原著論文

楽曲に対する感性評価と音響ゆらぎ特徴の対応付け

伊藤 雄哉, 山西 良典, 加藤 昇平, 伊藤 英則

名古屋工業大学大学院工学研究科

A Method of Relating Sensibility Evaluation about Music with Acoustic Features Based on Fluctuation

Yuya ITO, Ryosuke YAMANISHI, Shohei KATO and Hidenori ITOH

Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

Abstract : This paper relates audio features based on fluctuation of music with sensibility evaluation about music as a component technology for an automatic song select system demanded on human sensibility evaluation. When people listen to music, they select a song considering both their own feelings at the time and sensibility evaluation about the song. We consider that sensibility evaluations of songs are influenced by the features based on fluctuation about both volume and pitch of the songs. Thus, we focus on features of fluctuation that contain a dynamic idea on music, and extract thirty six features of fluctuation about both volume and pitch from each songs using Fast Fourier Transform. On the other, we prepare a subjective experiment for plural songs using Semantic Differential method, and obtain the sensibility evaluation about each song. Then, we study the relationships between extracted features and sensibility evaluation about the songs with multiple discriminant analysis. As a result, high accurate discriminant hit-rates and low discriminant error are shown, therefore we suggest that audio features based on fluctuation of songs influence sensibility evaluation about the songs. Furthermore, we confirm the especial parameters related with sensibility evaluation about music while considering canonical variates which construct discriminant spaces.

Keywords : Acoustic Features, Fluctuation, Sensibility Evaluation, Semantic Differential Method, Multiple Discriminant Analysis

1. は じ め に

現代社会において、音楽は人の文化的な生活をより豊かにするメディアの一つとして注目されている。近年の携帯音楽再生機器の普及からも、音楽が人間の生活にとって重要な役割を担っていることが伺える。メディア情報技術の革新により、音楽再生機器が大容量メモリを搭載可能となったことで、ユーザは膨大な楽曲データの保有と携帯が可能となった。ユーザは大量の楽曲の中からあらかじめ楽曲に付与されたテキスト情報（曲名、アーティスト名、ジャンル名、アルバム名等）を基に検索を行う。しかし、楽曲データベースが膨大になるほどテキスト情報を用いた選曲は困難であるため、より直感的な要求に基づく選曲手法が必要と考える。

人は音楽を聴取することで様々な感性評価を抱く。この感性評価が聴取時の心理状況と一致することで心理、生理的な効用を享受する。音楽療法の分野においても、同質の原理に基づいて心理状況と同質の感性評価を持つ楽曲を聴取することが、人に癒し効果や運動パフォーマンスの向上効果を与える事例が報告されている [1-4]。このことから、人の感性的要求に応じた楽曲検索技術の実現は、楽曲データベースから直感的な選曲を可能にするとともに、人に対する効用面においても有効と考えられる。また、ユーザが未知の楽曲に対しても容易に選曲が可能となるため、映画やCM等の様々な心理効果を求められる場面においても、BGM選択の支援手法として有効であると考えられる。人の感性情報に基づいた

楽曲選曲システムを実現させるためには、楽曲データベースに存在する楽曲群に対し、人が楽曲を聴取することで受ける楽曲印象を事前に付与させておく必要がある。本研究では、人の感性情報に基づいた楽曲選曲システムの実現をさせるための要素技術として、人が楽曲を聴取することで受ける感性評価と楽曲特徴（ゆらぎ特徴）との対応付けを行う。

2. 関 連 研 究

これまでに、MIDI形式データの楽譜情報から抽出した楽曲特徴を用いて楽曲印象との相関を分析した研究が多く報告されている [5-10]。しかしMIDI情報を用いた場合、抽出される楽曲特徴はメロディ、リズム、ハーモニーなどの基礎的な楽譜情報にテンポや強弱記号などが付加されたものに限られる。また、MIDI情報では、楽曲の持つ音響情報が大量に失われており、MIDI情報に対する印象が本来の楽曲に対する印象と必ずしも一致しない。本研究では、Waveform Audio Format形式（WAV形式）データに対して音響信号解析を行うことで、ビブラートやトレモロなどのMIDI情報からでは抽出することが困難な楽曲特徴を抽出する。

音響信号を扱った研究として、平江ら [11] はMFCC等の音響信号に着目しクラシック音楽に対して感性に基づいた分類をしている。この研究では、MFCCの時間平均値や標準偏差値等の静的な特徴量を特徴化しており、時間変動を捉えた特徴量は扱われていない。本研究では、音楽において重要な要素である時間変化に着目し、周波数解析により得た音

