance from a database of 47,000 patterns to estimate the musical score of an inputted performance. A posterior probability of each pattern in the database is then calculated by multiplying the adaptation probability by a prior probability, where the database is used to obtain the prior probability from the occurrence frequency of each pattern. The pattern with the highest prior probability is used as the estimated musical score. Experimental results showed that an F-measure of .93 was obtained, indicating that the proposed system is an effective means of estimating musical scores.

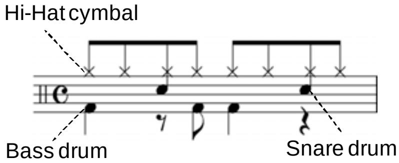
1. INTRODUCTION

Musicians face many challenges when practicing through self-learning, one of which is that they cannot obtain objective evaluations of their performances. Several researchers have previously developed systems to automatically evaluate the proficiency of musical performances [1-3], but their systems require users to input musical scores, which is often difficult for beginners. Estimating the performance proficiency of arbitrary performances is thought to be effective in terms of self-learning because users do not have to input the musical score they want to play. We have therefore developed a system that can evaluate proficiency of drum repetition without inputting a musical score.

To evaluate the performance proficiency of users' arbitrary performances, we need to estimate the musical score of the performance with a previous method for evaluating the proficiency of playing particular musical scores [2]. Our method is therefore based on the Bayesian method.

2. TARGET STYLE OF PRACTICE

We used a style of performance called "drum-loop performance" (DLP), which drummers usually use in practice (Fig. 1). It consists of one or two measures under a given tempo. The DLP musical score is primarily denoted by only a bass drum, snare drum, and closed hi-hat close cymbal, which is consistent with those used in actual performance. Figure 1 shows two typical patterns of DLP: 8-beat and 16-beat. DLP is one of the most basic musical scores for practicing the drums and is usually used in the training by not only amateur but also professional drummers. Consistent practice of the DLP is said to allow amateur drummers to acquire consistent skill for playing a variety of rhythm patterns.



(a) 8 beat rhythm

(b) 16 beat rhythm

Figure 1. Examples of rhythm patterns used in DLP

3. ESTIMATION OF A MUSICAL SCORE BASED ON THE BAYESIAN METHOD

3.1 Introduction

When playing drums, musicians often assume a musical score for the performance being done. We refer to such a score as an "intended score." On the other hand, an observed performance is called a "recorded performance", each note of which is called a "performed note". The estimated score obtained by the proposed method is called an "estimated score". The rhythm patterns of drums in the database are called "drum patterns". The relationships between these terminologies are shown in Fig. 2.

In this study, we assumed that the recorded performance is obtained via MIDI recording, so it is thus denoted as MIDI information. Therefore, the obtained information is thought to be a non-quantized performance, although conventional MIDI recordings represent the performance using ticks for time accuracy for example, 480 ticks per quarter note. It is important to note that human performance contains not only artistic but also non-artistic aspects, such as deviations or mistakes based on the player's habits and skill level, aspects that can be clarified by comparing them with the intended score. This means that the intended score and estimated score are not always consistent with each other.

To estimate a musical score from recorded performance, each performed note is extracted so as to estimate an intended score based on the maximum a posteriori (MAP) estimation. If the estimated score matches the intended score, the score estimation has been a success.

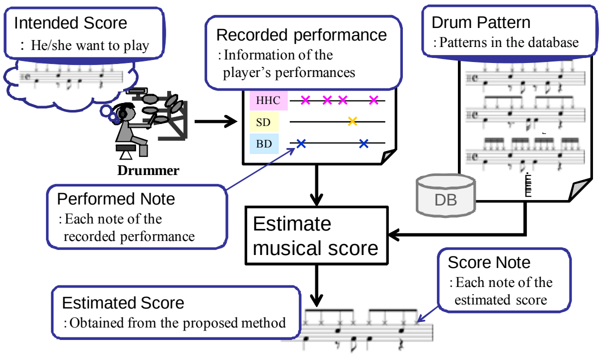


Figure 2. Terminology list

3.2 Database for estimating musical score

The proposed system uses a database constructed in a previous study [4] which in this study we call the "drum-pattern database". It stores timing and occurrence frequencies of drum patterns and consists only of a four-four meter within a measure. The frequencies are derived from the MIDI data. The drum patterns are quantized under the accuracy of demisemiquaver (8 divisions on a beat). The drum-pattern database is constructed using only a bass drum, snare drum and closed hi-hat cymbal.

