

# オノマトペを利用した Web 上からの楽曲動画検索システム モデルの提案

佐藤 静香<sup>1</sup> 児玉 英一郎<sup>1</sup> 王家宏<sup>1</sup> 高田 豊雄<sup>1</sup>

**概要:** 近年のインターネット普及に伴い、Web 上には多数の楽曲動画が存在するようになった。そのため、ユーザが求める楽曲動画を効率よく検索することが重要となってきた。また最近では、歌声合成ソフトウェアである初音ミクを用いた CGM 型の楽曲動画が数多く作成され、動画投稿配信サイトの YouTube など公開されている。

従来のアーティスト名や曲名を利用した検索方法では、初音ミクを用いた楽曲動画など無名なものを検索する際、曲名や制作者がわからないことも多く、検索が困難である。このため、「初音ミク」をキーワードとして検索し、投稿順にランキング形式で表示された検索結果の中から 1 曲ずつ視聴し、ユーザが気に入ったものを見つけないといった場当たりの手法に頼らざるを得ない。

本論文では、「カンカン」などの擬音語や、「のほほん」などの擬態語の総称であるオノマトペを利用し、Web 上から無名な CGM 型の楽曲動画をも効率よく検索するシステムのモデルの提案を行う。

また、本モデルの有用性確認のための、本モデルの実装系を利用した評価実験についても報告する。

## 1. はじめに

近年のインターネット普及に伴い、Web 上には多数の楽曲動画が存在するようになった。そのため、ユーザが求める楽曲動画を効率よく検索することが重要となってきた。楽曲動画を検索するシステムとしては、iTunes[1] や YouTube[2] が知られている。また最近では、歌声合成ソフトウェアである初音ミクを用いた CGM(Consumer Generated Media) 型の楽曲動画が数多く作成され、前述の YouTube など公開されている。

一方、従来からの楽曲動画検索の手法は、アーティスト名や曲名での検索が基本となっており、初音ミクを用いて制作された無名な楽曲動画の検索においては有用でない状況となっている。

そこで本研究では、「カンカン」などの擬音語や、「のほほん」などの擬態語の総称であるオノマトペを利用し、Web 上から無名な CGM 型の楽曲動画をも効率よく検索するシステムを提案する。

## 2. 楽曲検索の現状

### 2.1 検索方法

Web 上の楽曲に関するサイトとして iTunes[1] などの

音楽ダウンロードサイトや、YouTube[2] やニコニコ動画 [3] などの動画投稿配信サイトが多くユーザに利用されている。このようなサイトでは、アーティスト名や曲名での検索が基本となっている。

他にも、ハミングによる楽曲の検索手法に関する研究 [4] や米国の Melodis Corporation が運営する、楽曲の一部を歌うことで検索できる Web サイト [5] も存在しているが、主流な検索方法とはなっていない。

### 2.2 CGM 型コンテンツ

最近では、動画投稿配信サイトにおいて多くの CGM 型楽曲動画が投稿されている。CGM 型楽曲動画とは利用者が作成した楽曲動画を指し、歌声合成ソフトウェアである初音ミク [6] を用いて作成されたものが代表的である。初音ミクは、発売元が宣言した「キャラクター利用のガイドライン」[7] によって権利を開放されており、自由に楽曲動画を創作し発表することができるようになっている。YouTube、ニコニコ動画、Google[8] の各サービスで「初音ミク」を検索したときのヒット数を表 1 に示す。

表 1 キーワード「初音ミク」の検索ヒット数 (2012 年 12 月 8 日現在)

サービス名	ヒット数
Google	61,300,000
YouTube	171,000
ニコニコ動画	179,471

<sup>1</sup> 岩手県立大学大学院  
Graduate School of Software and Information Science,  
Iwate Prefectural University

