

## 歌詞情報を利用した Web 画像・楽曲連動 スライドショー自動生成システム

舟澤慎太郎<sup>†</sup> 石先広海<sup>††</sup> 帆足啓一郎<sup>††</sup>  
滝嶋康弘<sup>††</sup> 甲藤二郎<sup>†</sup>

本研究では、ユーザが指定した楽曲に対し、その歌詞情報を基に検索した Web 画像を用いて、スライドショーを自動で生成するシステムを提案する。楽曲の聴取と共にその内容に合ったスライドショーを鑑賞することで、より印象深い音楽体験を実現することができる。本稿では、歌詞情報を基に取得した Web 画像検索結果から、最終的に用いる画像の選定法を提案する。これにより、歌詞全体の印象に合った画像を選定することができ、かつ、ノイズとなるタグが多く付与されている画像を排除することができる。そして、評価実験によりその有効性を示す。更に、スライドショーにおける画像を切り替えるタイミングの改善手法を検討し、被験者評価実験によりその有効性を確認する。

### Proposal of Automated Music Slideshow Generation Method using Web Images

Shintaro Funasawa<sup>†</sup> Hiromi Ishizaki<sup>††</sup> Keiichiro Hoashi<sup>††</sup>  
Yasuhiro Takishima<sup>††</sup> Jiro Katto<sup>†</sup>

We propose a system to automatically generate slide shows for music, by utilizing Web images retrieved by queries extracted from song lyrics. In this paper, first, we propose a method to select images to compose the slide show from the result of web image retrieval based on lyrics. In this method, a system selects matching images with the whole impression of lyrics and removes images with many social tags that have no relation to the image content. Second, we propose a method to improve the image switching interval of the slide show.

### 1. はじめに

映画やテレビ番組、ミュージックビデオのような作品では、映像と音楽を効果的に融合することで、その作品の価値を高めている。例えば、テレビドラマの別れのシーンに BGM として悲しい音楽を流したり、ミュージックビデオにおいて歌詞のストーリーに沿った映像を提示したりすることで、それらの映像や音楽を単独で鑑賞する以上の印象を生み出すことができる。このような視覚と聴覚を刺激する効果を、普段の音楽鑑賞に適用することで、より印象深い音楽体験が実現できる。しかし、楽曲に合った映像作品を制作する場合、構成する素材の収集・選択、構成法の考慮など、様々な作業が必要となる。このため、映像作品の制作に慣れていないユーザが、個人で所有する楽曲を対象に、それに合う映像を制作するには多くの労力を必要とする。このような背景から、上記のような音楽の楽しみ方を誰でも手軽に実現するために、自動で楽曲に合った視覚コンテンツを生成するシステムが望まれる。

そこで我々は、ユーザが指定した楽曲に対して、自動でその楽曲に合ったスライドショーを生成するシステムを提案する。スライドショーを構成する素材は、楽曲の歌詞情報を基に検索した Web 画像を用いる。Web 画像は豊富で多様性のある素材であり、また、手軽に入手可能なため、Web 画像を用いることで、スライドショーを生成するための素材をユーザ自らが収集する手間を省くことができる。また、歌詞情報は、楽曲の内容を直接的に表現する特徴であるため、それを基に画像の検索を行うことで、より楽曲の内容に即したスライドショーを構成できると考えられる。このように、楽曲に合ったスライドショーを自動で生成し、楽曲と同期再生することで、手軽に印象深い音楽体験が実現できる。

我々は文献[1]において、提案システムにおけるスライドショーを構成する候補となる画像を Web から検索する処理についての提案を行った。その際の被験者アンケートにて、「複数の楽曲スライドショーにおいて特定の画像が頻出する」、「スライドショー全体の統一感に欠ける」、「画像を切り替えるタイミングが適切でない」、「間奏区間でも画像を切り替えるべきである」との問題点が指摘された。そこで本稿では、これらの問題点を解決すべく、前者 2 項目に対しては、歌詞全体の印象との適合度に基づく画像選定法を提案する。本手法では、楽曲毎に選定の指標が異なるため、より様々な画像を用いたスライドショー生成が期待される。そして、評価実験を通じて本手法の有効性を検証する。一方、後者 2 項目に対しては、画像切り替えタイミングの再構成による検討を行い、被験者評価実験によりその有効性を検証する。

<sup>†</sup> 早稲田大学理工学術院

Waseda University

<sup>††</sup> 株式会社 KDDI 研究所

KDDI R&D Laboratories Inc.

