Melody Generation System based on a Theory of Melody Sequence

Sakurako Yazawa
Graduate School of
Systems and Information Engineering,
University of Tsukuba,
Tsukuba, Japan
sakurako@music.iit.tsukuba.ac.jp

Masatoshi Hamanaka
Department of Clinical System Onco-Informatics
Kyoto University, Kyoto, Japan

Takehito Utsuro
Faculty of Engineering, Information and Systems,
University of Tsukuba, Tsukuba, Japan

***Abstract*—We propose a melody generation system based on the Implication-Realization Model (IRM) of music theory. The IRM is a music theory, which was proposed by Eugene Narmour. The IRM abstracts music. It then expresses music according to symbol sequences based on information constituting the music pitch, rhythm, and rests. Previous melody generation systems are mostly based on tone transition models, which do not have function of abstracting melodies observed in training data. In those previous systems, generated melodies do not reflect tone sequences that do not exist in training data. However, it is obviously required that a melody generation system is able to abstract melodies in training data and to output certain melody, which is rarely observed in the training data. Our melody generation approach properly abstracts melodies in training data based on the IRM. The IRM expresses contexts of melodies using symbol sequences. Our melody generation system consists of two models; that of symbol sequence transition and that of generating tones from symbols. With the former model, the symbol transition probability model is trained with the results of the IRM analysis. The system then generates an optimal symbol sequence according to the probability model. Then, from a set of tones, each symbol sequence generates a melody. We evaluated the proposed system through subjective human judgment and the results showed that our system properly generated melodies.**

Keywords music theory, melody generation, Implication-Realization Model (IRM)

I. INTRODUCTION

The goal of this paper is to develop an automatic melody generation system based on the Implication-Realization Model (IRM) of music theory[1],[2]. In the IRM, a symbol sequence can be expressed according to music components such as rests, pitch, and rhythm. One of fundamental assumptions on which the IRM is based on those listeners of music consciously or unconsciously predicts next melody when listening to music.

Previous approaches to melody generation can be summarized as below: One approach is to learn the probability model of the transition of pitch and then the next coming pitch is predicted following the melody input given by a user [3]. Another approach is to generate based on evolutionary computation [4]. In those previous approaches, melody generation systems do not have function of abstracting melodies observed in training data. Thus, they are not able to output certain melody, which is rarely observed in the training data.

Our melody generation approach, on the other hand, is based on the IRM and properly abstracts melodies in training data. It analyses and classifies the melody in this abstraction process [5],[6],[7]. We first developed an automatic melody analyser based on the IRM in our proposed melody generation system. The IRM can analyse atonal and tonal music. The "Generative Theory of Tonal Music (GTTM)"[15] analyser is an automatic melody analyser based on a music theory [8]. It can analyse tonal music but not atonal music. Since our automatic melody analyser is based on the IRM, it can analyse tonal and atonal music. We also developed a melody generation method based on the IRM. In the IRM, it is assumed that listeners of music consciously or unconsciously predict the next melody when listening to music, which means that the IRM expresses the melody context. Our melody generation method can be regarded as a combination of a probabilistic model and a music theory.

Our melody generation system consists of two models; that of symbol sequence transition and that of generating tones from symbols. With the former model, the symbol transition probability model is trained with the results of the IRM analysis. The system then generates an optimal symbol sequence according to the probability model. Then, from a set of tones, each symbol sequence generates a melody.

II. RELATED WORK

This section overviews previous approaches to the melody generation technique. As we introduced in the previous section, previous approaches [3],[4] to melody generation systems do not have function of abstracting melodies observed in training data. One issue when designing a melody generation system is to what extent a user interacts with the melody generation system. One extreme approach is the case where the melody generation system generates the whole melody fully automatically. Another approach is the case where a user provides the system with an existing melody, while the system generates a new part such as an arrangement and an accompaniment or adds acoustic effect to the original melody [8],[9],[10]. In those previous approaches, the users can interact with the melody generation system and the system can generate a new melody as requested by the users.

