

印象に基づく楽曲検索：自由かつ自然な印象表現の意味理解

熊本 忠彦[†] 太田 公子[†]

[†] 独立行政法人 通信総合研究所 けいはんな情報通信融合研究センター

〒 619-0289 京都府相楽郡精華町光台 2-2-2

E-mail: †{kuma,kimiko}@crl.go.jp

あらまし 大量にある楽曲データの中から自分の判断基準（例えば嗜好、感性、心的状態など）に合致するものを見つけ出すための手段として、印象に基づく楽曲検索方式が研究されている。この検索方式では、通常、システムが提示する少数の印象語に対する何らかの評価（例えば7段階評価など）という形で印象の入力が行われるため、入力可能な印象語の数が増えると、入力に要するコストも増大してしまう。その一方で、入力可能な印象語の数が限られていると、見つけたい楽曲の印象を正確に表現できない可能性が生じてくる。ある印象語を入力できるようにするためには、その印象語を検索条件に変換するためのルール（印象語解釈ルール）が必要とされるが、そのようなルールの作成は、通常、楽曲の聴取を伴うため、高コストであり、多くの印象語を対象とすることは実質的に困難であった。そこで、本稿では入力できるようにしたい印象語を既存の印象尺度（印象語の対からなる評価尺度）で評価することにより、印象語解釈ルールを作成するという方式を提案する。この方式は、印象語同士を楽曲を媒介としないで対応付けようという試みであり、楽曲を媒介として対応付ける従来の方式に比べ、誤差は大きくなるものと予想されるが、その反面、より多くの印象語を入力可能にすることができ、自由かつ自然な印象表現（印象語を含む文表現）を受理できる自然言語インタフェースの実現に資するものと考えられる。

キーワード 自然言語理解、コンテンツ検索、感性検索、自然言語インタフェース

Generating Rules to Understand Spontaneous User Input to an Impression-Based Music-Retrieval System

Tadahiko KUMAMOTO[†] and Kimiko OHTA[†]

[†] Keihanna Human Info-Communication Research Center, Communications Research Laboratory

2-2-2, Hikari-dai, Seika-cho, Kyoto, 619-0289 Japan

E-mail: †{kuma,kimiko}@crl.go.jp

Abstract Impression-based music retrieval is the best way to find pieces of music that suit the preferences, senses, or mental states of users. In a music-retrieval system of this kind, the system presents a relatively short set of words that denote impressions; the user selects words from this set for input to the system. With this approach, increases in the numbers of words presented lead to increase the cost to the user in terms of the time required to enter impressions. On the other hand, a set of words that is too limited will often leave users unable to accurately represent their target impressions. In order for the system to understand impression words, the system must have rules for transforming the words into a suitable form, i.e., into retrieval conditions. However, generating these rules is costly, since this process requires that many subjects listen to and evaluate many pieces of music against the target words. Increases in the numbers of words presented lead to increases in the cost of this process, such that the process is impracticable with large numbers of words. In this paper, we propose a new approach to the rule-generation process, in which the rules for use in deriving retrieval conditions from impression words are obtained by evaluating target impression words against impression scales that we have devised.

Key words natural language understanding, impression-based music retrieval, natural language interface

1. ま え が き

携帯電話やインターネットなどの情報通信技術の発展によりコンピュータ音楽が身近なものとなっている。音楽情報にあまり詳しくない人でも、着メロ（携帯電話等の着信メロディ）やBGM（例えばインターネット上のホームページやグリーティングカードサービスで利用されている）などをたくさんの楽曲データの中から検索し、選択する機会が増えている。音楽情報に詳しい人ならば、好きな曲や好きな作曲家、好きな演奏家など具体的な情報を用いて、容易に検索できるかもしれないが、そうでない人にとっては、非常に困難で非効率的な作業といえる。ハミング（鼻歌）や歌声などの音響情報や歌詞などのテキスト情報といった内容情報を用いた検索手段[1]～[4]も研究されているが、漠然とした要求しかなく、具体的な情報を入力できないような場合には、不向きと言える。ある特定の楽曲を探し出すというよりもユーザが持つ何らかの判断基準（例えば嗜好や感性、心的状態など）に合致するものを見つけないかという場合には、印象に基づく楽曲検索[5]～[12]が有効であり、他の検索手段との統合によって「モーツァルトの悲しくて、ゆったりとした曲」といったより高度な検索も可能となる。

