

# RSS 情報を利用した楽曲選択システムの構築

高木 恵美 (指導教員: 増永 良文)

## 1 はじめに

現在, iTunes や SonicStage, Winamp などの音楽プレーヤソフトには, アーティスト名やアルバム名など, 様々な条件による楽曲検索機能が備わっている。また, 先行研究において長澤は, 歌詞による楽曲検索を可能とした [1]。しかしそれらは, ユーザが明確な条件を持たない場合には検索が不可能である。そのため, 大量の楽曲を持っているにもかかわらず, いつも同じ楽曲ばかり聴いてしまうことも多い。そこで本研究では, ニュースなどの RSS 情報を利用することで日々異なる楽曲を抽出し, プレイリストを自動作成するシステムを構築することとした。

このシステムが実現すれば, 特に決まった条件がない場合でも, 大量の楽曲データの中から, 日々様々な楽曲を抽出して聞くことが可能となる。

## 2 歌詞検索システム ~ 先行研究 ~

先行研究として, 歌詞を検索条件とした楽曲検索システムの構築 [1] がある。そのシステムを概念図 1 に示す。このシステムでは, 楽曲一曲につき一つのテキストファイルを準備し, その楽曲の歌詞そのものをテキストデータとして格納しておく。歌詞データの取得には iTunes と連携させて自動で歌詞を取得することが出来る, フリーソフトの LyricsMaster を使用している。iTunes 内にも歌詞データを入れておくことができるが, 検索の効率化のため, あらかじめ歌詞データのあったファイルを歌詞データベースに格納しておく, インデックスを作成する。そして, 問合せインタフェースからユーザが入力したキーワードを歌詞を含む楽曲を検索し, 検索結果として楽曲名とアーティスト名を表示させる。表示された結果からユーザが聴きたい曲を選択し, iTunes で再生する。

## 3 楽曲選択システムの概要

本研究で開発する楽曲検索システムとは, 多くの楽曲の中から RSS 情報を利用して楽曲を選択することのできるシステムである。音楽プレーヤソフトに多くの楽曲を持っていても, アーティスト名やアルバム名など検索方法が限られてしまうため, 日々同じ曲を聴いてしまったりシングル CD から保存した楽曲が聞きにくかったりというユーザも多いと思われる。そこで,

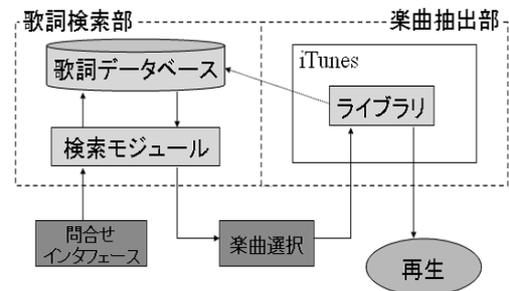


図 1: 楽曲検索システムの構成

日々変化する情報をもとに楽曲を抽出することでこれらの難点は解消されたと考え, 本研究では天気や季節に関する情報を含む『お天気豆知識』という RSS を用いることにした。

### 3.1 RSS からの情報取得

RSS とは, Web サイトの見出しや要約などのメタデータを構造化して記述する XML ベースのフォーマットのことであり, 主にサイトの更新情報を公開するのに使われている。RSS で記述された文書には, Web サイトの各ページのタイトル, 見出し, 要約, 更新時刻などを記述することができる。RSS 文書を用いることで, 構造を使って容易に文書を切り出すことが出来る。しかし本研究で用いた『お天気豆知識』の RSS 文書では, 本文の書き出しのみ記述されている形であり, そこから Web ページにアクセスすることで本文が表示されたため, Web ページから情報を取得することとした。”<http://weather.livedoor.com/mame/日付.html?r=rss>” という URL であったため, 日付を取得しそれをもとに Web ページを自動取得する。取得したページから HTML タグによって本文を切り出し, 残りの不要なタグも除去して文章のみをテキストファイルに保存する。その際, 文字コードを扱うために Jcode.pm というモジュールを使用した。

### 3.2 システム構成

本システムの構成を図 2 に示す。特定の Web ページから情報を取得し, 本文のみ切り出して保存する。そのファイルを Chasen を用いて形態素解析し, 名詞と形容詞をキーワードとして抽出する。キーワードをもとに歌詞検索システムを用いて楽曲を抽出し, iTunes 内にリストとして作成する。