3.3 Flowchart of musical score estimation

The score estimation flowchart is shown in Fig. 3, where the MAP estimation is used as the estimation method. This estimation considers a parameter θ for maximizing a probability distribution $P(\theta \mid X)$ of observed signal X as estimated value $\hat{\theta}$. The estimated value $\hat{\theta}$ is calculated as follows.

$$\hat{\theta} = \arg\max_{\theta} P(\theta \mid X) \cdot \tag{1}$$

We propose a score estimation method based on MAP estimation; in other words, the Bayesian method. The details are summarized as follows.

3.3.1 Algorithm for matching performance and drum pattern

To estimate the musical score of a recorded performance, we first need an algorithm that matches the recorded performance to a drum pattern. In our method, which uses only onset time and intensity information, the matching process is done with each percussion instrument. A conceptual flow of the matching algorithm is shown in Fig. 4.

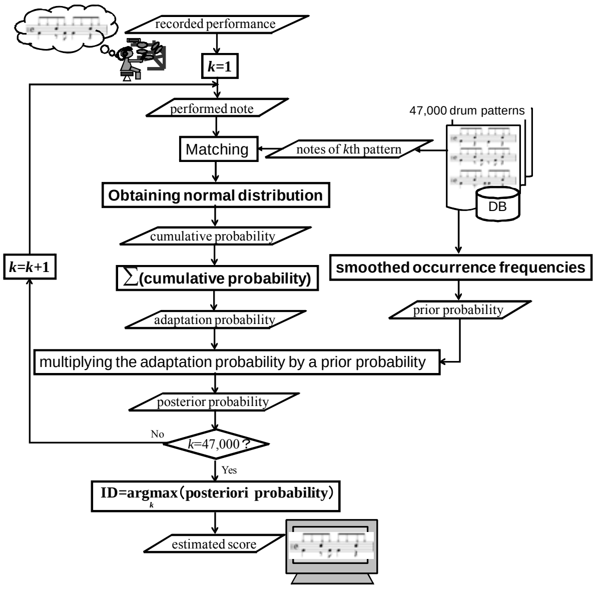


Figure 3. Score estimation flowchart

In this study, we assumed that users practice with a metronome. The drummer's performance is obtained via MIDI sequencer as recorded performance, and the first measure (defined by the metronome) is used to estimate a musical score. Adjusting the first performed note on the first measureto the first note for a drum pattern to be referred allows us to obtain the onset time R_m^k of mth note for a drum pattern, represented as a msec (not MIDI tick), where k is a drum pattern ID. A performed note is then matched with a score note that has a relatively small onset deviation from the performed note, as shown in Fig. 4 (a), where the maximum deviation on matching is set to 45 (msec). If several performance notes are located near a note in a drum pattern within 45 (msec), it is assumed that they correspond to the note in the drum pattern. To determine which performed note corresponds to the drum pattern's note, the performed note with the largest MIDI velocity among the notes within 45 (msec) is matched and considered the performed note for the drum pattern's note. On the other hand, drum pattern notes that do not match any performed notes are treated as missing notes. This processing is conducted in the order the notes appear in the drum pattern. The matching results are shown in Fig. 4 (b).

