Lyric Retrieval System Based on Word's Intimacy Value

Kazuyuki Matsumoto[1, a], Manabu Sasayama[2, b], Kohsuke Miyauchi[1],
Minoru Yoshida[1,c] and Kenji Kita[1,d]

[1]Faculty of Science and Technology, Tokushima University, Tokushima, Japan,
[2]Kagawa National College of Technology, Kagawa, Japan

[a]matumoto@is.tokushima-u.ac.jp, [b]sasayama@di.kagawa-nct.ac.jp, [c]mino@is.tokushima-u.ac.jp,
[d]kita@is.tokushima-u.ac.jp

Keywords: Lyric retrieval, query expansion, word intimacy value.

Abstract. In this study, we focused on fragment of wrong lyrics that are heard or remembered wrongly and entered as query for lyric search. Concretely, we aimed to improve search accuracy by query conversion considering noise of word. We also tried to calculate word intimacy degree that indicates how much the word is in common use so that we can replace the wrong lyric word in the query into the correct lyric word.

1 Introduction

Recently, there are music retrieval systems based on various data such as names of singers, composers or lyric writers, song titles, lyrics, singing voices, audio fingerprinting, etc. Among them, names of singers or composers, song titles and lyrics are the most popular queries. Such queries do not include much amount of information, therefore they tend not to include errors. For example, *Denmoku* is a digital index book that is often used in karaoke shops. This system primarily uses artist name, composer name or song title as query for song search.

On the other hand, in song retrieval from the content of the lyric, there tends to be more errors in the queries that are caused by mishearing or wrong memorization. As a result, the search accuracy would be decreased. Listening to the music, the listeners tend fail catching some sentences or some phrases where there is stretch of voice or change of musical pitch. Therefore, the false recognition of lyric such as mishearing or wrong memorization is caused. In this paper, we proposed the word conversion method by considering mishearing and wrong memorization.

2 Related Works

In previous works, Xu et al. [1] proposed a method to calculate acoustic distance for the purpose of quantifying the degree of misrecognition between syllables from the error of voice recognition. This method obtained the positive result that could retrieve the lyric more accurately than the method for usual document retrieval. However, because the lyrics were split by syllable unit, there must be lack of semantic information of each word. This causes a problem that the amount of information in a query decreases.

Sasayama et al. [2] investigated the error type of search query for the lyric retrieval system. They also proposed a lyric retrieval system by expanding query based on semantic information. They expanded query by using web n-gram corpus. Matsumoto et al. [3] proposed a lyric retrieval system that used the characteristic expressions in lyric as important phrase. Their system re-ranked the retrieval results by weighting the singer's document frequency value.

3 Definition of Lyric Noise

In this paper, the target lyrics are automatically collected from web. We defined the character string in the search query that is not included in search target as noise. This noise is inserted in the search query from various reasons. This section explains noise type.

1. Wrong notation
   This noise occurs with higher possibility when the lyric is not confirmed with lyric card.
2. Wrong word
   If the word is unknown word for the listener or misheard, this noise occurs. On the other hand, the problem of melody or acoustic quality might cause this noise.
3. Deletion or Addition
   This often occurs from mishearing by the listener. Although this is a rare case, the artist occasionally does not sing a song with accurate lyric.
4. Permutation
   This often occurs from mishearing or input error by the listener.

Among these noises, we focused on 1, 2 and 4 because we considered that the correct word could be estimated from the noise words. We also thought that these noise data would be easily obtained. By focusing on these noises, we would be able to limit the experimental condition.

4 Proposed Method

The flow of the proposed method is shown in Fig.1. The intimacy degree of the noise word in the input query is analyzed by referring to word intimacy degree database [4].

