Automatic classification of drum-rhythm patterns employed in popular music

Yuki MURAKAMI[*1] and Masanobu MIURA[*2]
[*] *Faculty of Science and Technology, Ryukoku University*
[*1]Yuki.MURAKAMI4290@gmail.com, [*2]miura@rins.ryukoku.ac.jp

ABSTRACT

The various styles of drum-rhythm patterns that are often employed in popular music can roughly be classified into "basic rhythm" and "fill-in" patterns. Compared to the basic rhythm pattern, the fill-in is an improvised pattern and is usually employed to fill in blanks in accompanying melodies. However, no criteria for objectively classifying these patterns have been reported to date. This study describes a classification method based on similarity in neighborhood. We propose a method of classifying drum-rhythm patterns by automatically adjusting the criteria based on the context of excerpts. The method deals with symbolic representations, as provided by MIDI data. Each sequence of one-measure-length contained in the given excerpt is evaluated as corresponding to either a "basic-rhythm" or "fill-in" pattern. The characteristics of sequences for drum-rhythm patterns are obtained by classifying them. In addition, we implemented the method as a visual application system that plays back MIDI data and indicates fill-in patterns.

I. INTRODUCTION

Drums are very popular musical instruments and they play a fundamental role in popular music. The rhythm patterns played on drums contribute greatly to determining the features of popular music in terms of rhythm. There have been some previous studies of drum-rhythm patterns. For example, Ellis and Arroyo, (2004) attempted to develop a system for automatic classification of musical genres, using the sequence of the beats in the drum-rhythm pattern. Since a characteristic of musical genres is that there may be multiple interpretations according to the context of an excerpt from the music, it is difficult to establish an appropriate interpretation for each individual excerpt. Here, we focus on the role of the drum-rhythm pattern. It is assumed that drum-rhythm patterns can be distinguished into "basic-rhythm" and "fill-in" patterns. The fill-in pattern is an improvised pattern compared to the basic-rhythm pattern, and is usually used to fill in blanks in the accompanying melody. In contrast, the basic-rhythm pattern repeats a similar and constant pattern. To the best of our knowledge, no criteria for objectively classifying these patterns have been reported to date. Therefore, this study proposes a classification method based on similarity in neighborhood. An outline of the classification procedure is shown in Fig. 1.

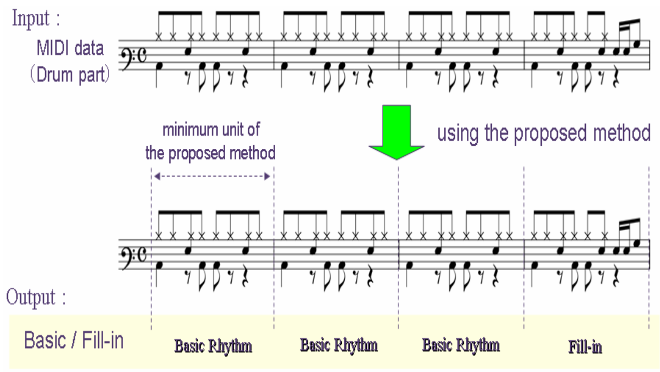

Figure 1. Outline of the classification procedure

II. SCOPE OF THIS STUDY

A. Description of drums

The term "drums" may refer to a "drum set" or "drum kit" made up of several percussion instruments, such as various types and sizes of drums and cymbals. Figure 2 shows a typical configuration of a set of drums that includes several percussion instruments. This configuration may be used independent of the musical genre or the player's level of proficiency. Figure 3 shows an extended configuration compared to that in Fig. 2. Usually, a wide variety of percussion instruments are used to play fill-in patterns' and the drum set shown in Fig. 2 does not cover all the instruments potentially used. Therefore, Fig. 3 shows the configuration of drums used in this study, in which several percussion instrument were added based on the tone name contained in the GM (General MIDI).

