A Cross Domain Lyrics Recommendation from Tourist Spots Reviews with Distributed Representation of Words

Yihong Han*, Ryosuke Yamanishi†, Yoko Nishihara† and Kenta Oku‡
*Graduate School of Information Science and Engineering, Ritsumeikan University, Shiga 525-8577, Japan
†College of Information Science and Engineering, Ritsumeikan University, Shiga 525-8577, Japan
Email: {is0387ps@ed, ryama@media, nisihara@fc}.ritsumei.ac.jp
‡ Department of Media Informatics, Ryukoku University, Shiga 520-2194 Japan
Email: okukenta@rins.ryukoku.ac.jp

***Abstract*—In this paper, we propose a system for recommending lyrics similar to the context of the reviews of a tourist spot. The system is based on the technique of distributed representation of words. Instead of the metadata, such as genres and artists, the proposed approach takes the listening environment into consideration. By using the listening environment, it is possible to recommend music lyrics that fit the atmosphere of a tourist spot when the user enjoys the sightseeing. In this paper, the proposed system recommends the music that fits the atmosphere of a tourist spot by sharing the distributed representation beyond the domains of lyrics and reviews. The system uses a lyrics corpus to build the distributed representation model, and the reviews' vectors are calculated with the model. As a result, the tourist spot reviews are assumed to be types of lyrics. Based on the lyric-like vector representation of reviews, the similarity between reviews and lyrics can be calculated.**

Keywords–Music Information Retrieval; Lyrics; Context Aware Music Recommendation; Cross Domain Search.

I. INTRODUCTION

The user-system interaction for listening to music has dramatically changed with the use of web services. Subscription services for music enable us to bring almost an infinite amount of music everywhere. Users do not need to select the music before going out. We previously enjoyed music in places designed especially for music, such as live houses and concert halls. However, nowadays, we can listen to music in many places, such as when driving, being on a flight, or trekking. That is to say, music has now become more of a co-entertainment while doing some other activities though it was previously the main entertainment for places designed especially for music.

Based on this background, context-aware music retrieval can be a new style to enjoy music: listening to music while interacting with the environment surrounding the user. In this paper, we take music in tourism in consideration as a listening context for the music. Perhaps, some of us might experience listening to music which includes the name of a tourist spot in the lyrics. The experience of visiting a tourist spot is more impressive while listening to music related to it, e.g., "*San Francisco*" by Glantis in San Francisco, USA and "*Lovers in Japan*" by Coldplay in Osaka, Japan. It is expected that even listening to the music without the name of a spot in the lyrics also enhances the impressions toward the trip if the sentiment for the music corresponds to the atmosphere of the spot: for example, listening to "*Perfect*" by Ed Sheeran which is a relaxed love song in the airy and relaxed park in Vancouver "Stanley Park" and listening to "*Toxicity*" by System of a Down in an exciting city like Kabuki-Cho, Japan. It is reasonable to say that such experiences are similar to listening to the background music for each scene in movies. This paper can be positioned in the location-aware music recommendation field [1] [2].

In this paper, the goal is to enrich the tourism experiences with listening to music for the tourist spot. We propose a method to recommend music that has lyrics suitable for the tourist spot. In the proposed method, the reviews for the tourist spots are assumed as the general evaluation or the collective intelligence of experience toward the tourist spot. By using the reviews of the tourist sport as the query, the proposed method retrieves the lyrics for the spot: that is, cross-domain retrieval between tourist spots and music. The distributed representations of words are modeled using a lyrics corpus. As the reviews of tourist spots are vectorized with the distributed representation model, the reviews of tourist spots should be assumed as the lyrics. The lyrics that have higher vector similarity with the reviews of the tourist spots are retrieved: the lyrics retrieval with tourist spots should become enable.

Section II will introduce the related work. The description of the proposed method will be in Section III. Section IV will show the experiment results and the evaluation of the proposed method. Finally, Section V will summarize this paper.

II. RELATED WORK

Music Information Retrieval (MIR) has been widely researched. Music retrieval with humming [3] and music genre classification [4] [5] are typical research topics in the MIR field.

