

2.2 音楽データベースの検索における 自動インデクシング法に関する研究

1. はじめに

今日、電子図書館に代表されるようにマルチメディアデータベースが大規模に構築されるようになってきており、音楽のデータも大量に蓄積されるようになってきている。そこで、大量の音楽データの中から目的の音楽を検索する技術に対する要望が高まっている。

音楽検索には、現在大きく分けて二種類ある。1つは、書誌情報検索である。書誌情報とは、曲のタイトルやアーティスト名などテキストで表現できる情報である。この書誌情報を入力して、目的の曲を検索する。一例に、昨今巷をにぎわしている napster が挙げられる。しかし、書誌情報検索では、曲を知っているだけでは検索できず、不便である。例えば、CMなどで耳にしたことがあってメロディを知っていても曲名などを知らない場合、書誌情報検索ではどうしようもない。

そこで、メロディ検索というものが考えられている。メロディ検索は、検索者が問い合わせとして、目的の曲のメロディの一部をマイクやキーボードで入力しそのメロディを含む曲を検索するものである。検索結果としては、問い合わせに一致する曲の書誌情報を検索結果とする。メロディの一部を入力するだけで、曲に関する書誌情報を知らなくても検索できる。

2. メロディ検索システム

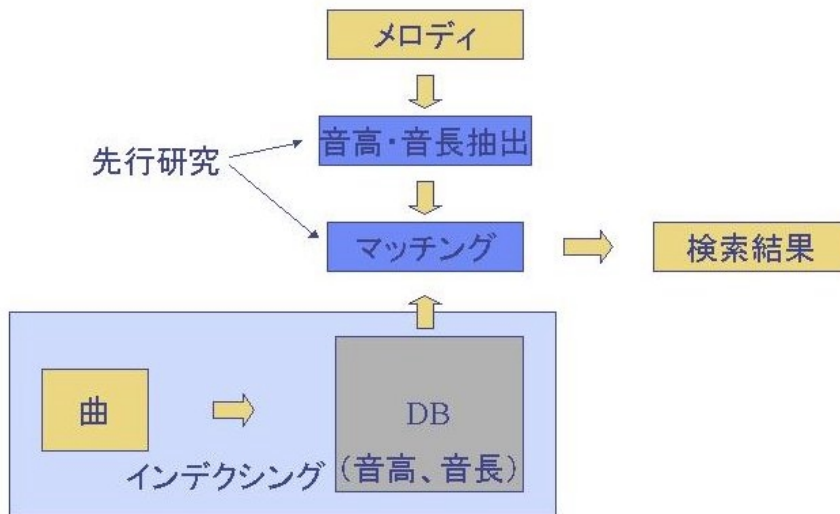


図 1 メロディ検索システム

メロディ検索システムを説明する(図1)。まず、検索者は、問い合わせとして、目的の曲のメロディの一部を、マイクやキーボードを用いて入力する。その、メロディから音高と音長を認識する。一方、データベースには音高と音長が記憶されている。

入力されたメロディの音高と音長とマッチする曲をデータベースから検索する。その結果をユーザーに提示する。音高、音長の認識、マッチングは先行研究があり、かなり深く研究されており、鼻歌で音楽を検索するシステムも構築されているが、データベースの規模が小さいものを対象としており、検索の効率に関してはあまり考えられていない。

ところが、現在データベースの大規模化が進んでおり、効率的に検索するためのインデクシング技術が重要となってきている。

3. インデクシング

インデクシングとは、インデクス(索引)を作成することである。インデクスを作成すれば、検索時、インデクスのみを検索すればよいので検索の効率を上げることができる。

インデクシングを行うには、曲をなんらかの単位に分割しなければならない(図2)。この、単位によって検索精度、検索速度が大きく変わる。そこで、音楽に最適なインデクス単位を考えなくてはならない。



図 2 インデクシング

4. 音楽のインデクシング単位

一般に、インデクスの単位が大きいと検索は速くなり精度は高くなる。なぜなら、単位が大きければ大きいほどインデクスの数が減り検索対象が少なくなるからである。精度が高くなる（検索結果に検索目的以外のものが引っかかりにくい）のは、一つのインデクスが持つ情報量が増え特定性が高まるからである。

