

## Lyricon: A Visual Music Selection Interface Featuring Multiple Icons

Wakako Machida  
 Ochanomizu University  
 Tokyo, Japan  
 Email: matchy8@itolab.is.ocha.ac.jp

Takayuki Itoh  
 Ochanomizu University  
 Tokyo, Japan  
 Email: itot@is.ocha.ac.jp

**Abstract**—This paper presents "Lyricon", a technique that automatically selects multiple icons of tunes block-by-block, and effectively displays the icons. Here, Lyricon selects icons based on not only musical features, but also lyrical keywords. In other words, Lyricon can reflect not only the features of the tunes but also the story of lyrics on its icon selection. Users can understand both impression of the sounds and the content of the lyrics, and they can choose songs which is suitable for their feeling based on the visual impression of the icons. Besides, embedding Lyricon on GUIs of music players is convenient to play specific parts of songs.

**Keywords**-Visualization, Icon selection, Lyrics, Music player

### I. INTRODUCTION

Today, people can enjoy their favorite songs easily and freely due to the evolution of portable music player products and free video sharing Web sites. We often demand smoothly working song selection mechanisms because we often store large number of tunes. Sometimes user cannot remember the melody of the song and the lyrics, only by looking the titles or artist names of the tunes on the song selection panels. We think that "music visualization" is useful to solve the problem.

Lyrics are very important on recent popular hit songs. There are many big sale songs which lyrics are key points, such as answer songs those lyrics reply to another song, and songs on compilation albums which collect multiple artists' songs following particular themes (e.g. "sea", "love"). As a feasibility study for lyrics, we asked the following two questions to 86 students in our university:

- "Are you usually conscious of lyrics while listening to the music?"
- "Do you often choose the music based on lyrics?"

We got results that 66 students answered "yes" for the former question, and 42 students answered "yes" for the latter question. These results indicate that lyrics may be informative for many people while selecting the songs they want to listen to.

This paper presents "Lyricon", a music visualization technique that represents musical structure by multiple icons, taking lyrics into account. Lyricon automatically selects multiple icons of songs block-by-block, and provides a user interface to effectively display the icons. Lyricon selects

icons based on not only musical features, but also lyrical features. We have designed Lyricon to represent musical and lyrical features by multiple icons, not by a single icon, because story of lyric may be too complex to explain by a single picture. We can apply this idea to selection of larger pictures as well as icons; however, in this paper we focus on icon selection, because icons are smaller than pictures, and therefore suitable to display in limited sizes of GUIs.

Figure 1 shows examples of selected icons. Users can understand both impression of the sounds and the content of the lyrics when they look at the icons selected by Lyricon, even before listening to the song. This paper discusses the detail in Section III. Moreover, embedding Lyricon on GUIs of music players is convenient for partial play of tunes. This paper presents potential user interfaces of Lyricon in Section IV.

### II. RELATED WORK

There have been several novel techniques for icon generation or selection. Setlur et al. presented Semantics [1] that synthesizes small pictures to generate semantics-matched icons. It can represent various semantics by single icon; however, the technique is not music-specific. Music Icons [2] by Kolhoff et al., and MIST [3] by Oda et al., generate/select icon pictures according to musical features. However, these techniques do not take lyrics into account. Moreover, these techniques assign only one icon for each tune, and therefore it may be often difficult to represent changes or structures of tunes.

There have been several novel techniques for visual representation of lyrics-based information. Xu et al. presented a technique to create slide shows according to the story of lyrics [4]. Framework of Lyricon is also applicable to slide show generation; however, we preferred to develop an icon selection technique, because we think we can understand the features and story of the songs more quickly by looking at sequences of icons. Neumayer et al. presented a technique to visually represent clusters of songs taking both features and lyrics into account [5].