Other related work include an automatic composition system Orpheus[11] which generates a melody considering rhythm within lyrics provided by a user and a melody generation system which determines specific tones based on a Bayesian network model given a chord [12]. Another related work on automatic composition are those in terms of assisting a user when performing a musical instrument [13], [14].

Compared with those previous approaches, our approach differs in that it is strictly based on a music theory, and it generates melodies fully automatically.

III. MUSIC THEORY "IMPLICATION-REALIZATION MODEL"

The IRM is a music theory, proposed by Eugene Narmour. The IRM abstracts music. It then expresses music according to symbol sequences based on information constituting the music pitch, rhythm, and rests (Figure 1).

When analysing melody using the IRM, we have the following two steps. The first step is to enclose the tones successively with a bracket. The bracket is an important structure when abstracting melodies. In the procedure of bracket abstraction, first, a large note column group is created in order to detect the location where the bracket is interrupted. A bracket containing three successive tones is then formed from the beginning to the end of the group. A set of three tones can not form a bracket, if there are only one or two tones. In such a case, we re-structure the tones sequence and then form a bracket. The second step of analysing a melody using the IRM is to assign a symbol to each bracket. Tones enclosed in brackets are assigned a symbol and are called "basic structures".

There are two important points in assigning symbols. The first point is the pitch of the current two to three consecutive notes. The second point is the interval direction.

There are ten basic structures in the IRM; eight types of symbols include three notes in a bracket (Figure 2), one "dyad" includes two notes in a bracket, and one "monad" includes one note in a bracket (Figure 3).

For example, "IP" in Figure 1 includes three tones in a bracket and is assigned the feature of "down sound up a narrow pitch". The tonal row with tones enclosed in various brackets is analysed based on the basic structures.

We developed our melody generation system by using these analysis results.

IV. SYSTEM OVERVIEW

Figure 4 shows the overview of the proposed melody generation system. Our system analyzes input melody by using the IRM. The system builds a symbol transition probability model from the results of the IRM analysis. The system then generates several output symbol sequences based on the probability model. Finally, the system generates a melody by assigning a symbol to each tone set.

The system consists of two models, melody analyzer and melody generator. The melody analyzer analyzes the input melody using the IRM for basic structures. Each basic structure has information; tones value, intervals and so on. The melody generator consists of two modules; probabilistic model training and symbol sequence generation.

The probabilistic model training module estimates the symbol transition probability $P(S_{i+1}|S_i)$ of two consecutive symbols S_i and S_{i+1} from an input melody. Then, when generating a melody, the probabilistic model is first provided with a randomly generated starting symbol S_0. It then generates a sequence $S_0,\ldots,S_8$ of nine symbols, where, following the preceding symbol S_i, as the subsequent symbol, S_{i+1} which simply maximizes the transition probability $P(S_{i+1}|S_i)$ is selected.

Next, the symbol sequence generation module generates a combination of the tones of one symbol. The combination of tones is generated based on the information of each symbol type. The symbol sequence generation module adds the tone information to each symbol in the symbol sequence. Here, we examine two approaches to adding the tone information to each symbol. In the first approach, for each symbol type, average of tones observed in the input melody is used and is referred to as "ave" in the following sections on experimental evaluation. In the second approach, on the other hand, for each symbol type, a tone which is randomly selected from the input melody is used and is referred to as "ran" in the following sections on experimental evaluation.

V. EXPERIMENTS

We conducted preliminary experiments to evaluate the proposed system. We randomly selected two songs from the Essen Folksong database [16] as the training songs for the evaluation.

We had once, 15 times, 60 times, 100 times as the number of training. In this section, we describe the following two types of evaluation: (i) whether the generated melody can be analyzed according to the IRM, and (ii) whether the elements of the symbols included in the generated melody are the same as that of the training melody.