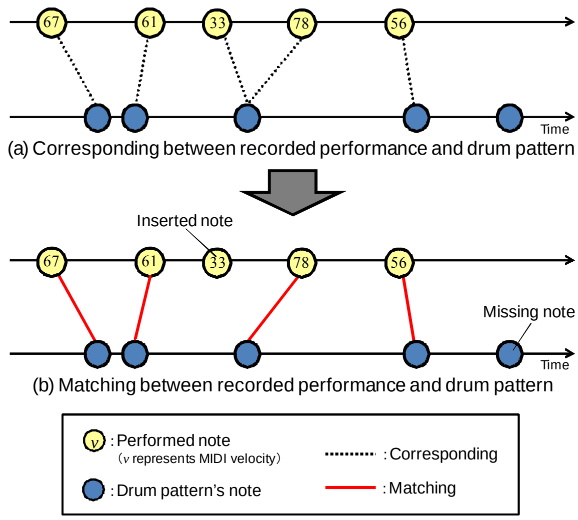


Figure 4. Outline of algorithm for matching performance and drum pattern

3.3.2 Calculating a normal distribution

A normal distribution T_n with parameters μ_n and σ is calculated as

$$T_n(t) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(t-\mu_n)^2}{2\sigma^2}\right), \qquad (2)$$

where t is a time, μ_n is onset time of a nth performed note and σ is a standard deviation. The standard deviation σ is substituted with 15 (msec) because we consider the inevitable deviation of the MIDI system. The T_n is used to estimate a musical score.

3.3.3 Calculating posterior probability for recorded performance

After matching the recorded performance with a drum pattern in the database, a cumulative probability p_n for the nth performed note is calculated from the nth onset time μ_n of the performed note and the mth onset time R_m^k of the kth drum pattern's note. Finally, the cumulative probability p_n is calculated as

$$p_n = \begin{cases} \int_{R_m^k}^{\infty} T_n(t)dt & (R_m^k > \mu_n) \\ \int_{-\infty}^{R_m^k} T_n(t)dt & (R_m^k \le \mu_n) \end{cases} . \qquad (3)$$

The more the onset deviations of μ_n from R_m^k increase, the more the cumulative probability p_n decreases. Therefore, the performed note with the smallest onset deviations to a drum pattern's note is considered more important in terms of estimating a musical score. A procedure for calculating a cumulative probability is shown in Fig. 5. After a performed note is matched with a drum pattern's note (Fig. 5 (a)), the cumulative probability p_n is calculated by considering an area enclosed by R_m^k and $T_n(t)$ (Fig. 5 (b)).

An adaptation probability $P(X \mid A_k)$ is then calculated by subtracting coefficient c from multiplying all cumulative probabilities p_n, shown as

$$P(X \mid A_k) = \prod_n p_n - c \quad (c = wI + vM), \qquad (4)$$

where X is a recorded performance actually observed, A_k is kth drum pattern, c is a coefficient, I is the number of inserted notes, M is the number of missing notes, w is a weighted coefficient for inserted notes and v is a weighted coefficient for missing notes.

A posterior probability $P(A_k \mid X)$ of each pattern is then calculated by multiplying the adaptation probability $P(X \mid A_k)$ by a prior probability $P(A_k)$, shown as

$$P(A_k \mid X) = P(A_k)P(X \mid A_k) . \qquad (5)$$

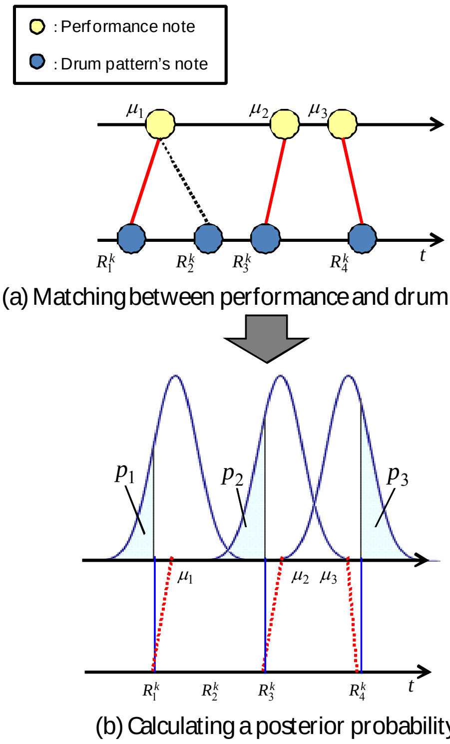


Figure 5. Procedure for calculating a cumulative probability

3.3.4 Estimating a musical score

After all the steps in 3.2.1 are performed for all drum patterns, the drum pattern with the highest posterior probability $P(A_k \mid X)$ is introduced as the estimated score. This score is obtained by the following calculation.