### III. MULTIPLE ICON SELECTION

This section describes our multiple icon selection technique consisting of the following technical components:

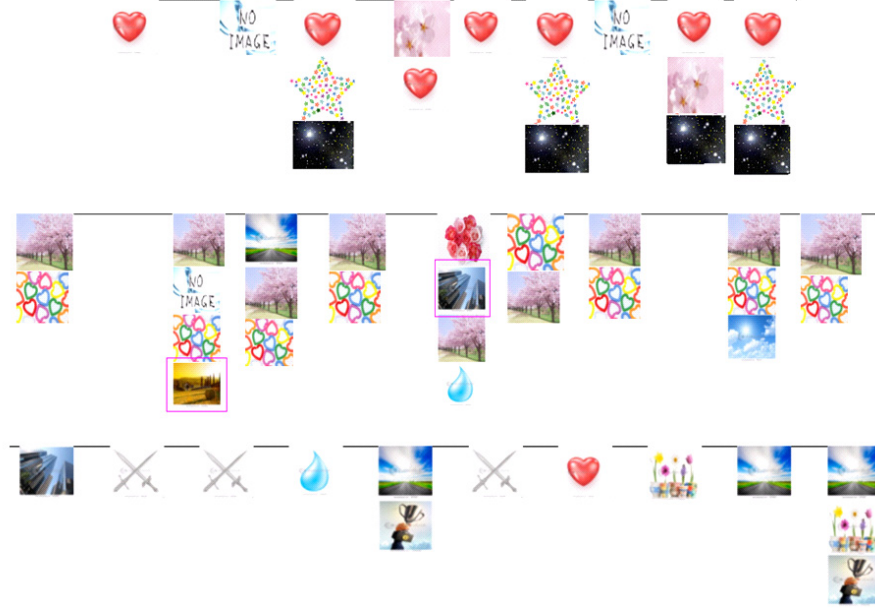


Figure 1. Examples of icon selection results of three songs.

- 1) Preparation during system development:
  - a) Category and keyword selection
  - b) Musical feature selection
  - c) Icon selection
- 2) Process during input of new songs:
  - a) Lyrical processing for icon candidate selection
  - b) Musical feature processing for final icon selection

Following sections describe these processing units. Section III-A to Section III-C describes preparation processes. Section III-B describes steps 1 to 3. Section III-D and III-E describes lyrical and musical processes for icons election. Note that our implementation supports songs those lyrics are written in Japanese, but the mechanism of Lyricon is not limited to Japanese songs.

#### A. Preparation(1): Category and keyword selection

Our implementation prepares "categories", and sets multiple keywords and icons into the categories. A category contains a set of related keywords that can be the main theme of the lyrics and its synonyms. At the same time, the category contains a set of multiple icons which bring different impression. Let us formulate categories, keywords, and icons, as following:

- Categories as  $C = \{c_1, \dots, c_{N_c}\}$ , where  $N_c$  is the number of categories.
- Keywords belonging to  $c_i$  as  $K_i = \{k_{i1}, \dots, k_{iN_{ki}}\}$ , where  $N_{ki}$  is the number of keywords belonging to  $c_i$ .
- Icons belonging to  $c_i$  as  $X_i = \{x_{i1}, \dots, x_{iN_{xi}}\}$ , where  $N_{xi}$  is the number of icons belonging to  $c_i$ .

- Adjectives of icon  $x_{ij}$  as  $A_{ij} = \{a_{ij1}, \dots, a_{ijN_{aij}}\}$ , where  $N_{aij}$  is the number of adjectives of  $x_{ij}$ .

To define categories, we first asked the following question to 86 students in our university:

"What kind of topics or themes do you occasionally want to use to select songs?"

We got many topics as the result of the question, and used 23 topics such as "Love", "Summer", and "Christmas", as categories. Then, we extracted synonym of the categories from Japanese thesaurus dictionary [6], and selected frequently used words as keywords. We scanned the lyrics of randomly selected a lot of Japanese hit songs, and divided the lyrics into words. We used "Chasen" [7] for this process. Then, we matched the words extracted from the lyrics with the synonym extracted the dictionary, and finally selected frequently matched synonym as keywords of the categories. Our current implementation prepares 23 categories and 248 keywords.