従来の印象に基づく楽曲検索において、印象の入力は、(1) 複数個の印象語の中から一つあるいは複数個を選択する[5], [6], (2) 複数個の印象語に対する多段階評価（例えば5段階評価など）[7], [8], (3) 複数個の印象尺度（印象語の対からなる評価尺度）に対する多段階評価[9]～[12]、といった形式で行われる。そのため、入力可能な印象語の数が増えれば増えるほど、印象入力に要する時間と労力が増えてしまい、ユーザフレンドリーとは言えなくなる。その一方で、入力可能な印象語の数が限られていると、ユーザ自身の言葉で入力することができず、見つけたい楽曲の印象を正確に入力できない可能性が生じてくる。従って、ユーザフレンドリーな検索を実現するためには、ユーザの自由かつ自然な入力を可能にする自然言語インタフェースの実現が課題となる。なお、画像検索の分野の先行研究において、自然言語インタフェースを有した、印象に基づく画像検索システム[13]が提案されているが、入力可能な印象語の数は20対の印象尺度に対応した40語に過ぎず、自然言語インタフェースの最大の利点といえる自由かつ自然な入力は実現されていない。

ある印象語を入力できるようにするためには、その印象語を検索条件に変換するためのルール（本稿では印象語解釈ルールと呼ぶ）が必要であり、そのようなルールの作成は、通常、楽曲の聴取、印象語や印象尺度による楽曲の評価、評価結果に対する多変量解析（主に因子分析）、といった手順で行われる。しかしながら、印象語の数が増えれば、より多くの楽曲を聴取する必要性が生じてくるし、印象語の評価に要する時間も増えてしまう。また、1回の楽曲聴取において、評価すべき印象語の数が多いと、聴取した楽曲の印象がぼやけてくる恐れもある。

そこで、本稿では、入力できるようにしたい印象語を我々が提案している印象尺度（10対20語の印象語で構成されている）[14]で評価することにより、印象語解釈ルールを作成する

表1 我々が提案している印象尺度[14]

Table 1 Impression scales proposed by the authors.

番号	印象尺度を構成する印象語
1	静かな — 激しい
2	落ち着いた — 忙しい
3	爽やかな — 重苦しい
4	明るい — 暗い
5	荘厳な — 軽々しい
6	ゆったりとした — 窮屈な
7	綺麗な — 綺麗でない
8	楽しい — 悲しい
9	気持ちが落ち着く — 気持ちが高揚する
10	心が癒される — 心が傷つく

という方式を提案する。この方式は、印象語同士を楽曲を媒介としないで対応付けようという試みであり、楽曲を媒介として対応付ける従来の方式[5]～[12]に比べ、誤差は大きくなるものと予想されるが、その反面、より多くの印象語を入力可能にすることができ、自由かつ自然な印象表現（印象語を含む文表現）を受理できる自然言語インタフェースの実現に資するものと考えられる。

以下、2.で我々が開発した印象尺度ベースの楽曲検索システムについて述べ、印象語解釈ルール作成のための前提条件について考察する。3.で印象に基づく楽曲検索システムが受理すべき印象語を選定し、4.で印象語解釈ルールを作成する。5.でこのルールを用いた意味理解（印象表現の検索条件への変換）の実行例を示し、最後に6.で本稿のまとめと今後の課題について述べる。