$$\text{ID of Estimated Score} = \operatorname*{argmax}_k \{P(A_k \mid X)\} . \qquad (6)$$

3.4 Calculating a prior probability

One way to calculate the prior probability of A_k that can be used to estimate a musical score is directly using the occurrence frequencies of each pattern in the database. However, it is not appropriate to do so because the similarity among patterns is not considered when obtaining the occurrence frequency. This means that a drum pattern with high occurrence frequency could be inaccurately introduced as an intended score even though a user played another pattern with a lower occurrence frequency.

To prevent such inaccuracies, we use another method that smoothes its occurrence frequency, so as to calculate the prior probability for A_k . Several surrounding occurrence frequencies are added to the occurrence frequency of A_k , as shown in Fig. 6.

This method we used is based on the concept of the Parzen window [5]. A prior probability $P(A_k)$ calculated by using the smoothed occurrence frequency is used when estimating a musical score. The Parzen window is a well-known method of determining a probability density function non-parametrically and is calculated as

$$p_N(x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{V}\varphi\left(\frac{x-x_i}{h}\right), \qquad (7)$$

where $p_N(x)$ is a probability density function, N is the number of samples, V is a volume of a hypercube centering on x, φ is a window function and h is a smoothing parameter.

The Parzen window can smooth a probability density function, by observing the probability of patterns of neighborhoods. Three examples of probability density function when using the Parzen window are shown in Fig. 7, where h is interpreted as an index for considering neighborhoods. A probability density function is too sharp or too flat when h is an inappropriate value (Fig. 7 (a)), whereas another probability density function is a desirable one when h is an appropriate value (Fig. 7 (b)).

