

## Lyricon: 複数アイコンによるビジュアルな 音楽選択インタフェース

町田 和嘉子<sup>†1</sup> 伊藤 貴之<sup>†1</sup>

本報告では、歌謡曲を複数アイコンで表現し、それを音楽プレイヤーのプレイリスト等に表示することで、ビジュアルから直感的に聴きたい音楽を選択させるユーザインタフェース“Lyricon”を提案する。旧来から歌詞がヒット要因となった歌謡曲は多数存在することから、著者らは歌詞に着目してアイコン自動選択機能を実装している。具体的には、あらかじめ登録されたキーワードを歌詞から抽出して、それに対応するアイコン候補画像を列挙し、楽曲特徴に基づいてその候補画像の一つをアイコンに選ぶ。以上の処理をAメロ、Bメロ、サビなどのブロックごとに適用することで、複数のアイコンで1曲の歌謡曲を表現する。このアイコン選択結果を見ることでユーザは、歌謡曲を聴かなくても曲の雰囲気や歌詞内容を把握できるようになり、大量の曲の中からその日の気分にあった曲の選択を視覚的に実現できると考えられる。あわせて著者らは、Lyriconのアイコン自動選択結果を縦に並べて表示したビジュアルなプレイリストを、WindowsPCとAndroid携帯の2つのプラットフォーム上で実装した。

## Lyricon: Visual Music Selection Interface with Multiple Icons

WAKAKO MACHIDA<sup>†1</sup> and TAKAYUKI ITOH<sup>†1</sup>

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

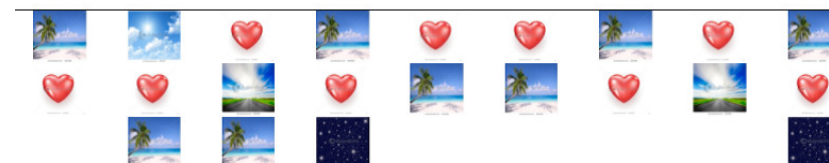


図1 アイコン自動選択結果 (1) “SUMMER TIME” song by NEWS

icons. Besides, embedding Lyricon on GUIs of music players is convenient to play specific parts of songs.

### 1. 概要

近年のインターネット上のMP3形式音楽配信の急速な普及により、携帯音楽プレイヤーなどで大量の音楽を持ち歩くことが可能になっている。しかし所有する曲の曲調や歌詞内容を全て把握するのは難しく、特に昔登録した曲やあまり聞いていなかった曲の認識が困難であり、曲選択の際に間違いが起りやすいと考える。また事前調査として86人の学生に対し「普段歌詞を意識して音楽を聴くことはあるか」というアンケートを実施したところ、66人がYesと答える結果が得られた。このことから歌謡曲において、歌詞は重要な存在であると考えられる。

そこで本報告では、楽曲の印象や歌詞の内容に合ったアイコンを、Aメロ、Bメロ、サビなどのブロックごとに選択し、再生時刻順に並べて表示させることで、曲の印象や内容の時間変化を視覚的に表現するユーザインタフェース“Lyricon”を提案する。Lyriconは、歌詞から抽出されたキーワード、楽曲特徴量の両面からブロックごとにアイコンを選択し、図1~4のように一覧表示する。そのためアイコンの一覧表示結果を見れば、実際に曲を聴かなくてもその曲の雰囲気や歌詞内容の展開を想像でき、音楽選択が行いやすくなると思われる。

