An Onomatopoeia-Based Web Music Video Searching System and Its Performance Evaluation

Shizuka Sato[1], Eiichiro Kodama[2], Jiahong Wang[2], and Toyoo Takata[2]

[1] Graduate School of Software and Information Science, Iwate Prefectural University, Japan
[2] Faculty of Software and Information Science, Iwate Prefectural University, Japan
crowz1547@gmail.com

Abstract. In recent years, as use of the Internet became widespread, numerous music videos became available on Web. In Japan, many of these music videos are the CGM (Consumer Generated Media) that are created using a singing synthesis software called Hatsune Miku, and published on YouTube and other similar Web sites. Existing Web sites, however, support only the search methods based on music video title and artist name, which could not be effectively used to search for the unknown music videos such as the CGM ones. This paper presents a system model for effectively searching for the unknown music videos, which is characterized by the use of the onomatopoeia. The system model consists of a music video collecting engine for collecting pairs of music video URL and its tags, an onomatopoeia assigning engine for assigning onomatopoeias to music videos, and an onomatopoeia retriever for presenting users the music video URLs satisfying their onomatopoeia requirements. We have implemented a prototype system of the proposed system model and conducted experiments to study its performance. It has been found that with the proposed system model, a precision ratio of 66.82%, a recall ratio of 56.36%, and an F-measure of 61.14% could be achieved.

Keywords: Onomatopoeia, Web Search, Music Video.

1 Introduction

In recent years, as use of the Internet became widespread, numerous music videos became available on Web. As a result, it becomes more and more important to meet users' requirements for more effectively searching for the music videos. It is found that in Japan, many of these music videos are the CGM (Consumer Generated Media) that are created using a singing synthesis software called Hatsune Miku, and published on YouTube[1] and other similar Web sites.

In this paper, we present a system model for effectively searching for the unknown music videos from Web, which is characterized by the use of the onomatopoeia. An onomatopoeia is a word that sounds like the common sound of the object it is describing, and an example of onomatopoeia is a train being called a choo choo.

C. Stephanidis (Ed.): HCII 2014 Posters, Part I, CCIS 434, pp. 246–251, 2014.
© Springer International Publishing Switzerland 2014

2 Problems of Existing Search Methods

Existing Web sites support only the search methods based on the title and artist name of a music video, which could not be effectively used to search for the unknown Hatsune Miku music videos since users could not know their titles and artist names. Using the existing techniques, users could only use the retrieval keyword "Hatsune Miku" to search for Hatsune Miku music videos, and find their favorite ones by checking the music videos given in the resulting list, which is generally sorted in publishing time order, one by one.

Considering the problem stated above, we think that a more effective search method for the unknown Hatsune Miku music videos is necessary.

3 Related Work

Two existing work are closely related to this research. In paper [2], the subject of searching for sound data from a static sound database with some onomatopoeic words was addressed. In paper [3], the subject of music retrieval from static music database using classic sensitive words was addressed.

For searching for the music videos, it has not been known which one of the sensitive word-based methods and the onomatopoeia-based methods is more effective. In fact, it may be too difficult to give a definite declaration concerning which one is more useful. We do not think these two methods are exclusive, but we think that to meet users' varying requirements, it would be important to investigate the possibility of using solely the onomatopoeia to support the retrieval of music videos.

4 Search for Unknown Music Videos Using Onomatopoeia

The proposed system model for searching for the unknown music videos from Web using the onomatopoeia is given in Fig. 1.

4.1 Music Video Collecting Engine

The *music video collecting engine* is used to collect the pairs of music video URL and its tags from the music video distribution sites such as YouTube [1] and Niconico [4]. Note that music videos in these sites have been assigned some attribute words such as "Hatsune", "Miku" and "Append", which are called the *tags* and are used to describe the music videos and make their retrievals easier. The collected data will be stored in the *target music video database* temporarily, and processed by the *onomatopoeia assigning engine*.

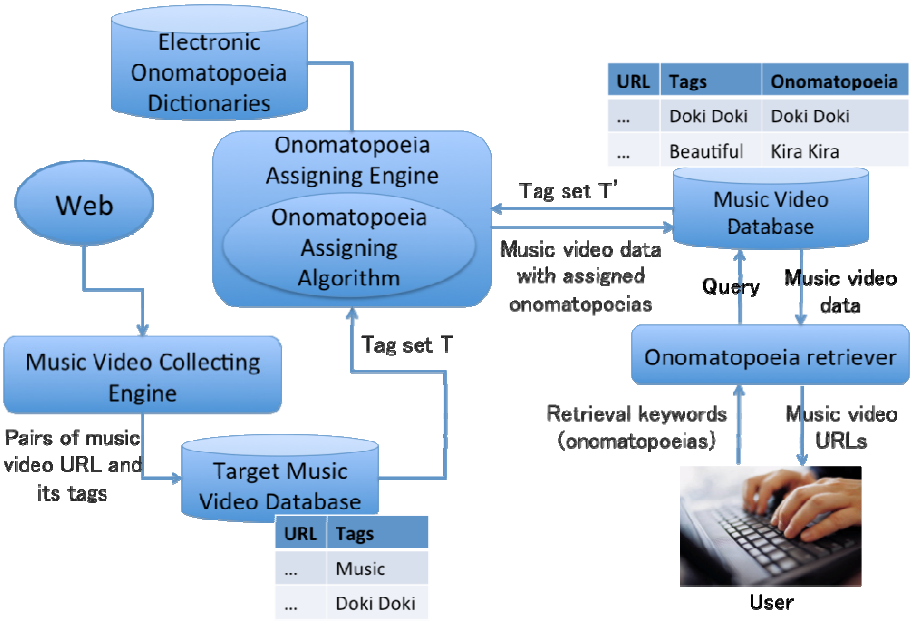

Fig. 1. Onomatopoeia-based Web music video searching system model

4.2 Onomatopoeia Assigning Engine

By applying the tag sets of *target music video data* and the tag sets of *music video data* stored in the *music video database*, onomatopoeia assigning engine assigns onomatopoeias to the target music video data using an *onomatopoeia assigning algorithm* given below. The resulting target music video data will be moved into the music video database. It is assumed that before the algorithm is started, there have been accumulated at least n records in the music video database, where n is a predefined system parameter.