As for the first type of evaluation, we confirmed that the generated melodies can be analyzed according to the IRM. We also confirmed that the generated melody better reflects the tone sequences of the original song as the number of training increases. We confirmed that elements contained in the generated melody is the same as that of the training song. So, the system generated a melody which has elements of the training song.

Based on the results of preliminary experiments, we further evaluated the proposed system through subjective human judgment. The system generated a total of twelve melodies; four melodies from one training song. Two of the four melodies were generated using the average value of the tonal information contained in the symbol of the training song and the other two melodies were generated based on the occurrence probability of the symbols in the training song. For each of the three training songs, we generated ten melodies. Out of the ten melodies, we selected those ending with consonant sound but not with discordance, which amount to about 50%. Finally, we randomly selected four melodies ending with consonant sound to be used in the subjective human judgment.

Fifteen subjects participated in subjective human judgments evaluation. Each subject listened to all queries and song pairs in random order without duplication. Every time he/she listened to them, he/she was asked as "how similar was the query to the training song?" and was requested to rank with a 5-point scale; 1) very similar, 2) similar, 3) neutral, 4) different, and 5) very different.

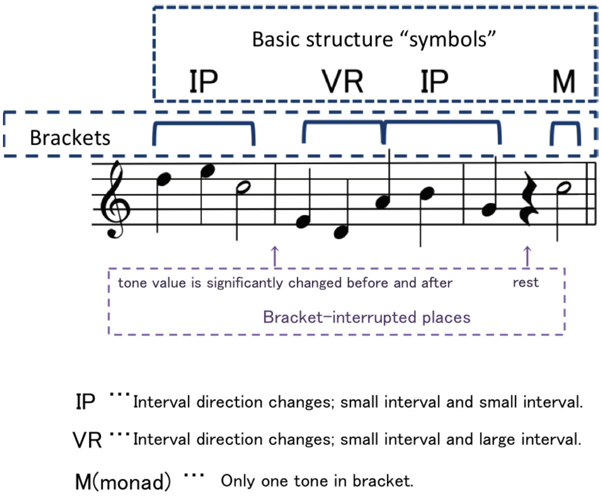

Figure 1. An Example of Analysing Melodies according to IRM

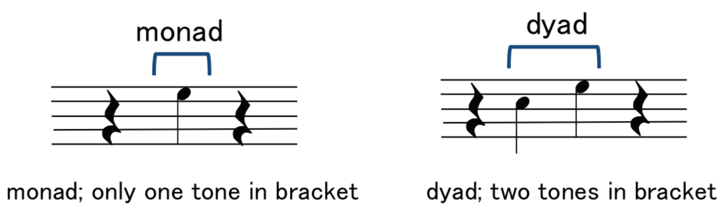

Figure 3. "dyad" and "monad"