量, および音高のゆらぎ情報を解析に用いることで, 楽曲における時間的な概念を捉える。

辻ら [12] は, MIDI 信号に含まれる旋律パターンを特徴化し, 類似曲検索を行っている。旋律の局所パターンは, 楽曲の時間的な推移に着眼したパラメータであるが, 多くの楽曲は複数の音色や旋律によって構成されており, 旋律の局所パターンでは楽曲の時間変化を十分に捉えることは出来ない。本研究では, ゆらぎ情報を扱うことで, 複数の周波数における時間変化を扱う。

楽曲におけるゆらぎ情報を扱った例として, 伊藤ら [13] は, 楽曲の音量のゆらぎ情報を用いたジャンル分類 (インストルメンタル, クラシック, ポップス, 演歌の4種) の有効性を報告している。しかしながら人は, 必ずしもテキスト情報として与えられるジャンルに基づいて選曲せず, むしろ心理状況に応じた感性評価に基づいて選曲するものであると考えられる [14]。本研究では, 人の感性評価に基づいた自動選曲システムの実現を目的とし, 楽曲におけるゆらぎ特徴と楽曲に対する人の感性評価の対応付けを行う。

3. ゆらぎ特徴

人が楽曲を聴取し知覚し得る音楽特性は様々である。中でも音量, および音高は, 音楽を表現する上で重要な要素であり [15], 音量, 及び音高の時間変化は人の感性評価に影響を与えると考えられる。また, 音楽における時間的な概念は, 認知学においてもその重要性が述べられている [16]。本研究では, 音量, 音高が時間的に変動する様を表すゆらぎ情報に基づいた楽曲特徴 (以下, ゆらぎ特徴と示す) の抽出を行う。

3.1 楽曲のゆらぎ情報

音楽を構成する3大要素としてリズム, メロディ, ハーモニーが存在する [17]。これらの要素は, 音量, 及び音高の時間構造として決定され, それぞれの差異が人間が感じる楽曲印象に影響を与えると考えられる。リズムは, 音量及び音高の時間的な規則性によって決定される。メロディは, 音量および音高の時間的な変化によって構成される。ハーモニーは, 同時に鳴る音の相対関係によって構成され, 時間とともに経過していく中で一定の機能を持つ。楽曲のゆらぎ情報は, 楽曲を時間的な側面から解釈し, 音楽を構成する要素を捉えることが可能である [18, 19]。そこで, 本研究では音量及び音高のゆらぎ情報に着目し, 音響信号解析により得られるスペクトルからゆらぎ特徴を抽出する。

3.2 特徴抽出

本研究では, 人間が楽曲を聴取し知覚し得る音響特徴を正確に抽出するため, WAV形式で保存された楽曲を用いた。WAV形式データに対し音響信号解析を施すことで, 楽曲の時間-周波数構造を可視化することが可能である。解析に用いる全ての楽曲は音質を揃えるために, CDを用いてサンプリング周波数44100Hz, 16ビットリニア量子化, ステ

レオのデータとして保存した。また本研究では, low, middle, high, ultrahighの4種類の周波数帯域を設定する。これは, 人が感じる印象に影響を与える楽曲の特徴が, 全周波数帯域における特徴だけでなく, 様々な帯域毎に存在すると考えたためである。本研究では, lowを200Hz未満, middleを200Hz以上800Hz未満, highを800Hz以上2000Hz未満, ultrahighを2000Hz以上とした。以下, 3.2.1, 3.2.2節で音量, および音高のゆらぎ情報 (ゆらぎスペクトル (例: 図1)) の算出方法 [13, 20] を示し, 3.3節でそれぞれのゆらぎスペクトルから解析に使用する特徴量の抽出について述べる。

3.2.1 音量のゆらぎ情報

時間分解能 (25ms) 毎にシフトしながら, 窓幅2048点のハニング窓 (およそ46ms) に対してフーリエ変換を施し, 各時刻における周波数のパワースペクトルを求める。各時刻において, 得られた全周波数帯域 (all) におけるパワースペクトルおよびlow, middle, high, ultrahighの4つの周波数帯域におけるそれぞれのパワースペクトルの積分値を算出する。このとき得られたパワースペクトルの積分値をそれぞれall, low, middle, high, ultrahighにおける音量とする。各時刻の音量を時系列化することで音量のゆらぎ情報が生成される。得られた音量のゆらぎ情報に対して, 再度高速フーリエ変換 (FFT) を施すことで音量のゆらぎスペクトルが生成される。このとき得られた周波数, 及びパワースペクトルは対数で表す。