例えば、日本語の文章に対するインデクスを考える。日本語には、おおざっぱに言って単語、文節、文という単位がある。文をインデクスの単位として用いたとする。文は単語にくらべ数が少ない。したがって、検索は速くなり、特定性は高くなる。

しかし、検索する時に、文章を入力する人がいるであろうか？また、その入力が目的の文章と正確に一致したものを入力するであろうか？

一般に、日本語のインデクスとしては、単語が用いられている。それは、検索時に単語が単語で入力されるからである。また、単語単位で十分な速度と精度が得られるからである。

つまり、インデクスを用いて検索を速く、精度よくしようと思うとき、やみくもにインデクスの単位を大きくすれば良いというものではなく、検索時に入力される単位を考慮に入れなければならない。

音楽では、フレーズが検索単位に一致することが、単純なメロディ検索システムを構築し検索実験を行ったことからわかった。また、フレーズは曲を特定するのに十分な情報を持っている。従って、音楽に対するインデクスの単位はフレーズが適切であると思われる。フレーズとは、一つのまとまりある音楽の流れのことを言う。以下にフレーズがインデクスの単位として良い根拠を列挙しておく。

- 検索者の記憶に一致している

単純なメロディ検索システムを構築し検索実験を行ったところ検索者はフレーズ単位で問い合わせを入力することがわかった

- 曲の特徴を十分に表現している
フレーズは曲を特定するのに十分な情報をもっており検索の精度を高めることができる
- フレーズ単位で検索するので高速

5. インデクシングの自動化

音楽データベースの大規模化が進み、自動インデクシング技術がますます重要になってきている。現在のメロディ検索システムではインデクシングの作業はほとんどのものが手作業で行っている。自動のインデクシング技術として N-Gram が提案されているが、N-Gram ではインデクスの単位が 1 音符であり検索精度が悪く、検索速度が遅いなどデメリットが多い。フレーズによるインデクシングは、検索精度、検索速度共に高く、フレーズの認識を自動化できれば、大規模データベースにおいても、よいインデクスを自動で作成できる。

本研究の目標としては、フレーズ単位でのインデクシングを自動化することにより、大規模データベースにおいても、高精度、高速な検索を実現することである。

6. フレーズ単位のインデクシングの自動化手法

今日、フレーズを自動で認識しようという試みがなされている。それでは、フレーズの認識に音楽の情報のみを用いている。しかし、フレーズは、非常に複雑な音楽的特徴を持っており、音楽の情報からのみでは、その認識率は、7割程度にとどまっている。つまり、この技術を用いてインデクスを作るとデータの 70% のみしか検索対象にならない。これでは、検索もれが多く検索に用いるには不十分である。

我々は、歌詞付きの音楽について、歌詞とフレーズの切れ目に関係があることを見出した。音楽の情報と歌詞の情報を併用すれば、音楽のみで認識するよりフレーズの認識精度が高まると考えられる。ここでは、我々は、歌詞から得られるフレーズの情報と、音楽から得られるフレーズの情報とを併用してフレーズを認識し、自動インデクシングを実現することが有効であるか実験を行い確認した。

7. 歌詞とフレーズの関係

実験に歌詞のついた曲を多数用いたことから歌詞とフレーズの関係に注目し

フレーズと歌詞の切れ目が一致するのではないかという結論を得た。そこで、歌詞を文節を用いて分割したものとメロディをフレーズ単位に分割したものを比較し関連性の分析を行った。その結果、「フレーズの切れ目が歌詞の切れ目と一致する」ことがわかった(図3)。つまり、文節からフレーズの切れ目の候補を見つけ出すことができるということである。この情報から正しいフレーズの切れ目を認識するには、得られた候補から絞り込む必要がある。

