

形容詞対を用いた音楽データベース検索システムに関する研究

3G-9

池添 剛 岡西 正 梶川 嘉延 野村 康雄

関西大学工学部電子工学科

1. はじめに

計算機が人間の感情を理解するシステムの一手法として、感性語による検索システム^{[1][2]}がある。しかし従来のシステムでは、例えば同じ「悲しい」と表現されるものでも「激しい悲しさ」や「穏やかな悲しさ」といった違いを表現できない。そこで、入力に人間の感情を表す形容詞対を数値で与えると、その感情に見合う曲を自動的に出力するシステムの構築を行う。このシステムにより、同じ「悲しい」という言葉で表される曲であってもさらに細かいニュアンスで検索を行うことができる。

人間の感情や曲の印象を数値化するために、セマンティック・ディファレンシャル法（以下SD法^[3]）による試聴実験を行い、その結果に因子分析^[3]を施した。なお、その結果を本研究では音楽感性空間^[4]と名付けている。また、検索方法は実際の音楽感性空間座標値（以下実際値）と音楽感性空間座標の子測値（以下子測値）との比較で行っている。

2. システム概要

形容詞対による音楽データベース検索システム（以下検索システム）の構成図を Fig.1 に示す。以下に Fig.1 の説明を行う。

「音楽感性空間の生成」では、人間の感情や曲の印象を数値化する。その方法はSD法による試聴実験を行い、その結果から主因子法^[3]により因子を求め、さらに基準化バリマックス回転^[3]を施す。その結果得られた座標値を本研究では「音楽感性空間座標値」と呼び、それを検索空間としている。

「検索システム」では、曲の検索および検索結果の出力を行う。その方法は、まず重回帰分析^[5]あるいはニューラルネットワーク（以下NN）を用いて、試聴実験で用いた形容詞対に対する子測値を求める。次にその子測値から距離の小さい実際値の曲から順に第1, 2候補として検索システムが出力する。

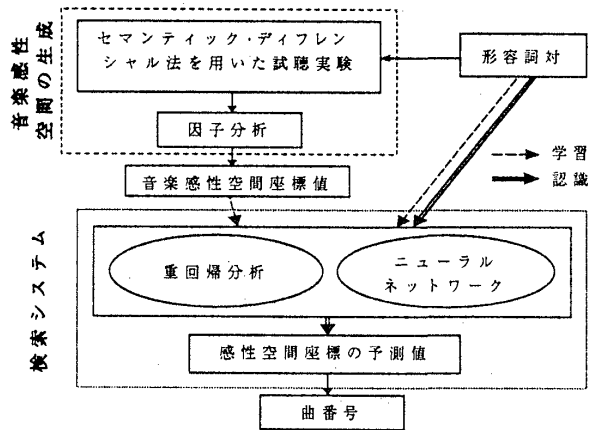


Fig.1 Organization of music database retrieving system.

3. 検索手順

ここでは、未知曲にも対応した検索システムの検索手順について述べる。

3.1 音楽感性空間の生成

実際のリスナーが演奏を聴取したときの印象を定量的に解析することによって初めて計算機でそれらを扱うことが可能となる。そこで、まずSD法を用いて実際のリスナーによる心理聴取実験を行う。SD法における指標は、「安定-不安定、滑らか-粗い、軽い-重い、柔らかい-硬い、透明-不透明、明るい-暗い、力強い-弱々しい」のそれぞれに対しての7段階評価である。また、被験者は関西大学交響楽団員40名、関西大学軽音楽部部員40名、合計80名であり、聴取曲は「ブルグミュラー25の練習曲」より25曲を用いた。次に試聴実験結果に因子分析を施す。その結果、第3因子まで求ま

A Study on Music Database Retrieving System
by Pair of Adjective
Takeshi Ikezoe, Tadasu Okanishi,
Yoshinobu Kajikawa, Yasuo Nomura
Department of Electronics, Faculty of Engineering,
Kansai University
3-3-35 Yamate-cho, Suita-shi, Osaka, 564-8680, Japan