3.2.2 音高のゆらぎ情報

音信号から音高を測定する手法の1つに, 零交差法がある。零交差法は, 高い精度でピッチを検出する手法であり, 楽曲等の複数の周波数で構成される音信号のピッチ推定に広く用いられる [21-23]。本研究では, 音量のゆらぎ情報の導出と同様の手順で各時刻 (窓幅: 2048点, 時間分解能: 25ms) において, 零交差法に基づいてピッチを推定し, これを音高とする。時刻 t における音高 $Pitch_t$ は下式で与えられる。

$$Pitch_t = \frac{smp \times count_0}{2}$$

ここで, smp はサンプリング周波数, $count_0$ は時刻 t の窓内における零交差の回数である。各時刻における音高を時系列化することで音高のゆらぎ情報が生成される。得られた音高のゆらぎ情報に対してFFTを施すことで音高のゆらぎスペクトルが生成される。このとき得られた周波数およびパワースペクトルは音量と同様に対数で表す。

3.3 楽曲特徴

3.2.1節, 3.2.2節にて生成された各ゆらぎスペクトルから楽曲特徴を算出する。ゆらぎスペクトルを $fluc(\omega)$ (ω : 周波数), スペクトルの個数を n とすると, 以下の(1)~(6)で算出される特徴量から楽曲特徴が構成される。表1に特徴番号と特徴量を示す。

表1 楽曲特徴

特徴番号	特徴量
[1]-[6]	allにおける音量のゆらぎスペクトルの特徴量(1)~(6)
[7]-[12]	lowにおける音量のゆらぎスペクトルの特徴量(1)~(6)
[13]-[18]	middleにおける音量のゆらぎスペクトルの特徴量(1)~(6)
[19]-[24]	highにおける音量のゆらぎスペクトルの特徴量(1)~(6)
[25]-[30]	ultrahighにおける音量のゆらぎスペクトルの特徴量(1)~(6)
[31]-[36]	音高のゆらぎスペクトルの特徴量(1)~(6)

(1) スペクトル積分値 S

$$S = \int fluc(\omega) d\omega.$$

(2) 周波数重心 f_c

$$f_c = \frac{\int \omega \cdot fluc(\omega) d\omega}{S}.$$

(3) スペクトル最大値 $fluc^{max}$

$$fluc^{max} = \max_{\omega > 0.1} (fluc(\omega)).$$

(4) スペクトル最大の周波数 f_{max}

$$f_{max} = \arg \max_{\omega > 0.1} (fluc(\omega)).$$

(5) スペクトル回帰直線の傾き α

$fluc(\omega)$ の回帰直線の傾きを線形最小2乗法によって算出する.

$$\alpha = \frac{n \sum \omega \cdot fluc(\omega) - \sum \omega \sum fluc(\omega)}{n \sum \omega^2 - \left(\sum \omega \right)^2}.$$

(6) スペクトル回帰直線の切片 β

(5) と同様に, $fluc(\omega)$ の回帰直線の切片を算出する.

$$\beta = \frac{\sum \omega^2 \sum fluc(\omega) - \sum \omega \cdot fluc(\omega) \sum \omega}{n \sum \omega^2 - \left(\sum \omega \right)^2}.$$

図1にゆらぎスペクトルの一例を示す. ここで, 特徴量(1), (3)は, 楽曲における時間変化のパワーに関する特徴量, 特徴量(2), (4)は, 楽曲における時間変化の周波数成分に関する特徴量, 特徴量(5), (6)は, 楽曲における時間変化の様々な周波数成分がどのように構成するかを表す特徴量と考えられる. 以上, 36個のゆらぎ特徴量を楽曲特徴とする. また, 全楽曲において平均値が0, 標準偏差が1となるように楽曲特徴毎に正規化を行った.

4. 楽曲印象

人が楽曲から受ける印象を抽出するため, 複数の形容詞対を尺度とするSD法に基づいた主観評価実験を行った.

4.1 呈示楽曲

本研究では, 楽曲に対して様々な嗜好を持つユーザに対しても対応することが可能な選曲システムを構築するため, 楽曲のジャンルを特定せず, 可能な限り多くのジャンル(ポップ, クラシック, ジャズ, ロック, ハウス等)の楽曲を用意し, その中から任意に選出した計150曲を使用した. また一般に, 終始一定の印象を持つ楽曲は少なく, 一楽曲内でも, 前奏・本編・間奏で印象は異なることが多い. そこで本実験では, 楽曲全体ではなく, 楽曲中の10秒程度の任意の1フレーズを主観評価の対象とした. 以後, この1フレーズに対する印象を楽曲の印象として扱う.