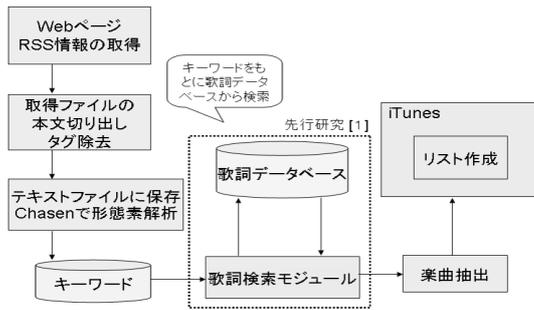


図 2: 楽曲選択システムの構成



図 4: 楽曲選択システムの実装結果

2月	2ツキ	2月	未熟語		
は	ハ	は	名詞-一般		
			助詞-係助詞		
			記号-読点		
(まか)まか	ボカボカ	(まか)まか	副詞-一般		
陽気	ヨウキ	陽気	名詞-形容動詞語幹		
に	ニ	に	助詞-副助詞		
な	ナ	なる	動詞-自立	五段・ラ行	連用タ接続
な	ナ	な	助詞-接続助詞		
薄着	ウスギ	薄着	名詞-サ変接続		
に	ニ	なる	動詞-自立	五段・ラ行	連用タ接続
な	ナ	な	助詞-副助詞-一般		
たか	タカ	たか	動詞-特殊・タ		
か	カ	か	助詞-副助詞/並立助詞/終助詞		
と	ト	と	助詞-格助詞-用		
と思	オモ	思う	動詞-自立	五段・ワ行促音便	仮定形
え	エ	え	助詞-接続助詞		
ば	バ	ば	助詞-接続助詞		
			記号-読点		
EOS					
寒	サム	寒い	形容詞-自立	形容詞・アウオ段	ガル接続
さ	サ	さ	名詞-接尾-特殊		
が	ガ	が	助詞-格助詞-一般		
ぶりかえし	ブリカエシ	ぶりかえし	動詞-自立	五段・サ行	
			記号-読点		

図 3: 形態素解析結果

### 3.3 Chasen によるキーワード抽出

キーワード抽出には Chasen を使用した。Chasen とは日本語を単語分割する形態素解析プログラムツールである。実際に取得した『2月、ぽかぽか陽気になって薄着になったかと思えば、寒さがぶりかえし、一度脱いだ着物をさらに着なおすので「衣」を「更」に重ね「着」するから「衣更着(きさらぎ)」になったという説。』ファイルから Chasen を用いて形態素解析を行なうと、図3のような結果が得られる。本研究では、情報の特色を表しやすい名詞と形容詞のみ取り出し、頻度の高い順にキーワードとしていくこととする。

## 4 楽曲選択システムの実装

### 4.1 実装環境

- リスト作成インタフェース: Visual Basic.NET
- キーワード抽出: Chasen
- 歌詞検索: Namazu
- RSS 情報取得, 楽曲抽出: Perl
- 音楽プレーヤソフト: iTunes 7.0.2

### 4.2 ライブラリからの楽曲抽出法

本システムでは検索により楽曲を抽出し, iTunes 内で新規のプレイリストを作成することを実現している。iTunes 内でのプレイリスト作成には, 音楽再生ソフト用開発ツールである iTunes COM Windows SDK(Apple 社)を使用した。以下に使用した関数について示す。

- CreatePlaylist(): 新規のプレイリストを作成する

- SelectPlaylist(): プレイリストを選択する
- ItemByName(): 指定されたタイトル名の楽曲を抽出
- AddTrack(): 楽曲をプレイリスト内に加える

まず CreatePlaylist() によって新規のプレイリストを作成し, SelectPlaylist() で選択する。そのプレイリストに, ItemByName() によってライブラリ内から抽出した楽曲を AddTrack() によって加える。

### 4.3 実装結果

図4はこれまで述べてきた手順で実装した楽曲選択システムによる検索結果の一例である。RSS 情報をもとに日々違ったプレイリストが作成され, iPod を同期すればプレイリストがそのまま移され, 出かける場合にもその日限定のプレイリストを聞くことができる。

この楽曲選択システムを用いて, 2007年1月18日から24日の1週間, 10曲程度ずつ選択し, 重複の割合の調査を行なった。楽曲の前日との重複度は最大38%, 最小9%, 平均して26%であり, この1週間で iTunes 内のライブラリに入っている楽曲163曲中, 56曲を聴くことが出来た。この結果から, 本研究の楽曲選択システムを用いることで, iTunes 内の楽曲を偏りなく聞けるようになったと言える。

## 5 まとめと今後の課題

本研究では, 多くの楽曲を聞きやすくするために, 日々変化する情報をもとに楽曲を抽出するシステムを提案し, 天気情報を用いて実装を行なった。今後は, より季節感にあった楽曲の抽出を行うために, 検索の対象を歌詞のみにとどまらず, 曲調やテンポからも考察する必要がある。

### [謝辞]

本研究を進めるにあたりご助言・ご助力いただいた, 本学大学院の長澤槇子先輩に深く感謝致します。

### 参考文献

- [1] 長澤槇子: "MP3 データの楽曲を対象とした検索システムの構築", お茶の水女子大学理学部情報科学科卒業論文要旨集, pp.73-74, 2006。