Googleでのヒット数には音楽やイラスト、テキストなどの作品が含まれている。このことから、初音ミクを利用した創作活動を始めとするCGMは、様々な取り組みがなされている活発な分野であるといえる[9]。また、Youtubeやニコニコ動画などの動画投稿配信サイトにおけるヒット数から、約170,000件ものCGM型楽曲動画が投稿されていることがわかる。しかし、利用者が作成したCGM型楽曲動画は無名なものも多く、曲名や制作者が不明であるものがほとんどである。膨大な検索結果の中からユーザが求める楽曲動画を見つけるためには、従来の検索手法では不十分な状況と考える。

### 3. オノマトペ

オノマトペとは、カンカンなどの擬音語や、ノホホンなどの擬態語の総称であり、犬の鳴き声や雨が降る様子など、我々の日常会話などで頻繁に使用されている。オノマトペは、具体的な事象を的確に表現できる語彙であり、コミュニケーションを図る上で重要なものである[10]。日常会話では、漠然としたイメージを表現する際に用いられることが多く、物事の様子がより伝わりやすくなるといった効果がある[11]。

オノマトペを収録した辞書としては日本語オノマトペ辞典[12]や現代擬音語擬態語用法辞典[13]などがあり、収録語数はそれぞれ、4500語、1064語となっている。また、オンラインオノマトペ辞書としては、オノマトペディア[14]が知られている。

オノマトペと同種の感性を表現するものとして、印象語や感性語が挙げられる。これらは人の感性を扱う多数の研究において利用されているが、このうちのどれが感性情報を表現するのに最も適しているかは明らかになっていない。そこで、これら感性を扱う研究で用いられている印象語[15]、感性語[16][17][18][19]、及びオノマトペ[10][11][20][21]を抽出し、その語数の比較を行った。本結果を表2に示す。

印象語と感性語では、共通する語が41語抽出された。こ

表2 感性を表現するものの語数比較

感性を表現するもの	語数
印象語	230
感性語	94
オノマトペ	129

れは、どちらも形容詞や副詞などを含んでいることが理由として挙げられる。また、印象語や感性語を扱う辞書は今のところ存在せず、その全体数を把握することは難しい状況であるといえる。一方、抽出したオノマトペの語数は印象語より少ない129語という結果であったが、前述した日本語オノマトペ辞典の収録語数は4500語となっている。

以上のことから、印象語、感性語及びオノマトペのうち、感性を表現するものとしてどれが良いかは一概に決定でき

ない状況と考えられる。

### 4. 現状の問題点

従来のアーティスト名や曲名を利用した検索方法では、初音ミクを用いた楽曲動画など無名なものを検索する際、曲名や制作者がわからず検索が困難である。このため、「初音ミク」をキーワードとして検索し、投稿順にランキング形式で表示された検索結果の中から1曲ずつ聴き、ユーザが気に入ったものを見つけるといった場当たりの手法に頼らざるを得ない。さらに、時間経過後、再度同じ曲を聴こうと思った際には、投稿順が変動していることに加え、曲名や制作者などの情報を忘れてしまう可能性もあり、検索できない場合もある。

以上の問題から、無名な楽曲動画を効率的に検索する方法が求められていると考える。よって、本研究では、物事の漠然としたイメージの表現に用いられるオノマトペを利用し、楽曲動画を効率的に検索するシステムを提案する。

### 5. 関連研究

関連研究として、オノマトペの一部である擬音語のみを利用した、静的な音響データベースからの音響データ検索に関する研究[22]や、古典的な感性語を用いた静的な楽曲データベースからの楽曲検索に関する研究[23]が知られている。

擬音語を用いた音響データの検索に関する研究では、オノマトペの一部である擬音語を用い、効果音などの音響データのみを対象に、静的データベースからの検索を行っているのに対し、本研究では、擬音語に加え擬態語も対象としたオノマトペ全体を利用し、楽曲動画データを対象に、Web上から検索を行おうとしており、この点が異なる。