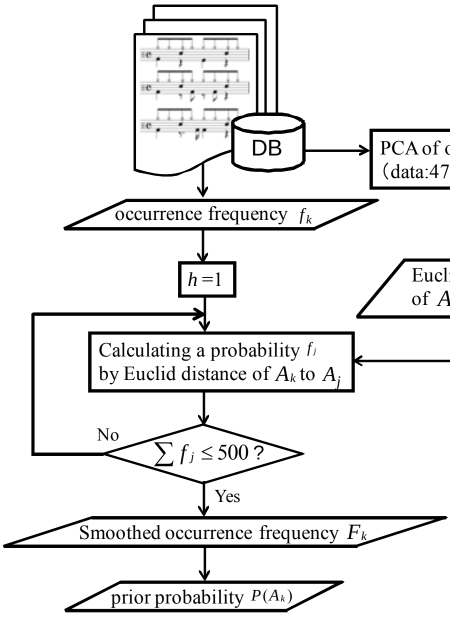


Figure 6. Flowchart of smoothing a posterior probability

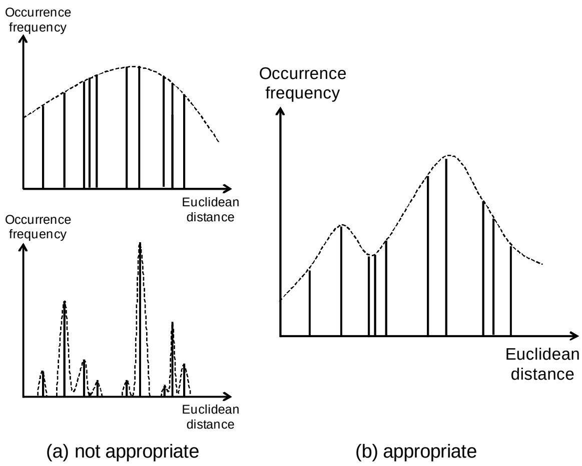


Figure 7. Estimated value of Parzen window

3.4.1 PCA of onset times in database

Before obtaining the smoothed occurrence frequencies, we used principal component analysis (PCA) on the drum pattern database. The matrix size for performing PCA is 47,000 × 96, for 47,000 patterns and 96 bits representing whether or not each of three instruments is played (bass drum, snare drum and closed hi-hat cymbal) at 32 onsets. First, the 49 principal components where the cumulative contribution ratio is more than 80% are obtained. The row matrix is then compressed to 49 parameters. Second, a Euclidean distance between each pattern is calculated using the score of a principal component for each pattern. Euclidean distance $d(k, j)$ between A_k and A_j is calculated as

$$d(k,j)=\sqrt{\sum_{i=1}^{49}(k_i - j_i)^2}\,, \qquad (8)$$

where A_k and A_j are kth and jth drum patterns, respectively, and k_i and j_i correspond to the 49-dimensional vectors for A_k and A_j, respectively.

3.4.2 Smoothing an occurrence frequency and calculating a prior probability

To obtain smoothed occurrence frequencies, a probability is obtained by converting the Euclidean distance between A_k and A_j to a probability, where A_k is the target drum pattern and A_j is the drum pattern near A_k and f_k and f_j are the occurrence frequencies of A_k and A_j, respectively. The probability represents the degree of similarity between them. Second, the smoothed occurrence frequency F_k is obtained by adding the score obtained by multiplying f_j by the probability defined by the Euclidean distance shown as

$$F_k = f_k + \sum_j \left[f_j \times \exp\left(-\frac{\|d(k,j)\|^2}{2}\right)\right], \text{ where } \sum f_j \le 500 \qquad (9)$$

F_k is evaluated for $\sum f_j \le 500$; that is, 1% of all samples in the drum pattern database.

After obtaining F_k for all drum patterns, we obtain prior probability $P(A_k)$ by dividing F_k by the total of smoothed occurrence frequencies for all drum patterns in the database, calculated as

$$P(A_k) = \frac{F_k}{\sum_l F_l}\,. \qquad (10)$$

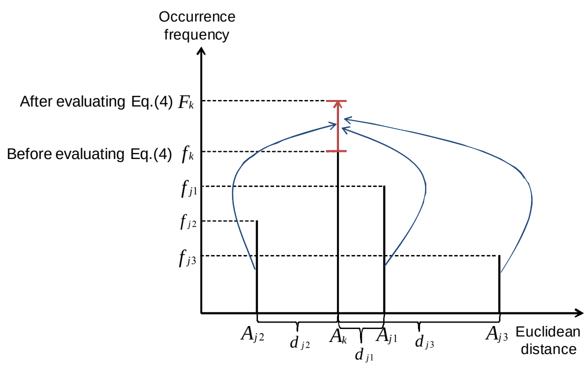


Figure 8. Calculating F_k

4. EVALUATING PLAYER PROFICIENCY

We used a method of evaluating proficiency based on one developed in a previous study [2]. A proficiency score is calculated from the relationship between an estimated score and the recorded performance.

First, we obtain 31 feature parameters for a player's performance classified into 18 parameters p_t for onset-time deviations and 13 parameters p_v in intensities. Second, principal components *P* are calculated by PCA. A proficiency score is then obtained from the relationship between the principal components *P* and the performance proficiencies evaluated by individuals using multiple regression analysis.

5. EVALUATION EXPERIMENT

5.1 Consideration of weighted coefficient based on the gradient method

An adaptation probability is calculated by subtracting coefficient *c* from the product of multiplying all cumulative probabilities, as shown in Eq. (4). Coefficient *c* is an index obtained from both the number of inserted notes and that of missing notes calculated for an adaptation probability. To find the best value of *w* and *v* in Eq. (4), we estimate the weighted coefficient *w* and *v* based on the gradient method.
The gradient method uses a recording obtained from the performances of two amateur drummers executing ten different drum patterns (shown in Fig. 9).