Music with singing can be considered as multimedia art and consists of acoustics and linguistics. That is, the affection towards music can be caused by the combination of "listening to acoustics" and "understanding lyrics." Let us focus on the related work for lyrics, which is the target of this paper. Tsukuda *et al.* [6] developed *Lyrics Jumper* that recommends artists whose lyrics have similar topics. Cai *et al.* [7] have proposed *MusicSense* that recommends the music while reading a document on the Web. In the work by Cai *et al.*, the affective words extracted from both lyrics and documents on the Web are used to relate the two types of domains with each other. Our proposed system does not focus on some specific words, but the overall similarity between lyrics and reviews by using the word distributed representation model.

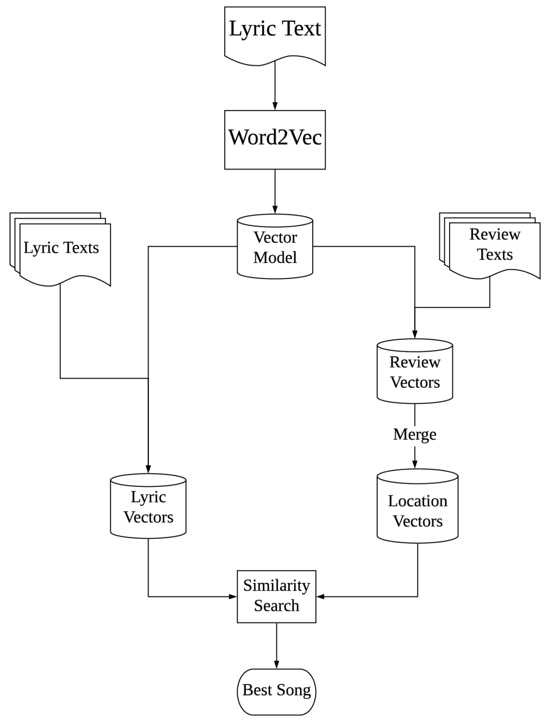

Figure 1. Flowchart of proposed method.

The music and location have been related to each other in several existing research works. Kaminskas *et al.* have proposed a location-aware music recommendation system while using a tag-based approach [8] and a knowledge-based approach [9]. In their tag-based approach [8], music and Point Of Interest (POI) are related to each other based on the tag given to those. Their knowledge-based approach [9] constructs the graph that semantically relates music with POI based on the knowledge with DBPedia and ranks the songs for a given POI. Moreover, a hybrid approach of a tag-based and knowledge-based approach has been proposed [10]. Our proposed method is considered as the lyrics-based approach in the above research context.

III. Proposed Method

In this paper, we propose a lyrics recommendation method based on the vector similarity between lyrics and reviews of tourist spots. The proposed method uses distributed representations of words [11] to quantify lyrics and reviews of tourist spots. We use an English lyrics corpus to build distributed representations as a model. The model is used for calculating the distributed representations vectors of lyrics and reviews of tourist spots. The mean vectors of a single tourist spot are calculated from all of the vectors concerning the reviews for the tourist spot. The method calculates the vector similarities between lyrics and tourist spot reviews. The lyric with the highest similarity to the given tourist spot is output as the recommendation result for the tourist spot. The concept of our method is, the texts of reviews for tourist spots are assumed as "pseudo lyrics." So, the lyrics and reviews can be unified into the same dimension and become comparable. Figure 1 shows an outline of the proposed method. Each process of the proposed method will be detailed in the next two sections.

A. Words Distributed Representations Model

We model a distributed representations with the lyric data fetched from a lyrics site "azlyrics" [12]. This paper focuses on English content, so we just choose English lyrics as the text corpus. All lyrics written in non-English language are omitted from the lyrics dataset. After the cleansing, there are 94,451 English lyrics remaining in the dataset. We use Word2Vec (Skip-Gram) [13] [14] framework to model word distributed representations as the primal study, though there are so many types of frameworks. We use this lyrics corpus on Word2Vec framework as the training data and get a distributed representation model that represents every word in the corpus as a 300 dimmensions vector. During the training, the parameter setup of Skip-Gram is as follows: $size = 300$, $window = 10$, $min\_count = 2$, $workers = 8$, $iter = 10$.