2. 印象語解釈ルール作成のための前提条件

今回作成される印象語解釈ルールは、入力された印象語を検索条件に変換するためのものであり、我々が開発した印象尺度ベースの楽曲検索システム[11], [12]に実装される。従って、この印象語解釈ルールに基づいて生成される検索条件は、我々のシステムで用いられている検索条件と同じ仕様でなければならない。そこで、まず、このシステムにおいて、どのような検索条件が用いられているかを示すことにする。

さて、我々のシステムへの印象入力は、表1に示した10対の印象尺度の中から1対以上を選択し、7段階評価するという形で行われる。すなわち、「悲しい」曲を見つけない場合は、印象尺度8の「とても楽しい（7点）—楽しい（6点）—少し楽しい（5点）—どちらとも言えない（4点）—少し悲しい（3点）—悲しい（2点）—とても悲しい（1点）」の中から「悲しい」を選択すればよい。このとき、検索条件は、選択されなかった印象尺度に便宜的に0点を付与することにより、

(0 0 0 0 0 0 0 2 0 0)

というベクトル形式で記述される。同様に、「静かでゆったりとした」曲を見つけない場合には、「静かな」と「ゆったりとした」を評価すればよく、その結果、

(6 0 0 0 0 6 0 0 0 0)

という検索条件が生成される。

以上のことから、印象語解釈ルールに基づいて生成される検索条件への要求仕様は、次のようになる。すなわち、(1) 検索条件は 10 次元のベクトルとして記述される。(2) このベクトルの各要素はそれぞれ表 1 に示された印象尺度に対応している。すなわち、各要素の意味は対応する印象尺度を構成する印象語の意味によって定義付けられ、値は印象尺度の評定スケール (例えば、「とても楽しい (7 点) ~ どちらとも言えない (4 点) ~ とても悲しい (1 点)」) に準じたものとなっている。(3) 「評価しない」印象尺度に対しては 0 点が付与される。(4) 2 つ以上の検索条件を合成し、1 つにすることができる。

3. システムに受理されるべき印象語の選定

ユーザフレンドリーな自然言語インタフェースを実現するためには、ありとあらゆる印象語を解釈できるようにしたいところであるが、実際問題としては時間と労力の制約から入力可能な印象語の数を限定する必要がある。そこで、システムに受理されるべき印象語を選定するに当たり、以下の 3 つの基準を導入する。

基準 1) 楽曲検索時における利用頻度が高い印象語

印象に基づく楽曲検索システムへの入力となる印象表現がアンケート調査 (303 名) により収集され、印象表現データベース [15] として Web 上で公開^(注1)されている。このデータベースから音楽作品の感情的性格を表す印象語 401 語と楽曲聴取者の感情状態の変化を表す印象語 95 語の計 496 語を抽出し、その利用頻度を調べた。その結果、印象語の異なり数 148 語のうち 62 語が頻度 2 以上であり、全体の 82.7% (410 語/496 語) を占めていることがわかった。そこで、頻度 2 以上の 62 語及びその否定語の計 124 語を受理されるべき印象語として選定することにする。但し、本稿では肯定形 (例えば「うるさい」) に対してはその否定形 (「うるさくない」) を、否定形 (例えば「うるさくない」) に対してはその肯定形 (「うるさい」) を否定語と呼ぶことにする。ここで参考のために、頻度 5 以上の印象語を表 2 に示す。

基準 2) 我々が提案している印象尺度を構成する印象語

表 1 に示された印象尺度は、基準 1 に基づいて抽出された印象語 496 語をベースに設計されたものであり、印象尺度の左側に現れる印象語は高頻度語の中から選ばれている。これに対し、印象尺度の右側に現れる印象語のうち 5 語は、頻度 1 の印象語であったり、設計者による自作であったりしたため、基準 1 で抽出された印象語の中には含まれていない。そこで、この 5 語とその否定語の計 10 語を受理されるべき印象語として選定することにする。

基準 3) 感情価測定尺度を構成する印象語

音楽の感情的側面に関する認知心理学的研究の分野において、音楽作品の感情的性格を測るための尺度として感情価測定尺度 [16] が提案されている。この尺度は表 3 に示された印象語 24 語からなるが、このうちの 9 語は基準 1 で抽出された印象語

表 2 利用頻度 5 以上の印象語

Table 2 Impression words that were used five or more times.