First, an estimated score is obtained from the player's performances by using the proposed method. Second, the F-measure is calculated between the intended score and the estimated score in order to adjust weighted coefficients *w* and *v* based on the gradient method. The estimated score is then again calculated using weighted renewal coefficients. This operation is repeated either for the difference between F-measure $^{g-1}$ when F-measure ^{g} is less than 10^{-3} or for observing the repetition of alternation of two consecutive F-measures, as well as when a value fluctuates. Renewal equations for *w* and *v* are shown below.

$$w^{g+1} = w^{g} - \alpha I((\mathrm{F-measure}^{g-1}) - (\mathrm{F-measure}^{g})), \tag{11}$$

$$v^{g+1} = v^{g} - \alpha M((\mathrm{F-measure}^{g-1}) - (\mathrm{F-measure}^{g})), \tag{12}$$

where w^g and v^g are the estimation parameters that obtain the repetition of *g*th.

The F-measure is defined as

$$\mathrm{F\text{-}measure} = \frac{2P_{recision}R_{ecall}}{P_{recision} + R_{ecall}}. \tag{13}$$

The precision is defined as

$$P_{recision} = \frac{H}{S}, \tag{14}$$

where *H* is the number of corresponding notes of the intended score to notes of the estimated score and *S* is the number of notes of the estimated score.

The recall is shown as

$$R_{ecall} = \frac{H}{T}, \tag{15}$$

where *T* is the number of notes of the intended score.

A weighted coefficient is thus obtained based on the gradient method: the value of *w* is 1.01 and *v* is 1.04. We can then use these values to estimate a musical score.

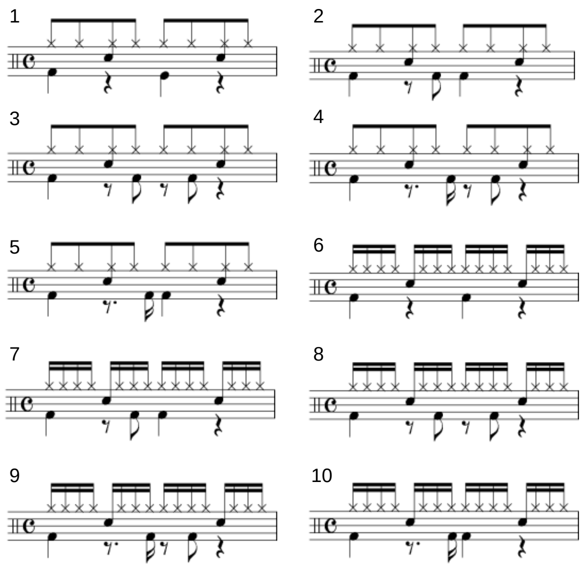


Figure 9. Drum patterns used in experiment

5.2 Experiment method

5.2.1 Experimental data

We used two recordings of ten different drum patterns as intended scores (shown in Fig. 9). One recording featured performances by both professional and amateur players, and the other featured simulated performance data with deviations and time shifting. The first recording included performances by one professional and three amateur players, while the second one included data generated by giving deviations and/or shifting the onset time of notes. A recorded performance without onset deviations is created from the drum patterns in Fig. 9 to obtain the simulated performance data. This data is obtained by adding onset deviation *s* (msec) to all notes as well as by adding a performance note to random onset deviation *r* (msec). In other words, the onset deviation from obtaining a normal random number of median *s* (msec) and standard deviation *r* (msec) is added to the recorded performances from the obtained drum patterns shown in Fig. 9, where *s* means the amount of shifting. We made seven types of simulated performance data, where *s* is 0, ±10, ±20 or ±30 and *r* is a constant value of 20 (msec). Using the simulated performance data enabled us to evaluate the proposed system's accuracy considering the effect of only onset deviations because no other deviation except for the onset deviations are contained in the simulated performance data.

