Orpheus: Automatic Composition System Considering Prosody of Japanese Lyrics

Satoru Fukayama[1], Kei Nakatsuma[1], Shinji Sako[2], Yuichiro Yonebayashi[1],
Tae Hun Kim[1], Si Wei Qin[1], Takuho Nakano[1], Takuya Nishimoto[1], and
Shigeki Sagayama[1]

[1] Graduate School of Information Science and Technology,
the University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan
{fukayama, k-nakatsuma, yonebayashi, kim, qin, t-nakano,
nishi, sagayama}@hil.t.u-tokyo.ac.jp
[2] Graduate School of Engneering, Nagoya Institute of Technology,
Gokiso-cho, Showa-ku, Nagoya-shi, Aichi, Japan
sako@mmsp.nitech.ac.jp

Abstract. *Automatic composition techniques are important in sense of upgrading musical applications for amatuer 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-path 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 generated songs are also reported, indicating that Orpheus can help users to compose their original Japanese songs. Trial version is available at* http://orpheus.hil.t.u-tokyo.ac.jp/

Key words: automatic composition, probabilistic modeling, prosody, dynamic programming, web application, copyright-free songs

1 Introduction

Recently, there has been wide interest in automatic composition algorithms which can help amatuer musicians to compose their original tunes. Although considerable research has been done on automatic composition[1][2][3], much less is done on composing songs from the lyrics, and a question remains that what information in the lyrics should be exploited for generating songs. Guido d'Arezzo invented a method to compose a melody by choosing the notes which correspond to the vowels in the lyrics[4]. Hayakawa used syntactic information of Japanese lyrics[5].

Musicologists argue that there are considerable correlations between music and prosody. For instance, it is well known that Beethoven composed a string quartet based on the prosody of German sentence: "Muss es sein? Es muss sein". In case of composing songs, prosody plays a more important role. However, no system that uses prosody has yet been attempted.

2   Song Composition Algorithm Considering Prosody

We present an algorithm for song composition using prosody of Japanese lyrics. Since Japanese is a "pitch accent" language, listener's apprehension is strongly affected by the pitch motions of the speaker. For example, the meaning of Japanese word "ha-shi" changes with the pitch. It means "bridge" with an upward pitch motion, and "chopsticks" with the motion inversed. A melody attached to the lyrics cause an effect similar to the pitch accent. Therefore we can assume that pitches of Japanese lyrics give constraints on pitch motions of the melody. Furthermore, chord progression, rhythm and accompaniment give constraints on the transitions and occurrences of the melody notes. If a certain melody for the lyrics were obtained, the melody would satisfy these constraints. Conversely, we can compose a song by finding the melody which optimally meets the condition.

Melody can be represented as a path as shown in Fig. 1. Constraints on the melody are comprehensible as transition probability and occurrence probability of each note which forms the melody path. Given the pitch series of the melody $X = \{x_0, x_1, \cdots, x_n\}$, the probability $P(X)$ will be calculated as follows:

$$P(X) = p_{oc}(x_0) \prod_{t=1}^{n} p_{tr}(x_t|x_{t-1})p_{oc}(x_t) \tag{1}$$

where $p_{oc}(x_t)$ is the occurrence probability, and $p_{tr}(x_t|x_{t-1})$ is the transition probability of notes. Now we can formalize the melody composition as finding the optimal X^* which maximizes the log-likelihood $\log P(X)$:

$$X^* = \underset{X}{\operatorname{argmax}} \log P(X). \tag{2}$$

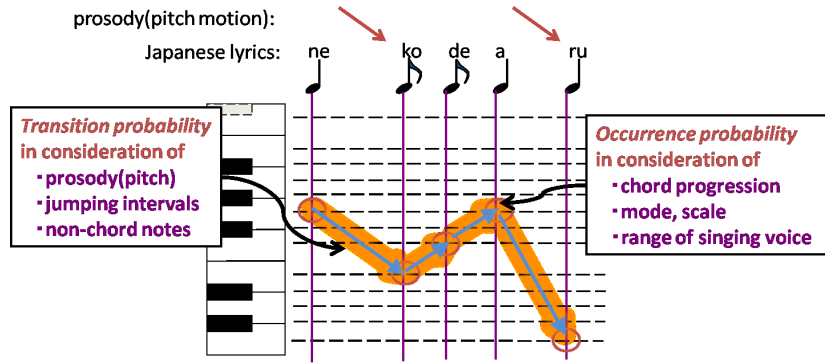We can obtain $X^* = \{x_0^*, x_1^*, \cdots, x_n^*\}$ using dynamic programming[6].

