

既存の音楽を手掛かりとする音楽検索システム

5G-3

門田 暁人† 松本 健一† 井上 克郎†, ‡

†奈良先端科学技術大学院大学 情報科学研究科

‡大阪大学大学院 基礎工学研究科

1 はじめに

映画などのマルチメディア作品の製作においては、場面に応じたバックグラウンド・ミュージックを入れる場合がある。製作者は、自分の頭の中のイメージに合う曲を見つけてくる必要があるが、それは必ずしも容易ではない。製作者が今までに聞いたことのある曲（既知の曲）の中に、自分のイメージに合う曲がない場合がある。また、イメージに合う曲が一つ見つかったとしても、作品中で同じ曲を繰り返し使えない場合には、似たイメージを持つ別の曲を探す必要がある。

本研究の目的は、既知の曲を手掛かりとして、似たイメージを持つ、あるいは、全体的には似たイメージを持つが一部が異なるような未知の曲を検索するシステムを開発することである。本システムでは、手掛かりとなる曲のイメージと検索したい曲のイメージの間の、類似点や相違点を検索条件として指定できる。例えば、手掛かりとなる曲について「明るさ、力強さが似ており、もう少し重たい感じのする曲」というような検索条件を指定できる。検索者は、既知の曲のうち自分のイメージに比較的近いと思われる曲を手掛かりとなる曲として選び、検索を進めていくことになる。自分のイメージ通りの曲が得られなかった場合でも、検索の結果得られた曲を新たな手掛かりとし、さらに検索を続けていくことで、最終的にイメージに合う曲が得られると期待される。

以上のようなシステムを実現するために、本研究では、それぞれの曲に対するイメージを、「明るさ」

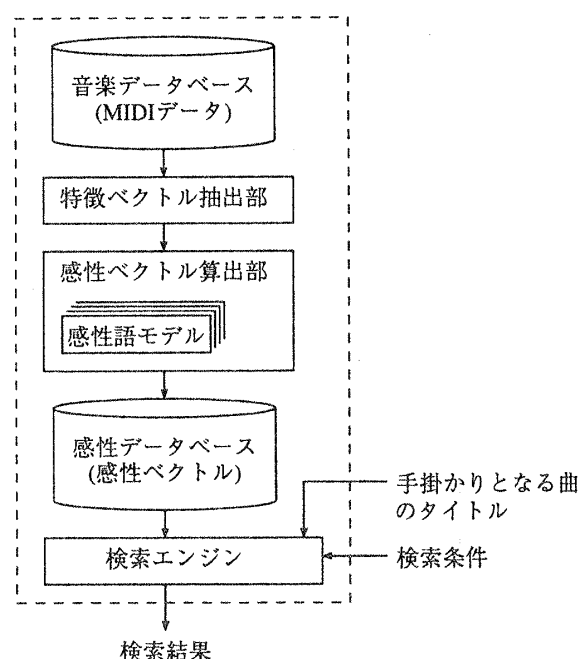


図1: 音楽検索システムの構成

「力強さ」「軽やかさ」などの複数の感性語に対する評価値の組（以下、感性ベクトルと呼ぶ）で表現する。例えば、ある曲に対する感性ベクトルを、(明るさ 3.2, 力強さ 2.8, 軽やかさ 1.2, 激しさ 4.0, 滑らかさ 1.2) というような数値の組で表す。2曲間の明るさ、力強さ等の類似の度合は、2曲間の感性ベクトルの各要素の差によって決定される。

以降、本報告では、検索システムの構成、および、感性ベクトルの算出方法について述べる。また、システムの実現の第一歩として、曲の「明るさ」の評価を試みた実験結果についても述べる。

2 音楽検索システムの構成

音楽検索システムの構成を図1に示す。システム各部の機能を以下に示す。

音楽データベース 検索候補となる曲の集合である。各曲は、標準 MIDI ファイル (SMF: Standard MIDI

A music retrieval system which uses known music as a reference key

Akito Monden†, Ken-ichi Matsumoto† and Katsuro Inoue†, ‡

†Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma-shi, Nara 630-0101, Japan

