

音楽データの印象の時間的推移を扱う印象メタデータ自動生成方式

伊地智 麻子[†] 清木 康^{††}

マルチメディアデータベースの分野では、メディアデータ検索のためのメタデータ自動生成の実現が、重要な研究課題となっている。特に、楽曲のように、時間軸に沿って印象が推移するメディアは、時間的推移に対応するメタデータ生成方式の実現が重要となる。本論文では、クラシック楽曲を、音楽形式・楽章などにより時間軸上でセクション（時間軸に沿って区切られた部分）に区切り、さらに各セクションごとの印象をもとに、楽曲全体の総合的な印象メタデータを生成する方式を提案する。本方式の特徴は、各セクションごとの印象について、隣り合う前後のセクションから生じる印象の差に応じて重みを算出し、各セクションの印象に重み付けすることにより、前後関係から生じる各セクションの印象の強度を印象メタデータへ反映させる点にある。また、本論文では、本方式により生成される印象メタデータの応用可能性を示すために、本方式によって生成される印象メタデータを対象として、すでに提案されている意味的連想検索方式を適用する感性楽曲検索の実現方式を示す。さらに、提案方式の検証実験、および、生成された印象メタデータの妥当性の検証実験結果を示し、有効性を検証する。

An Automatic Metadata Creation Method Dealing with Temporal Transition for Impression of Music Data

ASAKO IJICHI[†] and YASUSHI KIYOKI^{††}

In the research field of multimedia database systems, one of the most important issues is to realize automatic metadata creation for media data retrieval. This paper presents a metadata creation method for temporally separated classical music by "An Automatic Metadata Creation Method for Music Data," and a metadata aggregation function according to impression of each interval. We also present its application to semantic associative search based on the Mathematical Model of Meaning which has been proposed before. By applying the extracted metadata to semantic associative search, we realize a retrieval environment where users can access the highly related classical music to their impressions given as queries. This paper shows several experimental results of the metadata creation, and also shows experimental results of semantic associative search to clarify the feasibility and effectiveness of our method.

1. はじめに

近年、多種多様なメディアデータのデジタル化が進み、デジタル化された多くのメディアデータが、広域ネットワーク環境上に置かれるようになった。これにともない、広域ネットワークを介して多種多様なメディアデータを入手可能となっている。

アクセス可能なメディアデータ群の中から、要求に応じて適切なメディアデータを抽出するシステムの実現は有効であり、散在する膨大なメディアデータ群から適切なメディアデータを抽出する方法の確立は重要

な課題となっている。特に、音楽データについては、1) 本来、デジタル形式で流通している、2) MPEG4 など高品質な圧縮方式が適用されているなどの理由から、ネットワーク上においてオリジナルデータとほぼ同質のメディアデータを獲得することが可能であり、ネットワークを通じた音楽の送受信の有効性は高い。したがって、現在、ネットワーク上に大量の音楽データが置かれ、かつそれらは多種多様化の傾向にある。

このような状況において、利用者の目的に合致した音楽を検索可能なシステムの実現は有効であり、その実現は重要な研究課題となっている。しかし、連続メディアデータである音楽データの内容、表現、印象、および、その推移を対象とした検索環境の実現は確立されていない。

データベースの分野において、メディアデータを検索する方法には直接検索と間接検索の2通りの方法が

[†] 慶應義塾大学政策・メディア研究科
Graduate School of Media and Governance, Keio University

^{††} 慶應義塾大学環境情報学部
Faculty of Environmental Information, Keio University

ある。前者は、メディアデータ自身のデータのパターンにより検索する方法である。後者は、メディアデータにあらかじめ与えられている抽象的な情報（以下メタデータ）を介して検索する方法である。

音楽の内容に関する直接検索の方法として、メロディを与え、それとパターンが合致する音楽データを検索する方法がある。また、音楽が持つ属性である作曲家、題名といった属性情報を対象として検索する方法がある。音楽を対象とした直接検索では、ハミングを用いた音楽検索システム⁷⁾が研究されている。音楽の内容に関する間接検索として、我々は、発想標語を用いたクラシック音楽メタデータ生成方式³⁾を提案した。また、音楽データを対象とした間接検索としては、音楽のメロディパターンから印象メタデータを自動抽出し、その印象メタデータに対して間接的に検索する楽曲メタデータ自動生成方式^{4),10),11)}を提案した。この楽曲メタデータ自動生成方式においては、楽曲を最小単位として、その楽曲に対応する印象メタデータが生成される。この方式は、楽曲全体を単位（楽曲が始まってから終わるまでがアトミック）として、楽曲全体の総合的な印象メタデータ生成する方式であり、単一曲内での時間的な変化に応じた印象抽出は対象としていない。また、他研究では、統計値により定められた感情価を対象とした方法⁹⁾が示されている。

