

BGM 選択のための楽曲の心理的な印象空間へのマッピングによる 直感的な楽曲検索支援システム

A background music retrieval system supporting intuitive selection and visualization by mapping psychological impression space

小松住人*1
Sumito Komatsu

西尾典洋*2
Norihiro Nishio

杉山岳弘*1
Takahiro Sugiyama

*1 静岡大学情報学部
Faculty of Informatics, Shizuoka University

*2 静岡大学創造科学技術大学院
Graduate School of Science and Technology, Shizuoka University

We propose a background music retrieval system to support selecting music intuitively and visualizing on psychological impression space. The mapping of BGM in the space of impression words uses foundation data by the psychological experiment. An effect to discover new background music heuristically was shown with the trial system.

1. はじめに

膨大な音楽データの中から求める楽曲を探し出す作業には、多大な労力と時間を必要とする。音楽データを扱う場合の問題点として、まず、1曲1曲試聴しながら、試行錯誤を繰り返す必要がある。また、画像データのように一覧表示ができないため、直感的にデータの特徴を把握することが難しいという点がある。

また、既存の楽曲検索システムは、どれも音楽のみに焦点を当てたものが多く、映像と対応つけて音楽データを検索するものは少ない。映像編集における BGM 選択では、如何に映像とマッチしている音楽を探してくるかということが重要である。

音楽検索の先行研究では、ユークリッド距離[熊本 04]やマハラビス距離[土橋 06]等を利用して、求める音楽データに最も近いものを抽出する事例が多くある。しかし、映像に合う BGM を探す場合は、必ずしも自分が求めるイメージに近い必要はない。すなわち、自分が予想しない意外な音楽データや、新奇な音楽データを提示する方法も、BGM 選択の際には重要となる。

本研究では、心理実験により音楽データと印象語とを空間上で対応付けし、これをもとに、検索された音楽データを印象の軸で構成される空間上にマッピングして視覚化することで、直感的な楽曲検索支援を目指す。図 1 に本研究の枠組みを示す。

2. BGM 選択作業の分析

2.1 ビデオ分析とインタビュー調査

映像編集者の BGM 選択作業を分析するために、映像編集経験初級者・中級者・上級者、計 5 名の被験者に対して、普段 BGM を選択している方法で作業してもらい、そのプロセスをビデオで撮影した。その過程で、どのような思考のもとでその動作をしたのかなどを問う、インタビュー調査を並行して行った。その結果、各被験者に共通して見られた点は、どちらも映像をベースにした印象に着目していた点である。

2.2 BGM 選定の判断基準の形容詞化とクラスタリング

次に、映像編集者に BGM 選定の理由を形容詞化してもらった。そこで得られた 48 個の形容詞対(明るい-暗い、楽しい-つまらない等)を、著者が独自に、主観的に類似性を基にして 12 個のグループに分類した。これらを利用して、音楽と印象語とを

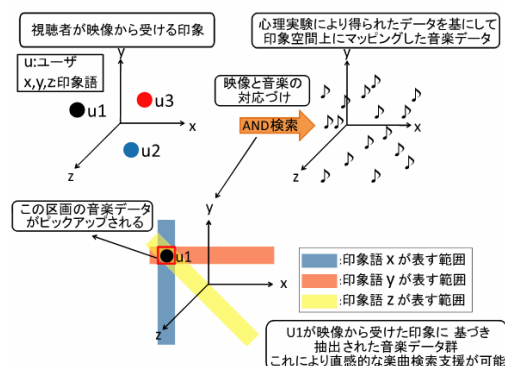


図 1 映像と音楽を対応つけて楽曲検索を行う枠組み空間上で対応付けするため、以下のような心理実験を行った。

- 形容詞対: 60 個 (客観データ: 48 個, 主観データ: 12 個)
- 被験者: 10 人 (男性: 7 人, 女性: 3 人, 21~24 歳)
- 音楽データ: 100 曲

この心理実験によって得られたデータを、まずクラスタ分析した。クラスタ間の距離グラフを基に、クラスタ数を 7 とし、表 1 のようにラベル付けした(クラスタの持つ形容詞対は、感情: しっかりと落ち着いた-高揚感のある、シーン: オープニング-エンディング等)。さらに、7 つのクラスタの中から BGM 選択の際に有効と思われるクラスタを絞り込む。そこで、音楽のイメージと印象語との関係性が強い形容詞対を多く含んでいるクラスタを選出するために、各クラスタにおける、印象語の尺度(範囲 1~5)による標準偏差を算出した。

表 1 各クラスタにおける標準偏差

	緊張感	ジャンル	感情	上品さ
標準偏差	1.242632	1.277571	0.9117	0.899602

	テンポ	対象年齢	シーン
標準偏差	1.045815	0.940364	0.996104

標準偏差の値において上位 4 つのクラスタ(感情、上品さ、対象年齢、シーン)を、BGM 選択の際に有効と思われるクラスタと判断した。

3. 解決策の提案と BGM 検索システムの試作

試行錯誤を減少させるため、2 種類の検索手法を提案する。

また、音楽ファイルの一覧性の欠如を解決するため、印象空間上に音楽データをマッピングし、音楽データの視覚化を提案する。なお、ここで言う印象空間とは、2.2 で説明した実験において、音楽と印象語とが対応づけられた空間を指すこととする。