‡Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan {akito-m, matsumoto, k-inoue}@is.aist-nara.ac.jp

File) 形式で保存される。

特徴ベクトル抽出部 各 MIDI データから、曲の構造の特徴づける特徴ベクトルを導出する。ベクトルの各要素としては、音の分布、テンポ、コード進行、リズムパターン、音色、符割り、スケール等を数値化したものが考えられる。

感性ベクトル算出部 感性ベクトルの各要素の値を、特徴ベクトルを説明変数とする感性語モデルにより算出する。重回帰分析による線形モデルやニューラルネット等の非線形モデルが考えられる。

感性データベース 各曲の感性ベクトルの集合である。
検索エンジン 手掛かりとなる曲のタイトル(曲名)、および、検索条件を検索者から受け取り、検索結果を出力する。

3 感性ベクトルの算出

本システムの特徴は、各曲から抽出した特徴ベクトルをそのまま曲の類似度の算出に使うのではなく、曲のイメージを表す指標である感性ベクトルに変換したものをを使う点にある。感性ベクトルは、音楽のイメージを表す感性語(明るい、力強いなど)に対する評価値の組で表されるため、人間の直観に基づいた検索が可能となる。

各感性語は、表 1 に示すように、対になる形容詞を持っている [1]。例えば、「明るい」という感性語に対しては反意語として「暗い」という形容詞が対になる。明るさを 1~5 の 5 段階で評価する場合には、人間がある曲を非常に明るいと感じる場合に「明るさ 5」となり、非常に暗いと感じる場合に「明るさ 1」となる。

表 1: 感性語対の例

明るい	5	⇔	1	暗い
力強い	5	⇔	1	弱々しい
軽やかな	5	⇔	1	重々しい
激しい	5	⇔	1	穏やかな
滑らかな	5	⇔	1	粗い

現時点では、どのような感性語と特徴ベクトルが適当であるかは明らかでない。複数の人間が利用する検索システムの構築のためには、複数人の間で評価値がばらつかないような感性語を選ぶ必要がある。

また、精度のよい感性語モデルが構築できるような特徴ベクトルを選ぶ必要がある。

4 実験

実験では、曲の途中で転調していない日本のポップスの MIDI データを 50 個用いた。あらかじめ長調の曲は全てハ長調に移調し、短調の曲はハ長調の並行調であるイ短調に移調した。ベースパートにおける 12 音階 (ド, ド♭, レ, …, シ) の各音の出現頻度を 12 次元の特徴ベクトルとした。また、特徴ベクトルを説明変数とし、「明るさ」の評価値を目的変数とする、重回帰分析による感性語モデルを作成した。

4 名の被験者に、各曲の明るさを SD 法を用いて 5 段階で評価してもらった。50 曲のうちからランダムに選んだ 30 曲を用いて感性語モデルを構築し、残りの 20 曲に適用してモデルの性能を評価した。

50 曲についての、各人の評価値と平均評価値との相関係数は、それぞれ 0.82, 0.89, 0.75, 0.78 となり、ある曲を明るいと感じるかどうかにについて各人が似た傾向を示すことが分かった。「明るさ」は曲の印象を表す一般的な感性語として、有力な候補であると言える。

また、評価用の 30 曲に対する、各人の評価値と感性語モデルによる評価値(モデルの出力値)との標準誤差は、0.80~0.92 程度となった。実験で用いた 12 次元の特徴ベクトルは、曲の明るさをある程度説明できると言える。

5 おわりに

既存の音楽を手掛かりとする音楽検索システムの概念とその構成について述べた。また、曲の「明るさ」の評価を試みた実験結果について述べた。

実験の結果、「明るさ」は曲の印象を表す一般的な感性語として、有力な候補であることが分かった。今後、他の感性語についても検討する予定である。

謝辞 本研究は、奈良先端科学技術大学院大学 98 年度情報科学研究科研究育成基金の補助による。

参考文献

- [1] 坂本崇, 梶川嘉延, 野村康雄: 音楽感性空間における非線形判別分析を用いた曲印象別グループの分割, 情処論, 40, 4, pp.1901-1909 (1999).