#### B. Preparation(2): Musical feature selection

At the same time, Lyricon uses several musical feature values. Our current implementation uses MIRtoolbox [8], working on MATLAB, for musical feature calculation. We had a feasibility study of features calculated by MIRtoolbox by applying randomly selected 26 Japanese hit songs. We subjectively estimated that 10 features were especially meaningful for our purpose. We then selected 3 features from the 10 features by the following procedure. We assigned pairs of inverse meaning of adjectives for each feature. For example, we assigned "fast" and "slow" for the feature "Tempo". At the same time, we calculated the 10 feature

Ex "LOVE" = {"love", "heart", "like", "fancy", "Amor", "charity" ...}

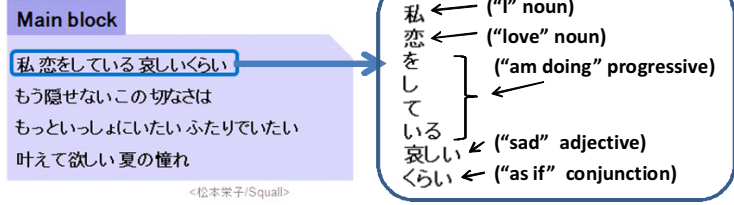
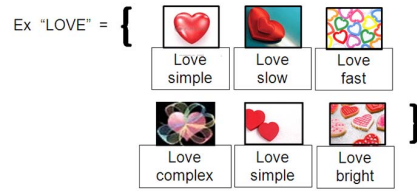


Figure 2. (Left) Keywords and icons of a category. (Right) Morphological analysis.

values for the songs as shown in Figure 2(Right). We then applied the following procedure for each feature. We asked 6 examinees to subjectively divide the 26 songs into two groups according to one of the pairs of adjectives, such as "fast" and "slow". At the same time, we divided the songs according to an automatically determined threshold value of the feature corresponding to the pair of adjectives, such as "Tempo". We then calculated the average concordance rate between the subjective and automatic division. Finally, we determined the feature can be used for Lyricon, if the two division results were sufficiently similar.

As a result of above mentioned process, we selected the following three features:

- *Tempo*. "slow" and "fast" are used as adjectives.
- *Percentage of high-tone range*. "simple" and "rich" are used as adjectives.
- *Percentage of inharmonic tones*. "primitive" and "complex" are used as adjectives.

Here, the word "simple" may point songs which sound naively because of small number of musical instruments or less sound effects. On the other hand, the word "rich" may point songs which feature many musical instruments or use various sound effects. The word "complex" may point songs which apply complicated chords or tensions like Jazz music. On the other hand, the word "primitive" may point songs which do not apply complicated chords or tensions.

### C. Preparation(3): Icon selection

Lyricon supposes to prepare several icons for each category. It also supposes that one or more adjectives, slow, fast, simple, rich, primitive, or complex are assigned to every prepared icon. The assigned adjectives are referred to select icons. Figure 2(Left) shows an example of a category "love", which contains 9 keywords, and 6 icons. Lyricon firstly specifies the category corresponding to the contents of song by keyword matching between each category and lyrics, and then selects the most adequate icon in the specified category based on musical features, referring the adjectives of icons.

During our experiments, we prepared enough number of icons for each category. We then showed the icons and adjectives assigned to the icons to 12 examinees, and asked

if the adjectives matched to the icons. We did not use icons which less than half of examinees agreed that they matched.

### D. Lyrical analysis for icon candidate selection

Since Lyricon assigns icons block-by-block, we would like to use lyrics divided based on blocks of the songs. We used "Lyric Master" [9] to obtained lyrics of Japanese hit songs which are divided block-by-block.

Lyricon then analyzes morphologic of each block and divides the block of the lyric into words by using "Chasen". Figure 2(Right) shows an example of the morphological analysis.

Let us describe a set of words in a block as  $W = \{w_1, \dots, w_N\}$ . If a word  $w_k$  completely matches to the keyword  $k_{ij}$ , Lyricon determines that the block is related to the category  $c_i$ . In this case Lyricon treats the set of icons  $X_i$  as the candidates to be assigned to the block, and finally one of the icons  $x_{ij}$  is assigned to the block.