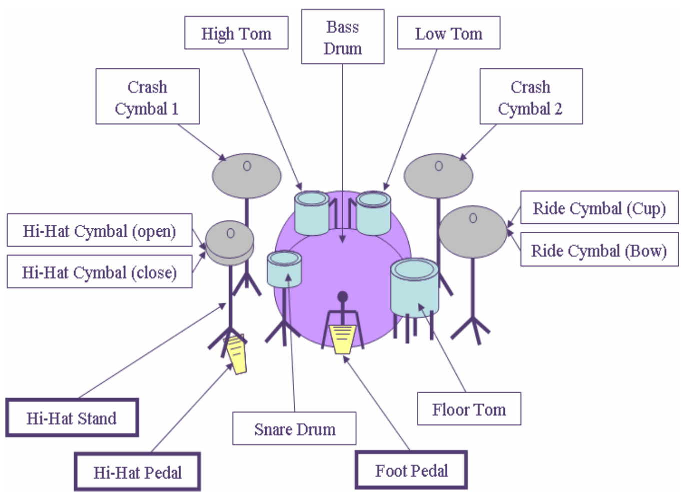

Figure 2. Popular configuration of the drums

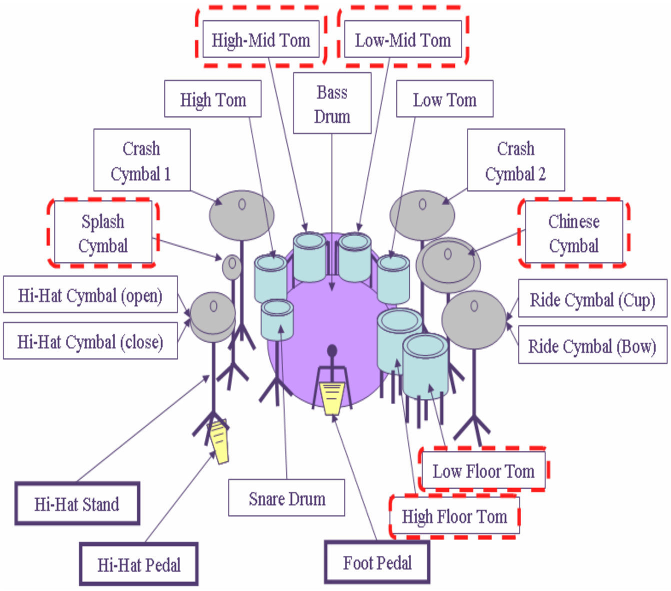

Figure 3. Extended configuration of the drums Configuration of extended drum set

B. Outline of MIDI data dealt with in this study

This study dealt with symbolic representations of drum-rhythm patterns, using MIDI data. Here, the XF format (Yamaha, 2000), which is an extended version of the SMF (Standard MIDI File) format, was used as the information included in the XF format was considered to be appropriate for our purpose.

Performance information for other musical instruments, such as piano, guitar, and so forth, included in the MIDI data were not dealt with. Specifically, only a subset of the performance data, in terms of the program number and onset time, was dealt with because the drum-rhythm patterns employed in music can be constructed by using only these data. The only objective beat used in this study was 4/4. One measure length was employed as the basic unit of a drum-rhythm pattern.

C. Concept of the representative note number

Percussion instruments are sometimes assigned as other instruments in MIDI data. For example, the bass drum is assigned as "Acoustic Bass Drum" and "Bass Drum 1". It is thought that these types of bass drums share a common role except for small differences in acoustics. Therefore, in this study, percussion instruments that have the same note number were grouped into a "representative note number". Here, the note numbers of the percussion instruments included are called the "included note number". The relations between them are shown in Table 1. The relationship between the note number and the percussion name is the same as those in the XG format (Yamaha, 2000). The differences in program number, based on the differences in the "Bank" in the XG format, are assumed to be the same as in the note number defined in GM. For example, terms such as "Kick Short" and "Kick Dry Mute" are considered to be the same as "Bass drum". Sixteen types of note numbers, as listed in Fig. 3 and Table 1, were dealt with in this study.