4.2 主観評価実験

本研究では, まず事前に先行研究[8-10, 24-26]を参考に100対の形容詞対を用意した. 次に, 用意した形容詞対に対して音楽経験者10名により楽曲の感性評価としての必要性を評価させる予備実験を実施した. 表2に, 予備実験により必要性が高いと判断された8対の形容詞対を示す. 主観評価実験では, 表2に示した8対の形容詞対についてSD法(Semantic Differential Method)に基づき7段階で評価した.

図1 ゆらぎスペクトルの一例

表2 印象形容詞対

印象番号	7段階評価						
	-3	-2	-1	0	+1	+2	+3
1	明るいー暗い						
2	なめらかなー荒い						
3	豊かなーやせた						
4	澄んだー濁った						
5	はげしいーおだやかな						
6	重いー軽い						
7	やわらかいー硬い						
8	厚いー薄い						

例えば、形容詞対「明るい-暗い」の場合、「非常に明るい」(-3),「とても明るい」(-2),「やや明るい」(-1),「どちらでもない」(0),「やや暗い」(+1),「とても暗い」(+2),「非常に暗い」(+3)の7段階で評価する。被験者は心身ともに健康な20代前半の男女18名(男15名,女3名)とした。主観評価実験には、WEBブラウザ上に用意した楽曲の主観評価用インターフェースを用いて実施した。このインターフェースでは、windows media playerを用いて、被験者が自由に楽曲の再生・巻き戻しが出来るよう準備した。選曲システムの応用場面を想定した時、日常的な場面における被験者の感性評価が必要であると考えたため、被験者は各自使用している計算機を用いて主観評価実験を行った。ただし、被験者各自が使用する計算機に搭載された音楽再生に関する周辺機器(サウンドカードおよびスピーカ)は全て統一した。また、楽曲聴取による被験者の疲労を軽減するため、150曲の楽曲を50曲から成る3つの楽曲群に分け、1日に1楽曲群ずつ聴取させた。直前に聴取した楽曲の影響を避けるため、楽曲群は被験者ごとにランダムな順序で呈示した。

4.3 楽曲印象値

前節で得られた評価値から1200個(150楽曲×8形容詞対)の楽曲印象値を求める。この時、各楽曲の印象値には、全被験者による評価値の最頻値を扱う。各楽曲の印象値に全被験者の最頻値を扱うことで、各被験者の嗜好によらず、一般的な印象値の計測が可能であると考えた。全被験者の7段階評価における分散が2以上である評価値については、楽曲印象値として不適切と判断し、分析対象から取り除いた。各印象形容詞対において、楽曲印象値として適切と判断された楽曲数を表3に示す。

5. 判別分析

分析対象とした楽曲を表4に示す。印象番号8(厚い-薄い)における「+3」(非常に薄い)と評価された楽曲が存在しない。また一般的に、楽曲数が3個未満となるクラスについては、等分散性が満たされず判別に適さない[27]。そこで、楽曲

表3 適切であると判断された楽曲の数

印象番号	1	2	3	4	5	6	7	8
楽曲数	131	122	82	125	133	91	108	91

表4 印象項目ごとの評価値の内訳

印象番号	7段階評価						
	-3	-2	-1	0	+1	+2	+3
1	7	39	29	3	32	17	4
2	5	26	39	4	21	22	5
3	2	22	29	11	13	4	1
4	9	26	43	3	27	12	5
5	9	25	40	3	22	21	13
6	5	6	20	5	28	22	5
7	9	18	29	5	29	14	4
8	4	14	41	14	15	3	0

が3曲未満となるクラスについては分析の対象から除外し、印象3, 8はそれぞれ5クラス, 6クラスによる判別を行った。

5.1 特徴選択

表1に示す36次元の楽曲特徴から、楽曲印象の項目毎に判別に寄与する特徴を選択した。特徴選択には変数増減法を用い、評価関数にはFisherの線形判別関数を採用した。表5に、各印象項目において選択された楽曲特徴 x_i (i : 特徴番号(表1参照))を示す。

5.2 正準判別分析による分析結果

表3に示した楽曲を用いて、前節で選択された楽曲特徴を説明変数として、各楽曲印象に対して正準判別分析を行った。

5.2.1 判別空間についての考察

各楽曲印象において、全学習データのうち印象実測値(観測値)と印象予測値(理論値)が一致した割合を示した判別正答率と、印象実測値(観測値)と印象予測値(理論値)との誤差を示した平均判別誤差をそれぞれ表6に示す。全楽曲印象において、無作為に判別を行った場合の理論的な正答率に対して2倍以上高い判別正答率が得られた。中でも、印象番号3, 5, 8の判別正答率は、比較的良好な結果が得られ、適切な判別が行われたと考えられる。