Then, if the system can find the synonym word whose intimacy degree is higher than that of the noise word, the noise word will be converted into that word. When the system find the synonym word, Japanese WordNet [5] synonym database is referred to.
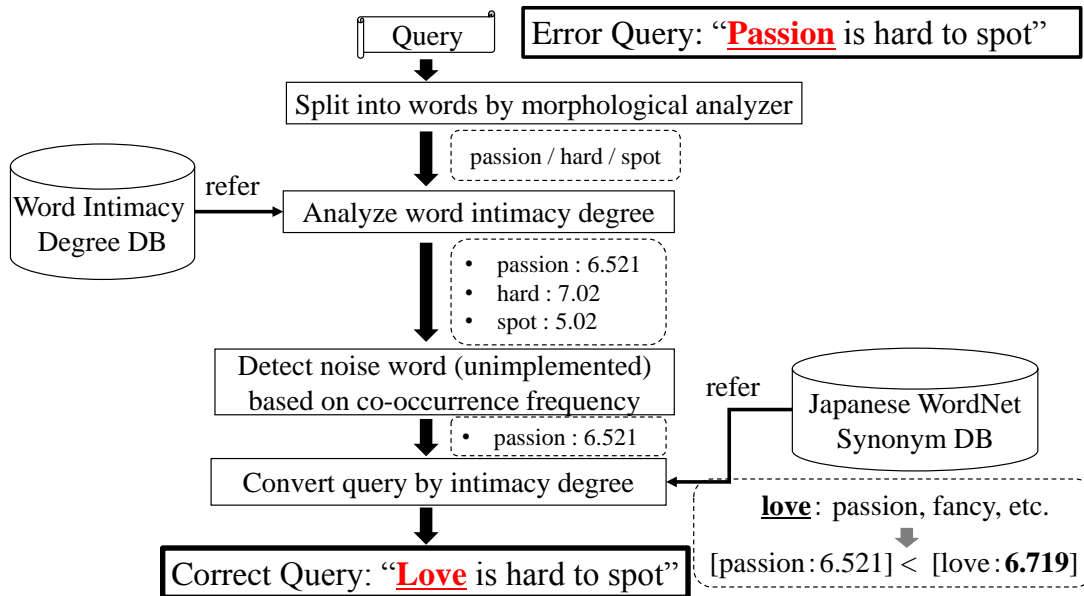

Fig.1. Flow chart of query conversion

5 Questionnaire

The questionnaire was conducted for acquiring noise examples. We used 68 phrases of the lyrics that included various noises. Four test subjects answered the questionnaire. A part of the questionnaire are shown in Table 1. The blank space surrounded by bracket '()' indicates the answer section. The options were indicated to the subjects. The underlined words are wrong answers chosen by the subjects. The words indicated in boldface are correct words.

Table 1: A part of the questionnaire

No.	Question	Song title [artist]	Options (synonyms)
1	( ) *no hate made zutto*	The song of love [Seiko Matsuda]	A). *Chi*, B). *Ochido*, C). Area, D). ***Chihei***
2	( ) *wo susumeyou Catcheve*	*Hoshi ni negaiwo*[chin para]	A). Step, B). ***Ayumi***, C). *Ashidori*, D). *Ashitsuki*
3	( ) *no kanashimi ha kokoro ni kakari*	*Kanashimi no tobari* [Sachiko Kobayashi]	A). Root, B). **Road** C). *Kouro*, D). *Keiro*
4	( ) *no watashi* Moonlight	*Don't Stop* [Tiara]	A). ***Hyojo***, B). *Gyousou*, C). *Kao*, D). *Sougou*
5	( ) *furimukasetai*	*DA KA RA* [Maki Ooguro]	A). *Koyoi*, B). *Touya*, C). *Konban*, D). ***Konnya***
6	( ) *wa amari medatanai kedo*	*People* [Kyoko Fukada]	A). *Ai*, B). *Hyojo*, C). Love, D). ***Aichaku***

6 Experiment and Discussions

In this section, we conduct the evaluation experiment to confirm the effectiveness of our method. As the baseline method, we used the retrieval system based on the suffix array algorithm. Some suffixes were extracted from the inputted lyric fragment. The test data included wrong queries that were obtained from the questionnaire.