## 2. 関連研究

楽曲の歌詞情報を基に検索した Web 画像を素材として、ミュージックビデオを自動生成するシステムが提案されている[2][3]. [2]では、歌詞に出現する単語を検索クエリとし、一般的な Web 画像検索エンジンや写真共有サイトを用いて取得した画像により、ミュージックビデオを生成している. [3]でも同様に、歌詞に出現する単語により Web から画像を検索し、その中から、人の顔の写っている写真、かつ、外で撮られた写真を優先し、更に、スライドショー全体の統一感を表現するために、楽曲全体のムードに合った画像を最終的に選出している. なお、この研究では、歌詞に出現する単語の中で、名詞、名詞句、人名、地名を検索クエリとして用いている.

これらの研究では主に、Web から検索した画像群から最適なものを選定する処理に重きを置いており、検索に与えるクエリの選定には、stop word を排除する、品詞を特定のものに限定する、という処理しか行っていない. しかし、歌詞に出現する単語にも、画像検索に有用なものとそうでないものが存在する. 例えば、“今”や“誰”という単語は歌詞における出現頻度の高い単語であるが、これらをクエリとして画像検索を行った場合、その単語を的確に表現する画像を取得することは難しい.

この課題に対し、我々は[1]において提案したスライドショー自動生成システムにて、Web 画像検索に与えるクエリの選定を行っている. [1]では、ユーザが指定した楽曲の歌詞の行毎に Web 画像を 1 枚ずつ検索することでスライドショーを構成しており、図 1 に示すように 3 つの処理から構成される. まず、歌詞に出現する名詞と歌詞全体を表現する印象語（以下、全体印象語、4.1 参照）を基に画像検索クエリを選定し、写真共有サイト Flickr[4]において画像に付与されているソーシャルタグを基に検索を行う.（候補画像検索）. そして、その検索結果ランクの最も高い画像を用いてスライドショーを構成する（画像選定）. 最終的に、生成したスライドショーと楽曲を同期再生し、図 2 に示すような形で、歌詞の行単位で画像を切り替え、出力する（同期再生）.

## 3. 問題点

[1]にて実施した被験者アンケートにて、「本システムの改善すべき点があればご記入をお願いします。」との質問を行ったところ、以下のような回答が得られた.

- a) 曲調と異なる画像がある
- b) 画像を切り替えるタイミングが適切でない
- c) 間奏区間でも画像を切り替えるべきである
- d) 複数の楽曲スライドショーにおいて特定の画像が頻出する
- e) スライドショー全体の統一感に欠ける
- f) 画質により印象が変わる
- g) 楽曲構造を考慮すべきである

b) に関しては、スライドショーにおける画像の切り替えを歌詞の行単位で行うことにより生じる問題である. つまり[1]では、行の切り替わりと同期して画像を切り替えているため、画像を表示する時間が歌詞の行の長さに依存し、表示が極端に短い、又は、長い画像が生じている.

一方、d) の原因の 1 つとして、Web 画像検索結果から最終的に用いる画像を選定する処理において、Flickr におけるランキングをそのまま利用していることが挙げられる. 即ち、Flickr のランキングは、検索クエリが同一であれば結果も同一となるため、異なる楽曲において同一のクエリが検索に用いられた場合、最終的に用いられる画像も同一のものが選択される.

また、e) に関して、[1]では、画像検索クエリを構成する要素の候補として、全体印象語を用いているため、スライドショー全体の統一感を考慮していると思われるが、全体印象語はあくまでクエリを構成する一候補に過ぎず、全ての画像検索において用いられてはいない. 実際、[1]の評価実験で使用した楽曲に対して生成したスライドショーでは、クエリの構成要素として全体印象語が選択されたケースは、全体の約 60% であった. 即ち、約 40% の画像は、全体の印象を考慮していないことになる.