静かな (51), 落ち着いた (30), 明るい (29), 気持ちが落ち着く (26), 心が癒される (22), ゆったりとした (16), 爽やかな (14), 荘厳な (12), 穏やかな (12), 楽しい (12), 軽快な (10), リラックスできる (9), 気持ちが安らぐ (7), 壮大な (7), 優しい (7), 綺麗な (6), 重厚な (6), 単純な (6), 軽やかな (6), 透明な (5), 心地よい (5), 柔らかな (5)

(括弧内の数字は利用頻度を表す)

表 3 感情価測定尺度 [16] を構成する印象語

Table 3 Affective value scale of music (AVSM).

沈んだ, 哀れな, 悲しい, 暗い, 陽気な, うれしい, 楽しい, 明るい, 優しい, いとしい, 恋しい, おだやかな, 強い, 猛烈な, 刺激的な, 断固とした, きまぐれな, 浮かれた, 軽い, 落ち着きのない, 厳肅な, おごそかな, 崇高な, 気高い

表 4 被験者の年齢構成

Table 4 Ages of the 100 subjects.

年齢	被験者数
～19	18
20～29	54
30～39	25
40～49	3
50～	0
合計	100

の中に含まれているため、ここでは残りの 15 語とその否定語の計 30 語を受理されるべき印象語として選定することにする。

以上の結果、全部で 164 個の印象語 (否定語を含む) が選定された。なお、自然言語インタフェースの実現に際しては、表記上の揺れ (例えば「綺麗な」、「奇麗な」、「キレイな」、「きれいな」といった揺れや「楽しい」、「愉しい」といった揺れ) や多様な言い回し (例えば「優雅な曲がいい」と「優雅な感じの曲を聴きたい」など) といった問題を前処理等において解決することにより、より自由かつ自然な印象表現を入力可能にすることができる。

4. 被験者実験に基づく印象語解釈ルールの作成

印象語解釈ルールを作成するに当たり、それぞれの印象語の意味をどう解釈し、どのような検索条件に変換すべきか、ということを調べるために、以下のような被験者実験を行った。すなわち、被験者に 3. で選定された印象語と我々の印象尺度を紙ベースで提示し、その印象語で表現される曲 (例えば「神秘的な曲」) はどのようなイメージの曲か、ということと各印象尺度に対する 7 段階評価という形で記述してもらった。但し、印象語の提示順序として 10 パターンを用意し、提示順序と所要時間による被験者への影響を分散させた。なお、該当するイメージがない場合は無理に評価しなくてもよいと教示した。

被験者は、男性 50 名、女性 50 名の計 100 名であり、プロレベル (演奏家としての収入があるような人) 2 名、セミプロレ

(注1): <http://www2.crl.go.jp/jt/a133/resource.html>

表5 被験者による印象評価実験の結果 (一部)

Table 5 Some results of the experiments.