### 2. 関連研究

Lyriconと同様に、楽曲に合ったアイコンを生成または選択する技術は、既にいくつか報

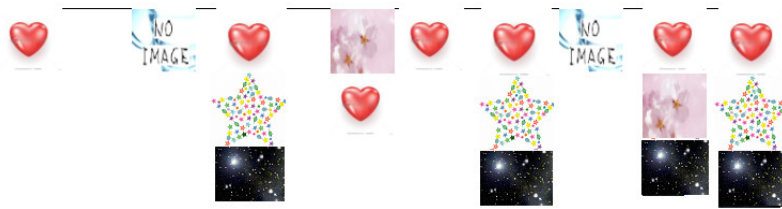


図2 アイコン自動選択結果 (2) "CHE.R.R.Y" song by YUI

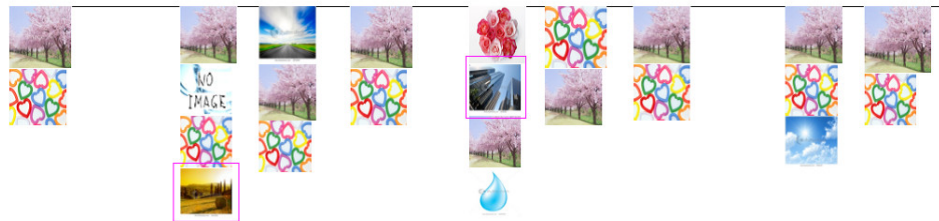


図3 アイコン自動選択結果 (3) "SAKURA" song by いきものがかり



図4 アイコン自動選択結果 (4) "alive" song by Mr.Children

告されている．Kolhoff らによる Music Icons<sup>1)</sup> と小田らによる MIST<sup>2)</sup> は，楽曲の特徴量に基づきアイコンを生成または選択する．しかしこれらの手法は歌詞を考慮していない上に，1 曲あたり 1 つのアイコンしか割り当てない．

また，文脈に基づくアイコン表現技術とし，Setlur らは Semanticons<sup>3)</sup> を提案している．Semanticons はあらかじめ用意された小さな絵を結合することで，任意の意味を表現するアイコンを生成する．ただしこの技術は音楽に特化していない．

アイコンに限定せず視覚情報と音楽情報を連携した技術は Lyricon の他にも多数研究さ

れている．初期の手法として Foote<sup>4)</sup> は音響情報の自己相似性を可視化する手法を提案している．最近の研究として藤澤<sup>5)</sup> らは音楽のムードを色彩で表現する手法を提案し，後藤ら<sup>6)</sup> は楽曲のメタファに「流す」「くっつける」といった動作を加えた手法を提案している．また，Chen<sup>7)</sup> らは感情（高レベル情報）と特徴量（低レベル情報）の両者に基づいてマッチングされた写真により楽曲を表現する手法を提案している．これらの手法はいずれも楽曲の特徴やメタデータなどと視覚情報を関連付けたものであり，歌詞を直接扱うものではない．

歌詞情報を視覚的に表現する技術としては Xu<sup>8)</sup> らによる歌詞内容に基づくスライドショー作成技術や，Newmayer<sup>9)</sup> らによる音楽特徴と歌詞による音楽クラスタの視覚表現技術などがあげられる．また Lyricon はこれらの表現をアイコンに置き換えたものと考えられる．

以上の手法とは全く逆に，視覚情報に合わせた楽曲を推薦する手法もいくつか提案されている．大山<sup>10)</sup> らは指定された写真の印象に合った音楽アレンジを提供する手法を提案している．また桐本<sup>11)</sup> らは指定された風景画像から抽出されたコンテキストに近い楽曲を推薦する手法を提案しており，この中でコンテキストとのマッチングに歌詞を参照している．

### 3. Lyricon によるアイコン自動選択

本章で説明する Lyricon のアイコン自動選択の処理過程は，箇条書きにすると以下の通りである．

- システム開発中の準備処理：
  - カテゴリ・キーワードの選定
  - 音楽特徴量の選定
  - アイコンの選定
- 楽曲入力時の処理：
  - 歌詞の処理・アイコン候補の列挙
  - 特徴量の処理・アイコンの最終選択

なお，著者らの現時点での Lyricon の実装では，日本語の歌詞を持つ歌謡曲のみを対象としている．

#### 3.1 準備処理 (1): カテゴリ・キーワードの選定

Lyricon ではまず“ カテゴリ ”を用意する．著者らはカテゴリを歌詞のメインテーマになり得る単語と定義する．また Lyricon では，各々のカテゴリに対して，関連する複数のキーワードを持たせるとともに，そのカテゴリを表す複数のアイコンを持たせる．著者らは Lyricon においてカテゴリとキーワード，アイコンを以下のように定式化する．

- カテゴリを  $C = \{c_1, \dots, c_{N_c}\}$  と定義する． $N_c$  はカテゴリの総数である．

- $c_i$  に属するキーワードを  $K_i = \{k_{i1}, \dots, k_{iN_{k_i}}\}$  と定義する.  $N_{k_i}$  は  $c_i$  に属するキーワードの総数である.
- $c_i$  に属するアイコンを  $X_i = \{x_{i1}, \dots, x_{iN_{x_i}}\}$  と定義する.  $N_{x_i}$  は  $c_i$  に属するアイコンの総数である.
- アイコン  $x_{ij}$  を表す形容詞を  $A_{ij} = \{a_{ij1}, \dots, a_{ijN_{a_{ij}}}\}$  と定義する.  $N_{a_{ij}}$  は  $x_{ij}$  に割り当てられた形容詞の総数である.