We evaluated the result of the proposed method by recall, precision and F-measure. The calculation formulas are shown in Eq. 1, 2, 3. S indicates the total number of phrases. C indicates the number of phrases that were outputted correctly. N indicates the total number of the outputted lyrics. The target lyric database included approximately 100,000 lyrics (Japanese pop music songs). We evaluated the top 20 lyrics which outputted by the lyric retrieval system. (i.e. the output was counted as correct answer if top 20 lyrics included the target lyric.)

$$\text{Recall} = \frac{C}{S} \tag{1}$$

$$\text{Precision} = \frac{C}{N} \tag{2}$$

$$\text{F} - \text{measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{3}$$

The result of the experiment is shown in Table. 2. From this result, we found that the proposed method could retrieve correct lyric by converting word based on word's intimacy degree. However, 16 queries were converted into wrong words. Because the word intimacy degree was not for specific person, our proposed method sometimes failed to judge the standard of conversion correctly. Then, table 2 shows an example of retrieval result.
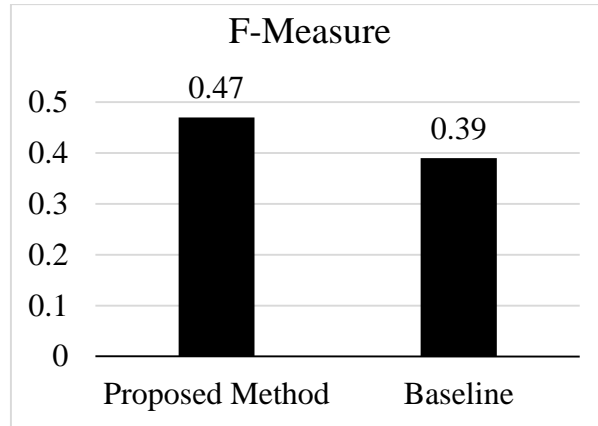
Fig. 2: The experimental result (F-measure)

Table 2: Example of the retrieval result (Input query: "*Kimi **wa hajimete***")

Rank	Artist / Song title	Score	Matched phrase
1	GODIEGO / "START SINGING AGAIN"	75.00	*Ima **wa hajimete** no stage omoidashite*
1	LOVE LOVE LOVE / "*Konna Kimochi wa*"	75.00	*Konna kimochhi **wa hajimete** dattari*
5	PYG / "*Hajimete no namida*"	56.25	*Sore **wa hajimete** no namida*
5	BONNIE PINK / "Hurricane"	56.25	*Konna itami **wa hajimete***
5	BOO / "Instant Love"	56.25	*Konna kimochi **wa hajimete***

7 Conclusions

In this paper, we proposed a method to convert the noise word into the correct word based on the word intimacy degree. As the experimental result, our method could obtained higher precision than the baseline method. In future, we would like to propose a method to detect noise in query by using co-occurrence frequency of the words. We also would like to conduct the evaluation experiment by using larger test data.

Acknowledgement

References

[1] X. Xu, M. Naito, T. Kato and H. Kawai: "Robust and fast lyric search based on phonetic confusion matrix." In Proc. The 10th International Society for Music Information Retrieval Conference (ISMIR2009), 417–422. (2009)

[2] M. Sasayama and K. Matsumoto: "Query expansion using semantic information for lyric search system." In Proc. Human Communication Group Symposium (HCG2013), pp.51-59. (2013)

[3] K. Matsumoto, M. Sasayama, Q. Xiao, M. Yoshida and K. Kita: "Reranking the search results for lyric retrieval based on the songwriters' specific usage of words", In Proc. the 4th International Conference on Electronics, Communications and Networks (CECNet2014), pp. 1045-1052. (2014)

[4] N. Amano, K. Kondo, S. Sakamoto and Y. Suzuki, NTT psycho-linguistic databases "Lexical Properties of Japanese", 1999, San-seido2011.

[5] Japanese WordNet: http://nlpwww.nict.go.jp/wn-ja/.