印象尺度		1	2	3	4	5	6	7	8	9	10
哀れな	$\bar{x}$	4.99	4.99	2.48	2.13	4.81	4.21	4.18	1.75	3.77	2.55
	σ	1.23	1.09	0.90	0.95	1.08	1.53	1.60	0.83	1.48	1.37
	N	81	72	86	94	67	78	73	93	56	86
懐かしい	$\bar{x}$	5.01	5.40	4.91	4.47	4.37	5.37	5.58	4.72	5.45	5.72
	σ	0.98	0.86	1.07	1.10	0.98	0.77	1.14	1.34	1.53	1.08
	N	85	86	74	81	65	84	80	81	91	94
古典的な	$\bar{x}$	5.42	5.56	3.59	3.47	5.57	5.15	5.51	3.58	5.06	5.09
	σ	0.83	0.95	1.27	0.99	0.77	1.28	0.93	0.94	0.95	0.86
	N	83	84	74	77	90	78	80	67	81	75
心地よい	$\bar{x}$	5.30	5.61	5.52	5.28	4.57	5.66	6.01	5.21	6.03	6.13
	σ	1.07	1.08	1.05	0.82	1.12	0.82	0.99	0.93	1.18	0.99
	N	92	94	95	90	74	89	92	85	95	92
優しい	$\bar{x}$	5.49	5.79	5.62	5.27	4.45	5.62	6.01	5.10	5.85	6.16
	σ	0.93	0.77	0.83	0.80	1.13	0.77	0.67	0.86	1.10	0.76
	N	95	95	85	88	67	92	93	81	93	96
力強い	$\bar{x}$	2.13	3.38	3.42	4.58	5.50	4.46	4.64	4.49	2.38	4.49
	σ	1.09	1.64	1.33	1.27	1.25	1.20	1.19	1.21	1.27	0.99
	N	96	79	74	73	88	68	59	61	92	53

表6 印象語と検索条件との対応表 (一部)

Table 6 Part of the table of correspondences between impression words and retrieval conditions.

印象尺度	1	2	3	4	5	6	7	8	9	10
哀れな			2.48	2.13				1.75		
懐かしい	5.01	5.40		4.47		5.37				5.72
古典的な	5.42	5.56		3.47	5.57		5.51		5.06	5.09
心地よい	5.30	5.61	5.52	5.28		5.66	6.01	5.21	6.03	6.13
優しい	5.49	5.79	5.62	5.27		5.62	6.01	5.10	5.85	6.16
力強い	2.13									2.38

ベル (音楽大学などで専門的に勉強したような人) 6 名, アマチュアレベル (バンドやオーケストラ, 合唱団などに入っているような人) 24 名, 趣味レベル (以上の条件には該当しないけれども一応演奏できるような人) 39 名, 未経験者 (ほとんど演奏できないような人) 29 名と音楽経験が豊かでない人も多数含まれていた。印象に基づく楽曲検索は, 音楽経験の豊富な人というよりも, そうでない人に対して特に有効な検索手段であり, そういう人の印象語に対する言語感覚を反映したデータを利用することは重要なことと言える。なお, 表4に被験者の年齢構成を示す。

以上の実験の結果, それぞれの印象語に対し, 印象尺度毎に評価結果の平均値 ($\bar{x}$) と標準偏差 (σ), データ数 (N) が得られた。このデータの一部を表5に示す。なお, データ数とは当該印象語を当該印象尺度で7段階評価した被験者の数であり, 「評価しない」を選択した被験者の数が除かれている。

さて, この実験結果は, 100 名の被験者がそれぞれの印象語の意味をどう解釈したかを示しており, 印象語と検索条件との対応関係を表している。しかしながら, データの中には標準偏差 (すなわち被験者による変動) が大きいものや, 「評価しない」とした被験者が多いものもある。そこで, 被験者 X 人中何人がその平均値 $\bar{x}$ を支持しているかを示す指標として, 支持率

p という関数を定義し, $p \geq a$ を満足するデータだけを採用することにする。

まず, 被験者による得点分布を正規分布と仮定すると, 区間 $(\bar{x} - \sigma, \bar{x} + \sigma)$ の間には約 0.683 X 人の被験者が含まれており, 同様に区間 $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$ の間には約 0.954 X 人の被験者が含まれていることになる。そこで, 区間 $(\bar{x} - 1, \bar{x} + 1)$ の間に含まれている被験者の数 n を計算し,

$$p = n/X$$

と定義することにする。なお, 区間 $(\bar{x} - 1, \bar{x} + 1)$ とは, 印象尺度の評定スケール上では, 「静かな (6 点)」を平均値 ($\bar{x}$) とすると, 「少し静かな (5 点)」より高得点で, 「とても静かな (7 点)」より低得点であることを意味している。

