LyricsRadar: A Lyrics Retrieval System Based on Latent Topics of Lyrics

Shoto Sasaki† Kazuyoshi Yoshii‡ Tomoyasu Nakano* Masataka Goto* Shigeo Morishima†
† Waseda University ‡ Kyoto University * National Institute of Advanced Industrial Science and Technology (AIST)

BACKGROUND

Lyrics is very important! But it’s difficult to find songs with my favorite lyrics.

We usually enjoy the lyrics only after listening to its song.

GOAL

To assist listeners to encounter songs with unfamiliar but interesting lyrics

CONVENTIONAL SYSTEM

Conventional lyrics retrieval system based on simple word search is used after listening to its song.

In addition, those system often fail to reflect user’s intention behind a query.

E.g., ‘tear’

a drop of salty liquid that comes out of your eye when you are crying
to damage something such as paper by drawing it hard or letting it touch something sharp

APPROACH

LyricsRadar analyzes the lyrics topics by using latent Dirichlet allocation (LDA) and visualizes those topics to help users find their favorite lyrics interactively.

The LDA can estimate various lyrics topics five typical topics common to all lyrics in a given database were chosen.

OVERVIEW

LyricsRadar has two visualization functions.

- the topic radar chart
- the lyrics map

We use favorite lyrics as a query

FUNCTIONALITY

Let’s browse lyrics

Interactively browse lyrics

Directly change the topic radar chart

Retrieve lyrics by songwriter

EVALUATION

When the lyrics of a song are selected, relative location to other lyrics of the same artist or songwriter in the space is investigated.

The lyrics of a song are selected at random in the space as basis lyrics and also target lyrics of four songs are selected to be compared according to the following conditions.

We used the results of LDA for the lyrics of the 6902 J-POP songs.

17 All Japanese speakers
age: 21 - 32

Presentation order was random.

The score of (1) was the closest to 1.0, showing our approach to be effective.

As the subjects’ responses about the reason of decision, we obtained such responses as a sense of the season, positive-negative, love, relationship, color, light-dark, subjective-objective, and tension.

CONCLUSION

We propose LyricsRadar, an interface to assist a user to come across favorite lyrics interactively.

Our main contribution is to visualize lyrics in the latent meaning level based on a topic model by LDA.

Future Work

- the visualization method that can reflect more numbers of topics by keeping an easy-to-use interactivity
- improvements to topic analysis by using hierarchical topic analysis