広域ネットワーク上で、膨大なデータベース群を対象としたシステムを実現するためには、自動的、かつ、客観的なメタデータ生成が重要となる。特に、間接検索において、メタデータを対象として検索を行うために、適切なメタデータの生成は、適切な結果を導くための重要なプロセスである。

音楽データを対象とした検索において、利用者の発行する検索キーワード（印象語）に合致した音楽を検索可能とすることによって、利用者は未知の音楽を得ることも可能となる。したがって、印象メタデータを対象とした検索においては、メディアデータへの適切な印象メタデータ生成方式の実現が重要である。特に、音楽の印象が時間的に推移する場合には、音楽の印象の時間的な推移を反映する印象メタデータ生成方式が必要となる。

本論文では、楽曲を時間軸上でセクション（時間軸に沿って区切られた部分）に区切り、その各セクションごとの印象をもとに、楽曲全体の総合的な印象を表す印象メタデータを生成する方式を提案する。本方式の特徴は、各セクションごとの印象について、隣り合う前後のセクションから生じる印象の差に応じて重みを算出し、各セクションの印象に重み付けすることに

より、前後関係から生じる各セクションの印象の強度を印象メタデータへ反映させる点にある。本方式により、印象の時間的な推移により生まれた印象の際立ちを反映する印象メタデータ生成が可能となる。

本方式では、たとえば、静かな楽曲の中に急に大きな音が鳴る場合、この大きな音は際立つと考える。この際立つという感覚は、絵画メディアにおいて、黒い背景の中に白い点がある場合、この白い点が際立つというようなことと類似する。本論文では、この際立ちを2つの部位間に生じる大きな差がもたらす現象としてとらえ、それを反映する印象メタデータ生成方式を示す。本方式では、際立つという現象をメディアの時間的な印象変化に適用し、時間軸に沿って推移する印象の前後差に応じた重みを算出し、その重み（つまり印象を強調する重み）を後のセクションの印象へ反映させる。この方式により、前後の印象変化が大きいセクション間では、後のセクションの印象が強調されることになる。

本方式は、時間的な印象の推移による印象の強調を反映した印象メタデータを生成し、検索者が問合せとして与える印象語に合致するメタデータを有する音楽データの検索を実現する方式として位置付けられる。

提案方式の実現方法として、楽曲メタデータ自動生成方式により、音楽形式・楽章で客観的に区切られたセクションごとの印象生成を行う方法を示し、さらに、各セクションの印象に前後の印象差による重みを波及させる方法を示す。さらに、この実現方法により、クラシック楽曲データを対象として抽出された印象メタデータを意味的連想検索^{5),6)}へ適用する方法を提案する。意味的連想検索方式は、形容詞を用いて表現される印象に基づくメディアデータの意味的連想検索方式であり、データ間の意味的な関係を動的に計算する意味の数学モデル^{5),6)}を拡張したメディアデータの意味的連想検索方式として、すでに提案されている方式である。

本提案方式の有効性を示すために、例として、クラシック楽曲を対象とした実験結果を示す。クラシック楽曲には、1曲を時間軸上で客観的に区切る方法として、音楽形式、楽章が示されている。

また、楽曲の印象形成は、楽譜・演奏者・聞き手の3要素が重要である。楽譜に依存する印象生成を phase-1、演奏者に依存する印象生成を phase-2、聞き手に依存する印象生成を phase-3 とすると、本提案方式は、phase-1 を対象とした印象メタデータ生成方式として位置付けられる。phase-2 については、本方式によって得られる印象メタデータ群を対象とした学習の

phase として扱うことなどが考えられる．phase-3 については，phase-1・phase-2 によって得られた印象メタデータ群を対象とした個人化の phase として扱うなどが考えられる．

2 章では，提案方式である印象メタデータ生成方式を示し，3 章では提案方式の実現システムのアーキテクチャとその概要，4 章では提案方式の実現システムの検証実験，提案方式により生成された印象メタデータの妥当性の検証実験とその結果，考察を述べ，5 章において本研究のまとめと今後の課題について述べる．

2. 提案印象メタデータ生成方式

本章では，楽曲を時間軸上でセクション（時間軸に沿って区切られた部分）に区切り，その各セクションごとの印象を対象として，楽曲全体の総合的な印象メタデータを生成する方式を提案する．本方式の特徴は，各セクションごとの印象に，隣り合う前後のセクションから生じる印象の差に応じて重みを算出し，各セクションの印象に重み付けることにより，前後関係から生じる各セクションの印象の強度を印象メタデータへ反映させる点にある．本方式により，印象の時間的推移により生じる印象の際立ちを波及させた印象メタデータ生成が可能となる．