さて, n の値の計算方法であるが, 区間 $(\bar{x} - 1, \bar{x} + 1)$ は区間 $(\bar{x} - (1/\sigma)\sigma, \bar{x} + (1/\sigma)\sigma)$ と変形できるので, n の値は, 標準正規分布の面積 [17] を用いて, 計算することができる。すなわち, 区間 $(\bar{x} - k\sigma, \bar{x} + k\sigma)$ の面積を $f(k)$, 標本数 (データ数) を N とすると,

$$n = f(k)N$$

という式が与えられる。従って, $X = 100$, $k = 1/\sigma$ であるこ

とから、

$$\begin{aligned} p &= n/X \\ &= f(k)N/100 \\ &= f(1/\sigma)N/100 \end{aligned}$$

という変形が成り立つ。ここで、被験者 100 人の過半数に相当する 51 人が支持しているか否かを閾値として採用することにし、 $\alpha = 51/100 = 0.51$ とする。すなわち、

$$p = f(1/\sigma)N/100 \geq 0.51$$

を満たすデータだけを支持されるデータとして採用することにし、他のデータは除去した。その結果の一部を表 6 に示す。なお、除去された部分には便宜的に 0 点を付与し、「評価しない」と同じ扱いをする。

また、 $p \geq 0.51$ という条件では、すべてのデータが除去されてしまった印象語が 5 語あった。このような場合には、最も支持率の高いデータだけを採用することにした。表 7 にこの 5 語をその支持率とともに示す。

以上の操作の結果得られた印象語と印象尺度の評価結果との対応表は、印象語をどのような検索条件に変換すべきかを示したものとなっており、この対応表が印象語解釈ルールとしてシステムに実装されることになる。

5. 自由かつ自然な印象表現の意味理解

自由かつ自然な印象表現の意味理解、すなわち検索条件への変換は、図 1 のような流れで行われる。まずはじめに、入力された印象表現は日本語形態素解析システム juman [18] により形態素に分解される。このとき、各形態素には基本形情報、品詞情報、品詞細分類情報、活用形情報といった情報が付与される^(注2)。例えば、「愉しくて優しい感じの曲がいい」という印象表現からは表 8 のような解析結果が得られる。

次に、この形態素解析の結果から 3. で選定された印象語が抽出される。具体的には、表 9 のような基本形情報と印象語との対応表を用い、この対応表に登録された基本形が形態素解析結果の中に見つかれば、対応する印象語が抽出されるという仕組みであり、このような対応表を用いることにより、表記上の揺れ（「綺麗な」、「奇麗な」、「キレイな」、「きれいな」や、「楽しい」、「優しい」など）も解消される。表 8 に示された形態素解析結果を例にとると、基本形「楽しい」と「優しい」が対応表の左側（基本形情報の項）に現れているので、その右側にある印象語「楽しい」と「優しい」がそれぞれ対応する印象語として抽出される。また、形態素解析結果から印象語だけを抽出するので、言い回しが多少異なっても、その文に含まれている印象語が同じなら、同じ検索条件が生成され、結果として言い回しの違いが解消されることになる。例えば、印象表現「優雅な曲がいい」と「優雅な感じの曲を聴きたい」は、言い回し

(注2)：juman からの出力には形態素の読み方や活用型に関する情報も含まれている。

表 7 支持率の低い印象語

Table 7 Impression words with low support rate

印象語	支持率
壮大でない	0.502
かっよくない	0.495
幻想的でない	0.473
メリハリのない	0.470
ダイナミックでない	0.449

自由かつ自然な
印象表現

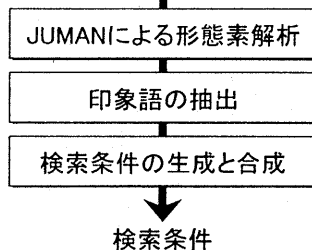


図 1 意味理解の流れ

Fig. 1 Flow for the generation of retrieval conditions.