また、印象番号1, 2, 3, 5, 8の判別誤差は1未満であり、大半の楽曲は誤判別されたとしても、隣り合う評価値(± 1)に判別されたことが示唆される。

5.2.2 正準変量についての考察

各楽曲印象に対して正準判別分析を行った結果、得られた各正準変量の寄与率を表7に示す。各楽曲印象において寄与率が上位2つである第1および第2正準変量を用いて生成した判別空間における楽曲散布図を図2に示す。

表5 選択された楽曲特徴

印象番号	楽曲特徴
1	x_7, x_{10}, x_{21}
2	x_{17}, x_{27}, x_{31}
3	x_2, x_5, x_9, x_{20}
4	x_6, x_7, x_{16}
5	$x_2, x_3, x_6, x_7, x_{14}, x_{15}$
6	x_{31}, x_{32}
7	x_2, x_{15}, x_{20}, x_3
8	$x_8, x_{10}, x_{11}, x_{17}, x_{32}$

表6 学習データに対する判別正答率と判別誤差

印象番号	判別正答率	判別誤差
1	0.493	0.95
2	0.475	0.95
3	0.569	0.68
4	0.44	1.10
5	0.556	0.74
6	0.473	1.02
7	0.472	1.01
8	0.593	0.68

表7 寄与率

印象番号	LD1	LD2	LD3	LD4	LD5	LD6
1	0.61	0.22	0.17	-	-	-
2	0.61	0.3	0.09	-	-	-
3	0.48	0.32	0.19	0.01	-	-
4	0.56	0.31	0.13	-	-	-
5	0.52	0.24	0.1	0.09	0.03	0.01
6	0.67	0.33	-	-	-	-
7	0.56	0.21	0.16	0.07	-	-
8	0.37	0.29	0.21	0.10	0.03	-

印象番号1, 2, 5では, 第1正準変量の0を基点として, 評価値の正負が判別されていることが確認できる. 特に, 印象番号5については, 各クラスの楽曲が第1正準変量に沿って連続的に分布していることが確認できる.

一方, 印象番号3, 8については, 判別正答率および判別誤差の結果が比較的良好であったにも関わらず, 各クラスの楽曲が混在している. これは, 印象番号3, 8の判別における第1, 第2正準変量の寄与率が比較的低く(表7参照), 第1, 第2正準変量だけでは楽曲の印象を説明することが困難であるためと考えられる.

6. 楽曲特徴と楽曲印象の対応付け

図2において, 第1および第2正準変量によって楽曲を評価値毎にうまく判別できた楽曲印象1, 2, 5について考察を行う. 表8に, 各楽曲印象の第1および第2正準変量における楽曲特徴と判別係数をそれぞれ示す. ここで, 各楽曲特徴の判別係数は, 正準変量に対するそれぞれの楽曲特徴の影響度として扱う.

・楽曲印象1「明るいー暗い」について

図2より, 「明るい」(-3, -2, -1) 楽曲が第1正準変量の負に多く, 「暗い」(+1, +2, +3) 楽曲が正に多いことから, 第1正準変量がこの楽曲印象の判別に強く寄与することが確認できる.

表8の楽曲印象1を見ると, 第1正準変量において, x_{21} (high帯域における音量のゆらぎスペクトルの最大値) の第1判別係数が負に大きい. high帯域のゆらぎ成分は, 女性ボーカリストの楽曲や比較的高い楽曲などに多く含まれる. また, x_7 (lowにおける音量のゆらぎスペクトルの積分値), x_{10} (lowにおける音量のゆらぎスペクトルの最大値の周波数) の第1判別係数が負であることが確認できる. low帯域には, バスドラムやベース等のロック調の楽曲に多く見られる重低音の影響が強く現れる. これらのことから, highにおける音量のゆらぎが激しく, かつ, ドラムやベースなどの重低音が激しく変化するロック調の楽曲が「明るい」と判別されたことが示唆される.

・楽曲印象2「なめらかなー荒い」について

図2より, 「なめらかな」(-3, -2, -1) 楽曲が第1正準変量の負に多く, 「荒い」(+1, +2, +3) 楽曲が正に多いことから, 第1正準変量がこの楽曲印象の判別に強く寄与することが確認できる.