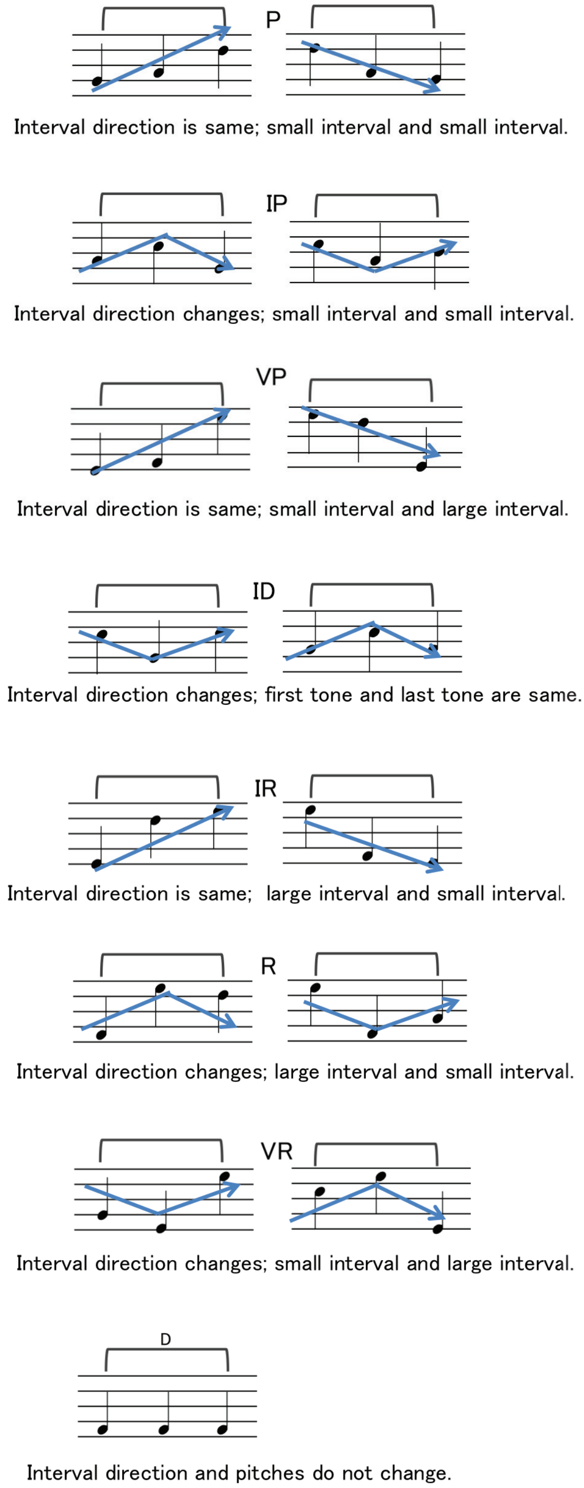

Figure 2. Basic structure symbols

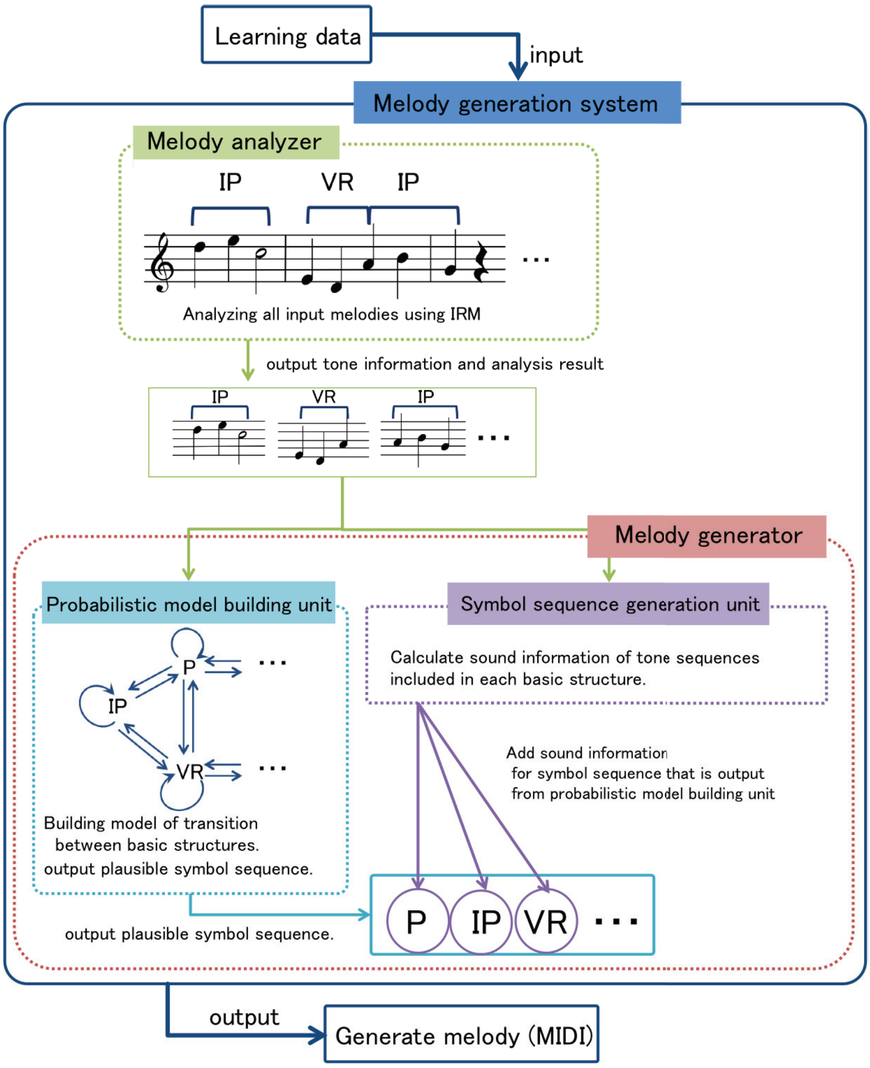

Figure 4. Processing flow of proposed melody generation system

VI. RESULT AND DISCUSSION

Figures 5, 6, 7 show evaluation results for the training songs No.1, No.2, and No.3, respectively. Figure 8 shows average of the evaluation results of the training songs No.1, No.2, and No.3. Experimental results showed that the melodies generated by the proposed system reflected the elements of the training song. The results showed that the generated melodies more reflected the melody of the training data when stochastically and randomly using note elements included in the training data (as shown with the scores of "ran" in the Figures 5, 6, 7, and 8) than when using the average of the note elements of the training data (as shown with the scores of "ave" in the Figures 5, 6, 7, and 8). This is mainly because the average of the note elements of the training data tend to remove characteristic melodies within the training data. On the other hand, when stochastically and randomly using note elements included in the training data, it is much more expected that the generated melody includes characteristic melodies that are observed within the training data.