表 8 形態素解析の例

Table 8 An example of morphological analysis by juman.

形態素	基本形	品詞	品詞細分類	活用形
愉しくて	楽しい	形容詞	*	タ系連用テ形
優しい	優しい	形容詞	*	基本形
感じ	感じ	名詞	普通名詞	*
の	の	助詞	接続助詞	*
曲	曲	名詞	普通名詞	*
が	が	助詞	格助詞	*
いい	いい	形容詞	*	基本形

表 9 基本形情報と印象語との対応表（一部）

Table 9 Part of the table of correspondences between basic forms and impression words.

基本形情報	印象語
楽しい、嬉しい、たのしい、楽しげだ、うれしげだ、たのしげだ、楽しさ、うれしさ、たのしさ	楽しい
優しい、やさしい、優しさ、やさしさ	優しい
哀れだ、憐れだ、あわれだ、哀れみ、憐れみ、あわれみ	哀れな
単純だ、たんじゅんだ、シンプルだ、しんぷるだ、単純、たんじゅん、シンプル、しんぷる	単純な

は異なるが、同じ印象語（すなわち「優雅な」）が抽出されるので、どちらの印象表現からも「優雅な」に対応した検索条件が生成されることになる。

なお、複数の形態素から構成される印象語（例えば「楽しくない」や「元気の出る」など）の抽出には、基本形情報の並び方と印象語との対応を示した表（一部を表 10 に示す）が用いられる。例えば、印象表現「うるさくない曲」は、その形態素解析結果において基本形「うるさい」と「ない」が並ぶどこか

表 10 基本情報の並び方と印象語との対応表 (一部)

Table 10 Part of the table of correspondences between sequences of basic forms and impression words.

基本情報の並び方	印象語
うるさい ない	うるさくない
楽しい ない	楽しくない
元気 出る ない	元気の出ない
暗い なる ない	気持ちが暗くならない

ら、印象語「うるさくない」が抽出される。

最後のステップでは、抽出された印象語が印象語と検索条件との対応表 (表 6 を参照) に基づいて検索条件に変換される。複数の印象語が抽出されている場合には検索条件の合成も行われる。この検索条件の合成は、それぞれの印象尺度における平均値の計算という形で行われるが、「評価しない」に該当する 0 点は計算の対象外として扱われる。

ここで、表 8 から抽出された印象語「楽しい」と「優しい」を例に、検索条件への変換および合成の流れを示す。まず、「楽しい」に対応する検索条件は、表 1 から、

(0 0 0 0 0 0 0 6 0 0)

であり (印象語と検索条件との対応表でも同じベクトルである)。「優しい」に対応する検索条件は、表 6 から、

(5.49 5.79 5.62 5.27 0 5.62 6.01 5.10 5.85 6.16)

である。従って、この 2 つの検索条件が合成され、

(5.49 5.79 5.62 5.27 0 5.62 6.01 5.55 5.85 6.16)

となる。

6. ま と め

現在、我々は、自然言語インタフェースを有する、印象に基づく楽曲検索システムを開発している。自然言語インタフェースを採用することにより最も期待されることは、見つけたい楽曲の印象をシステムが提示する少数の印象語の中から探すというのではなく、語彙的にも言い回し的にも自分の言葉で表現できるという点にあるが、従来のシステムではごく限られた数の印象語しか入力できなかった。ある印象語を入力できるようにするためには、多くの被験者による楽曲聴取実験が必要のため、非常に高コストであり、多くの印象語を対象とすることは実質的に困難であった。そこで、本稿では、「印象に基づく楽曲検索」タスクにおいて受理されるべき印象語 164 語を選定し、これらの印象語を既存の印象尺度で評価することにより、印象語解釈ルールを作成するという方式を提案した。この方式は、評価時に楽曲の聴取を伴わないので、比較的低コストであり、164 語という多くの印象語を入力可能にすることができた。