B. Quantifying Lyrics and Tourist Spot Reviews to Vectors

This section describes the general concept for the distributed representation for both lyrics and tourist spot reviews. Based on the distributed representation model described in Section III-A, the distributed representations vectors of lyrics and reviews for tourist spots can be obtained. In detail, for every word in lyrics and reviews, we fetch the word vector from the distributed representations model. Then, the mean vector of lyrics or reviews $\overline{\boldsymbol{V}}$ is calculated by summing all word vectors for each dimension and dividing the sum by the number of words in the lyrics or reviews (1) as follows:

$$\overline{\boldsymbol{V}} = \frac{\vec{v_1} + \cdots + \vec{v_i}}{N - \gamma}, \tag{1}$$

where, $\vec{v_i}$, N, and γ denote the vector of ith word, the number of words in a lyric or review text, and the number of words that only exist in review texts, respectively. Note, as the distributed representations model is trained from lyrics corpus, the words that only exist in review texts but do not appear in lyrics texts, will not get their vectors from the model. In the calculation of average vectors, we counted the number of these "not available words" and subtracted the number γ from the word number in the whole text. As a result, the number of "available words" is the divisor in the equation.

C. Merging Vectors of Reviews for Tourist Spot to Spot Vectors

In Section III-B, we obtain both vectors of lyrics and tourist spot reviews (hereafter, review vectors). Here, another process for the tourist spot reviews is detailed in this section. Tourist spot reviews are written by a human so they may include emotional expressions and subjective estimations. By merging reviews of the same tourist spot, the individuality and general properties of the tourist spot can be represented. Based on this consideration, we calculated the weighted arithmetic mean of each tourist spot. The mean vector of tourist spot s: $\overline{\boldsymbol{X_s}}$ is calculated using the following equation:

$$\overline{\boldsymbol{X_s}} = \frac{\omega_{s,1}\vec{x_{s,1}} + \cdots + \omega_{s,j}\vec{x_{s,j}}}{\omega_{s,1} + \cdots + \omega_{s,j}}, \tag{2}$$

TABLE I. EXAMPLES OF TOURIST SPOTS AND THE CORRESPONDING RECOMMENDED LYRICS. THE CONTENTS OF LYRICS WILL BE DETAILED IN TABLE II.

Tourist spots	Recommended lyrics ID
The Montcalm at the Brewery London City	84127
The Beekman A Thompson Hotel	84057
Conservatory Garden	55628
Riverside Park	74814
Hudson River Park	74814
Roosevelt Island	39663
Fort Troon Park	74814
New York Harbor	56837
Franklin D Roosevelt Four Freedoms Park	53520
Long Beach	54217

where, $\vec{x}_{s,j}$ and $\omega_{s,j}$ denote the vector of jth review and the number of words in jth review for tourist spot s, respectively.

A given tourist spot may have two types of reviews: short simple or long detailed reviews. By weighting the vectors of words using the number of words in the review, the contributions to the tourist spot expression should be differently evaluated depending on the length of the description. We suppose that the longer the review text is, the more information this review brings to the tourist spot expression. Hereafter, the mean vector of each tourist spot is named as "location vector."

D. Lyrics Recommendation Based on Similarity Between Lyrics and Spots

For a tourist spot s, we calculate the cosine similarity between its location vector and every lyrics vector. The lyric with the highest similarity for the tourist spot s is recommended as the lyrics toward the tourist spot s. The cosine similarity is calculated using the following equation:

$$\cos(\vec{X_s}, \vec{L_k}) = \frac{\vec{X_s} \cdot \vec{L_k}}{|\vec{X_s}| \times |\vec{L_k}|}, \quad (3)$$

where, $\vec{L_k}$ shows the vector of kth lyrics in the lyrics dataset. The proposed method recommends $\arg\max_k \cos(\vec{X_s}, \vec{V_k})$ as the recommendation result for the tourist spot s.

IV. Experiment

The effectiveness of the proposed method is subjectively discussed through the lyrics recommendation experiments. In the experiment, we evaluate some tourist spots randomly selected from "TripAdvisor [15]". We take some of the recommendation results as examples to be discussed in detail. Also, we discuss the overall tendency of the recommendations.

This paper is a working in progress, so the objective evaluation for the recommendation will be our future work. The discussion may lead to the direction for the idea of our future objective experiments.

A. Results

TABLE I shows the recommendation results of each tourist spot in lyrics ID. TABLE II shows the list of lyric URLs for each lyrics ID on lyrics site "azlyrics."

Figure 2 shows the number of tourist spots corresponding to each of the lyrics, where the horizontal axis is the lyrics ID and the vertical axis is the number of tourist spots. In the figure, from left to right, the lyrics are sorted in the descending order.

TABLE II. EXAMPLES OF RECOMMENDED LYRIC ID AND ITS URLS. THE URL WAS RETRIEVED ON MARCH 16, 2020.