### E. Musical feature analysis for final icon selection

Lyricon also calculates feature values selected in the preparation process. We selected three features, "Tempo", "Percentage of high-tone range", and "Percentage of inharmonic tones", as described in Section 3.1. Lyricon then selects the adjectives of the song according to the calculated feature values. Our current implementation selects at least one adjective from the selected 6 adjectives described in Section 3.1 by the following procedure.

Let the three feature values  $F_1$ ,  $F_2$ , and  $F_3$ , and these ranges  $[F_{1min}, F_{1max}]$ ,  $[F_{2min}, F_{2max}]$ , and  $[F_{3min}, F_{3max}]$ . Here, we define the relevance of the song to the two adjectives of the  $i$ -th feature value as  $R_{ia}$  and  $R_{ib}$ , calculated as  $R_{ia} = F'_i$ , and  $R_{ib} = 1.0 - F'_i$ .

When there are multiple icon candidates in a same category specified from a block of the lyric, Lyricon selects one of the icons which is assigned the adjective bringing the maximum  $R_{ia}$  or  $R_{ib}$  value.

## IV. USER INTERFACE

We implemented visual music selection interfaces that display the multiple icons selected by our technique on two platforms: Windows PC and Android OS.

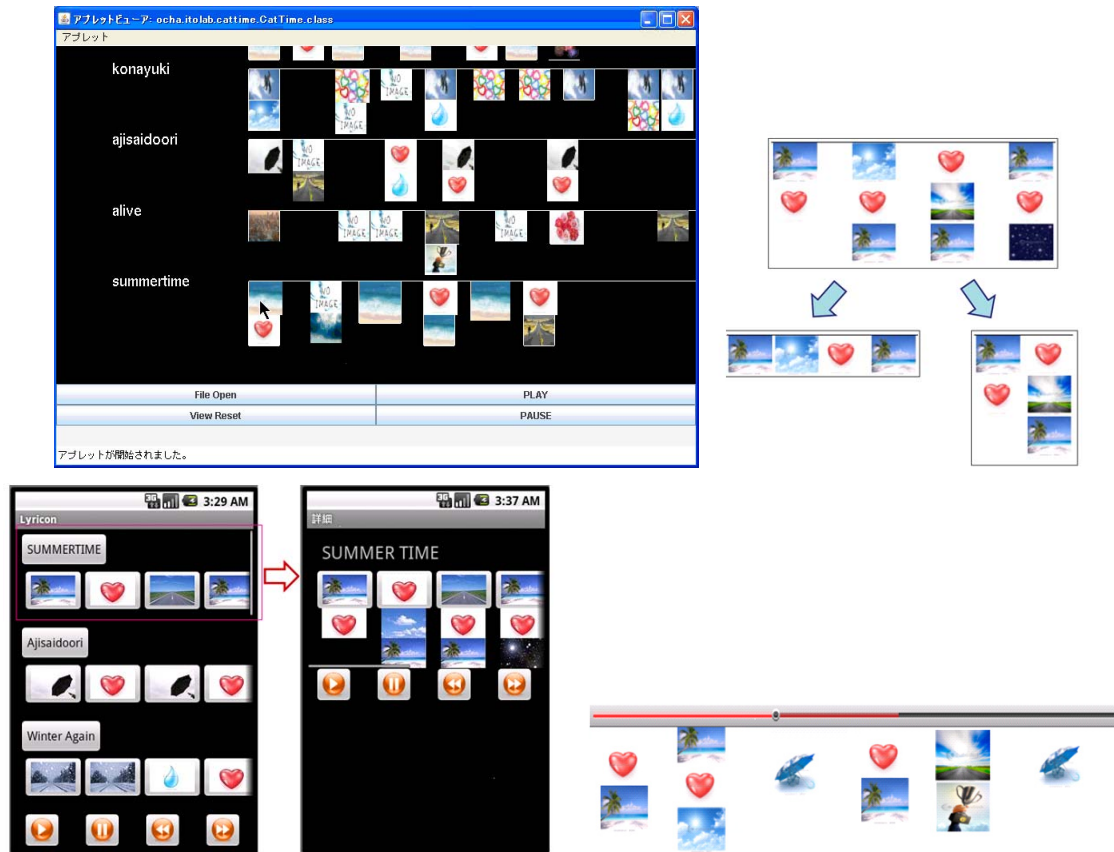


Figure 3. (Upper-Left) User interface implemented for Windows PC. (Upper-Right) Zooming user interface. (Lower-Left) User interface implemented for Android OS. (Lower-Right) A scroll-bar of a music player software