著者らはカテゴリの選定に際して, 86 人の学生を対象として, 聴きたい歌謡曲を選択する際に, どんなテーマの歌詞を選ぶことがあるか, というアンケートを採取した. その結果をもとに, 「恋」「夏」「海」「クリスマス」などの 26 単語をカテゴリとして採用した. さらに著者らは, 無作為選択した歌謡曲 50 曲の歌詞を, 形態素解析ソフトウェア茶筌<sup>12)</sup> を用いて単語に分割し, 日本語シソーラス辞書から抽出したカテゴリの類義語と照合し, 共通する 248 単語を各カテゴリに属するキーワードとして採用した.

### 3.2 準備段階 (2): 音楽特徴量の選定

Lyricon はまず各カテゴリを構成するキーワードと歌詞から抽出した単語のマッチングにより, 各ブロックに対応するカテゴリを決定し, このカテゴリが有するアイコンをアイコン候補とする. 続いて音楽特徴量に基づき, 各カテゴリに含まれるアイコンから最適なものを選択する.

ここで, アイコン選択のために参照する音楽特徴量を何を採用するかで, アイコン選択結果は大きく変わる. 著者らは以下の方法で, 音楽特徴量の種類を選定した. まず著者らは, 試験的に無作為抽出した 26 曲のサンプル曲について, 数値解析ソフトウェア MATLAB の上に実装された楽曲特徴分析パッケージ MIRtoolbox<sup>13)</sup> を用いて 10 種類の特徴量を抽出した. 続いて著者らは, その 10 種類の特徴量の各々に対して, 2 つの相反する意味を持つ形容詞対を定めた. たとえば, “Tempo” という, 曲のテンポを表す特徴に対しては, “速い” と “遅い” という形容詞を与えた.

続いて著者らは, 10 種類の特徴量の妥当性を検証した. 具体的にはまず, 6 人の学生被験者に先ほどの 26 曲のサンプル曲を聴いてもらい, 各曲が 10 種類の形容詞対のそれぞれどちらに該当するかを, 被験者らの感性によって選択させた. それと同時に,  $i$  番目の特徴量  $f_i$  の境界値を  $F_i$  とし,  $f_i$  値が  $F_i$  より大きいかが否かでサンプル曲を 2 分した. ここで,  $f_i$  に関する 2 つの形容詞 A, B のうち被験者が A, B を選択する確率をそれぞれ  $p_A, p_B$  とし, サンプル曲を 2 分する前のエントロピーを  $E_b(p_A, p_B)$ , 2 分した後の各々のクラスタにおけるエントロピーを  $E_{a1}(p_A, p_B)$ ,  $E_{a2}(p_A, p_B)$  とする. ここで著者らは, 以下の数式

表 1 採用した特徴と形容詞

| 特徴                  | 形容詞      |
|---------------------|----------|
| テンポ (tempo)         | 遅い・速い    |
| 高音域の割合 (brightness) | 単純な・輝かしい |
| 不調和音の多さ (roughness) | 素朴な・複雑な  |

$$E = E_b(p_A, p_B) - (E_{a1}(p_A, p_B) + E_{a2}(p_A, p_B))$$

で定義される情報利得  $E$  を最大とする  $F_i$  値を求めた. そして, この処理を 10 種類の特徴量の各々について適用し, 情報利得の最大値が特に高い特徴量を採用した. 著者が採用した 3 種類の特徴量と, それらに割り当てた形容詞を, 表 1 に示す.

ここで用いた “単純な” という言葉は, 伴奏楽器が少ないなどの理由で倍音成分が小さい曲, “輝かしい” とは, さまざまな楽器が使用されているなどの理由で倍音成分が大きく, 派手でにぎやかな印象を与える曲, “複雑” はジャズのテンションに代表される複雑な和音構成を含む曲, “素朴な” とは童謡のように音楽理論的にもシンプルな構成の曲, というニュアンスを含んでいる.

### 3.3 準備段階 (3): アイコンの選定

著者らは各カテゴリに対し, 前節で選定した形容詞に対応する 6 種類のアイコンを選定した. 選定したアイコンの妥当性を検証するために, 12 人の学生被験者に “アイコン画像から受けるイメージ” を回答させ, 著者らの意図するイメージと合致した結果を得たものを採用した. そして各アイコンには, 前節で選定した形容詞をタグとして付与する. 図 5 は “恋” と “夏” カテゴリのアイコン例である. “恋” カテゴリには “好き” や “愛する” など 48 種のキーワードが, “夏” カテゴリには “夏” や “パカンス” など 14 種のキーワードが割り当てられている.