本稿では b) の課題を解決すべく、画像切り替えタイミングの再構成を行い、更に、[1]では考慮されていなかった間奏区間での画像の切り替えも実現し、c) の課題の解決も図る. 更に、d) および e) の課題も解決すべく、全体印象語との適合度に基づく画像選定手法を提案する.

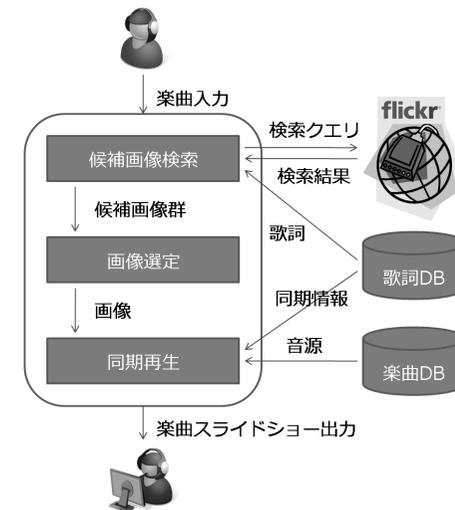


図 1 Web 画像・楽曲連動スライドショー生成システム処理概要図



図 2 システムインターフェース

#### 4. 歌詞全体の印象との適合度に基づく画像選定法

画像選定処理では、歌詞情報を基に取得した Web 画像検索結果の中から最終的に用いる画像を選定する。本節では、「複数の楽曲スライドショーにおいて特定の画像が頻出する」、「スライドショー全体の統一感に欠ける」という課題を解決するための画像選定法を提案する。

本処理では、各候補画像に対して、全体印象語との適合度を表現する score を算出し、その値が最も高い画像を選定する。score は、画像に付与されている全てのタグと全体印象語との間における関連の強さを基に算出する。こうすることで、全体印象語が直接タグとして付与されていない画像でも、“全体印象語らしさ”を表現することができ、それを基に画像を選定することで、スライドショー全体の統一感を生み出すことができる。また、楽曲に付与されている全体印象語の組み合わせにより score 算出の指標が変化するため、異なる楽曲にて同じ画像が用いられるケースが少なくなり、より多様性のあるスライドショーを構築することができる。

なお、全体印象語は本処理において用いるため、Web 画像検索のクエリとしては用

いない。つまり、検索クエリは歌詞に出現する名詞のみにより構成する。そして、その検索結果として最大 1500 枚の画像を取得する。

##### 4.1 全体印象語

歌詞全体の印象を表現する単語は、我々が以前行っていた、歌詞情報に基づく楽曲分類[5]における分類結果を基に付与する。具体的には、表 1 に示す季節・時間帯・天候に関するカテゴリに対して、楽曲が分類されたカテゴリのラベルを全体印象語として適用する。

表 1 全体印象語に用いるカテゴリラベル

概念	カテゴリラベル
季節	春, 夏, 秋, 冬
時間帯	朝, 昼, 夕方, 夜
天候	晴れ, 曇り, 雨, 雪, 虹

##### 4.2 score 算出法

score は、入力楽曲における全体印象語との関連の強いタグ（以下、関連タグ）が多く付与されている画像ほど、大きい値をとる。このため、まずは、全体印象語とタグとの関連の強さを示す関連度を、Flickr における共起確率を基に算出する。共起確率とは、2つのタグが同じ画像に付与されている確率であるため、2つのタグ間の関連の強さを表す指標となる。この共起確率は、Flickr における UF(User Frequency)を基に算出する。UF(W)とは、単語集合 W の要素全てをクエリとして AND 検索した際の、検索結果におけるユニークな画像投稿者数を示す。単純に画像ヒット数ではなく UF を用いるのは、Flickr では一定のユーザが同一のタグを付与した画像を大量にアップロードすることで、画像ヒット数が不当に高い値になることがあるため、その影響を緩和するためである。そして、全体印象語  $n_{all}$  に対するタグ  $t$  の関連度は、 $n_{all}$  における  $t$  との共起確率から、 $n_{all}$  と同じ概念（表 1 参照）に属する他の印象語における  $t$  との共起確率を引くことにより算出する。全体印象語とタグ間の共起確率だけでなく、同概念に属する他の印象語における共起確率も考慮することで、1つの印象語に対してのみ共起確率の高いタグを重要視することができる。そして、全体印象語毎に関連度の高いタグを関連タグ集合と定義し、それを基に score を算出する。画像の score は、全体印象語の関連タグが多く付与されているほど高くなり、逆に、全体印象語との関連の低いタグが多く付与されているほど低くなる。関連の低いタグを考慮するのは、Flickr におけるノイズタグへの対策である。ここでいうノイズタグとは、画像を検索されやすくするために付与する画像の内容とは無関係なタグを指す。このようにして score を定義することで、全体印象語との関連が高く、かつ、ノイズタグが少ない画像を選定することができる。