古典的な感性語を用いた、静的な楽曲データベースからの楽曲検索に関する研究では、「悲しい」や「楽しい」などの感性語を用い、静的な楽曲データベースから検索を行っているのに対し、本研究では、感性語ではなくオノマトペを利用し、楽曲動画データを対象に、Web上から検索を行おうとしており、この点が異なる。また、感性語に似た表現の印象語を用いた楽曲検索に関する研究[24]も行われているが、ここでいう感性語や印象語は、どちらも語の範疇に属しており、現状においてどちらを利用するかは好みの問題でしかない[25]。

楽曲動画の検索に際し、感性語を利用すべきかオノマトペを利用すべきかは、学術的に明確にはなっておらず、どちらが有用か、その判断は困難であるが、これらは排他的なものではなく、ユーザの多種多様な検索要求に対応するためにも、これまで研究されてこなかったオノマトペ全体を利用した楽曲動画検索について研究することは重要であると考えられる。

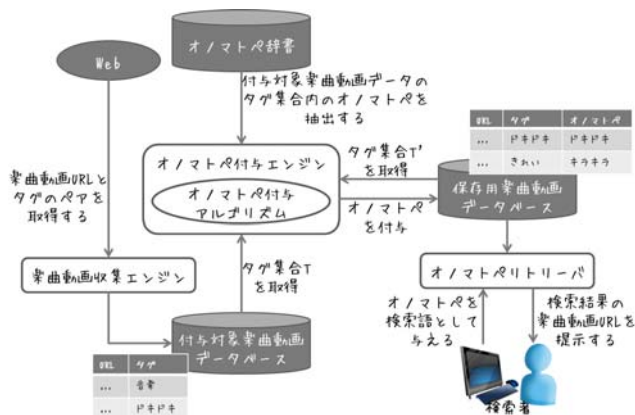


図 1 オノマトペを利用した Web 上からの楽曲動画検索システムのモデル

## 6. オノマトペを利用した Web 上からの無名な楽曲動画検索手法の提案

### 6.1 オノマトペを利用した Web 上からの無名な楽曲動画検索手法の概要

本提案のオノマトペを利用した Web 上からの楽曲動画検索システムのモデルを図 1 に示す。

以下、図 1 の各構成要素の詳細を示す。

### 6.2 楽曲動画検索エンジン

タグの集合 (初音, ミク, append など) 注釈や検索を目的として楽曲に付与されたキーワード (タグ) からなる集合, 以下, タグ集合と呼ぶ) が付与された動画投稿配信サイト (YouTube[2], ニコニコ動画 [3] など) を対象とし, 対象動画投稿配信サイトから楽曲動画の URL とタグ集合のペアを取得する。取得したデータは一時的にデータベースに登録し, オノマトペ付与エンジンへ渡される。

### 6.3 オノマトペ付与エンジン

オノマトペ付与対象の楽曲動画データのタグ集合と, 保存用楽曲動画データのタグ集合を利用し, 本研究独自のオノマトペ付与アルゴリズムに従い, オノマトペの付与を行う。以下に本オノマトペ付与アルゴリズムを示す。但し, ここでは事前に  $n$  件のレコードが保存用楽曲動画データに含まれているものとする。

#### 6.3.1 オノマトペ付与アルゴリズム

- (1) 付与対象の楽曲動画データのタグ集合  $T$  を取得する。電子化されたオノマトペ辞書を用い,  $T$  内に含まれるオノマトペを抽出し, 付与対象の楽曲動画データのオノマトペとして付与する。
- (2) 保存用楽曲動画データから 1 件のレコードを取り出し, そのレコード内のタグ集合  $T'$  を取得する。
- (3) 次式により,  $T$  と  $T'$  との類似度を計算する。

$$Jaccard(T, T') = \frac{|T \cap T'|}{|T \cup T'|}$$

- (4)  $n$  件の保存用楽曲動画データに対し, (2), (3) を繰り返す, 類似度が最大の保存用楽曲動画データのオノマトペを付与対象の楽曲動画データのオノマトペとして付与する。また, その最大の類似度が閾値  $\theta (0 < \theta < 1)$  以上の場合には, (1) で付与したオノマトペを, 類似度が最大の保存用楽曲動画データにも付与する。