表8の楽曲印象2を見ると第1正準変量において, x_{17} (middleにおける音量のゆらぎスペクトルの回帰直線の傾き) と x_{31} (音高のゆらぎスペクトルの積分値) が負に大きい. ゆらぎスペクトルの回帰直線の傾きは, 値の

図2 各印象ごとの第1, 第2正準変量による楽曲の散布図

大きさに伴い高周波（短周期）のパワースペクトル値が大きくなるという特徴を持つ。音量のゆらぎスペクトルの高周波は、ドラム等の周期の短いリズム楽器と密接に関係している。このことから、音高のゆらぎが激しく、かつ、middleにおけるドラム音などのリズム楽器によるゆらぎが激しくなるほど「荒い」楽曲であることが示唆される。

• 楽曲印象5「はげしいーおだやかな」について

図2より、「はげしい」(-3, -2, -1) 楽曲が第1正準変量の負, 「おだやかな」(+1, +2, +3) 楽曲が正に多い。また、各評価値が第1正準変量に沿って連続的に分布しており、第1正準変量がこの楽曲印象の評価値に強く影響を与えていると考えられる。

表8の楽曲印象5を見ると、第1判別係数において x_3 (allにおける音量のゆらぎスペクトルの最大値), x_7 (lowにおける音量のゆらぎスペクトルの積分値) が正に大きく、 x_{15} (middleにおける音量のゆらぎスペクトルの最大値) が負に大きいことがわかる。これは、全周波数帯域の音量の最大値が大きく、low帯域の音量が大きいほど、より「はげしい」楽曲である傾向があり、一方で、middle帯域の音量の最大値が大きいほど、より「おだやかな」楽曲である傾向があることを示す。これらのことから、この「はげしさ」に関する楽曲印象は周波数帯域別での音量のゆらぎの大きさに関する特徴に大きく影響を受けているということが示唆される。

表8 第1, 第2判別係数

楽曲印象	特徴	第1判別係数	第2判別係数
1	x_7	-0.412	0.28
	x_{10}	-0.46	0.69
	x_{21}	-1.05	-0.48
2	x_{17}	1.12	-0.41
	x_{27}	-0.01	0.86
	x_{31}	0.52	0.61
3	x_2	1.02	-0.24
	x_5	-0.51	1.27
	x_9	-0.79	-0.88
	x_{20}	-0.17	-1.18
4	x_6	0.42	-1.78
	x_7	0.93	1.55
	x_{16}	-0.4	0.11
5	x_2	0.77	-0.63
	x_3	-1.89	1.04
	x_6	-0.06	-0.69
	x_7	-0.80	-0.42
	x_{14}	-0.5	0.16
6	x_{15}	1.33	-0.59
	x_{31}	-0.79	0.83
7	x_{32}	0.94	0.79
	x_2	1.01	0.55
8	x_{15}	0.72	-1.12
	x_{20}	-1.1	-1.29
	x_{33}	-1.36	0.46
	x_8	-0.19	0.38
8	x_{10}	0.88	-0.71
	x_{11}	0.67	0.24
	x_{17}	-0.71	-0.37
	x_{32}	-0.48	-0.78

図2の結果から、楽曲印象3, 4, 6, 7, 8については、それぞれの判別正答率（表6）と比較して楽曲のクラスが混在している。このことから、第1, 第2正準変量だけでは印象のクラスを説明することが困難であることが示唆される。また、楽曲印象6については、特徴選択により選択された楽曲特徴数が2つと少なかった。このことから、楽曲印象6については、評価のクラス判別に寄与する特徴が本研究で採用した36次元の楽曲特徴内に存在しなかったことが考えられる。

7. 楽曲特徴の有効性

表5で示された楽曲特徴を用いて、各楽曲印象に対しLeave-one-out交差検定を行った。Leave-one-out交差検定法を用いることで、オープンテストに準じた頑健な最適化モデルを実現できる[28]。表9に、交差検定により得られた判別正答率と判別誤差を示す。同表の結果から、未知データに対しても判別正答率、判別誤差共に、良好な結果が得られており、選択された楽曲特徴が、印象値の判別に有効であることが確認された。

8. ま と め

本研究では、楽曲のゆらぎ情報に着目した楽曲特徴を抽出し、人が感じる楽曲印象との関係性を正準判別分析により解析した。全ての楽曲印象において、良好な判別正答率、及び判別誤差が得られ、ゆらぎ情報が楽曲印象に影響を与えることが示唆された。また、「明るいー暗い」、「なめらかなー荒い」、「はげしいーおだやかな」の印象値の判別に影響を与える楽曲特徴を確認した。本研究の今後の課題として次の3点が考えられる。

(1) 楽曲数の増加

表4の楽曲印象3「豊かなーやせた」のように、評価のクラスに属する楽曲数に偏りが見られた。今後はより多くの楽曲を用いて分析することで楽曲数の偏りを解決し、全ての楽曲印象に対して適切なクラス数で判別が行えるように改良する。