本方式は，図 1 に示すように，次の基本的前提に基づいている．

- 楽曲にはセクションという概念があり，それぞれのセクションが印象を与えている．
- さらに，それぞれのセクションの印象が相互に影響（強調・軽減）することにより，楽曲全体の印象を形成している．

これらの前提に基づく本印象メタデータ生成方式は，次の 5 ステップにより構成される．各ステップの概要を図 2 に示す．

Step-1 楽曲のセクション設定．

Step-2 セクションごとの印象生成．

Step-3 セクションごとの印象より，隣り合う前後のセクションから生じる印象の差に応じて重みを算出．

Step-4 セクションごとの印象の重み付け．

Step-5 重み付けされたセクションごとの印象より，選択および合成により，楽曲全体の総合的な印象メタデータを生成．

2.1 Step-1 楽曲のセクション設定

Step-1 は，図 3 に示すように，楽曲 M_d ($d = 1, l$) (d : 楽曲の識別子， l : 楽曲数の最大値) を，時間軸上で意味のある長さごとのセクション $S_{[d,j]}$ ($j = 1, m$)

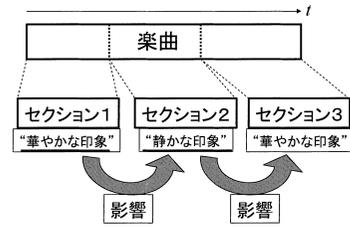


図 1 印象メタデータ生成方式の基本的前提

Fig. 1 The data structure to analyze impression of music data.

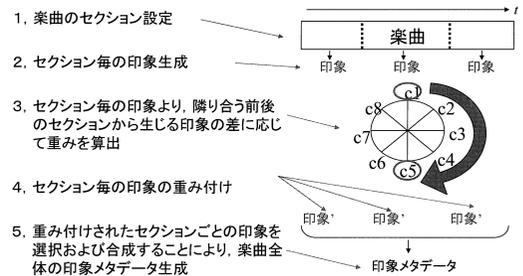


図 2 印象メタデータ生成方式の 5 ステップ

Fig. 2 The outline of a new metadata creation method for music data retrieval.

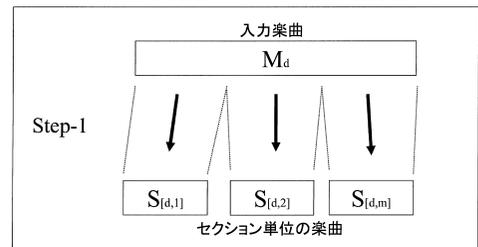


図 3 Step-1: 楽曲のセクション設定

Fig. 3 Step-1: Division of music into sections.

(j : セクションの識別子， m : セクション数の最大値) に区切る．セクション設定ポリシー（セクションの境界を決定する要素）を P_f ($f = 1, n$) (f : セクション設定ポリシーの識別子， n : セクション設定ポリシー数の最大値) として，Step-1 を式 (1) として定義する．

$$f_{divide}(M_d, P_f) \rightarrow \{S_{[d,1]}, \dots, S_{[d,m]}\} \quad (1)$$

2.2 Step-2 セクションごとの印象生成

Step-2 は，図 4 に示すように，セクション $S_{[d,j]}$ から，セクションごとの印象 $I_{[d,p]}$ ($p = 1, m$) (p : 印象の識別子) を生成する．印象を生成する方法を A_h ($h = 1, q$) (h : 印象生成方法の識別子， q : 印象生成方法数の最大値) とし，Step-2 を式 (2) として定義する．

$$f_{impression}(S_{[d,1]}, \dots, S_{[d,m]}, A_h) \rightarrow \{I_{[d,1]}, \dots, I_{[d,m]}\} \quad (2)$$

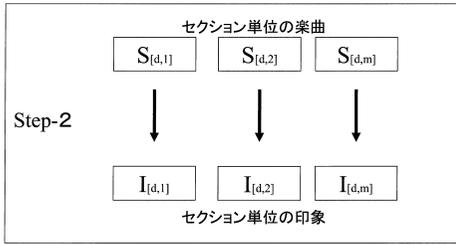


図4 Step-2: セクションごとの印象生成

Fig. 4 Step-2: Impression extraction for each section.

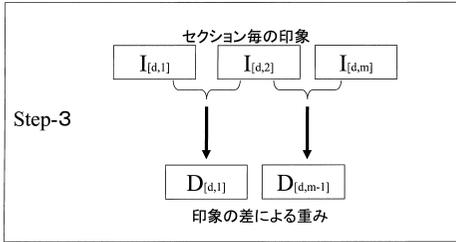


図5 Step-3: 印象の差に応じて重みを算出

Fig. 5 Step-3: Computation of impression-strength.