Figure 3(Upper-Left) shows our implementation of the user interface for Windows PC. It horizontally displays a set of icons for one song, and vertically aligns the sets of icons. Users can select their favorite songs by clicking the names of songs. They can start or stop of playing the songs by pressing the downside buttons. Also, users can click icons so that Lyricon can start the play of songs from the corresponding parts of the songs.

Here, this mechanism occupies large area of window space to completely show the selected icons of many songs. To solve the problem, we implemented a level-of-detail mechanism to control the number of displayed icons. It vertically reduces the number of displayed icons according to change of heights of windows. Also, it vertically reduces the number of displayed icons according to change of widths of windows. Figure 3(Upper-Right) illustrates the mechanism.

Figure 3(Lower-Left) shows our implementation of the user interface for Android OS. It features start, pause, next, and previous buttons as orange buttons. It also features horizontal and vertical scroll bars: users can browse icons through a song using the horizontal scroll bar, and many

songs using the vertical scroll bar. The user interface initially displays the most important icon for each block. When a user clicks a name of a song, the user interface zooms up so that all icons of the specified song are displayed.

We think that Lyricon can be also applied to icon indication along scroll-bars of media players. Figure 3(Lower-Right) shows an illustration of the application. Users can easily play the song from any block that they want to listen, by looking at the icons.

## V. RESULTS

### A. Examples

Figure 1 shows examples of the multiple icon selection results applied to real Japanese hit songs. Theme of the song displayed in Figure 1(upper) is love, and the icon selection result clearly represents the theme. Theme of the song displayed in Figure 1(center) is Japanese cherry, which blossoms during graduation and enrollment season in Japan. It displays an icon of country and an icon of city, because the central character of the song graduates from a school in a country and then moves to a city. The icon selection result narrates such story of the song. The former part of lyrics

displayed in Figure 1(lower) negatively contains keywords "fight" and "tear", but the latter part of the lyrics positively contains keywords "love" and "flower". The icon selection result well represents the change of the nuance.

These results demonstrate that the icon selection results by Lyricon clearly represents the theme of the songs, and story of the lyrics.

### B. Subjective evaluation of icon selection results

We showed printed icon selection results for examinees, and asked them several questions. Examinees were 13 female university students majoring computer science.

1) *Impression of songs associated by looking at icons:* We asked 50 examinees to look at sequences of icons which expressed the whole songs, and asked to answer their impression of the songs. We prepared 10 sequences of icons, and asked to freely write keywords imagined to be contained in the lyrics of the songs, and impressions imagined to be led from musical features of the songs. We extracted adequate keywords and impressions from their answers, and calculated adequate answer rates, which are the rates of the number of adequate answers against the total number of answers.

Table I shows the adequate answer rates for keywords and impressions. This result denotes that the answers were very adequate for several songs (e.g. icon set 2, 5, and 6). The icon set 2 is selected for a summer love song, which have fast and bright musical features. Figure 1(Upper) shows this icon set. Lyricon successfully selected icons of summer and love icons, and examinees adequately mentioned keywords including summer and love, and impressions including fast and bright.

On the other hand, we could not get high adequate answer rates for some of other songs. Icon set 7 was selected for a song of family love, but many examinees imagined a song of love between a man and a woman looking at icons of heart. We need to apply more sophisticated natural language processing techniques to distinguish between family love and man-woman love. Icon set 10 was selected for a song of urban life struggling against business and economics. We did not prepare adequate categories and keywords for such songs. We need to prepare wider categories and keywords for more variety of songs.

2) *Selection of icons from lyrics:* We asked 50 examinees to read lyrics of 7 songs, and then choose the best sequence of icons for the lyrics. We prepared 10 similar sets of icons for each song, and asked them to choose one of the icon sets for the song.