Lyric No.	Lyric URLs
59978	https://www.azlyrics.com/lyrics/mcfly/mcflythemusical.html
84127	https://www.azlyrics.com/lyrics/sunkilmoon/strangerthanparadise.html
92790	https://www.azlyrics.com/lyrics/whitestripes/littlecreamsoda.html
84057	https://www.azlyrics.com/lyrics/sunkilmoon/beautifulyou.html
80687	https://www.azlyrics.com/lyrics/slimdusty/themanfromthenevernever.html
55628	https://www.azlyrics.com/lyrics/lobo/whyisitme.html
35313	https://www.azlyrics.com/lyrics/gregoryalanisakov/fireescape.html
74814	https://www.azlyrics.com/lyrics/rodstewart/manhattan.html
56035	https://www.azlyrics.com/lyrics/lorettalynn/imshootinfortomorrow.html
39663	https://www.azlyrics.com/lyrics/idinamenzel/oneshortday.html
20655	https://www.azlyrics.com/lyrics/cowboyjunkies/arlington.html
56837	https://www.azlyrics.com/lyrics/luckyboysconfusion/likeratsfromasinkingship.html

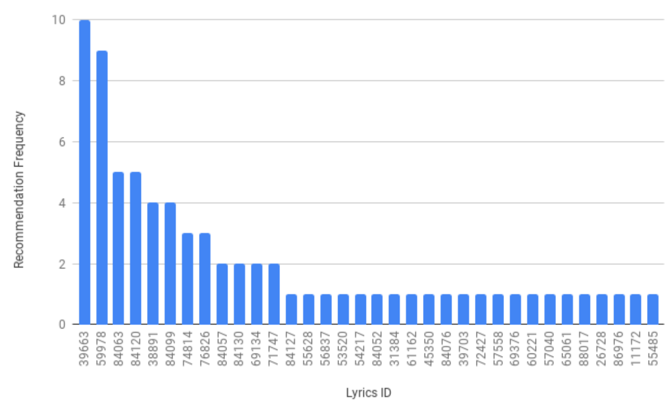

Figure 2. Number of locations corresponded to same lyric.

B. Discussions

From the experiment results, the tendency of the recommendations was found. Several tourist spots corresponded to the same lyrics. In TABLE I, the same lyrics "74814" was recommended to "Riverside Park," "Hudson River Park" and "Fort Troon Park." As these three locations are all parks, we supposed that the reason for this tendency was caused by the specific common features in the several tourist spots. To verify the assumption, we focused on the lyrics "39663" and "59978," which were recommended for the most spots. The tourist spots corresponding to the lyrics were studied in detail to find if there had been some similarity among them. The tourist spots recommended to lyrics "39663" and "59978" are each shown

TABLE III. TOURIST SPOTS CORRESPONDING TO LYRIC "39663"

Tourist spots	Locations
Roosevelt Island	State of New York, America
Bowling Green	Commonwealth of Kentucky, America
Governors Island National Monument	State of New York, America
SoHo	State of New York, America
West Village	State of New York, America
Meatpacking District	State of New York, America
Twin Peaks	State of California, America
Lincoln Park Conservatory	State of Illinois, America
Greenwich	London, England
Greenwich Park	London, England

TABLE IV. TOURIST SPOTS CORRESPONDING TO LYRIC "59978."

Tourist spots	Genre of location
Neue Galerie	Museum
Solomon R Guggenheim Museum	Museum
New York Historical Society Museum Library	Museum
Museum of Arts and Design	Museum
United Nations Headquarters	Organization
Broadway	Street
Radio City Music Hall	Theater
Le Puy du Fou	Theme Park
Westminster	Street

in TABLE III and TABLE IV, respectively.

As a result, it was suggested that the tourist spots corresponding to the same lyrics had specific common features. In TABLE III, there were a lot of spots located in the U.S.A., especially in the State of New York: for "Greenwich" and "Greenwich Park," the latter is inside of the former so substantially they should be almost the same spot. Generally, the tourist spots that corresponded to lyric "39663" were similar to each other in their location.

In TABLE IV, lyric "59978" was recommended to many museums and other historical places, such as "Radio City Music Hall" and "Westminster." Kaminskas mentioned that the recommendation result of music could be diversified with the matching of music and location information in his paper [16]: this is a common issue in this field. We will improve the method to recommend more diverse lyrics depending on the characteristics of tourist spots by using more specific features for each spot.