Fig. 1. Automatic composition as an optimal-path searching problem under constraints

3   Implementation and Experimental Results

Orpheus is an automatic composition system that we implemented using melody composition algorithm based on prosody. This system computes melody from the lyrics input with choices of chord progressions, rhythm patterns, and accompaniment instruments. Flow chart of the processes is shown in Fig. 2. 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].

We did two experiments to evaluate the system. 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. 3. 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. 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. 4. 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.
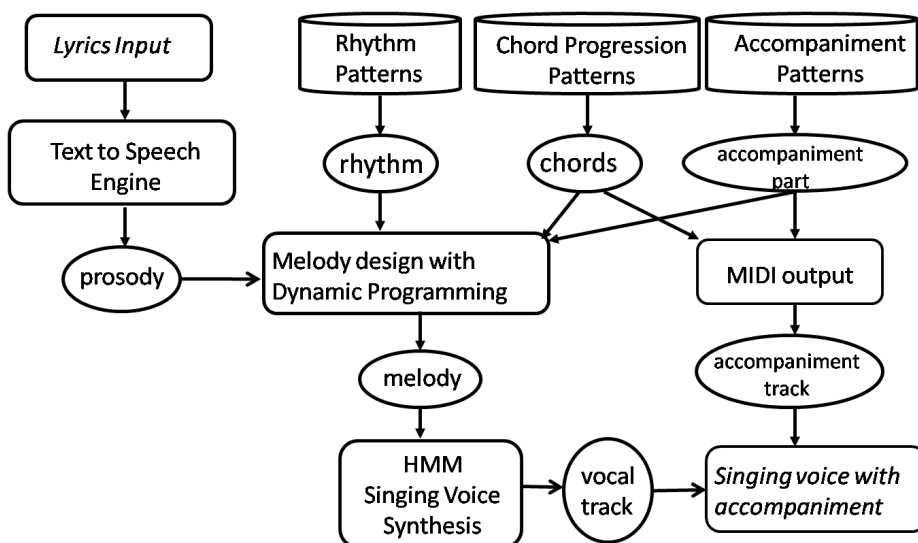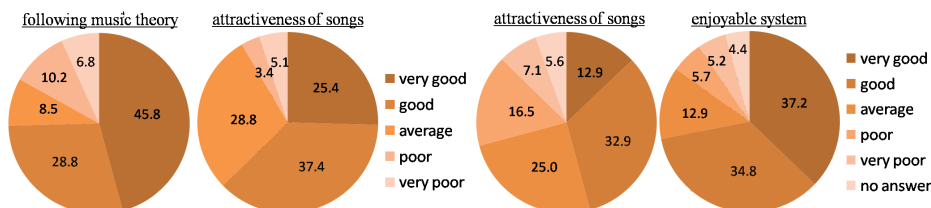

Fig. 2. Flow chart of processes: *Orpheus* generates songs with the lyrics input and the choices of patterns.

---

[3] http://orpheus.hil.t.u-tokyo.ac.jp/automatic-composition/index.en.cgi

Fig. 3. Evaluation results on 59 songs by a classical music composer [%]

Fig. 4. Evaluation results on generated songs and the *Orpheus* by 1378 users [%]

4   Discussions and 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.

References

1. Hiller, L., and L. Isaacson.: Experimental Music. McGraw-Hill, New York (1959)
2. Xenakis, I.: Formalized Music. Revised edition. Pendragon Press, New York (1992)
3. Cope, D.: Computers and Musical Style. A-R Editions. Madison, Wisconsin (1991)
4. Loy, D. G.: Composing with computers — a survey of some compositional formalisms and music programming languages. Current Directions in Computer Music Research. pp. 291–396. The MIT Press, Cambridge, Massachusetts (1989)
5. Hayakawa, K., Inagaki, H., Tanaka, K.: Automatic song composer from phrase structure of lyrics., 57th IPSJ Annual Convention, pp.11–12. (1998) (in Japanese)
6. Bellman, R. E.: Dynamic Programming. Princeton University Press, Princeton, New Jersey (1957)
7. Galatea Project, `http://hil.t.u-tokyo.ac.jp/~galatea/`
8. Sako, S., Miyajima, C., Tokuda, K., Kitamura, T.: A Singing Voice Synthesis System Based on Hidden Markov Model., Transactions of IPSJ, pp.719–727. (2004) (in Japanese)
9. CrestMuse Project, `http://www.crestmuse.jp/index-e.html`