5.2.2 Experimental condition

We estimated a musical score by using simulated and actual performance data under three different conditions.

Condition Q (Quantization):

A recorded performance is quantized under the accuracy of demisemiquaver.

Condition NP (Not using Prior Probability):

Although a recorded performance is used to estimate a musical score, a prior probability is not used.

Condition PP (using Prior Probability):

A smoothed occurrence frequency is applied to a recorded performance as a prior probability (as described in Section 3.4) to estimate a musical score.

5.3 Result and discussion

We estimated a musical score for both the simulated and actual performance data. The F-measure is calculated from the relationship between the estimated score and the intended score (shown in Eq. (13)) to determine the proposed method's accuracy.

5.3.1 Actual performance data

The F-measure is calculated from each player's performance. F-measure averages are shown in Table 1, where P is a professional player and A1, A2 and A3 are amateur drummers. These averages were obtained from each condition for all actual performance data. Condition Q gives an F-measure of .88, condition NP gives .91 and condition PP gives .93, indicating that condition PP is better than conditions Q or NP for estimating a musical score using the actual performance data.

Table 1. F-measure averages calculated from performance of each player

Subj.	Cond.Q	Cond.NP	Cond.PP
P	.85	.90	.92
A1	.97	.99	.99
A2	.94	.97	.98
A3	. 76	.79	.82
Avg.	**.88**	**.91**	**.93**

5.3.2 Simulated performance data

The F-measure is calculated from each shift range. F-measure averages are shown in Table 2. These averages were calculated from each condition for all types of simulated performance data, where condition Q gives an F-measure of .85, condition NP gives .93 and condition PP gives .94, indicating that condition PP is better than conditions Q or NP for estimating a musical score using the simulated performance data. The F-measure ratio increases most when the shift range is −30 (msec), which suggests the proposed method can handle larger onset deviations.

Table 2. F-measure averages calculated from performance of each shift range

s[msec]	Cond.Q	Cond.NP	Cond.PP
0	.93	.97	.99
−10	.94	.97	.98
10	.95	.98	.99
−20	.83	.93	.95
20	.87	.96	.96
−30	.65	.81	.84
30	.77	.87	.88
Avg.	**.85**	**.93**	**.94**

6. PROPOSED SYSTEM FOR PRACTICING DRUMS

6.1 Overview of proposed system

An overview of the proposed system for practicing drums is shown in Fig. 10. Drummers input only the tempo in which they want to practice, enabling them to play an intended score using MIDI drums in sixteen measures while listening to a metronome. Player performance is obtained as a recorded performance. When a drummer has finished playing, a musical score is estimated from the recorded performance by using the proposed system, and a proficiency score is then calculated based on the relationship between the estimated score and the recorded performance [2]. The proposed system also displays visually estimated musical and proficiency scores as well as an instantaneous condition such as deviation of onset or velocity for each note.

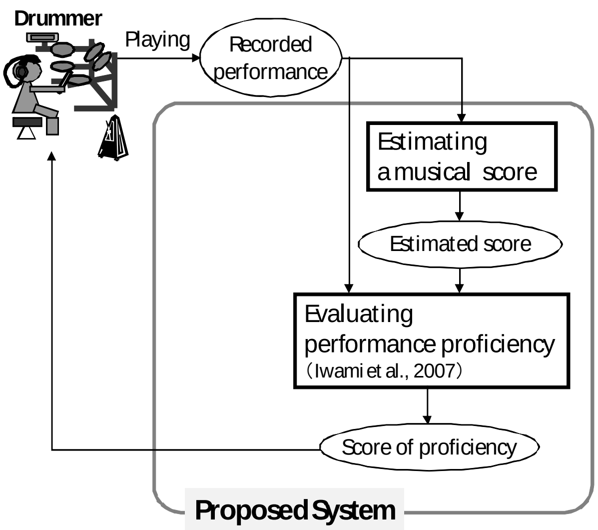


Figure 10. Flowchart of proposed system

6.2 Implementation of the proposed system

6.2.1 Setup of practice content

The main window of the proposed system is shown in Fig. 11. Users only have to input the desired tempo. Clicking the Start button initializes a metronome playback by a MIDI synthesizer. The user plays the MIDI drums in sixteen measures and the system records the user's performance as MIDI data.