## 6.4 オノマトペリトリバ

ユーザから与えられた検索語のオノマトペと, 楽曲動画に付与されたオノマトペとを比較し, 合致する楽曲動画の URL を提示する。

現在の, 楽曲動画検索のためのユーザインタフェースのイメージを図 2 に示す。例えば, タンタンと入力すると図 3 のように結果が表示される。楽曲動画を視聴する際には, URL 部分をクリックすることでブラウザが起動し, 再生することができる。

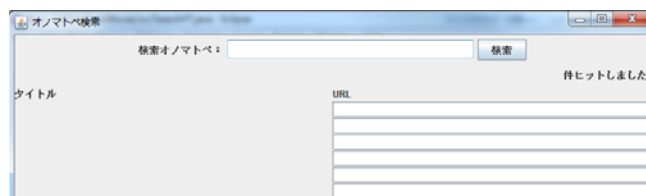


図 2 検索ユーザインタフェースのイメージ

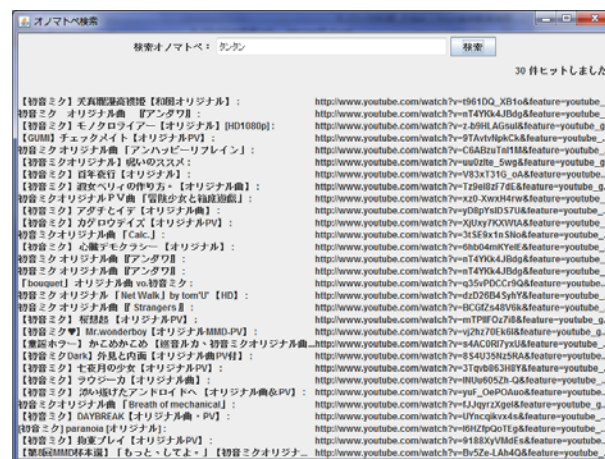


図 3 検索結果表示イメージ

## 7. 評価

### 7.1 構築実験

本提案手法に基づき, オノマトペによる楽曲動画検索システムの構築を行った。本実装環境を表 3 に示す。

### 7.2 適合率

本モデルの有用性確認のため, 適合率の評価を行った。初期データとして, 100 件の教師用データと付与対象楽曲

表 3 実装環境

OS	Windows 7
CPU	Intel®Core™i3-2357M CPU @1.30GHz 1.30GHz
メモリ	8GB
使用言語	Java
保存用楽曲動画データ数	100
付与対象楽曲動画データ数	100

動画データ 100 件を用いて、提案アルゴリズムに基づきオノマトペの付与を行った。被験者 12 名を用意し、楽曲動画に付与されていた全オノマトペ 48 語を用いて行った。検索語ごとに検索結果のうち被験者が正解と判定した数を求め、被験者 12 名の平均値を平均正解数とした。

実験結果の抜粋を表 4 に示す。表 4 において、各オノマトペごとの適合率は、被験者 12 名の適合率の平均値とし、全オノマトペの適合率の平均値を提案システムの適合率とした。表 4 から、本提案システムの適合率 67.57% が得られた。

表 4 本実装環境での適合率 (抜粋)

検索オノマトペ	検索による提示数	提示数に含まれる平均正解数	オノマトペごとの適合率
ガチャガチャ	1	0.92	91.67%
カラカラ	3	2.08	69.44%
ジャカジャカ	7	6.58	94.05%
シャンシャン	4	2.75	68.75%
タンタン	26	20.50	78.85%
ノホホン	1	0.67	66.67%
ピコピコ	27	15.42	57.10%
ビュンビュン	1	0.92	91.67%
ユルユル	1	0.83	83.33%
ルンルン	1	0.75	75.00%
ワクワク	1	0.75	75.00%
合計/平均	226	150.42	67.57%

