

個人差を考慮した感性評価値のファジィ化による楽曲の印象推定

新田 秋平[†] 姜 東植[‡]琉球大学大学院 理工学研究科[†] 琉球大学 工学部情報工学科[‡]

1. はじめに

近年、コンピュータの性能の向上、及びインターネットの普及に伴い、人々が手に取ることの出来るメディア情報量は増大している。音楽情報も例外ではなく、今やユーザーは自身の持つコンピュータや携帯端末に膨大な音楽データベースを持つことが可能となっている。ユーザーは膨大なデータの中から、主にアーティスト名やジャンル名などの情報によって選曲を行っているものの、これらの情報によってユーザーが思い描く楽曲を選曲することは、アーティストやジャンルが持つ傾向を知らなければ困難である。

そのため、人が持つ感性的なイメージに基づいて直感的に選曲するシステムは利便性において有用であると考えられる。また、ユーザーの好みに沿う楽曲を推薦することはマーケティングにおいても注目されており、楽曲の販売サイトでは顧客の購入履歴を基に楽曲を関連づける事で似た好みを持つであろう顧客に効果的な楽曲を推薦するよう取り組んでいる。しかし、購入履歴に基づいた楽曲推薦は顧客のその時々の直感的な欲求に必ずしも合致しているとは限らない。顧客が直感的に求めている楽曲を推薦する事が可能になればより有効なマーケティングが可能となると考える。

更に、音楽は医療の場でも用いられており、患者のその時々の精神状態に合致した音楽を聴かせることで、音楽が患者にとって有効に作用するという同質の原理⁽¹⁾が発見されており、神経症の治療やリハビリテーション等に音楽が重要な役割を果たしている。患者の気分合うような選曲をすることは、より効率的な治療に繋がるであろうと考える。以上の背景から、本研究では人が楽曲を聴いた時に感じる印象を自動的推定し、直感的な選曲を可能とするシステムの開発を目指す。

感性に基づいた楽曲検索システムへの取り組みとして、池添ら⁽²⁾は楽曲の MIDI 形式データから特徴を抽出し SD 法と因子分析によって作り出した感性空間に楽曲を配置する事で感性語による検索システムを開発した。しかし、MIDI 形式のデータから得られる情報は楽譜情報にテンポ等の情報を付加したものに限られるため、実際に人の耳に届く音響信号の持つ波形の情報も必要であると考えられる。音響信号の波形を扱った研究として、伊藤雄哉ら⁽³⁾は楽曲の WAVE データから音の時間構造を考慮した音響情報であるゆらぎ情報を抽出して楽曲印象の推定を行った。しかしながら楽曲を聴いた時に感じる印象は、傾向はあれ聴き手によって個人差が存在すると考えられる。そのため、推定された印象と個人が感じ取る印象との間に誤差が発生する。そこで本研究では、個人差がある事を考慮した楽曲印象の推定手法を提案する。聴き手によって楽曲から感じる印象に個人差があることを踏まえ、推定値に対してファジィ化を行う。

2. 特徴量

楽曲の音響信号の波形から、その楽曲を印象づける音楽的な要素を特徴量として抽出する。本稿では、音楽の時間構造を表現する特徴量としてゆらぎ情報⁽³⁾を導入する。

ゆらぎ情報は、楽曲の WAVE データに対して、窓幅 2048 点のハニング窓を 24ms ずつシフトしながらフレームを取り出し、各フレームに対して FFT を施すことで 24ms 毎の周波数パワースペクトルを計算し、各時間におけるスペクトルの特徴を算出していくことで得られるスペクトル特徴の時系列データである。

こうして得られた音量・音高・リズムのゆらぎ情報に対して、更に FFT を施してスペクトルを導出する。得られたスペクトルから、積分値・重心・スペクトル最大値・スペクトル最大の周波数・線形最小二乗法による回帰直線の傾きと切片の 6 つの特徴量を各ゆらぎ情報から抽出する。

Estimation of Musical Impression by Fuzzification Musical Impression Value in Consideration of Individual Difference

[†] Graduate School of Engineering and Science, University of the Ryukyus

[‡] Faculty of Engineering, University of the Ryukyus

	-3	-2	-1	0	1	2	3	
静かな	にぎやかな
軽量感のある	重量感のある
スローテンポ	アップテンポ
暗い	明るい
日常的な	ドラマティックな
やわらかい	かたい
厚い	薄い
しんみりした	うきうきした

Fig.1 使用する形容詞対

3. 楽曲印象のファジィ化

楽曲印象はSD法に基づいた7段階評価によって数値化する。本研究で用いた8つの形容詞対を図1に示す。楽曲に対する印象は聴き手によって個人差がある事を踏まえ、ファジィ化された印象値であるファジィ評価値を推定する。