(2) 高次元の正準変量による判別空間の解釈

今回は、第1, 第2正準変量から構成される判別空間に基づいて考察を行ったが、印象によっては第3以上の

表9 Leave-one-out交差検定法による判別正答率と判別誤差

印象番号	判別正答率	判別誤差
1	0.47	0.98
2	0.46	0.98
3	0.49	0.79
4	0.42	1.13
5	0.53	0.79
6	0.44	1.02
7	0.43	1.08
8	0.56	0.75

正準変量が判別に大きく寄与していることもあり、本来の判別空間を捉えきれたとは言えない。今後は得られた正準変量にバリマックス回転 [29] を適用するなどしてこの問題を解決する。

(3) 自動選曲システムへの応用

本研究では、楽曲印象の判別におけるゆらぎ特徴の有効性を判別正答率と判別誤差を用いて評価した。しかしながら、自動選曲システムに対する実用性がこれらの結果と必ずしも一致しない。人間の感性評価では「明るい (-2)」楽曲を、自動選曲システムが「やや明るい (-1)」や「非常に明るい (-3)」と判別し選曲したとしても許容される可能性がある。そこで、システムを実装後、ユーザビリティを評価する予定である。

今後は、上記3点の課題に取り組むとともに、人間が満足する自動選曲システムの実装を目指す。

参 考 文 献

- [1] 中村貴展, 加藤修一, 佐川泰広, 山本佐代子, 武田昌一: Experiments on EEG Based on the Principle of Homogeneity in Music Therapy, 日本音響学会研究発表会講演論文集, Vol. 1, pp.575-576 (2003).
- [2] 伊藤孝子, 岩永誠: 音楽に対する同質感がリラクゼーションに及ぼす影響, 広島大学総合科学部紀要理系編, Vol. 25, pp.141-150 (1999).
- [3] 古賀弘之: 気分と同質・異質の音楽聴取に伴う気分変化の分析, *Bulletin of the Graduate School of Education, Hiroshima University. Part.1, Learning and curriculum development*, Vol. 54, pp.91-97 (2005).
- [4] Altshuler, I. A.: A Psychiatrist's Experiences With Music as a Therapeutic Agent, *Music and Medicine*, Vol. 1, pp.270-271 (1948).
- [5] 熊本忠彦, 太田公子: 印象に基づく楽曲検索: 楽曲印象値の自動付与, 情報処理学会研究報告, Vol. 41, pp.89-96 (2002).
- [6] 熊本忠彦, 太田公子: 印象に基づく楽曲検索: システムの実装と評価, 情報処理学会研究報告, Vol. 63, pp.37-42 (2002).
- [7] 普天間淳, 工藤博章, 松本哲也, 竹内義則, 大西昇: 人間の情感と楽曲パラメータとの対応付け, 電子情報通信学会技術研究報告, Vol. 100 (713), pp.51-56 (2001).
- [8] 佐藤聡, 菊地幸平, 北上始: 音楽データを対象としたイメージ検索のための感情値の自動生成, 情報処理学会研究報告, Vol. DBS-118, pp.57-64 (1999).
- [9] 佐藤聡, 菊地幸平, 北上始: 感情に基づく音楽作品検索システムの実現に向けての検討, 情報処理学会研究報告, Vol. 2001 (16), pp.51-56 (2001).
- [10] 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グループの分割, 情報処理学会論文誌, Vol. 40 (4), pp.1901-1909 (1999).
- [11] 平江遼, 西隆司: 感性に基づくクラシック音楽の分類, 日本音響学会誌, Vol. 64 (10), pp.607-615 (2008).
- [12] 辻康博, 星守, 大森匡: 曲の局所パターン特徴量を用いた類似曲検索・感性語による検索, 信学技報, Vol. 16, pp.17-24 (1997).
- [13] 伊藤直樹, ヨハンライフ, 徳弘一路, 西口磯春, 高澤嘉光: 楽曲のゆらぎ情報を用いた自動音楽ジャンル分類システム, 電子情報通信学会技術研究報告, EA, Vol.102, pp.1-6 (2002).
- [14] Donald, M. and Joseph, P.: *Music Therapy In Principle And Practice*, Charles C Thomas Pub Ltd (2005).
- [15] Lu, L., Liu, D. and Zhang, H.-J.: Automatic Mood Detection and Tracking of Music Audio Signals, *IEEE Trance on Audio, SPEECH, And Lang Processing*, Vol. 14 (2006).
- [16] Meyer, L. B.: *Emotion and Meaning in music*, University of Chicago Press (1956).
- [17] 三浦信一郎: 西洋音楽思想の近代-西洋近代音楽思想の研究, 三元社 (2005).
- [18] 山田真司, 井村和孝, 新井裕子 and 小田満理子, 西村英樹: 音楽演奏者の時間的制御能力について, 情報処理学会研究報告 [音楽情報科学], Vol.95 (46), pp.21-28 (1995).
- [19] 佐々木實, 石川智子, 山田真司: 等時価からなる楽譜のピアノ演奏における時間的ゆらぎ: 芸術的対メトロノームの演奏の分析, 日本音響学会誌, Vol. 45 (10), pp.743-747 (1989).
- [20] 武者利光: ゆらぎの科学, Vol. 1-9, 森北出版 (1991-1999).
- [21] 三井実, 石川智治, 小林幸夫, 宮原誠: 音のゆらぎ情報と時間伸び縮み歪について, 電子情報通信学会技術研究報告, EA, Vol. 99 (260), pp.9-16 (1999).
- [22] 北原鉄朗, 後藤真孝, 奥乃博: 音高による音色変化に着目した楽器音の音源同定: FO 依存多次元正規分布に基づく識別手法, 情報処理学会論文誌, Vol. 44 (10), pp.2448-2458 (2003).
- [23] 田中伸和, 野口健太郎, 田所嘉昭: サンプリング制御に基づくノッチフィルタを用いた多周波推定法の検討, 電子情報通信学会技術研究報告, Vol. 103 (719), pp.69-74 (2004).
- [24] 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 (12), pp.3201-3212 (2001).
- [25] 熊本忠彦, 太田公子: 印象に基づく検索のための印象語選定法の提案, 情報処理学会論文誌, Vol.44 (7), pp.1808-1811 (2003).
- [26] 安倍幸治, 小澤賢司, 鈴木陽一, 曾根敏夫: 音色表現語, 感情表現語及び音情報関連語による環境音評価, 日本音響学会誌, Vol. 54 (5), pp.343-350 (1998).