Onomatopoeia Assigning Algorithm

1. Get tag set T of the target music video data to which onomatopoeias will be assigned. Using electronic onomatopoeia dictionaries, onomatopoeias included in T are extracted, and the results are assigned to the target music video data.
2. Get the next record from the music video database, and then get the tag set T' from the record.
3. Calculate the similarity degree of T and T' by the following formula.

$$Jaccard(T, T') = \frac{|T \cap T'|}{|T \cup T'|}$$

4. Repeat steps 2 and 3 for all the music video records, and assign the onomatopoeias of the music video data of the maximal similarity degree to the target music video

data. In the case that the maximal similarity degree is greater than a predefined threshold θ $(0 < \theta < 1)$, the onomatopoeias obtained in step 1 are also assigned to the music video data of the maximal similarity degree.

4.3 Onomatopoeia Retriever

The *onomatopoeia retriever* presents users the music video URLs that satisfy their onomatopoeia requirements by comparing the onomatopoeias provided by users as the retrieval keywords with the onomatopoeias of the music videos.

Figure 2 shows user interface of the search engine. As an example, if the retrieval keyword "TanTan" is entered, we will have the results shown in Fig. 3. Users can enjoy a music video with a browser by clicking the URL.

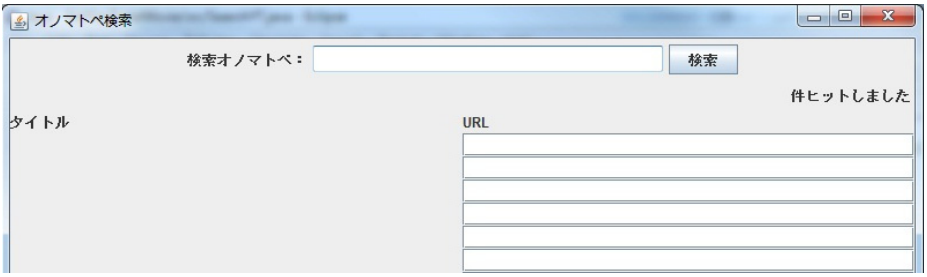

Fig. 2. User interface of the search engine

Fig. 3. Results for retrieval keyword "TanTan"

5 Performance Evaluation

We have implemented a prototype system of the proposed system model and conducted experiments to study its performance in terms of *precision ratio* and *recall ratio*.

Music video data used for the experiments consists of 100 initial music videos and 100 target music videos. The experiment for the precision ratio was conducted with 12 test subjects and 48 onomatopoeia words that have been assigned to the music videos. For each retrieval keyword, the ratio of correct answers, which are judged by the test subjects, included in the searching results of music videos are calculated, and the average of 12 test subjects is taken as the precision ratio.

For the recall ratio, first the music videos corresponding to the 48 onomatopoeia words, which are judged by 5 test subjects, are determined as the correct answers. Music videos are searched using the onomatopoeia words that have been assigned to them, and the ratio of correct answers included in the searching results of music videos is taken as the recall ratio.

A part of the experiment results are shown in Table 1, where the average of the precision ratios (resp. recall ratios) corresponding to onomatopoeia words is taken as the precision ratio (resp. recall ratio) of the prototype system. *F-measure* is calculated by the formula: 2(1/recall ratio+1/precision ratio).

Table 1. Results of Precision Ratio and Recall Ratio

Onomatopoeia	Precision ratio	Recall ratio
ガチャガチャ（GachaGacha）	91.67%	25.00%
カラカラ（KaraKara）	69.44%	100%
ジャカジャカ（JacaJaca）	94.05%	46.15%
シャンシャン（ShanShan）	68.75%	33.33%
タンタン（TanTan）	78.85%	83.33%
ノホホン（NohoHon）	66.67%	50%
ピコピコ（PikoPiko）	57.10%	50%
ビュンビュン（ByunByun）	91.67%	33.33%
フラフラ（FuraFura）	41.67%	100%
ルンルン（RunRun）	75%	0%
ワクワク（WakuWaku）	75%	50%
Average	66.82%	56.36%

6 Conclusion

In this paper we have proposed a new system model for searching for the music videos from Web, which is characterized by the use of the onomatopoeia. We have conducted experiments to evaluate performance of the proposed system model. Experiment results shown that with the proposed system model, a precision ratio of 66.82%, a recall ratio of 56.36%, and an F-measure of 61.14% could be achieved.

References

1. YouTube, http://www.youtube.com
2. Uota, K., Suzuki, K., Pitoyo, H., Hashimoto, S.: Sound Database Retrieved by Onomatopoeia and Sound, Technical report of Electronics, Information and Communication Engineers. HCS 104(446), 19–22 (2004)
3. Ikezoe, T., Hajikawa, Y., Nomura, Y.: Music Database Retrival System with Sensitivity Words Using Music Sensitivity Space. Information Processing Society of Japan Journal 42(12), 3201–3212 (2001)
4. Niconico, http://www.nicovideo.jp/