### 3.4 楽曲入力時の処理 (1): アイコン候補の列挙

アイコンを曲のブロックごとに選択するに先立ち, Lyricon は最初に歌詞をブロックごとに分割する. 著者らは歌詞のブロック分割に LyricMaster<sup>14)</sup> を用いている. 続いてブロックごとに分割した歌詞に形態素解析を適用し, 単語の最小単位に分割する. 著者らは形態素解析に茶筌<sup>12)</sup> を用いている. 例えば, 「私恋をしている 哀しいくらい」という歌詞 (松本栄子 “Squall” より引用.) は 「私/恋/を/し/て/いる/哀しい/くらい」というように分割される.

ここで, ある 1 つのブロックに含まれていた単語を  $W = \{w_1, \dots, w_N\}$  とする. もし単語  $w_k$  がキーワード  $k_{ij}$  と一致したら, Lyricon はこのブロックをカテゴリ  $c_i$  と関連付ける.

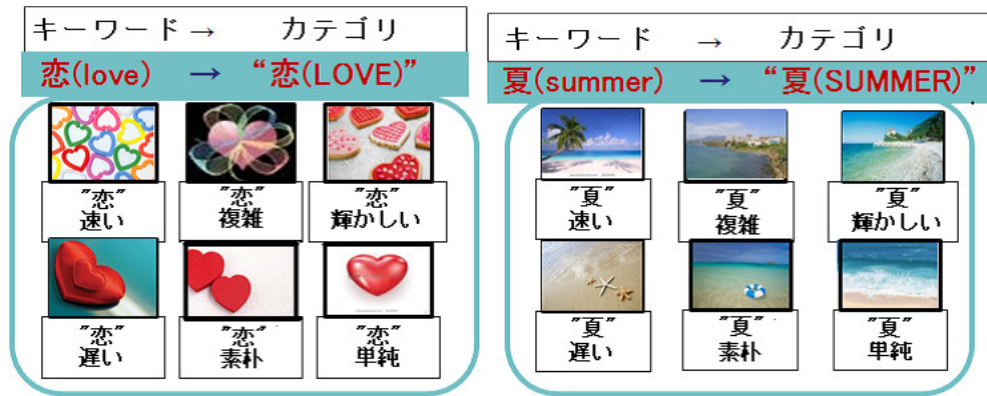


図5 アイコン候補

この場合、Lyricon はアイコン群  $X_i$  に含まれる複数のアイコンを表示用アイコンの候補とし、最終的に音楽特徴量によってこの中から一つのアイコンを決定する。前述の歌詞の例だとサビブロックの歌詞中に“夏”カテゴリに含まれるキーワード“夏”と、“恋”カテゴリに含まれるキーワード“恋”が存在する。その結果として、“夏”カテゴリに属するアイコンと“恋”カテゴリに属するアイコンがこの歌詞のサビブロックの候補となる。図5はそれらのアイコンの例である。

### 3.5 楽曲入力時の処理 (2): アイコンの最終決定

Lyricon は続いて、3.2 節で選定した“テンポ”、“高音域の割合”、“不調和な音の多さ”の3つの特徴量を算出する。この特徴量から、3.2 節で定めた6つの形容詞について、以下の手順でその曲への相応しさの順位を決定する。まず、ここで扱う数値を以下のように定式化する。

- 3.2 節で求めた境界値を  $F_i$  とする。  
( $i = \{tempo, brightness, roughness\}$ )
- 3.2 節で求めたサンプル曲の各特徴量の最大値と最小値の差を  $D_i$  とする。
- 算出した音楽特徴量を  $M_i$  とする。

以上の数値から Lyricon では、特徴の突出度  $R_i = \frac{|M_i - F_i|}{D_i}$  を求める。そして、 $R_i$  の値を最大にする  $M_i$  に対応する形容詞と、3.4 節で列挙されたアイコン候補のタグを比較し、同じ形容詞が設定されているアイコンを各ブロックにおける最終選択結果とする。