Table 1. Correspondence relation between "representative note number" and "included note number" for individual percussion instruments.

Subgroup	Instrument	Representative note number	Included note number
(Cymbal) Cymbal group	Hi-Hat Cymbal (Open)	46	
	Hi-Hat Cymbal (Closed)	42	
	Crash Cymbal 2	57	
	Crash Cymbal 1	49	
	Ride Cymbal (Cup)	53	
	Ride Cymbal (Bow)	51	51,59
	Splash Cymbal	55	
	Chinese Cymbal	52	
(Tom-Tom) Tom group	High Tom	50	
	Hi-Mid Tom	48	
	Low-Mid Tom	47	
	Low Tom	45	
	High Floor Tom	43	
	Low Floor Tom	41	
(S.D and B.D) Basic group	Snare Drum	38	31,38,40
	Bass Drum	36	33,35,36

D. Subgroups of drums

In this study, the drum-rhythm pattern was modeled as a combination of three subgroups, as shown in Fig. 4. The "basic" group comprised a bass drum and snare drum, and the "tom" and "cymbal" groups comprised tom-toms and cymbals, respectively. Details of the percussion instruments assigned to each of the three subgroups are shown in Table 1. The occurrence frequencies of the bass drum and snare drum were usually higher compared to those for the tom group and/or cymbal group in drum-rhythm patterns, so the basic group was assumed to play an important role in determining the features of the drum-rhythm pattern. The tom group, which comprised six types of toms, was found to have an important role in playing "fill-in" patterns. The cymbal group, which comprised eight types of cymbals, was assumed to be important in both basic and fill-in patterns.

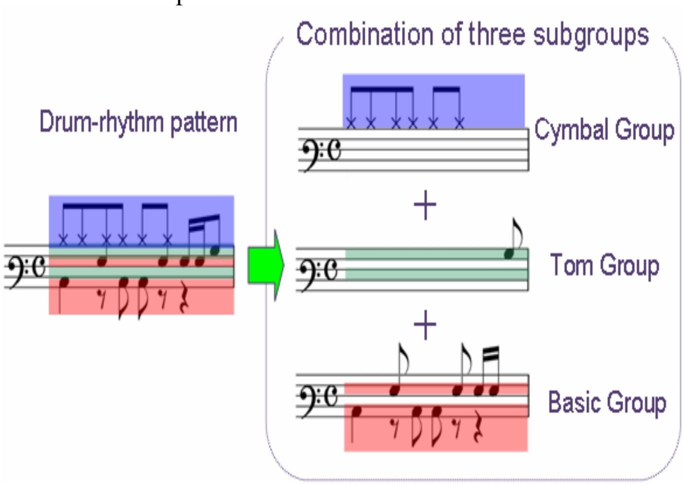

Figure 4. Outline of three subgroups

III. PREPROCESSING FOR PROPOSED METHOD

In this study, one measure length of a 4/4 beat was the minimum unit in terms of length. Other types of rhythm patterns, such as those excluding 4/4 beats and irregular meters, were not dealt with. All the music excerpts used in the study were analyzed to extract all the different rhythmic patterns. After finding an original rhythm pattern, the onset timing had to be quantized according to a common level for all patterns. The delta time for quarter notes in MIDI data varies between MIDI files in XF format, such as 96, 384, 480, and so forth. To collect the drum-rhythm patterns for all music in our database, it was necessary to use delta time as a unifying factor. In this study, a common delta time of 240 was used and care was taken to avoid overlaps between the drum-rhythm patterns that were collected. As a result, we collected the drum-rhythm patterns used in 5234 music excerpts in our database. Users can look for a specific pattern in all of these drum-rhythm patterns by specifying the drum ID.