タグ  $w$  に対するタグ  $t$  の共起確率  $P(t|w)$  を(1)式のように定義する.

$$P(t|w) = \frac{UF(t \cap w)}{UF(w)} \quad (1)$$

そして, 全体印象語  $n_{all}$  に対するタグ  $t$  の関連度  $R(t|n_{all})$  を(2)式のように定義する.

$$R(t|n_{all}) = P(t|n_{all}) - \frac{\sum_{w \in W, w \neq n_{all}} P(t|w)}{|W| - 1} \times weight \quad (2)$$

但し,  $W$  は  $n_{all}$  の属する概念に含まれる印象語集合である. 例えば,  $n_{all}$  = “春” の場合, “春” は “季節” という概念の印象語であるため,  $W = \{ “春”, “夏”, “秋”, “冬” \}$  となる.

このように定義した関連度を, それぞれの全体印象語と共起する  $UF(t) \geq 5$  を満たすタグ  $t$  を対象に算出し, 経験的に関連度が 0.024 以上のタグを関連タグとして判定した. 表 2 に抽出した関連タグと関連度の一部を示す. 表 2 では, “春”, “夏”, “朝”, “雨” に対しての関連タグ(tag)と関連度(value)を示しており, 関連度の高い順にソートしてある.

そして, 楽曲  $m$  が入力されたときの画像  $i$  における score を(3)式のように定義する.

$$score(i) = \sum_{w \in N_{all}(m)} \frac{\sum_{t \in T(i) \cap T_{related}(n_{all})} R(t|w)}{|T(i) \cap T_{related}(n_{all})|} \quad (3)$$

但し,  $N_{all}(m)$  は楽曲  $m$  における全体印象語集合,  $T(i)$  は画像  $i$  に付与されているタグ集合,  $T_{related}(n_{all})$  は全体印象語  $n_{all}$  における関連タグ集合,  $|T|$  はタグ集合  $T$  に含まれる要素数をそれぞれ示す.

以上のようにして, 各候補画像の score を算出し, 画像を選定する.

表 2 関連タグと関連度の抽出結果一例

春		夏		朝		雨	
tag	value	tag	value	tag	value	tag	value
春	0.919	夏	0.939	朝	1.000	雨	0.953
桜	0.636	tokyo	0.223	朝日	0.067	Rain	0.764
花見	0.150	花火	0.180	日の出	0.067	傘	0.069
さくら	0.136	向日葵	0.092	夜明け	0.062	台湾	0.069
梅	0.134	海	0.080	霧	0.057	台北	0.054
花	0.094	祭り	0.077	朝焼け	0.048	水滴	0.034

## 5. 画像切り替えタイミング再構成

スライドショーにおいて各画像を表示する時間を適切にするために, 画像切り替えタイミングの再構成を行う. 歌詞の行単位で画像を切り替えると, その歌詞の行の長さに画像の表示時間が依存し, 適切な表示時間とならないことがある. 例えば, 表示時間が短いと, その内容を十分に把握するのが困難にし, 逆に表示時間が長いと, 退屈感を生み出す. このような問題を解決すべく, 表示時間の短い行は周辺の行と結合し, 1 枚の画像を表示し続け, 一方表示時間の長い行は分割を行い, 複数の画像を表示する. 更に, 楽曲の画像表示時間の最頻値を基に, 間奏区間における画像切り替えタイミングも設定する.

具体的な手順を以下に示す. なお, 表示時間が短い, 又は, 長いと判定する閾値は, それぞれ 4[sec], 12[sec] と経験的に設定した.