3.1 言語的特徴検索

映像編集者は、映像の印象に着目して BGM を探す。既存の楽曲検索システムのように、映像全体の印象を形容詞化して検索する方法では、形容詞化しにくい映像の場合はうまく機能しない。そこで本手法では、映像における最も際立たせたい対象(主役)と、それを取り巻く場所(舞台)に着目する。主役から受ける印象を予め、印象語を登録するための連想ワードデータベースに登録することにより、主役の印象にマッチした BGM を検索するという点が、本検索手法の特徴である。加えて、例えば、ペンギンに焦点が当てられている映像において、それが南極での映像なのか、真夏の動物園での映像なのかで印象が異なる。そこで、映像における主役だけでなく、舞台から受ける印象も検索条件に利用し、映像にマッチした音楽データを抽出する。

3.2 印象語による絞り込み検索

2.2 で述べた手法によって得られた、感情・上品さ・対象年齢・シーンという4つの印象語に対して、表2に示すような、印象尺度を構成する印象語の対を用意した。これら4つの特徴を利用して、音楽データを絞り込む。図2に検索インタフェースを示す。上部に言語的特徴検索エリア、下部に、スライダで操作する印象語による絞り込み検索エリアを配置している。さらに、特に重視したい印象語を選択し、検索を開始する。

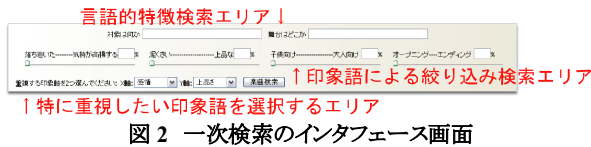


図2 一次検索のインタフェース画面

表2 印象語による絞り込み検索で利用する4つの印象語

特徴	印象尺度を構成する印象語の対
感情	しっとり落ち着いた — 高揚感のある
上品さ	泥臭い — 上品な
対象年齢	子供向けな — 大人向けな
シーン	オープニング — エンディング

3.3 視覚的空間選択

2.2 の心理実験で抽出した4つの特徴で構成される印象空間へ、検索結果をマッピングし、音楽データの可視化を図る。図3は実際に音楽データをマッピングした画面である。

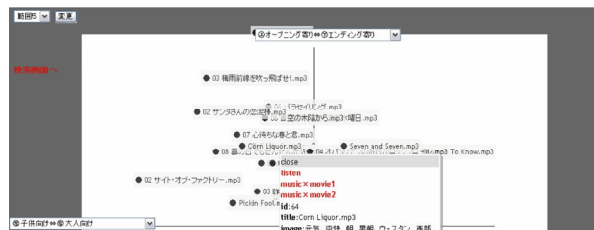


図3 心理的印象空間への音楽データのマッピング

図3では、X軸(横軸)に対象年齢を表す印象語が、Y軸(縦軸)にシーンを表す印象語が代入されている。各印象語における各音楽データの持つ特徴量を基にして空間上に音楽データは配置される。また、X軸・Y軸にそれぞれ4つの特徴のうち2つの特徴を代入することによって、音楽データが持つ特徴

量に応じて配置が変化する。これによりユーザは、視覚的に音楽データの特徴を空間的に把握でき、例えば、ある曲を基準として印象を相対的かつ直感的に楽曲検索が可能となる。

3.4 BGM 選択支援システムの試作

本システムは、データベースに対して各種の検索を行う検索部、データベースに対してデータを追記する登録部、検索結果をブラウザに表示するための表示部から構成される。検索部は、連想ワードに着目した言語的特徴検索と、各種データ分析により得た印象語による絞り込み検索から構成される。登録部は主として音楽データのメタデータを登録した音楽データベース、二段階の検索で利用する連想ワードデータベース、入力された各種特徴量のパラメータをシステムが処理できるように変換する特徴量変換データベースから構成される。表示部ではデータベースから抽出された検索結果を印象空間上にマッピングする。

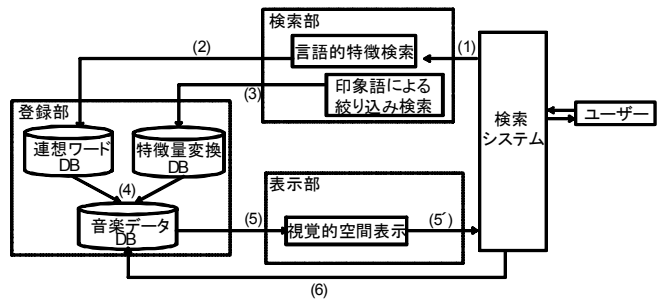


図4 システム構成図

4. 評価実験

本研究で試作した BGM 選択支援システムによって、提案した2種類の検索手法、並びに検索画面の操作性から検索結果画面の表示に至るまでのインタフェースの有効性を調査する。

被験者には、実験用に用意した2種類の映像コンテンツに対して付ける BGM を、試作した BGM 選択支援システムを利用した場合と利用しない場合の2パターンで選定してもらう。両パターンにおける作業時間を計り、選択作業のプロセスを比較する。結果、5名の被験者から、安定した時間で映像にマッチする音楽を取り出すことができ、純粋に BGM 選択を楽しむことができるといった、有効な結果を得ることができた。

5. まとめ

BGM 選択におけるボトルネックに対して、映像編集者の BGM 選択作業を分析した。これにより考案した言語的特徴検索、印象語による絞り込み検索、視覚的空間選択という3つの手法で解決を図った結果、それらのアプローチの有効性を確認することができた。

参考文献

- [熊本 04] 熊本忠彦: "印象に基づく楽曲検索のための個人適応手法の設計と評価," 第3回情報科学技術フォーラム, pp. 55-58 (2004.8)
- [土橋 06] 土橋佑亮, 片寄晴弘: "SOM を用いたベースラインからの音楽ジャンル解析," 音楽情報科学研究会 (SIGMUS) [情報処理学会], pp. 31-36 (2006.8)