IV. CATEGORISING DRUM-RHYTHM PATTERNS

Here, we introduce a parameter m_i, which means the amount of similarity between two drum-rhythm patterns including the *i*-th measure, obtained by calculating the score for matching the *i*-th measure to the other measures included in the *n*-length sequence of a drum-rhythm pattern. The suffix *i* in m_i means the ID of the measure in the current music excerpt. If the note number and onset time are exactly the same between subgroups contained in two current drum-rhythm patterns, the score is added and otherwise the score is not added. In other words, when the current subgroups are the same, the score is 1, and otherwise the score is 0. Here, a parameter P_i is defined as the combination of the subgroups S_{i1}, S_{i2}, S_{i3}., shown as;

$$P_i = S_{i1} + S_{i2} + S_{i3} \qquad (i = 1,2,\cdots,n) \quad (1)$$

An example of a subgroup is shown in Fig. 5. To measure the similarity between two rhythm patterns, a function $F_{match}(\ )$ is defined as follows;

$$F_{match}(P_j, P_k) = \sum_{l=1}^{3} \begin{cases} 1 & (S_{jl} = S_{kl}) \\ 0 & (S_{jl} \neq S_{kl}) \end{cases} \quad (2)$$

In case of $j = k$, $F_{match}(P_j, P_k) = 0$

This calculation is done for all measures of *n*-length, in which comparisons are made between the *i*-th measure and *k*-th measure $(i - n + 1 \le k \le i + n - 1)$. The number of comparisons is *n*(*n*-1) in total. The average thus obtained is equal to m_i. m_i is given by equation (3). An example of the procedure used for matching the *i*-th measure to the other three types of drum-rhythm patterns is shown in Fig. 6. Therefore, parameter m_i is defined as follows:

$$m_i = \frac{1}{n(n-1)} \sum_{j=i-n+1}^{i} \sum_{k=j}^{j+n-1} F_{match}(P_i, P_k) \quad (3)$$

The average of m_k in an *n*-length sequence of a drum-rhythm pattern is called threshold t_{ik}, in which *j* means the beginning ID of the *n*-length sequence. Threshold t_{jk} is shown as follows:

$$t_{jk} = \frac{1}{n} \sum_{k=j}^{j+n-1} m_k \quad (4)$$

If the m_i is less than t_{jk}, the *i*-th measure is evaluated as a possible fill-in pattern. After this procedure, a constant value (1, in this case) is added to the amount of the fill-in likelihood c_i.

$$c_i = \begin{cases} +1 & (if \quad m_i < t_{jk}) \\ +0 & (if \quad m_i \ge t_{jk}) \end{cases} \quad (5)$$

After calculating c_i for all *i*, if c_i is less than or equal to threshold *T*, the *i*-th measure is evaluated as a fill-in pattern. If c_i is more than *T*, the *i*-th measure is evaluated as a basic pattern. Here, *T* is set as equal to *n*, the length of the sequence observed. The core function of judging $F_{class}(\ )$ is represented as follows:

$$F_{class}(C_i) = \begin{cases} fill & (c_i = T) \\ not \quad fill & (c_i < T) \end{cases} \quad (c_i = 0,1,2,\cdots,M), \quad (6)$$

where *M* means the length of the current music excerpt. The procedure of listing all the possible patterns in the *n*-length sequence is continuously done from the beginning to the end of the current music excerpt. In the case of the beginning and/or ending of an excerpt, the comparison method does not perform as well, so the length of a sequence is determined to be shorter than it actually is. For example, *n* is equal to four in calculating m_i for the 1st, 2nd, and 3rd measures. This situation is called Ex-1. Another exception procedure is used in calculating the beginning and/or ending part of an excerpt, where the calculation is done by comparing only two patterns. In this case, m_i is less than or equal to 1, and both patterns are evaluated as possible fill-in patterns. However, the fill-in patterns shown for conjunctive two measures are quite rare, so the standards for evaluation are set as follows (Ex-2):

$$c_i = \begin{cases} +1 & (m_i \le 1) \\ +0 & (m_i > 1) \end{cases} \quad (i = 1,2 \quad \text{or} \quad M-1, M) \quad (7)$$

Moreover, when evaluating beginning and ending areas, the number of comparisons between two measures is less than the ordinal comparisons, so a constant value (1, in this case) is added to c_i (Ex-3). When performing all the procedures mentioned above, c_i for all *i* is commonly evaluated in comparison to the size of *T*. An outline of the procedure used for dealing with these exception patterns (Ex-1, Ex-2, and Ex-3) is shown in Fig.7.