Table II shows correct answer rates, which are the rates of the numbers of correctly answered examinees against the total number of them. This result denotes that the rates were totally good, and Lyricon works well to associate semantics of songs by looking at the sequences of icons. On the other hand, this result denotes that there were several mistakable

icon sets. For example, lyric 2 in Table II was a Christmas song, but some of examinees selected another icon set shown in Figure 1(upper). We assume that some of examinees associated the Christmas day from icons of hearts and starts from the wrong icon set. We would like to gather such mistakable examples and discuss how to improve in our future experiments.

## VI. CONCLUSION AND FUTURE WORK

This paper presented "Lyricon", a technique for automatically selecting multiple icons of tunes block-by-block. Lyricon firstly selects candidates of icons according to words of lyrics block-by-block, and then selects suitable icons from the candidates according to musical features. Lyricon also supports user interfaces to effectively and adaptively display the icons. This paper demonstrated the effectiveness of Lyricon with examples and subjective evaluation results.

As a future work, we would like to reexamine the icons and keywords. Section V-B discussed that several songs bring mistakable results, or consist of important words which are not prepared by our implementation. We would like to prepare more categories, keywords, and icons to support more variety of lyrics. Also, we would like to have more experiments to find more mistakable results and discuss how to improve.

Also, we need to improve the implementation. Section V-B also discussed that interpretation of lyrics is mistakable while Lyricon just extracts keywords. We would like to apply more sophisticated natural language processing techniques to solve the problem. Another issue is that Table I suggests adequate answer rates for impressions were a little bit worse than those for keywords. We think one reason may be selection of musical features, and therefore we would like to discuss how to improve the processes of musical features.

As another future work, we would like to extend Lyricon to allow users to edit the categories, and to add keywords and icons. For example, we assume it is effective if users can add their favorite photographs or original pictures which they can easily imagine the songs.

## REFERENCES

- [1] V. Setlur, C. Albrecht-Buehler, A. A. Gooch, S. Rossoff, B. Gooch, Semantics: Visual Metaphors as File Icons. Computer Graphics Forum (Proc. of Eurographics), Vol. 24, No. 3, pp. 647-656, 2005.
- [2] P. Kolhoff, J. Preub and J. Loviscach: Music Icons: Procedural Glyphs for Audio Files, Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'06), pp. 289-296, 2006.
- [3] M. Oda and T. Itoh: MIST: A Music Icon Selection Technique Using Neural Network, NICOGRAPH International, 2007.
- [4] S. Xu, T. Jin, F. C. M. Lau, Automatic Generation of Music Slide Show Using Personal Photos, 10th IEEE International Symposium on Multimedia, pp. 214-219, 2008.

Table I  
ADEQUATE ANSWERS RATES, WHERE ANSWERS ARE KEYWORDS IMAGINED TO BE CONTAINED IN THE LYRICS, AND IMPRESSIONS IMAGINED TO BE  
LET FROM MUSICAL FEATURES.

Icon set	1	2	3	4	5	6	7	8	9	10
Rate(keyword)	0.97	0.96	0.87	0.96	0.94	0.86	0.18	0.98	0.85	0.25
Rate(impression)	0.57	0.88	0.62	0.63	0.98	0.85	0.33	0.35	0.24	0.28

Table II  
CORRECT ANSWER RATES OF SELECTION OF ICON SETS FOR LYRICS OF SONGS.

Lyric	1	2	3	4	5	6	7
Rate	0.92	0.76	0.82	0.80	1.00	0.84	0.96

- [5] R. Neumayer, A. Rauber, Multi-Modal Music Information Retrieval - Visualisation and Evaluation of Clusterings by Both Audio and Lyrics. 8th International Conference on Computer-Assisted Information Retrieval, 2007.
- [6] Japanese WordNet, <http://nlpwww.nict.go.jp/wn-ja/>
- [7] Chasen, <http://chasen.naist.jp/hiki/Chasen/>
- [8] O. Lartillot, MIRtoolbox, <http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- [9] K. Maehashi, Lyric Master, <http://www.kenichimaehashi.com/lyricsmaster/>