## 4. アイコン選択例

実際に Lyricon を用いて歌謡曲のアイコンを自動選択した例を、図1から図4に示す。図1が示す歌謡曲の歌詞のテーマは“夏”であり、図2が示す歌謡曲の歌詞のテーマは“恋”である。それを反映して、アイコン自動選択結果に海やハートのアイコンが目立っており、歌詞をイメージしやすい結果となっている。図3では、大部分を占めている春を表す桜のアイコンの中に、歌詞の前半部分のブロックには田舎を表すアイコン、後半部分のブロックには都会を表すアイコンが含まれており、この歌謡曲の主題である“卒業シーズンに見られる思い出と旅立ち”のテーマを表しているといえる。図4は、曲の前半部分には“戦い”や“涙”といった暗いイメージのアイコンが目立っているが、後半に行くにつれ、“ハート”や“花”といった明るいイメージのアイコンが増え、初めはやさぐれているがだんだんと自分を奮い立たせていく、という歌詞を表現している。

このように Lyricon によるアイコン自動選択結果が、歌詞の明確な主題だけでなく、歌詞のストーリー性も表現していることがわかる。ただし、歌詞内に比喩表現が多い、当て字が多い、などの難解な文章表現を含む歌詞において、現段階では妥当なアイコン選択結果を得ることが難しい場合もある。

著者らは Lyricon によるアイコン自動選択結果を検証するために、50人の学生被験者に対し、以下の2つのユーザテストを行った。

1つ目は、10曲のアイコン選択結果を回答者に提示し、どのような歌謡曲を連想するかを回答させた。回答の一部を表2に示す。その結果、実際の歌詞の意味や曲調などに合致する回答が、図2と図3では60%、71%を占めた。他の多くの歌謡曲においても、同様に良好な結果を得た。しかし図4では23%と、あまり好ましくない結果が出た。また回答の種類も多岐にわたり、一定の印象を与えることが出来ていなかった。その原因について学生被験者からのコメントを検討したところ、曲調と歌詞の印象が異なる場合に間違った回答が多いことがわかった。

2つ目は、7曲の歌謡曲の歌詞を回答者に読んでもらい、用意した複数のアイコン選択結果の中からその歌詞を適切に表しているものを回答させた。結果を表3に示す。こちらは7曲中の平均正答率は87%（最低76%、最高100%）という良好な結果を得た。

## 5. ユーザインタフェース

著者らは、Lyricon によって自動選択されたアイコンを一覧表示するビジュアルなプレイ



リストを、Windows PC と Android 携帯の 2 種類のプラットフォーム上で実装している。現時点での著者らの実装では、アイコン画像には、無料動画投稿サイトの画面の 10 分の 1 程度の画素数 (32 × 26) のものを用いている。

### 5.1 Windows PC での実装

Windows PC での実装画面を図 6 に示す。この実装画面では、縦方向に楽曲群が、横方向に 1 曲のアイコン群が並んでいる。ウィンドウ下部には再生開始および再生停止ボタンが搭載されているのに加えて、アイコンをクリックすることで楽曲中の対応する時刻から途中再生を開始することも可能である。

また、この一覧表示は縦方向、横方向ともに大きな画面空間を占有する。この問題を解消するために著者らは、マウス操作による縦横独立のズーム表示を実装している。これによって、表示するアイコン数を自由に操作できる。図 7(上) のアイコン上で垂直にマウスを動かした場合、図 7(左下) のようにそれぞれのブロックごとにアイコン表示数を変えられる。また水平にマウスを動かすと図 7(右下) のように、アイコン表示の対象となるブロック数を変更できる

### 5.2 Android 携帯への実装

著者らは Lyricon のプラットフォームとして、携帯音楽プレイヤーやスマートフォンも

視野に入れている。その一環として著者らは、Android 携帯エミュレータを用いてユーザーインターフェースを実装している。図 8 にその表示例を示す。この実装では、縦方向のスクロールバーによって多数の曲を眺め、横方向のスクロールバーで曲全体を通したアイコン群を眺めることができる。初期画面では各曲のブロック内で一番重要度の高いアイコンをボタン形式で表示しており、興味を持ったアイコンを押すと、その箇所からの曲が再生される。もしユーザが曲の全体を表したアイコン一覧を表示させたい場合は、タイトルボタンを押すことにより、図 8(右) に示すように、その曲全体を表現するアイコンを表示する。また、本実装では実際の音楽プレイヤーと同様に、再生・停止ボタンのみならず、曲の頭出しボタン