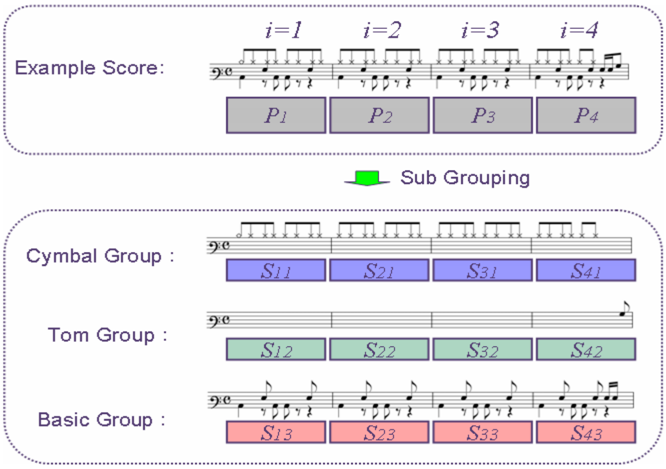

Figure 5. Example of subgroup

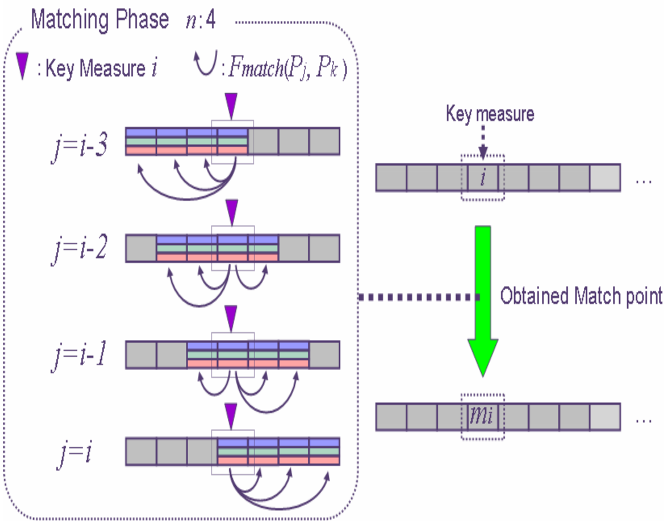

Figure 6. Example of matching procedure

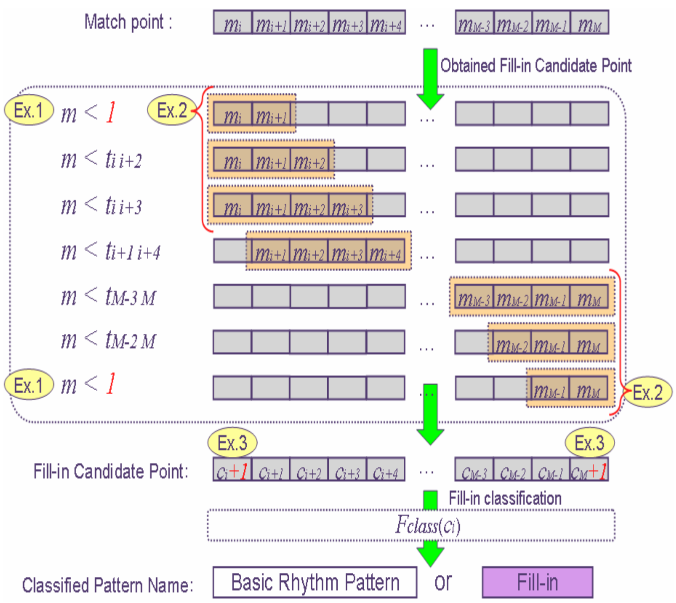

Figure 7. Example of procedure used for evaluating fill-in patterns compared with that used for exception patterns (Ex-1, Ex-2, and Ex-3)