V. Conclusion

In this paper, we proposed a cross-domain lyrics recommendation system based on the distributed representation of words. The vectors of tourist spot reviews were generated by using the distributed representation model with lyrics corpus. Then, the tourist spot reviews were assumed as a type of lyrics in the proposed system. The system merged the vectors of tourist spot reviews to location vectors and calculated the similarity between location vectors and lyric vectors. The system finally selected the lyrics with the highest similarity to the arbitrary tourist spot as the recommendation result. During the experiment, we found a tendency that several locations corresponded to the same lyric in the recommendation results. We confirmed some commonalities among those locations corresponding to the same lyric through the survey. This discussion will help the future work of this research to achieve better recommendations.

This paper is still a work in progress, so the objective evaluation of the recommendation result will be one of the tasks in the future. Also, we should carry out subjective evaluation experiments to verify usability in real use cases.

Acknowledgment

This work is in part supported by KAKENHI #16K21482, #15K12151, and #19K12567. The data of lyrics is collected by Assist. Prof. M. Yoshida at Toyohashi Institute of Technology, Japan. We show our best appreciation.

References

[1] Z. Cheng and J. Shen, "On effective location-aware music recommendation," ACM Trans. Inf. Syst., vol. 34, no. 2, April 2016. [Online]. Available: https://doi.org/10.1145/2846092

[2] M. Pichl, E. Zangerle, and G. Specht, "Improving context-aware music recommender systems: Beyond the pre-filtering approach," in Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval, ser. ICMR ' 17. New York, NY, USA: Association for Computing Machinery, 2017, p. 201 - 208. [Online]. Available: https://doi.org/10.1145/3078971.3078980

[3] S. Pauws, "Cubyhum: A fully operational query by humming system," in Proceedings of ISMIR 2002, 2002, pp. 187–196.

[4] D. Wang, T. Li, and M. Ogihara, "Are tags better than audio features? the effect of joint use of tags and audio content features for artistic style clustering," pp. 57–62, 2010.

[5] E. Pampalk, A. Flexer, and G. Widmer, "Improvements of audio-based music similarity and genre classification," in Proceedings of the 6th International Conference on Music Information Retrieval, 01 2005, pp. 628–633.

[6] K. Tsukuda, K. Ishida, and M. Goto, "A lyrics-based music exploratory web service by modeling lyrics generative process," in Proceedings of the 18th International Conference on Music Information Retrieval, 2017, pp. 544–551.

[7] R. Cai, C. Zhang, S. Wang, L. Zhang, and W.-Y. Ma, "Musicsense: contextual music recommendation using emotional allocation modeling," in Proceedings of the 15th ACM international conference on Multimedia, 2007, p. 553 - 556.

[8] M. B. abd Marius Kaminskas and F. Ricci, "Location-aware music recommendation," International Journal of Multimedia Information Retrieval, vol. 2, no. 1, 2013, pp. 31–44.

[9] I. Fernández-Tobias, I. Cantador, M. Kaminskas, and F. Ricci, "Cross-domain recommender systems: A survey of the state of the art," in Proceedings of the 2nd Spanish Conference on Information Retrieval, 2012, p. 187 - 198.

[10] M. Kaminskas, F. Ricci, and M. Schedl, "Locationaware music recommendation using auto-tagging and hybrid matching," in Proceedings of the 7th ACM conference on Recommender systems, 2013, p. 17 - 24.

[11] T. Mikolov, W.-t. Yih, and G. Zweig, "Linguistic regularities in continuous space word representations," in Proceedings of HLT-NAACL, 2013, pp. 746–751.

[12] "AZLyrics - Song Lyrics from A to Z," https://www.azlyrics.com/, last accessed on 01/28/20.

[13] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," CoRR, vol. abs/1301.3781, 2013.

[14] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in Advances in Neural Information Processing Systems 26. Curran Associates, Inc., 2013, pp. 3111–3119.

[15] "Tripadvisor: Read Reviews, Compare Prices & Book," https://www.tripadvisor.com/, last accessed on 01/28/20.

[16] M. Kaminskas and F. Ricci, "Emotion-based matching of music to places," in Emotions and Personality in Personalized Services, J. Fagerberg, D. C. Mowery, and R. R. Nelson, Eds. Springer, 2016, pp. 287–310.

 ISBN: 978-1-61208-761-0