### 7.3 再現率

適合率と同様に、初期データとして、100 件の教師用データと付与対象楽曲データ 100 件を用いて、再現率の評価を行った。あらかじめ被験者 5 名により 48 語のオノマトペに対応する正解の楽曲動画を決定しておき、これを正解楽曲動画とした。楽曲動画に付与されていたオノマトペ 48 語を使用して検索後、検索結果中の正解楽曲動画の数を求め、再現率を算出した。

実験結果の抜粋を表 5 に示す。表 5 において、各オノマトペごとの再現率は、被験者 5 名により決定した正解楽曲動画がシステムによってどれだけ提示されたかを示している。また、このオノマトペごとの再現率の平均を提案システムの再現率とした。表 5 に示すように、本提案システムの再現率は 54.61% である。

表 5 本実装環境での再現率 (抜粋)

検索オノマトペ	設定した正解数	提示数に含まれる正解の数	オノマトペごとの再現率
ガチャガチャ	4	1	25.00%
カラカラ	2	2	100.00%
ジャカジャカ	13	6	46.15%
シャンシャン	3	1	33.33%
タンタン	12	10	83.33%
ノホホン	2	1	50.00%
ピコピコ	12	6	50.00%
ビュンビュン	3	1	33.33%
ユルユル	4	1	25.00%
ルンルン	2	0	0.00%
ワクワク	2	1	50.00%
合計/平均	219	119	54.61%

#### 7.3.1 検索に要するユーザの時間的負荷の比較

従来からの、アーティスト名や曲名での検索システムによる検索時間と、提案手法による検索時間との比較を行い、検索の効率化が実現できたか検証するための実験を行った。

11 名の被験者を用意し、従来システムと提案システムそれぞれにおいて表 6 に示したオノマトペを用い検索させた。ここで、従来の楽曲動画検索システムとしては、YouTube[2] を利用した。各被験者が、キーワードを入力して提示された結果から 1 つずつ聞き、その楽曲動画が被験者の入力したキーワードに合致すると判断するまでの時間を計測した。

表 6 入力キーワードとなるオノマトペ

1	ノホホン
2	ガチャガチャ
3	タンタン
4	ピコピコ
5	ジンジン
6	キラキラ

従来検索システムを用いた場合の被験者の時間的負荷を表 7 に示し、提案システムを用いた場合の被験者の時間的負荷を表 8 に示す。結果の考察には、被験者 11 名の検索時間の平均値を用いた。YouTube はインターネット上でのサービスであり、多くの検索対象データを保有しているのに対し、本提案システムでの検索はローカルな環境であるし、また保有検索対象数も少ないものとなっている。検索時間の観点からは完全に公平な比較環境とはなっていないが、ここでは、ユーザ側から見た時間的負荷の比較ということで実験を行っている。

このような実験状況において、従来手法での被験者の時間的負荷の平均と、提案手法での被験者の時間的負荷の平均値を比べると、約 60 秒の軽減が図れたことがわかる。

また、検索時間調査の被験者 11 名にアンケートを行ったところ、そのときの気分に合わせて楽曲動画を選びたいと

表 7 従来手法を用いた場合の被験者ごとの時間的負荷 (秒)