V. CONSTRUCTION OF AN APPLICATION PROGRAM AND EXPERIMENTAL EVALUATION

A. Constructing an application

The proposed method was implemented as a visual application system. This system visualizes the result of the proposed method for each measure, for a given music excerpt. This visualization enables users to identify the result of the proposed method by listening and seeing. An example of a screen capture of the application system is shown in Fig. 8.

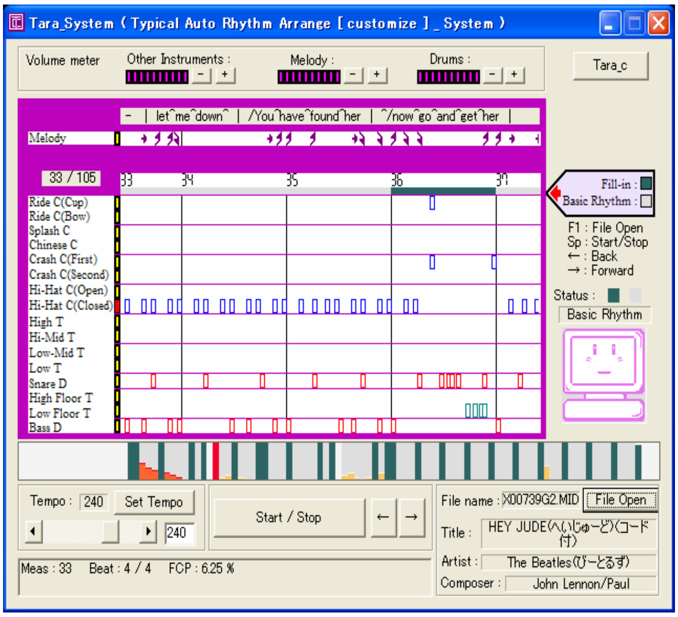

Figure 8. Example of visualization of results of application

B. Experiment evaluation

Ten subjects were asked to evaluate fill-in or basic-rhythm patterns for each sequence of one-measure-length contained in ten excerpts of Japanese popular songs. Five of the subjects were amateur drummers (D1-D5). Their experience ranged from three to eight years, with an average of approximately four years. Another five subjects were non-drummers (N1-N5). An instructive movie was shown to all the subjects to explain the nature of fill-in patterns (the explanation was not specific and a variety of examples were presented). All the measures contained in the excerpts had to be categorised as either basic or fill-in patterns. The excerpts came from ten Japanese pop songs, and were extracted from 3779 excerpts in total. Subjects were allowed to listen to the excerpts again and again, and were also able to change their judgments after further listening. Subjects were not allowed to discuss their categorisations during the experiments. among them about it during the experience.

To evaluate the performance of the proposed method, recall, precision, and the F-measure were used. The number of fill-in patterns evaluated by both the proposed system and all subjects is represented as H, and the number evaluated by only the proposed method is represented as *S*. So, the precision is represented as follows:

$$P = \frac{H}{S} \quad (8)$$

Here, the number of cases evaluated as fill-in by only one subject is represented as *T*, and R is represented as follows:

$$R = \frac{H}{T} \quad (9)$$

Therefore, the F-measure is represented as follows:

$$F\text{-}measure = \frac{2PR}{P+R} \quad (10)$$

Figure 9 shows the F-measure for each subject, in which the correct answer means the output of exact agreement between subjects D1 to D5. Here, 50R means 50% judgment with irrespective of the drum-rhythm and 25R indicates 25% were judged as fill-in. As a result, the proposed method yielded an F-measure of 76.0%, which was inferior to the scores for the amateur drummers, which ranged from 72.4 to 94.6%. However, these scores were almost equivalent to those of the non-drummers, which ranged from 69.7 to 86.7%. Therefore, the new method was found to accurately classify drum-rhythm patterns based on the musical context.