1. 楽曲の歌詞の各行における表示時間を算出し, それらの最頻値を基本表示時間  $I$  として定義する.
2. 段落の切り替わる箇所において, 段落の間における演奏時間が 4[sec] 以上ならば, その区間を間奏として抽出する.
3. 表示時間が 4[sec] 以下の行を次の行と結合する. 次の行がなければ, 前の行と結合する. 但し, 結合は同段落に属する行同士でのみ行う. このように, 行が結合された場合, 結合後の行に対して画像が 1 枚検索される.
4. 表示時間が 12[sec] 以上の行を等分割する. 但し, 分割後の表示時間が基本表示時間  $I$  に最も近くなるように分割数を調節する. このように, 行が  $n$  分割された場合, 分割前の行に対して検索された候補画像から上位  $n$  枚を選択し表示する. また, 間奏区間に対しても同様に分割し, 画像検索クエリには全体印象語を用いる.

以上のようにして, 画像を切り替えるタイミングの再構成を行う.

## 6. 評価実験 I: 画像多様性評価

### 6.1 実験概要

画像選定手法の違いによる, スライドショーに用いられる画像素材の多様性への影響を評価するための実験を行った. ここで挙げた多様性とは, 複数の楽曲に対してスライドショーを生成した際に, 特定の画像が多用されず, 様々な画像素材を用いてスライドショーを構成する特徴である. 複数楽曲のスライドショーを鑑賞する際, 一部の画像のみ頻出することは, スライドショーの新鮮さを損なう原因となるため, 用いられる画像が多様であることは重要な特徴である.

本実験では, 12 曲の楽曲に対して 2 つの画像選定手法によりスライドショーを生成し, その画像重複率を算出することで, 多様性を評価した. 画像重複率は, 複数楽曲

のスライドショーに対して重複して使用された画像の回数を、使用された全ての画像数によって割ることで算出した。更に、画像多様性がスライドショーに求められる特徴であることを確認するため、被験者アンケートを実施した。具体的には、「一部の画像だけでなく、様々な画像が使用される」という特徴はスライドショーにおいて重要か、との問いに対し、特に重要である、重要である、あまり気にしない、の選択肢から何れか1つを選択してもらった。本アンケートの対象被験者は、大学生20名である。

なお、比較対象は、scoreに基づく画像選定手法とFlickrのランキングに基づく画像選定手法とする。後者の手法では、Flickrにおける“interestingness”指標を用いて画像検索結果のソートを行い、そのランクの最も高い画像を選定する。

### 6.2 実験結果

それぞれの手法に対する画像重複率を表3に示す。この結果は、12曲のスライドショーを生成した場合、Flickrのランキングに基づく画像選定を行うと、約3回に1回は同じ画像が表示されるが、scoreに基づく画像選定を行うと、約10回に1回まで抑えられることを示している。以上より、scoreに基づく画像選定手法を適用することで、特定の画像が頻繁に利用される問題を緩和し、より多様なスライドショーを生成できているといえる。また、被験者アンケートの結果では、回答者20名中19名が画像多様性という特徴は、特に重要である、又は、重要である、と回答していた。よって、画像多様性がスライドショーにおいて重要な特徴であることが確認できた。

表3 各画像選定手法の楽曲12曲に対する画像重複率

手法	score-based	Flickr-based
画像重複率	10.8%	27.1%

## 7. 評価実験Ⅱ：画像切り替えタイミング再構成評価

### 7.1 実験概要

5章に示した手法による画像表示時間の改善効果の検証と、それによるスライドショー全体への影響を評価するために、被験者による評価実験を行った。被験者は、画像切り替えタイミングの再構成を行う前と後のスライドショーを、楽曲の1コーラス分鑑賞し、それぞれに対して以下の項目について5段階評価を行う。なお、提示する順序効果は考慮してある。

- Q1) スライドショーの完成度を評価して下さい。  
 5: 完成度が高い ~ 1: 完成度が低い
- Q2) 各画像を表示する長さ(時間)は適切でしたか?  
 5: 大抵の画像において適切だった ~ 1: 大抵の画像において不適切だった