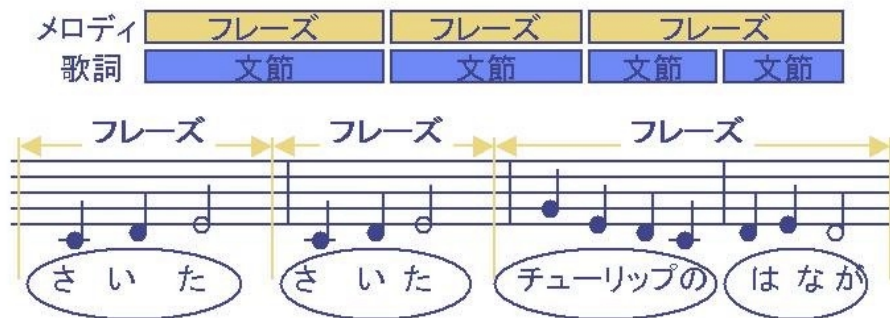


図 3 歌詞とフレーズの関係

文節を自動的に切り出すのに、京都大学で開発された自動係り受け解析器を用いた。KNPはJUMAN(形態素解析器(京都大学))からの情報に基づいて文節を切り出す。

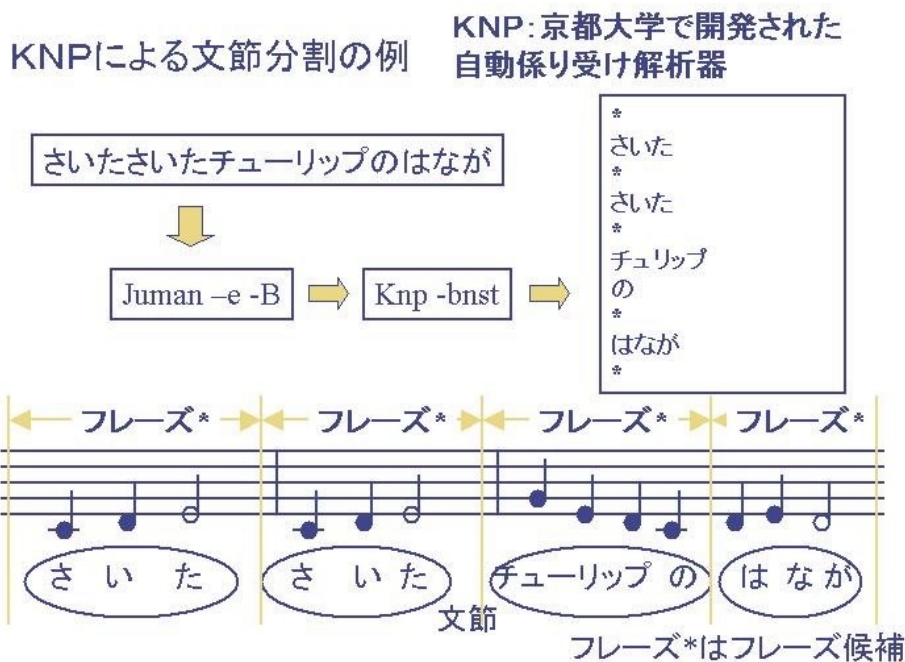


図 4 KNPによる文節分割の例

8. 音長によるフレーズ絞り込み

正しいフレーズの切れ目に絞り込むため、フレーズの切れ目の音楽的特徴を解析したところ、切れ目においては音長が大きいことがわかった。音長が大きい音符を「曲の平均音長より大きい音符」と定義し、文節より得られたフレーズの切れ目の候補から絞り込んだ(図5)。



図5 音長によるフレーズへの絞り込み

9. 実験

本手法を評価するために、実験を行った。手作業でフレーズを認識したもの(フレーズ)、手作業で文節を抜き出したもの(文節)、その文節から音長でフレーズを認識したもの(文節+音価)、自動文節切り出しプログラム KNP(京都大学)と音長でフレーズを認識したもの(KNP+音長)を用いて、これらと比較した。また、これらから得られたインデクスを用いて検索実験も行った。

結果

インデクス手法	認識数	一致数	作業
フレーズ	539	539	手動
文節	1439	534	手動
文節 + 音長	721	507	手動
KNP + 音長	575	369	自動

