Automatic composition techniques are important in sense of upgrading musical applications for amateur musicians such as composition support systems. In this paper, we present an algorithm that can automatically generate songs from Japanese lyrics. The algorithm is designed by considering composition as an optimal-solution search problem under constraints given by the prosody of the lyrics. To verify the algorithm, we launched Orpheus which composes with the visitor’s lyrics on the web-site, and 56,000 songs were produced within a year. Evaluation results on the generated songs are also reported, indicating that Orpheus can help users to compose their own original Japanese songs.

1. INTRODUCTION

Recently, there has been wide interest in automatic composition algorithms which can help amateur musicians to compose original tunes. Although considerable amount of research has been done on automatic composition[1][2][3], much less has been done on composing songs from the lyrics, and the question of what information in the lyrics should be exploited for generating songs remains. Syntactic information of the lyrics are used in some researches[4] to compose a song from the lyrics. Musicologists argue that there are considerable correlations between music and prosody. In case of composing songs, prosody plays a more important role. However, no system that uses prosody has yet been attempted.

2. MELODY COMPOSITION ALGORITHM EXPLOITING PROSODY OF JAPANESE LYRICS

Our objective is to develop a method to generate a melody automatically which satisfies the constraints given by the prosody of Japanese lyrics. We define melody composition here as generating a melody given the lyrics, patterns of rhythms (by “rhythm tree” which is described later in this section), and harmony sequence with specifications of tonality and scale.

2.1 Japanese Prosody and its Role in Composition

Japanese is said to “have a fixed shape consisting of a sharp decline around the accented syllable, a decline that is usually analysed as a drop from a H¹ tone to a L²”[5]. Furthermore, “the place of the accent is lexically contrastive, as in ka’mi ‘god’ vs. kami’ ‘paper’”[5]. (Fig. 1)

A melody attached to the lyrics cause an effect similar to the accent. Therefore we can assume that the prosody of Japanese lyrics imposes constrains on pitch motions of the melody.

2.2 Composition of Rhythm

In order to design a rhythm on the given lyrics, two problems have to be considered. The first problem is the allocation of lyrics on the melody, and the other problem is how to handle the unity of rhythm in the same song.
In order to solve the first problem, we assumed that melody consists of segments and the lyrics should be divided into these segments. For instance, 2 bars can be handled as a segment for a song with a length of 8 bars. This is because the lyrics input by the user are not always “formatted” as poem like the usual lyrics are.

Furthermore, in most classical Japanese songs, one syllable (mora) corresponds to one note in a melody. Thus the number of notes in each segment is determined by allocating the syllables. When we consider the constraints on dividing the lyrics, the following 3 criteria can be assumed:

- A similar number of syllables in each segments is preferred.
- The border of the segments should not be crossed over within a word.
- Too short lyrics should be iterated prior to allocation.

Under these constraints, we can solve the syllable allocation problem by using dynamic programming.

2.2.2 Keeping Unity of Rhythm in Melody

Even though the numbers of notes in each of the segments are decided, there still exist a large degree of freedom in rhythm. One possible way to put constraints on rhythm is to arrange that the generated rhythm belongs to a same “family” of rhythm. It is reasonable to assume that the “rhythm family” does not change in a relatively short song.

Here the second problem, that is, the requirement to keep the unity of rhythm in melody arises. To cope with this problem, we introduce a “rhythm tree” that is one rhythm has a similar feature when one can be derived by uniting or dividing the note just one on the another. In practice, a tree structured template of rhythm as shown in Fig. 2 are prepared by hand beforehand. Since the number of syllables in each segment corresponds to the number of notes, this tree structured template determine the rhythm in each segment.(Fig. 2)
and accompaniment instruments. Flow chart of the processes is shown in Fig. 3. We used Galatea-Talk[7] text-to-speech engine to analyze the prosody of Japanese lyrics, and HMM singing voice synthesizer[8] to generate the vocal part. We also implemented the system as a web-based application.

3. EVALUATION RESULTS

We did two experiments to evaluate the algorithm. Firstly, we asked a classical music composer to evaluate generated songs in five-grade evaluation. The results on 59 generated songs are shown in Fig. 5. These results indicate that 83.1% of the generated pieces satisfactorily follow classical music theory, and 91.6% of the songs were voted as attractive aside from musical theory. Example of generated song is shown in Fig. 4.

Secondly, we uploaded our system to get comments from a large number of users on the internet. During a year of operation, about 56,000 songs were generated by the users and 1378 people answered the questions about Orpheus and the generated songs. Summarization of answers in five-grade evaluation is shown in Fig. 6. Judging from the results, about 70.8% commented that the generated songs are attractive, and 84.9% of the users had fun trying this system.

4. DISCUSSIONS

The evaluations results by a composer indicate that most of the generated songs are able to be called “melody” at least in theory. Songs that are evaluated “very poor” had irregular usages of non-chord tones which rule cannot be described with the relationship between the current note and the previous note.

Evaluation by the users on the internet suggest that our composition system is an enjoyable solution for amateur

http://orpheus.hil.t.u-tokyo.ac.jp
to compose their original songs. One reason for the result could be the directability on generating songs. Users can generate various melody by typing arbitrary lyrics since the generated song will vary based on the prosody of the lyrics. This may enabled the user not only generating a song automatically but also to generate their original songs with their original lyrics.

5. CONCLUSION

This research attempted to design an algorithm to compose a song automatically from the lyrics using prosody information, which enables users to make their original songs easily. The results indicate that our method and implemented system Orpheus is an enjoyable solution for amateur musicians.

However, it should be noted that our algorithm can be applied to lyrics written in “pitch accent” languages only. As a next step, we plan to extend the composition algorithm to handle “stress accent” languages, such as English, by putting constraints on metric structure of the melody.

6. REFERENCES