り、その累積寄与率は 91.6%であった。そのため、この3つの因子により曲印象の大半が説明される。さらに、その解に基準化バリマックス回転を施した。

3.2 重回帰分析による音楽感性空間の予測

3.1 では既知の形容詞対の音楽感性空間を構成したが、ここでは未知の形容詞対が音楽感性空間のどこに位置するかを重回帰分析を用いて推定する方法について述べる。重回帰分析とは式(1)において誤差 ε の平方和が最小となる結合係数 b を求める手法である。まず、音楽感性空間座標値を目的変数、試聴実験で得られた7つの形容詞対の値を説明変数として重回帰分析を行う。次に、未知曲の形容詞対の値を説明変数に代入し、先ほど求めた結合係数との行列演算を行う。その結果求まる目的変数が音楽感性空間座標の予測値である。

$$\begin{bmatrix} x_{11} & x_{21} & \cdots & x_{m1} \\ x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (1)$$

x : 説明変数 b : 結合変数 ε : 誤差 y : 目的変数
 m : 説明変数の数 n : データの数

3.3 NN による音楽感性空間の予測

3.2 で述べた重回帰分析による方法は線形的予測である。また、重回帰分析は中間層のない NN と等価であると考えられる。そこで、重回帰分析による方法との比較対象として NN を用いた非線形回帰分析¹⁵⁾による予測値を求める。その方法は、NN の入力、教師信号にそれぞれ形容詞対の値、実際値を 0.05 から 0.95 に正規化したものを用いて学習を行う。次に、学習した NN に未知の形容詞対の値を入力信号として入力する。その結果、出力に予測値が現れる。

3.4 検索方法

3.2, 3.3 で求めた予測値に対して、全ての実際値との距離を求める。その結果、予測値からの距離が一番小さい曲を第1候補、次に小さい曲を第2候補として検索システムが出力する。

4. 実験結果

ここでは、3.で述べた手順を「ブルグミュラー25

の練習曲」の25曲に対して行った実験結果を示す。Table 1 は、試聴実験から得られた特定の1曲の形容詞対に対する予測値を他の24曲を用いて求めるという方法を25曲に対して行った結果である。なお、NN の学習は誤差逆伝播法を用い、学習係数 0.1、慣性係数 0.9 で行った。

Table 1 を見ると、重回帰分析による方法よりも NN を用いた方法の方が高い正解率を得ていることがわかる。また、NN の中間層の層数やユニット数を様々変化させても、第1候補で84%以上、第2候補までではほぼ100%の正解率を得ることができた。このことより、重回帰分析を非線形に拡張する有効性が確認できる。

Table 1 Result of retrieve

	第1候補		第2候補		
	正解数	正解率[%]	正解数	正解率[%]	
重回帰分析	17/25	68	22/25	88	
NN	5	21/25	84	25/25	100
	5-3	21/25	84	25/25	100
	30-15	21/25	84	24/25	96
	7	22/25	88	24/25	96
	15-10	22/25	88	25/25	100
	15-15	22/25	88	25/25	100
	20	23/25	92	24/25	96
	5-5	23/25	92	24/25	96
	10-10	23/25	92	25/25	100
	10	24/25	96	25/25	100
	7-7	24/25	96	25/25	100
	20-20	24/25	96	25/25	100

5. まとめ

今回、SD法および因子分析、重回帰分析、NNを用いて形容詞対による音楽データベース検索を行った。その結果、従来のシステムに比べ、より細かいニュアンスで検索を行うことができた。今後は、今回作成したシステムに自動インデクシングシステムを組み込む予定である。

【参考文献】

- [1] 栗田他: 情報処理学会論文誌, Vol.33, No.11, pp.1373-1383, Nov.1992.
- [2] 辻他: 電子情報通信学会技術研究報告書, SP96-124, pp.17-24, Mar.1997.
- [3] 岩下: 「SD法によるイメージの測定」, 川島書店, pp.43-129, 1983.
- [4] 坂本他: 平9関西連大, S13-2, Nov.1997.
- [5] 豊田: 「非線形多変量解析」, 朝倉書店, pp.145-166, 1996.