被験者	ノホン	ガチャガチャ	タンタン	ピコピコ	ジンジン	キラキラ	平均
1	5:55	3:01	3:51	0:55	2:30	4:06	3:23
2	2:55	2:03	3:55	1:29	4:43	1:40	2:47
3	2:27	1:46	0:22	0:17	1:22	1:49	1:21
4	3:53	1:00	3:45	1:32	1:59	0:47	2:09
5	6:00	1:06	0:44	1:41	1:30	1:27	2:05
6	1:35	2:00	0:50	0:30	3:30	4:05	2:05
7	0:53	2:19	1:13	0:45	1:55	1:30	1:26
8	1:07	1:40	1:27	0:42	2:48	3:17	1:50
9	0:45	3:06	0:32	0:18	1:20	0:12	1:02
10	2:22	0:58	1:33	0:27	0:22	0:11	0:59
11	2:07	0:30	0:37	0:15	0:35	0:22	0:44
平均	2:44	1:46	1:43	0:48	2:03	1:46	1:48

表 8 提案手法を用いた場合の被験者ごとの時間的負荷 (秒)

被験者	ノホン	ガチャガチャ	タンタン	ピコピコ	ジンジン	キラキラ	平均
1	1:40	0:30	0:40	0:20	0:15	0:45	0:42
2	1:07	1:49	1:00	1:47	2:58	2:21	1:50
3	0:51	0:28	0:45	0:17	1:03	0:38	0:40
4	0:20	0:30	1:13	0:41	1:10	1:22	0:53
5	0:36	0:32	0:41	1:00	1:01	0:53	0:47
6	1:02	1:00	0:24	0:26	1:29	2:10	1:05
7	0:30	0:33	1:11	0:34	1:13	1:09	0:52
8	0:53	0:24	1:04	0:23	1:20	0:47	0:48
9	0:28	0:22	0:17	0:20	0:57	0:17	0:27
10	0:19	0:18	0:29	0:18	0:36	0:16	0:23
11	0:19	0:34	0:34	0:17	0:43	0:48	0:33
平均	0:44	0:38	0:45	0:35	1:10	1:02	0:49

きに利用したいという意見が多数みられた (表 9)。従来システムでは検索が困難であった気分による検索が、本提案システムでは可能であるという被験者の意見が得られた。以上、被験者から得られた意見から、ユーザが選択できる検索インタフェースの幅を広げることができたと考える。

表 9 提案手法を用いた楽曲動画検索に関する感想・意見

提案手法 (オノマトペによる楽曲動画検索) を利用したい場面にについて
<ul style="list-style-type: none"> <li>聞きたい曲が特に決まっておらず、なんとなく新しい曲を聴いてみたいと考えたとき</li> <li>そのとき聞きたい気分の曲をすぐに見つけたいとき</li> <li>曲名が思い出せないとき</li> <li>自分好みの知らない曲を探したいとき</li> <li>そのイメージの曲が聴きたいと思ったとき</li> <li>なんとなくの気分で「こんな感じの曲が聴きたい」というとき</li> </ul>

## 8. おわりに

本研究では、従来のアーティスト名や曲名を利用した検索方法では、CGM 型楽曲動画などの無名なものを検索しようとした場合に、場当たりの手法に頼らざるを得ない状況であること、時間経過後に再度同じ楽曲動画を視聴しようとした際に、検索できない場合があることを問題点としてとらえ、この解決を図った。また、この問題点の解決のために、オノマトペを利用した Web 上からの無名な楽曲動画検索手法の提案を行った。本モデルの有用性確認のため本モデルの実装を行い、構築したシステムを利用して、適合率と再現率の評価を行った。その結果、適合率 66.82%、再現率 56.36%、 $F$  値 61.14%であることを確認した。また、従来手法との検索時間比較実験から、検索時間が約 60 秒短縮されることを確認した。