- [27] 奥野忠一, 久米均, 芳賀敏郎, 吉澤正: 多変量解析法, 日科技連(東京)(1981).
- [28] 竹内翔大, ツインタアレクトピラス, 木田学, 早川直樹, 川波弘道, 猿渡洋, 鹿野清宏: Leave-one-out 交差検定法に基づく質問応答データベースの最適化, 電子情報通信学会技術研究報告, Vol. 107 (116), pp.31-36 (2007).
- [29] 谷口るり子: バリマックス回転の正準判別分析への応用, 計算機統計学, Vol. 10 (1), pp.3-10 (1998).

伊藤 雄哉 (非会員)

2009 年名古屋工業大学知能情報システム学科卒業。現在, 同大学院工学研究科博士前期課程情報工学専攻在籍中。音響工学, 感性工学などに興味を持つ。感性工学では特に音楽情報に関する研究に従事。日本音

響学会会員。

山西 良典 (学生会員)

2007 年名古屋工業大学知能情報システム学科卒業。2009 年同大学院工学研究科博士前期課程情報工学専攻卒業。現在, 同大学院工学研究科博士後期課程情報工学専攻在籍中。音楽学, 生理学, 人間工学などに興味を

持つ。感性工学では特に音楽情報と感性および生理学に関する研究に従事。情報処理学会会員。

加藤 昇平 (正会員)

1998 年同大学院工学研究科博士後期課程電気情報工学専攻修了。同年豊田工業高等専門学校助手, 1999 年同講師, 2002 年名古屋工業大学講師, 2003 年同助教授。現在同大学院情報工学専攻所属, 准教授。博士(工学)。知能・感性ロボティクス, 知識推論・計算知能, ヒューマンインタラクションなどに興味を持つ。2006 年日本感性工学会技術賞。2010 年日本知能情報ファジィ学会論文賞。情報処理学会, 人工知能学会, 電子情報通信学会, 日本ロボット学会, 日本知能情報ファジィ学会, 電気学会, IEEE 各会員。

伊藤 英則 (正会員)

1974 年名古屋大学大学院工学研究科博士課程電気電子専攻満了。工学博士号取得。1974 年日本電信電話公社横須賀研究所勤務。1985 年(財)新世代コンピュータ技術開発機構出向。

1989 年名古屋工業大学教授。2010 年同大学名誉教授。この間, 数理言語理論, 計算機ネットワーク通信, OS, 人工知能などの研究開発に従事。2006 年日本感性工学会技術賞。2010 年日本知能情報ファジィ学会論文賞。