Another impressive result is that when comparing the minor and major training songs, the generated melody tend to include characteristic melodies within the minor training song (the training song No.3 with the evaluation result shown in Figure 7) than those within the major training song (the training songs No.1 and No.2 with the evaluation results shown in Figures 5 and 6). We estimate that this is probably because the minor songs tend to have more characteristic melodies than the major songs.

VII. CONCLUSION

Our melody generation approach properly abstracts melodies in training data based on the IRM. The IRM expresses contexts of melodies using symbol sequences. Our melody generation system consists of two models; that of symbol sequence transition and that of generating tones from symbols. With the former model, the symbol transition probability model is trained with the results of the IRM analysis. The system then generates an optimal symbol sequence according to the probability model. Then, from a set of tones, each symbol sequence generates a melody. We evaluated the proposed system through subjective human judgment and the results showed that our system properly generated melodies. We plan to apply the proposed system to the task of training with more than one training songs.

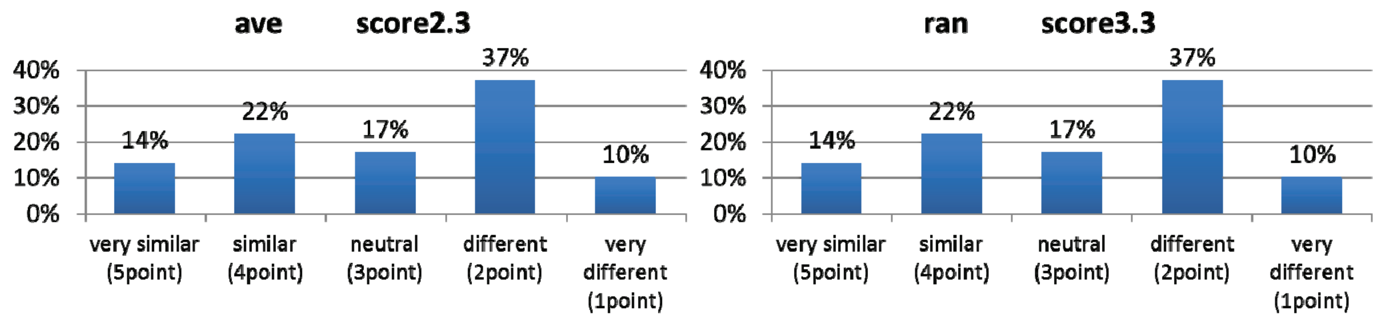

Figure 5 Experimental results (for the training song No.1)