表 2 アイコンからの歌謡曲連想の回答例 (カッコ内は回答数)

| 曲名             | 正答                                     | 誤答                 |
|----------------|--|--------------------|
| CHE.R.R.Y(図 2) | 女の子の恋 (28)/春に聴きたい (4)/かわいい (3)         | クリスマス (2)/天体観測 (1) |
| SAKURA(図 3)    | 春 (40)/恋、片思い (18)/新学期 (6)/途中で一度衰しい (1) | 部活 (1)/空 (1)       |
| Alive(図 4)     | 都会 (2)/応援 (2)/葛藤 (1)                   | 戦い (8)/喧嘩 (2)      |

表 3 アイコン選択結果と歌詞の結びつけ正答率

| 曲名  | 正答率 (%) |
|-----|---------|
| 曲 1 | 92      |
| 曲 2 | 76      |
| 曲 3 | 82      |
| 曲 4 | 80      |
| 曲 5 | 100     |
| 曲 6 | 84      |
| 曲 7 | 96      |



図 6 Windows PC での実装画面



図 7 縦横方向のズーム操作

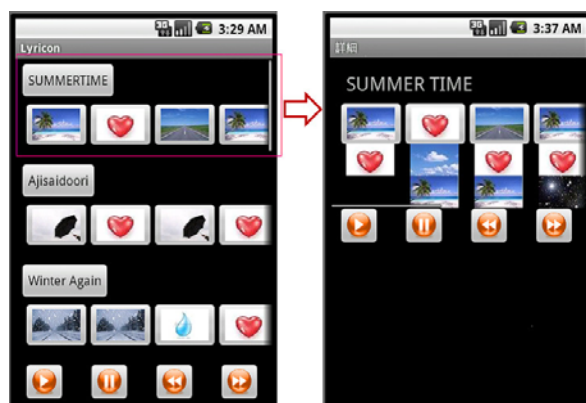


図 8 Android 携帯エミュレータでの実装画面

や次の曲へ進めるボタンをディスプレイ下方に搭載している。

## 6. 結論と今後の課題

本報告では、歌詞と楽曲の特徴を反映したアイコンをブロックごとに選択して一覧表示する、ビジュアルな音楽選択インタフェース Lyricon を提案した。これを用いることにより、曲を聴く前に歌詞内容や楽曲雰囲気をつかえる事が出来るため、自分の聴きたい曲を間違わずに聴くことができ、時間の節約につながる、また、場違いな曲を間違えて流し、気まずい空気を作る、といったことが無くなると考える。今後の課題として、以下2点を挙げる。1つ目として、歌詞解釈の精度を上げたいと思う。具体的には、修飾関係や係り受けなども考慮し、同じ“恋”の曲でも失恋を歌ったものか、両想いを歌ったものか、ということ判断できるようにしたいと思う。2つ目として、使用アイコンの再検討をしたい。現在は著者らの主観が入ったアイコン画像を用いてしまっているため、Flicker などの公共タグが付いた画像などを使用する事を検討している、

## 参 考 文 献

1) P. Kolhoff, J. Preub and J. Loviscach: Music Icons: Procedural Glyphs for Audio Files, Brazilian Symposium on Computer Graphics and Image Processing (SIB-GRAPI '06), pp. 289-296, 2006.

2) 小田瑞穂, 伊藤貴之: MIST: 音楽に印象の合うアイコンを自動選択する一手法, 第22回 NICOGRAPH 論文コンテスト, 2006.

3) 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.

4) Jonathan Foote: Visualizing Music and Audio using Self-Similarity. seventh ACM international conference on Multimedia, 1999.

5) 藤澤隆史, 谷光彬, 長田典子, 片寄晴弘: 音楽のムードを色彩で可視化する. CrestMuse Symposium 2008.

6) 後藤孝行, 後藤真孝: Musicream: 楽曲を流してくっつけて並べることのできる新たな音楽再生インタフェース. 第12回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2004) 論文集, pp.53-58, December 2004.

7) 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.

8) 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.

9) Chin-Han Chen, Ming-Fang Weng, Shyh-Kang Jeng, and Yung-Yu Chuang: Emotion-based Music Visualization using Photos. Lecture Notes in Computer Science, 2008.

10) 大山喜冨, 伊藤貴之: DIVA: 画像の印象に合わせた音楽自動アレンジの一手法の提案. 芸術科学会論文誌 Vol. 6 No. 3 pp. 126 - 135

11) 桐本篤, 佐々木史織, 清木康: 風景画像を用いた環境状況コンテキスト対応型音楽推薦システムの実現. DEWS 2008.

12) 茶筌, <http://chasen.naist.jp/hiki/ChaSen/>

13) MIRtoolbox, <https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>

14) Lyric Master, <http://www.kenichimaehashi.com/lyricsmaster/>