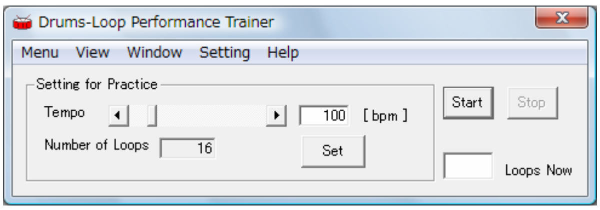


Figure 11. Main window of proposed system

6.2.2 Showing the estimation result

When the drummer has finished playing, the proposed method estimates the musical score from the recorded performance data and then displays the score in a pop-up window (Fig. 12) featuring cross marks (closed hi-hat cymbal), green circles (snare drum) and blue circles (bass drum).

A second window showing the proficiency score as well as advice for tackling weak points is also provided (Fig. 13). The proficiency score is between 0 and 100 and comments include “Great!”, ”Good”, “Needs Improvement” and so forth. Specific advice on the user’s performance is shown in the text box on the right.

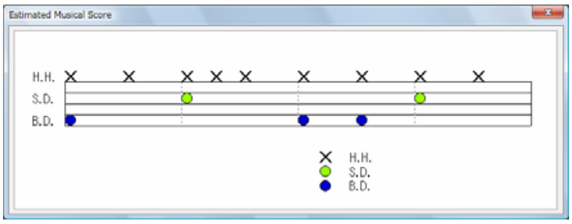


Figure 12. Window displaying estimated score

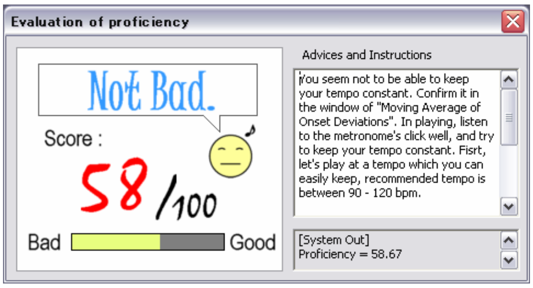


Figure 13. Window displaying proficiency score and advice

7. CONCLUSION AND FUTURE WORK

We developed a method that estimates the musical score of drum performances based on the Bayesian method. An estimated score is obtained from the drum pattern with the highest posterior probability. We described a method of smoothing the occurrence frequency so as to obtain a prior probability for drum patterns. Experimental results indicated that an F-measure of .93 was obtained when using the proposed system, indicating that it is superior to conventional methods. We also designed a support system for practicing the drums that estimates musical and proficiency scores based on the user’s performance.

In the future, we intend to analyze the trend of onset deviations in musical performance so as to the improve accuracy of estimating the musical score. We also intend to expand the method so that it is able to analyze the acoustic signals of drum patterns.

REFERENCES

[1] E. Schoonderwaldt, K. F. Hansen and A.Askenfelt, "IMUTUS – an interactive system for learning to play a musical instrument" Proc. of ICL, (2004).

[2] N. Iwami and M. Miura, "A support system for basic practice of playing the drums", Proc. of ICMC, pp. 364–367, (2007).

[3] S. Morita, N. Emura, M. Miura, S. Akinaga and M. Yanagida, "Evaluation of a scale performance on the piano using spline and regression models", Proc. of ISPS, pp. 77–82, (2009).

[4] Y. Murakami and M. Miura, "Automatic detection system for fill-in from drum patterns employed in popular music", Proc. of WESPAC X, CD-ROM, 0073, (2009).

[5] R. O. Duda, P. E. Hart and D. G. Stork, “Pattern classification”, pp. 164–174, Wiley, (2001).