(3-1) ファジィ評価値

ファジィ評価値はメンバシップ関数の最大値が1となる正規ファジィ数で、中心値 c と曖昧度 p を用いて

$$\tilde{X} = \langle c, p \rangle \dots \dots \dots (1)$$

と表す。このファジィ数の α レベル集合は次のような閉区間となる。

$$[\tilde{X}]_{\alpha} = [c - (1 - \alpha)p, c + (1 - \alpha)p] \dots \dots (2)$$

また、ファジィ数同士の加算とファジィ数の実数倍は次のように計算される。

$$\langle c_1, p_1 \rangle + \langle c_2, p_2 \rangle = \langle c_1 + c_2, p_1 + p_2 \rangle \dots \dots (3)$$

$$k\tilde{X} = \langle kc, k|p| \rangle \dots \dots \dots (4)$$

(3-2) ファジィ評価値の推定

楽曲の数を M 、評価に用いる形容詞対の数を N とする。この時、 K 人の評価者の楽曲評価データから、評価の平均値 y_{mn} と標準偏差 s_{mn} を用いて教師データとなるファジィデータを

$$\tilde{Y}_{mn} = \langle y_{mn}, s_{mn} \rangle \dots \dots \dots (5)$$

$$(m = 1, \dots, M), (n = 1, \dots, N)$$

とする。以下では1つの形容詞対に注目して添字の n を省略する。楽曲 m のファジィ評価値 \tilde{F}_m を t 個の特徴量とファジィ数の線形結合とする時、その中心値と曖昧度は、

$$c_m = a_0 + a_1x_{m1} + \dots + a_tx_{mt} \dots \dots \dots (6)$$

$$p_m = e_0 + e_1|x_{m1}| + \dots + e_t|x_{mt}| \dots \dots \dots (7)$$

となる。これを利用して、次の線形計画問題を解くことにより、ファジィ評価値を推定する。

$$\begin{aligned} & \text{Minimize } \sum_{m=1}^M p_m \\ & \text{subject to } \begin{cases} c_m - (1 - \alpha_1)p_m \leq \max\{-3, y_m - (1 - \alpha_2)s_m\} \\ c_m + (1 - \alpha_1)p_m \geq \min\{y_m + (1 - \alpha_2)s_m, 3\} \\ e_i \geq 0, i = 0, 1, 2, \dots, t \end{cases} \\ & (m = 1, 2, \dots, M) \end{aligned}$$

この線形計画問題により、 \tilde{F}_m の α_1 レベル集合が \tilde{Y}_m の α_2 レベル集合を内包するように a_0, \dots, a_t と e_0, \dots, e_t に制約条件を課した上で曖昧度を最小になるよう a_0, \dots, a_t と e_0, \dots, e_t の値を設定する事でファジィ評価値を推定する。また、ファジィ評価値は評価値の下界と上界までは包含する必要はないため、-3及び3を下限と上限として制約条件を緩和する。

4. 実験

20代の学生11名に対して楽曲の主観評価アンケートを行った。得られた評価データから1人のデータをテスト用データとして抜き出し、残り10人のデータを元にファジィ評価値を推定した。テスト用データの評価値の、推定されたファジィ評価値に対してのメンバシップ値を算出する事で、個人差を包含しているか評価する。これをテスト用データとして抜き出すデータを入れ替えながら行い、印象毎のメンバシップ値の平均値を取ったものを表1に示す。また、表2にテスト用データの値とファジィ評価値の中心地との平均推定誤差を示す。

TABLE. 1 メンバシップ値の平均値

Impression No	1	2	3	4	5	6	7	8
$\mu(x)$	0.466	0.536	0.6	0.579	0.682	0.602	0.348	0.6

TABLE. 2 平均推定誤差

Impression No	1	2	3	4	5	6	7	8
Average Margin of error	1.437	1.087	0.861	1.031	0.769	0.954	1.222	1.025

文 献

- (1) I.M. Altshuler, "A Psychiatrist's Experience with Music as a Therapeutic Agent", Nordic Journal of Music Therapy, Vol.10, Issue 1, pp.69-76, (2001)
- (2) T.Ikezo, Y.Kajikawa & Y.Nomura: "Music Database Retrieval System with Sensitivity Words Using Music Sensitivity Space", The Journal of the Information Processing Society of Japan, Vol.42, No.12, pp.3201-3212, (2001) (in Japanese)
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, No.12, pp.3201-3212, (2001)
- (3) Y.Ito, Y.Yamanishi, S.Kato: "Estimation of music impressions using musical fluctuation features", The Journal of the Acoustical Society of Japan 68(1), pp.11-18, (2011-12-25)
伊藤雄哉・山西良典・加藤昇平: "音楽ゆらぎ特徴を用いた楽曲印象の推定", 日本音響学会誌, Vol. 68, No.1, pp.11-18, (2012) (in Japanese)