対象の被験者は大学生20名で、楽曲データはJ-POP楽曲20曲を用い、1曲につき10名分の評価を収集した。

### 7.2 実験結果

図3、図4に、Q1、Q2に対する評価結果の一部をそれぞれ示す。両図では、対象楽曲の内、結合処理の多い楽曲5曲(ID: 51, 181, 476, 509, 1708)と、分割処理の多い楽曲5曲(ID: 39, 129, 339, 546, 1676)に対する評価値平均を示している。なお、これらの楽曲は、再構成前後における評価値の変化が比較的大きいものを中心に選出した。

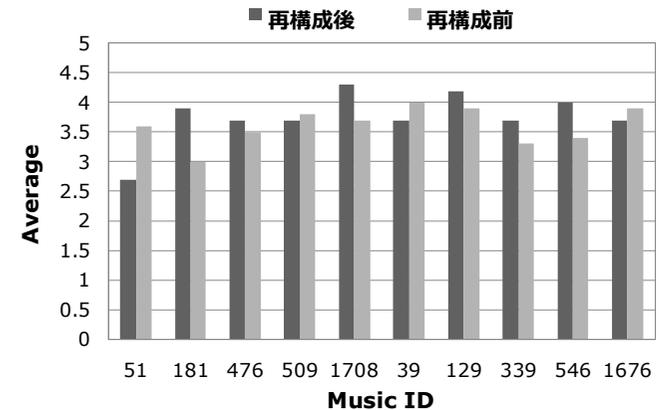


図3 各楽曲におけるQ1の評価値平均

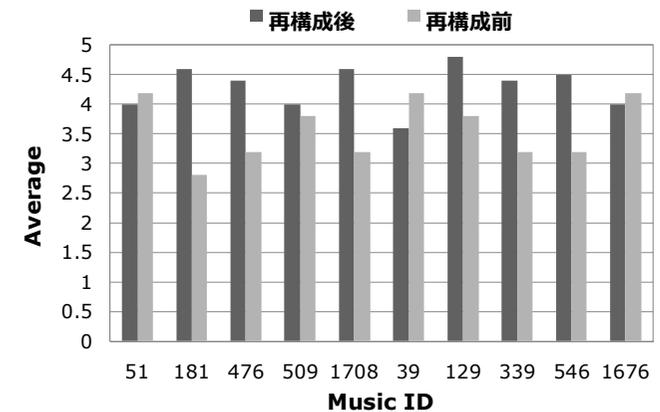


図4 各楽曲におけるQ2の評価値平均

まず、対象楽曲 20 曲に対しての Q1 の平均評価値は、再構成前は 3.69、再構成後は 3.85 となった。同様に Q2 に関しては、再構成前は 3.66、再構成後は 4.29 となった。更に Q2 では、t 検定 (有意水準 10%) の結果において、再構成後の評価値の方が有意に高いと判定された楽曲は 8 曲あり、逆に、再構成前の評価値が有意に高いと判定された楽曲はなかった。これらの結果より、本手法によって画像を切り替えるタイミングが適切に再構成されていることが示された。

また、図 3、図 4 より、Q2 において再構成後の評価値が高くなっている楽曲に関しては、Q1 でも同様に再構成後の評価値が高くなっている。これより、画像の表示時間を適切に設定することで、スライドショー全体の評価の向上に繋がるといえる。ここで、Q1 における改善度 (再構成前後間の評価値の差) が Q2 におけるそれと比較して低くなっているのは、スライドショーの完成度に影響する要因には様々なものがあり、画像表示時間の適切性はその中の一要因に過ぎないため、表示時間の改善度がそのままスライドショー全体の評価の改善度に直結していないのだと考えられる。

更に図 4 において、個々の楽曲を参照すると、ID:181 は、対象楽曲の中で最も再構成を行うことによる効果が表れている。この楽曲はラップが中心であり、再構成前の画像の平均表示時間が 2.52[sec]と、短時間で頻繁に画像が切り替わっていたため、再構成による効果が大きかったのだと考えられる。なお、同様の傾向が ID:129, 339 においても確認できた。このように、平均表示時間が極端に短い、又は、長い楽曲においては、再構成による改善効果が大きいといえる。