表 1 認識数と一致数の比較

インデクス手法	認識比	一致 (%)	作業
フレーズ	1	100	手動
文節	2.7	99	手動
文節 + 音長	1.3	94	手動
KNP + 音長	1.1	68	自動

表 2 認識数と一致数の比による比較

検索実験

インデクス手法	1 位	10 位内	時間(s)
フレーズ	99	100	0.40
文節	94	100	0.94
文節 + 音長	97	98	0.43
KNP + 音長	85	88	0.41

表 3 検索精度 (%) と検索時間

考察

まず、文節と音長の情報を用いることの有効性を判断する。文節 + 音長とフレーズを比較すると、文節 + 音長の認識数が若干多いが、一致（正しいフレーズと一致した認識フレーズ）したものが94%とほぼすべてのフレーズを認識することが出来ている。検索実験からも、文節 + 音長はフレーズと同程度の性能が実現されていることがわかる。したがって、文節 + 音長をインデクスに用いることは有効であると判断できる（表1）（表2）。

KNP + 音長は、検索実験において最低の結果をしめしている（表3）。これは、正しいフレーズを70%しか認識できていないためである。70%しか認識できていないのは、文節 + 音長と比較すればわかるように、文節の自動切り出しが原因である。これは、KNPの性能よりも形態素解析レベルでのあやまりであることが分析からわかった。

まとめると、文節 + 音長を用いることは有効である。しかし、形態素解析の精度が問題で自動ではインデクスとしての利用には耐えられない。

10. 今後の予定

歌詞を用いた自動インデクシングの実現のために、形態素解析の精度を上げなければならない。

今回の実験では、音符と一対一の対応を持つひらがなのみの歌詞に対して形態素解析を適応した。ひらがなのみでは、曖昧性が高いので精度が落ちる。

そこで、漢字混じりの歌詞に対して形態素解析を適応することで、形態素解析の精度の向上をはかり、利用に耐えるインデクスの自動化をめざしていきたい。

参考文献

- [1] 貝塚 智憲, 後藤 真孝, 村岡 洋一 歌声の旋律情報と歌詞情報をキーとした曲検索システム, 54 回情報処理学会全国大会, 1997
- [2] 蔭山 哲也, 高島 洋典 ハミング歌唱を手掛かりとするメロディ検索, 電子情報通信学会論文誌 D-II Vol. J77-D-II No. 8, pp. 1543- 1551, 1994.
- [3] 園田 智也, 後藤 真孝, 村岡 洋一 WWW 上での歌声による曲検索システム, 電子情報通信学会技術報告, SP97-103, 1998
- [4] 園田 智也, 村岡 洋一 DP マッチングを用いる検索の DB 索引作成方法と応用, 情報処理学会第 60 回全国大会

- [5] 西原 祐一,小杉 尚子,紺谷 精一,山室 雅司 時間正規化を用いたハミング検索システム, 情報処理学会研究会報告,99-MUS-30, 音楽情報科学,30-6,1999
- [6] 西原 祐一, 小杉 尚子,紺谷 精一,山室 雅司 ハミング検索用音楽データベースの分析～部分曲分割方法の改善に向けて, 情報処理学会研究会報告,99-MUS-2, 音楽情報学,32-2,1999
- [7] 柳瀬 隆史,高須 淳宏,安達 淳 音楽検索における自動インデクシング法, 情報処理学会研究会報告,DBS116-42,1998
- [8] Atsuhiko Takasu, Takashi Yanase, Teruhito Kanazawa, and Jun Adachi Music Structure Analysis and Its Application to Theme Phrase Extraction LNCS 1696, p. 92 ff.
- [9] 成田 智也,杉山 雅英 楽曲の高速検索手法の検討,電子情報通信学会技術研究報告,SP,音声, Vol.2000 No.16
- [10] YIP Chi Lap, Ben KAO, A STUDY ON N-GRAM INDEXING OF MUSICAL FEATURES ICME2000, TP2 Audio Retrieval