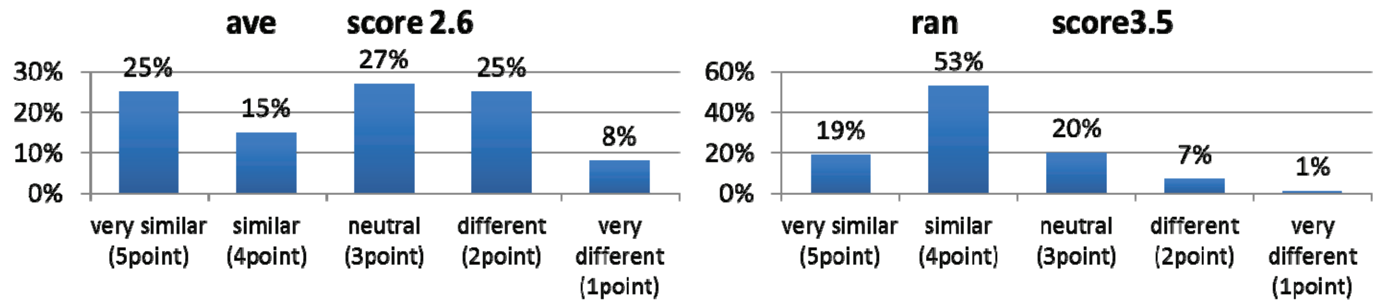

Figure 6 Experimental results (for the training song No.2)

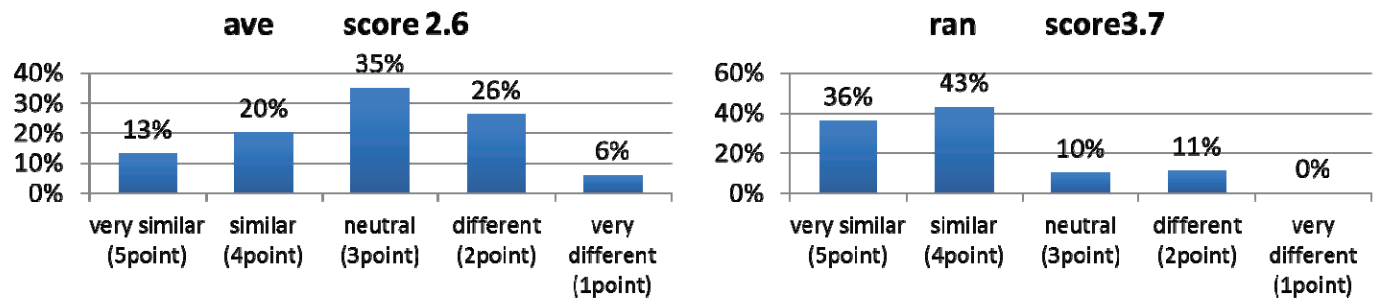

Figure 7. Experimental results (for the training song No.3)

