

F-26 MPEG-7を用いた多次元ベクトル空間法による音楽検索 Music retrieval by multidimensional vector space method using MPEG-7

比嘉 陽一†
Youichi HIGA

長山 格‡
Itaru NAGAYAMA

1. はじめに

現在、音楽検索は主に Web 等によるテキスト検索、鼻歌などを用いたハミング検索 [1], そしてキーワードによる検索 [2] など多種多様な検索手法が存在する。しかし、多種にわたる検索手法が存在しながら、それらを複合して検索結果に反映させる手法は未だ一般的ではないと言えよう。そこで本研究ではこれら 3つの手法で用いる尺度をベクトルで表現し検索結果に反映させる事が出来る複合検索手法を提案する。

2. 複合検索について

研究に際し、18歳~22歳までの男女49名に対して人間の音楽に対する意識調査を行った。質問は、

- 質問 1. 初めて聞いた曲をどうやって探しますか?
質問 2. 音楽のどこに注目して聴いていますか?
質問 3. 好みの曲をどのようにして相手に伝えますか?

の三つを自由形式で答えてもらった。質問 1 に対する回答の集計を図 1 に示す。これを見るのと分かるように、人に聞くという結果が一番多い。アンケートによる具体的な回答によると

- レコード店の店員に曲の雰囲気伝える (19歳, 男)
- 歌っている人の性別とかジャンルとかを店員に言って探す (20歳, 女)
- 友人にメロディーを自分で歌って聞く (18歳, 男)

などさまざまである。

表 1, 表 2 は質問 2, 質問 3 を集計したもので、複数回答を含む。これらの事例より、人間は音楽を一元的ではなく曲のメロディ、曲名、歌詞やキーワード等多次元的に認識していると考えられる。一方、従来の音楽検索では曲のメロディのみ、曲名のみ、キーワードのみのいずれかで検索する手法が主流であり、複合的な検索を行う手法は一般的ではない。そもそもこれらは独立した概念であり、単純には比較出来ない。しかし共通の尺度があればこれら異なった概念を比較する事が可能である。そこで本研究では曲データを多次元ベクトルで表現し、異なる概念を併合した検索を試みる。

3. 多次元空間検索と音楽への応用

通常のリレーショナルデータベースでは B+Tree 構造を用い次元の値を大小関係によって構造化している [3]。それに対して多次元空間索引とは、図 2 のように 2 次元以上の空間を構造化したものであり、データを数値化し多次元空間上に配置し、検索する際入力されたデータを

†琉球大学大学院理工学研究科情報工学専攻
Graduate school of Engineering, University of the Ryukyus.

‡琉球大学工学部情報工学科
Department of Information Engineering, University of the Ryukyus.

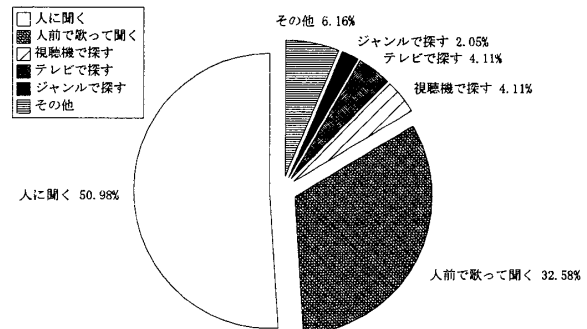


図 1: 質問 1. に対する回答の集計

表 1: 質問 2. 音楽のどこに注目して聴いていますか?

歌詞	雰囲気	声	ノリ	メロディ	サビ	その他
44.9%	14.3%	12.2%	8.2%	4.1%		12.2%

表 2: 質問 3. 好みの曲をどのようにして相手に伝えますか?

聴かせる	歌う	ジャンル	曲名	歌手名	雰囲気	その他
36.7%	22.4%	22.4%	18.4%	16.3%	14.3%	6.1%

ベクトルとして入力データに近いベクトルを求めるものである。多次元空間索引を用いることにより、高速な検索が出来る他、それぞれの検索キーを複合的に用いることが出来るという利点が考えられる。本研究ではこれを踏まえ、複数の検索キーをベクトルに変換しそれらを併せた値で音楽を表現する事を考える。すなわち、ある曲をメロディ、タイトル、曲のイメージといった複数の概念でそれぞれ表現し、それぞれの概念をベクトルとする。そしてこれらベクトルの集約を曲とする。このように曲そのものを多次元空間上に配置することにより、複数の概念を用いて曲を表現する事が可能となる。

