

音楽感性空間を用いた音楽データベース 検索システムの主観的評価に関する研究

池添 剛 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

我々はこれまで、感性語を用いた音楽データベース検索システムに関する研究を行ってきた^{[1][2]}。しかし、そのシステムの評価実験方法にはいくつかの問題点があった。そこで本稿では、これらの問題点を解決した方法でシステムの評価を行い、そのシステムの有効性を示す。

2. システムの評価実験における問題点とその解決法

我々が提案したシステムの評価実験における問題点には、以下の二点が挙げられる。

- (1) 試聴実験の条件が適切であったのか。つまり、用いた感性語、および、楽曲が適当であったのか。
- (2) 提案したシステムが、実際のユーザにとって有効であったのか。

そこで、それぞれの問題点を次のように解決する。

- (1) 音楽心理学の分野で行われている研究の結果を参考にして、感性語を選びなおす。また、用いる楽曲のジャンルを増やすと共に、曲数も増やす。
- (2) 提案した検索システムを実際のユーザに使用してもらう事により、検索システムの主観的評価を行う。

このようにする事により、提案したシステムの問題点を解決できると考えられる。

3. 実験結果

3.1 検索システムの再構築

ここでは、一つ目の問題点である、試聴実験の条件を改善する。そして、改善した場合においても、提案したシステムが有効であるのかを実験により確かめる。

今回行った試聴実験の条件を表1に示す。なお、今回用いた感性語は、文献[3]を参考にして選んだ。

A Study on Subjective Evaluation of Music Database Retrieving System with Music Sensibility Space.

Takeshi Ikezoe, Yoshinobu Kajikawa,

Yasuo Nomura

Department of Electronics, Faculty of Engineering,

Kansai University

3-3-35, Yamate-cho, Suita-shi, Osaka, 564-8680, Japan

また、使用する楽曲は SMF (Standard MIDI File) 形式のものであり、市販されているものやインターネットで入手可能なものを用いた。さらに、文献[4]において、「160人で行った実験結果と20人で行った実験結果とで明確な差が現れなかった」と述べられている。そこで、各曲について20人のデータを取れば、160人で行った場合と等価の結果が得られると考えられる。しかし、今回の実験では、感性語に対して7段階で評価を行っているため、20人の中央値を実験結果とすると、7段階評価にならない場合がある。そこで今回は、各曲について21人のデータが取れるように実験を行い、21人の中央値を実験結果とした。そして、この条件で実験を行った結果に因子分析を施す事により、検索空間となる音楽感性空間^{[1][2]}を作成した。その結果、第5因子までの累積寄与率は約95%であったので、求める因子数を5とした。なお、各因子の解釈を行った結果が表2である。

次に、この音楽感性空間を用いて、検索システム部を作成する。なお、以前の結果^[1]より、ニューラルネットワーク(以下、NN)を用いたシステムの方が、重回帰分析を用いたシステムよりも良い結果が得られるということがわかっている。そのため、ここではNNを用いたシステムのみを構築する。また、試聴実験を行った75曲のうち、50曲を学習曲、残

表1 試聴実験条件

被験者	関西大学交響楽団部員 50名
聴取曲	75曲(ジャンル: Classic, Jazz, Ragtime etc.)
SD法による評価指標	すべて7段階評価で、用いた感性語は、 ○明るいー暗い ○重いー軽い ○かたいーやわらかい ○安定ー不安定 ○澄んだーにごった ○滑らかー歯切れの良い ○激しいー穏やか ○厚いー薄い の8種類

表2 各因子の意味

第1因子	明度
第2因子	力量
第3因子	安定度
第4因子	質量
第5因子	感触

りの25曲を未知曲として振り分けた。そして、この条件で検索システム部を構築し、未知曲の音楽感性空間座標の予測値を求めた。その予測精度を表3に示す。なお、NNの規模において20-20とは、中間層が2層で、それぞれ20ユニットからなるネットワークを表している。表3を見ると、どの規模のNNを用いても、十分高い精度で予測できている事が確認できる。そこで、この予測値を用いて実際に検索を行ってみた。表4が第1候補での検索結果である。なお、試聴実験の結果得られた、特定の一曲の曲印象をシステムに入力した際に、その曲が検索されれば正解とした。また、不正解曲においてA→Bとは、本来、曲Aを出力するべきところで曲Bを出力した事を表している。表4を見ると、どの規模のNNでも、非常に高い正解率を得ている事がわかる。しかし、すべての規模で100%の正解率を得ているわけではない。そこで、その不正解であった曲について考察を行う。