逆に ID:39 では、再構成前の評価値が再構成後と大きく差をつけて高くなっている。この楽曲では、再構成前の平均表示時間が 14.2[sec]で、再構成後が 7.0[sec]である。本手法では、表示時間が長いと判断する閾値を 12[sec]と設定しているが、この評価結果より、ID:39 に対してはこの閾値が適切でないといえる。これは、ID:39 はスローテンポなバラード曲であり、表示時間が長くても楽曲の雰囲気と合っていると判断されたためであると考えられる。このように、設定した閾値が適切でない楽曲に対しては、曲調やテンポに応じて閾値を動的に設定することが必要となるが、今回使用した楽曲でこのような傾向が見られたのは ID:39 のみであったため、大抵の楽曲では今回設定した閾値でも十分適用できると考えられる。

また、ID:509 はラップが中心の楽曲であり、再構成前の平均表示時間が 2.74[sec]と短いにも関わらず、再構成を行うことによる改善効果がさほど大きくない。これは、楽曲内における画像の表示時間の変化が、楽曲の場面変化を捉えていると解釈されたことが 1 つの原因として挙げられる。ID:509 には、ラップの部分とそうでない通常の歌唱の部分があり、再構成を行わない場合、ラップの箇所では画像が頻繁に切り替わり、通常の箇所では比較的画像が長く表示される特徴がある。このような特徴を、“楽曲の雰囲気の変化を的確に反映している”と良い傾向に捉えることで、再構成前の評価値が若干高くなっていると考えられる。実際に、本実験の際に行った自由記述形式

のアンケートでも、上記のような特徴を重視する意見があった (20 名中 2 名) が、その一方で、画像が頻繁に切り替わる特徴を問題視する意見も見られた (20 名中 4 名)。このように、“画像の表示時間”か“楽曲の場面変化”のどちらを重視するかは、被験者によりまちまちであるため、切り替えタイミングを再構成することにより“画像の表示時間”を適切に設定した後、ズームやパンなどのスライドショーの表示における効果で“楽曲の場面変化”を表現するなど、両特徴を考慮する対策が必要となる。

## 8. おわりに

本稿では、ユーザの指定した楽曲に対して、歌詞情報を基に検索した Web 画像を用いたスライドショーを自動生成するシステムを提案した。更にその中で、Web から検索した画像群から最終的に用いる画像の選定法、スライドショーにおける画像を切り替えるタイミングの再構成法を提案し、評価実験を通じてそれらの有効性を示した。

更なるシステムの発展のために、今後は以下に挙げる改善策を検討していく。

- 歌詞情報における形容詞や英詞の考慮

本システムでは、歌詞に出現する名詞のみを考慮しているが、名詞以外にも視覚的に重要な意味をもつ単語が存在すると考えられる。例えば、現状では、“黄色い花”という歌詞が出現した場合、“黄色い”という情報までは考慮できない。よって、形容詞を考慮することで、更に歌詞の内容を詳細に表現することができる。また、邦詞だけでなく英詞も考慮することで、更なる改善が期待できる。

- スライドショーの表示法の工夫

本システムでは、スライドショーの表示に関する工夫として、画像切り替え時にフェード処理を施している。それに加えて、より多様な効果 (ズーム、パンなど) を、7 章の実験にて示したように、楽曲の盛り上がりなどに応じて効果的に加えることで、更に印象深いスライドショーが実現できると考えられる。

## 参考文献

- 1) 舟澤慎太郎, 石先広海, 帆足啓一郎, 滝嶋康弘, 甲藤二郎: 歌詞特徴を考慮した Web 画像と楽曲同期再生システムの提案, 第 8 回情報科学技術フォーラム, E-034 (2009).
- 2) D. A. Shamma, B. Pardo, and K. J. Hammond: MusicStory: a Personalized Music Video Creator, Proceedings of the 13<sup>th</sup> Annual ACM International Conference on Multimedia, pp.563-566 (2005).
- 3) R. Cai, L. Zhang, F. Jing, W. Lai, and W. -Y. Ma: Automated Music Video Generation Using Web Image Resource, Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing, 2:pp.737-740 (2007).
- 4) Flickr, <http://www.flickr.com/>
- 5) 舟澤慎太郎, 石先広海, 帆足啓一郎, 滝嶋康弘, 甲藤二郎: 歌詞の印象に基づく楽曲検索のための楽曲自動分類に関する検討, 第 71 回情報処理学会全国大会, 5R-2 (2009).