

印象に基づく楽曲検索のための自然言語インタフェースの設計と実装

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

1. まえがき

大量にある楽曲データの中から、ある特定のもの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、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 [1]-[4]。これら先行研究において、印象の入力は、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）という形で行われる。そのため、入力可能な印象語の数が多いと、入力に要する時間と労力が増大してしまい、少ないと、検索可能な印象のバリエーションが限られてしまう。そこで、本稿では、ユーザの自由かつ自然な印象表現（印象語を含む文）を受け取れる自然言語インタフェースを提案する。

2. 自然言語インタフェースの設計

自然言語インタフェースの役割は、ユーザが入力する自由かつ自然な印象表現から楽曲検索システムへの入力となる検索条件を生成することであり、その流れは図1のようになる。以下、各処理について述べる。

2.1 形態素解析部

形態素解析部には汎用の日本語形態素解析システムである juman[5] を採用する。juman は、入力された文を形態素の列に分解し、それぞれの形態素に基本形、品詞名、品詞細分類名といった情報を付与する。

2.2 キーワード抽出部

検索要求が同じでも「優雅な曲がいい」や「優雅な感じの曲を聴きたい」のように言い回しが異なる場合がある。そこで、このような言い回しの多様性を吸収するために、印象語や程度語をキーワードスポッティング的に抽出する方式を採用する。本稿では、抽出される語をキーワードと呼び、印象語には、印象に基づく楽曲検索時に高頻度で利用される語、音楽作品の感情的性格を測るための語（感情価測定尺度 [6]）、我々の楽曲検索システム [3][4] においてすでに入力可能な語（印象尺度、表1参照）の計 82 語（異なり数）とその否定語*1 82 語の計 164 語を、程度語には、シソーラス [7] から抽出した程度を表す副詞等の計 119 語を定義する。また、印象語、程度語とは別に、接尾辞「ない」のような印象語を構成する語や係り受けを決定するために必要な情報（助詞「は」など）もキーワードとして抽出する。

2.3 検索条件生成部

検索条件生成部は、印象語解釈ルール（一部を表2に示す）[8] を用いて、抽出された印象語から検索条件を生成する。

この印象語解釈ルールは被験者実験（男性 50 人、女性 50 人）に基づいて作成された。まず、各被験者に印象語と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、ということをも 1 対以上の印象尺度に対する 7 段階評価という形で回答してもらった。その結果、印象語と各印象尺度における 7 段階評価値との対応関係が得られたが、このデータの中には被験者による変動が大きいものや、不適切な印象尺度と考えた被験者が多いものもあった。そこで、それぞれの印象尺度において、データ

自由かつ自然な印象表現（印象語を含む文）

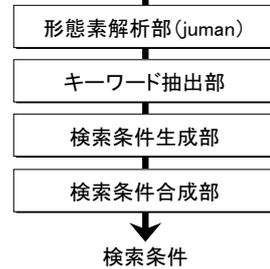


図1: 自然言語インタフェースにおける処理の流れ

表1: 我々が提案している印象尺度

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

表2: 印象語解釈ルール（一部）

印象尺度	1	2	3	4	5
哀れな	nil	nil	2.48	2.13	nil
心地よい	5.30	5.61	5.52	5.28	nil
優しい	5.49	5.79	5.62	5.27	nil
印象尺度	6	7	8	9	10
哀れな	nil	nil	1.75	nil	nil
心地よい	5.66	6.01	5.21	6.03	6.13
優しい	5.62	6.01	5.10	5.85	6.16

の平均値 \bar{x} と標準偏差 σ を求め、区間 $(\bar{x} - 1, \bar{x} + 1)$ の間に含まれている被験者の数（被験者による得点分布を正規分布と仮定した場合の理論上の数）を計算した。そして、この数が被験者の過半数（51 人）以上であるデータだけを採用することにし、他のデータは除去した（実際には記号「nil」を付与した）。その結果得られた対応表が印象語解釈ルールである。

一方、程度語が抽出されている場合には、程度語定数表（一部を表3に示す）[9] を用いて、印象語から生成された検索条件の大きさを拡大/縮小する。すなわち、ある印象語から生成された検索条件を $\vec{v} = [v_i]$ ($i = 1, 2, \dots, 10$)、その印象語を修飾している程度語の定数を a とすると、この検索条件は

$$\vec{v}' = [a(v_i - 4) + 4] \quad (i = 1, 2, \dots, 10) \quad (1)$$

と変換される。

この程度語定数表も被験者実験（男性 50 人、女性 50 人）に基づいて作成された。まず、各被験者に程度語によって修飾された印象語（例えば「わりと静かな曲」）と点数の基準表を紙ベースで提示し、その印象語で表現される曲が点数の基準表のどのあたりに位置するか、ということも回答してもらった。但

An Impression-based Music-Retrieval System with Natural Language Interface, Tadahiko Kumamoto and Kimiko Ohta, Communications Research Laboratory

*1 本稿では、肯定形に対してはその否定形を、否定形に対してはその肯定形を否定語と呼ぶ。

表 3: 程度語定数表 (一部)

最高に (静かな)	1.75	比較的 (静かな)	0.54
本当に (静かな)	1.57	少し (静かな)	0.5
とても (静かな)	1.5	やや (静かな)	0.47
割と (静かな)	0.63	かすかに (静かな)	0.18