VI. CONCLUSION

This paper proposed a method of automatically finding fill-in patterns by analyzing performance information for drums from music data based on the MIDI format. The results of an experimental evaluation showed that the proposed method was effective for automatic classification of drum-rhythm patterns. The proposed method was implemented as a visual application, based on the results of a questionnaire given to users. The results suggested that this was an effective method of presenting fill-in patterns to users, in implementing the proposed system, because users did not need to have any special knowledge in this field. In future work, we will consider the features of music excerpts based on different musical genres. Subjective comments on our subsystem indicated that it is useful for confirming types of drum-rhythm patterns among several subjects, so this system will be extended to provide a training system for playing the drums, and also to provide a new interface for playing back MIDI files. A database was constructed during this study, and using the data collected we will work on developing an automatic system of producing drum-rhythm patterns for given melodies.

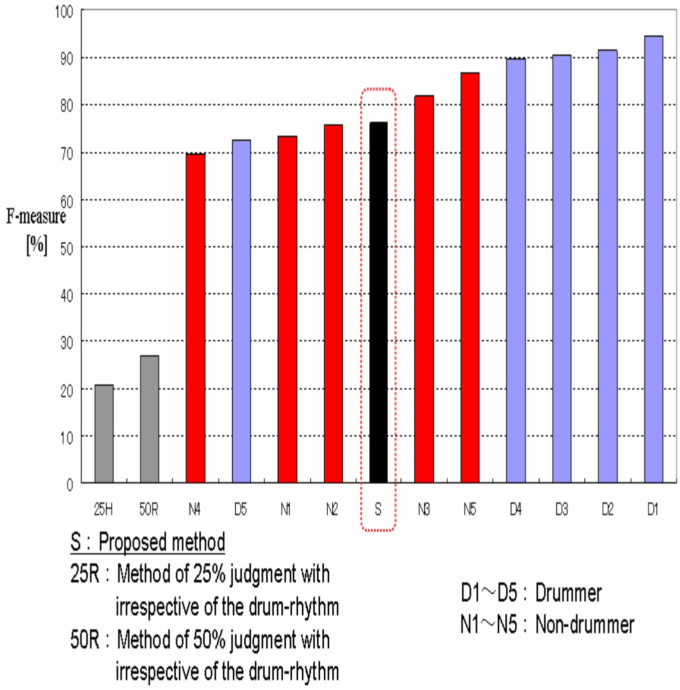

Figure 9. F-measure for evaluation of patterns by subjects and by proposed method

ACKNOWLEDGMENT

This study is partly supported by the "High-Tech Research Center" Project for Private Universities: matching fund subsidy from MEXT (Ministry of Education, Culture, Sprots, Science and Technology), 2006-2011. Grant-in-Aid for Scientific Research (195373).

REFERENCES

Daniel P.W. Ellis and John Arroyo (2004). Eigenrhythms: Drum pattern basis sets for classification and generation. Proceeding of ISMIR, pp.554-559.

YAMAHA. (2000). *XF format specification*, retrieved September 24, 2007, from http://www.yamaha.co.jp/product/syndtm/read/aoyama/pdf/xfspc.pdf

YAMAHA. (2000). *XG format specification*, retrieved October 9, 2007, from http://www.yamaha.co.jp/product/syndtm/read/aoyama/xgspec/pdf/xg_v135_j.pdf

N. Iwami and M. Miura (2006). Criteria for evaluating player's proficiencies in drum-loop performances. Proceedings of the Technical Meeting of the Musical Acoustical Society of Japan, MA2006-44.