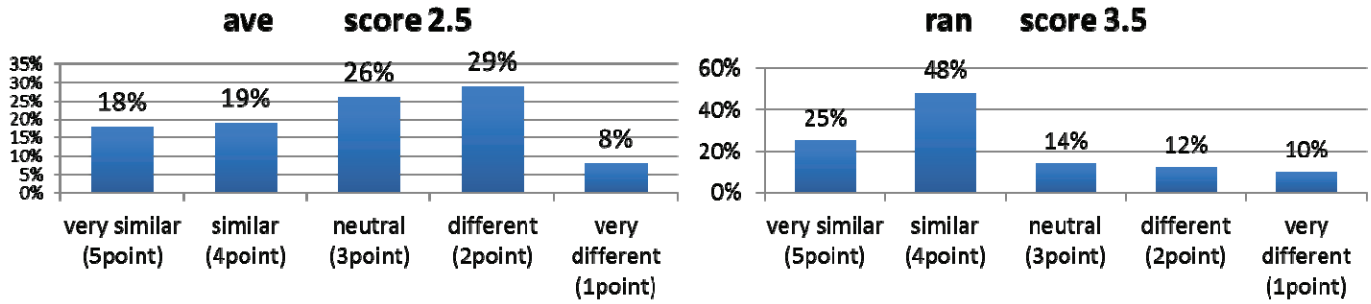

Figure 8. Experimental results (average of the training songs No.1, No.2, and No.3)

REFERENCES

[1] Eugene Narmour, "The Analysis and Cognition of Basic Melodic Structures", The University of Chicago Press, 1990.

[2] Eugene Narmour, "The Analysis and Cognition of Melodic Complexity", The University of Chicago Press, 1992.

[3] Ames, C., "The Markov Process as a Compositional Model: A Survey and Tutorial", Leonardo, Vol. 22, No. 2 (1989), pp. 175-187.

[4] Todd, P. M., Werner, G. M.: "Musical Networks", MIT Press, 1999, pp. 313-339.

[5] Sakurako Yazawa, Yuhei Hasegawa, Kohei Kanamori, Masatohsi Hamanaka: "Melodic Similarity based on Extension Implication-Realization Model", MIREX Symbolic Melodic Similarity Results, 2013

[6] Sakurako Yazawa, Hiroko Terasawa, Keiji Hirata, Satoshi Tojo, Masatoshi Hamanaka," Identifying the basic structures of Implication-Realization Model in a melody ", IPSJ special interest group on music and computer, 2010-MUS-87, No.1, Oct. 2010 (in Japanese)

[7] Sakurako Yazawa, Hiroko Terasawa, Keiji Hirata, Satoshi Tojo, Masatoshi Hamanaka," Analysis the chain structures of Implication-Realization Model in a melody ", IPSJ special interest group on music and computer, 2012-MUS94SLP90, No.34, Feb. 2012 (in Japanese).

[8] PGMusic: Band in a Box, http://www.pgmusic.com/index.html.

[9] UJAM: http://www.ujam.com/.

[10] Musicshake: http://eng.musicshake.com/about.

[11] Shigeki Sagayama, Shinji Sakou, Gen Hori, Satoru Fukayama: Automatic Song Composition from Japanese Lyrics Based on a Probabilistic Approach, Institute of Systems, Control and Information Engineers, Vol.56, No.5, pp.219-225 (2012) (in Japanese).

[12] Makiko Katsura, Tetsuro Kitahara, Haruhiro Katayose, Noriko Nagata: A Chord Voicing Reasoning System with Bayesian Network, IPSJ special interest group on music and computer, 2008-MUS-74, Vol.2008, No.29, pp.163-168 (2008) (in Japanese).

[13] Keiji Hirata, Tatsuya Aoyagi: Pa-Pi-Pun : Creativity-support Tool for Generating Jazz Harmony, Information Processing Society of Japan , Vol. 42, No. 3, pp. 633-641 (2001) (in Japanese).

[14] Benjamin Luke Evans，Nagisa Munekata，Tetsuo Ono: "CMY": an Automated Music Composing System based on the Harmonic Theory of Classical Music ,IPSJ special interest group on music and computer 2013(in Japanese)．

[15] Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo, "ATTA: Automatic Time-sapan Tree Analyzer based on Extened GTTM", Proc. ISMIR2005, pp. 358-395.

[16] [Essen Folksong database] " http://www.esac-data.org/"