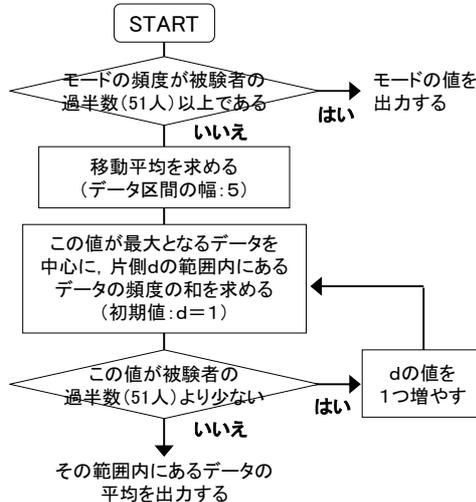


図 2: 程度語定数を決定するためのアルゴリズム

し、点数の基準表は、「どちらとも言えない」を原点(0点)とする数直線であり、「とても静かな曲(12点)」より静かな曲と「とても静かでない曲(-12点)」より静かでない曲を想定することにより、29段階の評価が可能であった。

以上の実験の結果得られた被験者の評価結果(点数の分布)には多義性に由来すると思われるばらつきが見られた。そこで、図2に示すアルゴリズムを導入し、程度語の意味を一意に決定した。すなわち、移動平均を用いて多義性を解消し、被験者の過半数以上の人が支持する範囲で局所平均を求めることにより平均的な意味解釈を決定した。このアルゴリズムの出力値を印象尺度の評定スケール(7段階評価値)に換算し、式(1)を用いて a の値を逆算することにより、程度語定数表を作成した。

2.4 検索条件合成部

複数の検索条件が生成された場合には検索条件の合成を行う。この操作は、印象尺度毎に平均値を計算することにより行われるが、「nil」は計算の対象外として扱われる。例えば、印象語「楽しい」と「優しい」からは、それぞれ

$$(nil \ nil \ nil \ nil \ nil \ nil \ 6 \ nil \ nil) \\ (5.49 \ 5.79 \ 5.62 \ 5.27 \ nil \ 5.62 \ 6.01 \ 5.10 \ 5.85 \ 6.16)$$

という検索条件が生成され、この2つが合成されると、

$$(5.49 \ 5.79 \ 5.62 \ 5.27 \ nil \ 5.62 \ 6.01 \ 5.55 \ 5.85 \ 6.16)$$

となる。

3. 楽曲検索システムへの実装

我々は提案の自然言語インタフェースを印象尺度ベースの楽曲検索システム [3][4] に実装した。本システムの検索画面を図3に示す。画面上部にあるテキスト領域はユーザとシステムの対話の履歴を表示するための領域であり、中段左側にあるテキスト領域はシステムからの要求もしくは応答を表示するための領域である。一方、その右側にあるテキスト領域が

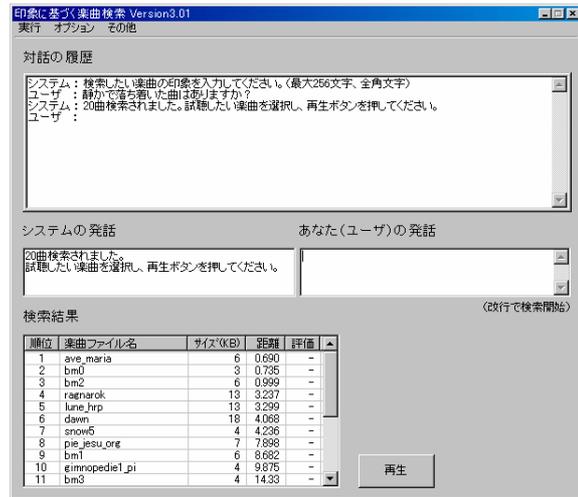


図 3: 自然言語インタフェース (検索画面)

ユーザが印象入力を行うための領域であり、印象表現を入力し、改行を行えば、文の解釈、検索条件への変換、検索が行われ、その結果が、画面左下の検索結果表示領域に表示される。

4. まとめ

本稿では、我々が開発した楽曲検索システムに対し、自然言語インタフェースを実装した。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、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。これに対し、我々が提案する自然言語インタフェースでは、「静かな」や「優雅な」のような印象語 164 語、「かなり」や「わりと」のような程度語 119 語を用いて、自由かつ自然な文形式で、検索したい楽曲の印象を入力することが可能である。

今後の課題としては、自然言語インタフェースの性能評価、自然言語対話による絞込み検索方式の開発が挙げられる。

参考文献

- [1] 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.
- [2] 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.
- [3] 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,” 情処研報, vol.2002-MUS-46, pp.37-42, Jul. 2002.
- [4] 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, no.LD-6, pp.63-64, Sept. 2002.
- [5] 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.
- [6] 谷口高士, “音楽と感情,” 北大路書房, 京都, 1998.
- [7] 大野晋, 浜西正人, “類語国語辞典,” 角川書店, 東京, 1985.
- [8] 熊本忠彦, 太田公子, “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,” 信学技報, vol.SP2002-134, pp.81-86, Dec. 2002.
- [9] 熊本忠彦, 太田公子, “印象に基づく楽曲検索: 自然言語インタフェースの実現,” 情処研報, vol.2003-NL-153, no.13, Jan. 2003.