今後の課題は、現在開発中の自然言語インタフェースを完成させることであり、その一環として、「少し」や「とても」、「かなり」といった程度表現や「もっと」や「もう少し」のような

比較表現が印象語に付されたときに、その印象語から生成される検索条件に対しどのような処理を施すべきか、といった問題に取り組んでいる。また、この自然言語インタフェースが本当にユーザフレンドリーであるのかを検証するために、被験者実験を行い、システムが受理すべき印象語として選定された 164 語の妥当性や、印象語から生成される検索条件の妥当性、複数の印象語から生成され、合成される検索条件の妥当性を検証する予定である。

文 献

- [1] S.G. Blackburn, and D.C. De Roure, "A tool for content based navigation of music," Proc. 6th ACM International Multimedia Conf., pp. 361-368, Bristol, UK, Sept. 1998.
- [2] 園田智也, 後藤真孝, 村岡洋一, "WWW 上での歌声による曲検索システム," 信学論 (D-II), vol.J82-D-II, no.4, pp.721-731, Apr. 1999.
- [3] 橋口博樹, 西村拓一, 張建新, 滝田順子, 岡隆一, "モデル依存傾斜制限型の連続 DP を用いた鼻歌入力による楽曲信号のスポッティング検索," 信学論 (D-II), vol.J84-D-II, no.12, pp.2479-2488, Dec. 2001.
- [4] 小杉節子, 小島明, 片岡良治, 串間和彦, "大規模音楽データベースのハミング検索システム," 情処学論, vol.43, no.2, pp.287-298, Feb. 1999.
- [5] 近藤央規, 宝珍輝尚, 都司達夫, "感性の主因子を用いたマルチメディアデータ検索システム," 信学技報, vol.DE96, no.3, pp.13-18, May 1996.
- [6] 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.SP96-124, pp.17-24, Mar. 1997.
- [7] 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," 情処学データベースシステム研報, vol.118, no.8, pp.57-64, May 1999.
- [8] 佐藤聡, 小川潤, 堀野義博, 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," 信学技報, vol.SP2000-137, pp.51-56, Feb. 2001.
- [9] 池添剛, 梶川嘉延, 野村康雄, "形容詞対を用いた音楽データベース検索システム," 情処学音楽情報科学研報, vol.33, no.2, pp.7-14, Dec. 1999.
- [10] 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," 情処学論, vol.42, no.12, pp.3201-3212, Dec. 2001.
- [11] 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," 情処学音楽情報科学研報, vol.2002, no.63, pp.37-42, Jul. 2002.
- [12] 熊本忠彦, 太田公子, "印象に基づく楽曲検索システム: N グラム統計量の利用," 情報技術レターズ (FIT2002), vol.1, no.LD-6, pp.63-64, Sept. 2002.
- [13] 原田将治, 伊東幸宏, 中谷広正, "感性語句を含む自然言語文による画像検索のための形状特徴空間の構築," 情処学論, vol.40, no.5, pp.2356-2366, May 1999.
- [14] 熊本忠彦, 太田公子, "印象に基づく楽曲検索: 検索ニーズに合った印象尺度の設計," 情処学自然言語処理研報, vol.2002, no.4, pp.35-40, Jan. 2002.
- [15] 熊本忠彦, 太田公子, "印象に基づく楽曲検索研究のための印象表現の収集," 情処学論, vol.43, no.10, pp.3231-3234, Oct. 2002.
- [16] 谷口高士, 音楽と感情, 北大路書房, 京都, p.75, Jan. 1998.
- [17] P.G. ホーエル, 初等統計学, 浅井晃, 村上正康 (訳), 培風館, 東京, p.295, Feb. 1989.
- [18] 黒橋 禎夫, 長尾 真, 日本語形態素解析システム JUMAN version 3.61, <http://pine.kuee.kyoto-u.ac.jp/nl-resource/juman.html>, May 1999.