3.1 多次元空間上での音楽の表現

多次元空間上で音楽を表現する具体的な手法を示す。本研究の検索対象となる例の曲 (きらきら星) の楽譜を図 3 に示す。また、図 4 は図 3 を音階の変化を数値で表現したもので、ここでは初めの音 (ド) を 1 としている。この音階の相対的な変化が検索対象のデータとなり、これをキーにして検索を行う。この時、音の長さは考慮しない。人間が音楽を認識する場合、元曲とは違う調であ

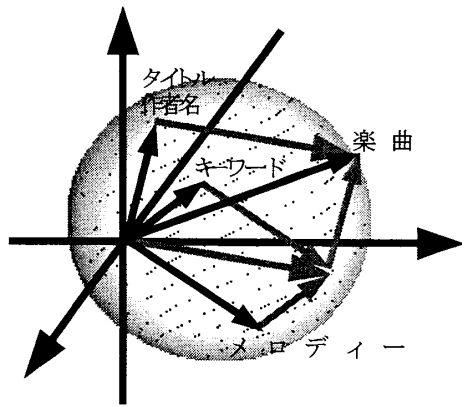


図 2: 多次元空間ベクトルの音楽への適用



図 3: きらきら星の譜例

ることがよくある。また、曲のテンポが人によってまちまちな事もある [4]。この形式は音の高低の変化のみに着目しているので調やテンポを無視することができ、利用者は違う調で入力しても検索できる。このように得られる数値を曲のメロディとしベクトル化する。歌詞や曲の雰囲気となるキーワード等もそれぞれを多次元空間上に配置しこれら要素を用いベクトルとして表す。

3.2 音楽の検索について

検索時に入力されたキー(メロディ、キーワード等)も上記の手法によりベクトル化し、これと候補となるベクトルとの内積を求める。この手法を用いる事により、候補となるベクトル同士の類似度を求める事が出来るというメリットがある。

4. MPEG-7

MPEG-7の正式名称はMultimedia Content Description Interface といい、通称MPEG(Moving Picture Experts Group)[5]として知られるISO/IEC JTC1 SC29/WG11において策定されたマルチメディア・コンテンツに対するメタデータの表記方法に関する国際標準規格である。通常MPEGといえばMPEG-1, MPEG-2, MPEG-4といった映像音声データの圧縮に関する標準化技術として知られている。しかしMPEG-7はこれまでのMPEGの規格と違い、データの内容を記述するための表記方法に関する国際標準規格である。MPEG-7ではXML Schemaをベースに、マルチメディア・コンテンツの特徴記述の際に必要な各種データ型などを追加することで記述定義言語を規定している [6]。そのため、一つのデータファイルに対し曲のメロディ、歌詞、作

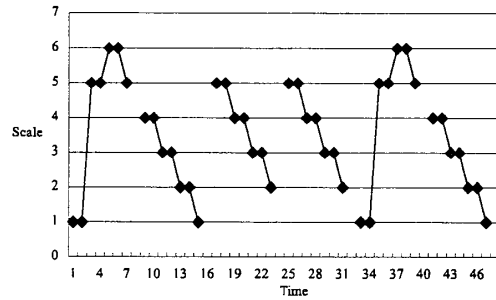


図 4: 譜面を数値化したグラフ

者や演奏者の情報、感情に関する記述等の情報を抽出しやすい形で定義できる。[7]これにより多様な検索を実現する事が可能となるためMPEG-7を活用する事は有効であると考えられる。

5. 今後の課題

本研究で提案する音楽検索手法では、WebServiceとして提供可能な機能を持たせるためSimple Object Access Protocol (SOAP)[8]の実装を検討している。SOAPとは、HTTPとXMLを使ってリモート・プロシジャー・コール(RPC/Remote Procedure Call)を実現することを目的としたプロトコルであり、ネットワークを通じて、公表、実行可能な自己完結したアプリケーションを目指すWeb Serviceの中核となる。この機能を持つ事により多種多様なアプリケーションでの実装が容易になる。

参考文献

- [1] 西村拓一, 橋口博樹, 関本信博, 張建新, 後藤真孝, 岡隆一: “始端特徴依存連続DPを用いた鼻歌入力による楽曲信号のスポッティング検索の高速化” 情報処理学会研究報告 Vol.2001.No.103 pp.7-13
- [2] 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.3201-3211(Dec.2001)
- [3] 北川博之: “データベースシステム” 昭晃堂, 2002
- [4] 古田典久, 野池賢二, 野瀬隆, 乾伸雄, 小谷善行, 西村恕彦: “曲検索システムのための音符列類似度の設計”, 情報処理学会第56回全国大会講演論文集, 4M-1, pp.2-40 - 2-41, 1998.
- [5] <http://mpeg.telecomitalia.com>
- [6] <http://www.itscj.ipsj.or.jp/mpeg7>
- [7] MPEG-7 Multimedia Description Schemes XM(Version3.1)
- [8] <http://www.w3.org/TR/SOAP/>