不正解曲に対して、本来出力すべき曲とシステムが出力した曲との音楽感性空間上での距離関係を調べたところ、それぞれの曲は非常に近い距離にあった。つまり、曲印象の近いところで誤認識していた。しかし、本システムの目的は、特定の曲を検索する事ではなく、入力された印象に見合う曲を検索する事なので、この誤認識は何ら問題がないものといえる。つまり、数値的には本システムは非常に有効であるといえる。

3.2 検索システムの主観的評価実験

ここでは、二つ目の問題点を解決するために、検索システムを実際のユーザに使用してもらった事により、検索システムの主観的評価を行った結果を示す。

今回行った実験の方法は、関西大学工学部電子工学科情報工学研究室の学生29名の各人に対して検索システムで10曲検索してもらい、各曲が入力した印象に対してどれだけ満足できるものであったのかを5段階で評価してもらおうというものである。今回行った主観的評価実験の結果を表5に示す。表5を見ると、約71%の曲で満足という評価を得ていることがわかる。この事より、我々が提案したシステムは、実際のユーザにとっても有効であるといえる。しかし、若干ではあるが約10%の曲で不満足という評価があった。これは、提案したシステムにおいて、個人差を一切考慮していないためだと考えられる。そのため今後は、個人差を考慮したシステムに改良する必要があると考えられる。

4. まとめと今後の方針

本稿では、我々が提案してきた検索システムが抱えている問題点の解決を試みた。その結果、その問題点が解決でき、検索システムの有効性も確認できた。

表3 未知曲の予測精度

NNの規模	相関係数	平均二乗誤差
5	0.983	1.24×10^{-3}
5-5	0.989	8.04×10^{-4}
7	0.989	8.73×10^{-4}
7-5	0.991	6.31×10^{-4}
7-7	0.990	7.41×10^{-4}
10	0.991	7.01×10^{-4}
10-5	0.992	6.03×10^{-4}
10-10	0.994	4.90×10^{-4}
15	0.989	8.09×10^{-4}
15-5	0.991	6.92×10^{-4}
15-10	0.992	5.80×10^{-4}
15-15	0.993	5.82×10^{-4}
20	0.988	9.17×10^{-4}
20-10	0.991	6.49×10^{-4}
20-20	0.991	6.83×10^{-4}

表4 検索精度

NNの規模	正解数	正解率 [%]	不正解曲
5	75/75	100.0	
5-5	75/75	100.0	
7	75/75	100.0	
7-5	75/75	100.0	
10	75/75	100.0	
10-5	75/75	100.0	
10-10	75/75	100.0	
15	74/75	98.7	21→59
15-5	75/75	100.0	
15-10	75/75	100.0	
15-15	75/75	100.0	
20	74/75	98.7	41→36
20-10	75/75	100.0	
20-20	75/75	100.0	

表5 主観的評価実験結果

満足度	曲数 [曲]	割合 [%]
非常に満足	89/290	30.7
比較的満足	117/290	40.3
どちらでもない	57/290	19.7
比較的不満足	25/290	8.6
非常に不満足	2/290	0.7

今後は、自動インデクシングシステムに対しても実験を行い、その有効性を確認していく。また、提案したシステムを、個人差も考慮したシステムに改良していく予定である。

【参考文献】

- [1]池添他:第58回情処全大,Vol2,pp.109-110,Mar.1999.
- [2]池添他:第59回情処全大,Vol2,pp.21-22,Sep.1999.
- [3]川原他:広島大学教育学部紀要第4部,26, pp.75-85,Sep.1977.
- [4]難波他:音の評価のための心理学的測定法,pp.107-133,1998.