## 参考文献

- [1] iTunes  
<http://itunes.apple.com/jp/>
- [2] YouTube  
<http://www.youtube.com>
- [3] ニコニコ動画 (原宿)  
<http://www.nicovideo.jp/>
- [4] 大石康智, 後藤真孝, 伊藤克寛, 武田一哉: 音声だけでシームレスにハミング検索と曲名検索が可能な楽曲検索システム, 情報処理学会研究報告, 音楽情報科学, Vol.2006, No.113, pp.3-8 (2006).
- [5] midomi  
<http://www.midomi.co.jp/>
- [6] クリプトン・フューチャー. メディア株式会社  
<http://www.crypton.co.jp>
- [7] キャラクター利用のガイドライン  
[http://piapro.jp/license/character\\_guideline](http://piapro.jp/license/character_guideline)
- [8] Google  
<http://www.google.co.jp/>
- [9] 特集「CGM の現在と未来: 初音ミク, ニコニコ動画, ピアプロの切り開いた世界」(解説 5 件), 情報処理学会誌, Vol.53, No.5, PP.464-494 (2012).
- [10] 浅賀千里, YusufMukarramah, 渡辺知恵美: オンラインオノマトペ用例辞典「オノマトペディア」における用例の意味により分類するための係り受け関係を考慮したクラスタリング手法, 電子情報通信学会 第 19 回データ工学ワークショップ, A3-4 (2008).
- [11] ラートサムアイパン カンウィパー, 渡辺知恵美, 中村聡史: オノマトペロリ: オノマトペを利用した料理推薦システム, 情報処理学会研究報告, Vol.2009-DD-73, No.6, pp.1-7 (2009).
- [12] 小野正弘: 擬音語・擬態語 4500 日本語オノマトペ辞典, 小学館 (2007).
- [13] 飛田良文, 浅田秀子: 現代擬音語擬態語用法辞典, 東京堂出版 (2002).
- [14] 浅賀千里, 渡辺知恵美: Web コーパスを用いたオノマトペ用例辞典の開発, 電子情報通信学会第 18 回データ工学ワークショップ, B9-2, (2007).
- [15] 印象を使った画像・音楽の検索  
<http://www.mdbl.sfc.keio.ac.jp/kansei/>
- [16] 椋木雅之, 田中大典, 池田克夫: 対義語対からなる特徴空間を用いた感性語による画像検索システム, 情報処理学

- 会論文誌, Vol.42, No.7, pp.1914-1921 (2001).
- [17] 木本晴夫: 感性語による画像検索とその精度評価, 情報処理学会論文誌, Vol.40, No.3, pp.886-898 (1998).
  - [18] 坂内祐一, 石澤正行, 重野寛, 岡田謙一: 感性語を媒介にした香りコミュニケーションモデル, 情報処理学会論文誌, Vol.47, No.12, pp.3414-3422 (2006).
  - [19] 杉原太郎, 森本一成, 河村知典, 黒川隆夫: 若年層向け音楽感性検索システムに使用するための感性語選択, 感性工学研究論文集, Vol.5, No.3, pp.127-134 (2009).
  - [20] 陳志文: 日本語教育におけるオノマトペの提出順序についての一提案—「2005年現代雑誌200万字言語調査語彙表」の考察から—, 日台研究支援事業報告書 (2007).
  - [21] 市岡健一, 福本文代: Web 上から取得した共起頻度と音象徴によるオノマトペの自動分類, 電子情報通信学会論文誌, Vol.J92-D, No.3, pp.428-438 (2009).
  - [22] 魚田慧, 鈴木健嗣, Pitoyo Hartono, 橋本周司: 擬音語と音響を用いた音響データベースの検索, 電子情報通信学会技術研究報告 HCS, Vol.104, No.446, pp.19-22 (2004).
  - [23] 池添剛, 梶川嘉延, 野村康雄: 音楽感性空間を用いた感性語による音楽データベース検索システム, 情報処理学会論文誌, Vol.42, No.12, pp.3201-3212 (2001).
  - [24] 熊本忠彦, 太田公子: 印象に基づく楽曲検索システム的设计・構築・公開, 人工知能学会論文誌, Vol.21, No.3, pp.310-318 (2006).
  - [25] 上田修一: 感性キーワードの発展とその限界, 三田図書館情報学会論文誌, 短報, No.41, pp.17-25 (1999).