2.3 Step-3 セクションごとの印象より、隣り合う前後のセクションから生じる印象の差に応じて重みを算出

Step-3は、図5に示すように、セクションごとの印象 I から、隣り合う前後のセクションから生じる印象の差に応じて重み $D_{[d,s]}$ ($s = 1, m-1$) (s : 印象の差の識別子) を算出する。Step-2で生成された、隣り合う前後のセクションの印象の意味的な距離(意味的な近さ)を計量する。隣り合う前後のセクションから生じる印象の差に応じた重みの算出方法を $C_{[h,i]}$ ($i = 1, t$) (i : 印象の差の算出方法の識別子, t : 印象の差の算出メソッド数の最大値)として、Step-3を式(3)として定義する。

$$f_{degree}(I_{[d,1]}, \dots, I_{[d,m]}, A_h, C_{[h,i]}) \rightarrow \{D_{[d,1]}, \dots, D_{[d,m-1]}\} \quad (3)$$

2.4 Step-4 セクションごとの印象に重み付け

Step-4は、図6に示すように、Step-3で算出された、隣り合う前後のセクションから生じる印象の差に応じた重み D をセクションごとの重みとして印象へ反映させ、重み付き印象 $I'_{[d,p]}$ ($p = 1, m$) (p : 重み付き印象の識別子) を生成する。セクションごとの重み付け方法を W_j ($j = 1, q$) (j : 重み付け方法の識別子, q : 重み付け方法数の最大値)とし、Step-4を式(4)として定義する。

$$f_{weight}(I_{[d,1]}, \dots, I_{[d,m]}, D_{[d,1]}, \dots, D_{[d,m-1]}, W_j) \rightarrow \{I'_{[d,1]}, \dots, I'_{[d,m]}\} \quad (4)$$

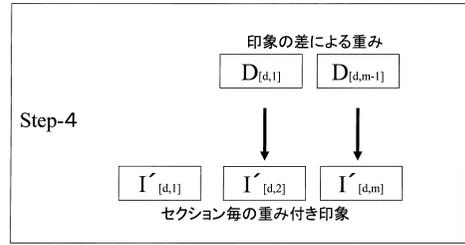


図6 Step-4: セクションごとの印象に重み付け

Fig. 6 Step-4: Weighting impression according to impression-strength.

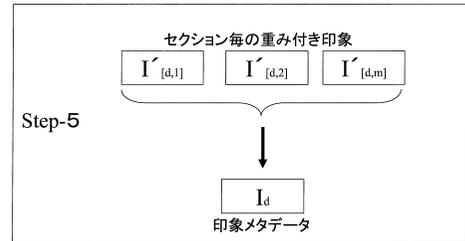


図7 Step-5: 楽曲の総合的な印象メタデータを生成

Fig. 7 Step-5: Combining impression for adjusting themselves to a query structure.

2.5 Step-5 重み付けされたセクションごとの印象より、選択および合成により、楽曲全体の総合的な印象メタデータを生成

Step-5は、図7に示すように、重み付けされたセクションごとの印象 I' から、楽曲全体の総合的な印象メタデータ I_d ($d = 1, l$) を生成する。楽曲全体の総合的な印象メタデータを生成する方法を X_k ($k = 1, r$) (k : 総合的な印象メタデータを生成する方法の識別子, r : 総合的な印象メタデータを生成する方法数の最大値)とし、Step-5を式(5)として定義する。

$$f_{aggregate}(I'_{[d,1]}, \dots, I'_{[d,l]}, X_k) \rightarrow \{I_d\} \quad (5)$$

3. 印象メタデータ自動生成実現方式

本論文では、2章に示した印象メタデータ生成方式による楽曲メディアデータベース検索を、次の2プロセスによって実現する。

プロセス1 本提案方式による楽曲データからの印象メタデータ生成

プロセス2 抽出された印象メタデータと問合せとの間の意味的連想検索

図8に、本方式による意味的連想検索システムの概観を示す。本提案方式は、このシステムにおいて、図8における太枠の“メディアデータのメタデータ生成モデル”の部分を実現する方式として位置付けられる。

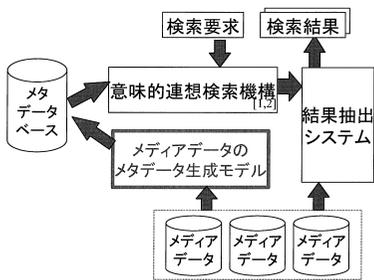


図 8 実現方式
Fig. 8 The implementation method.

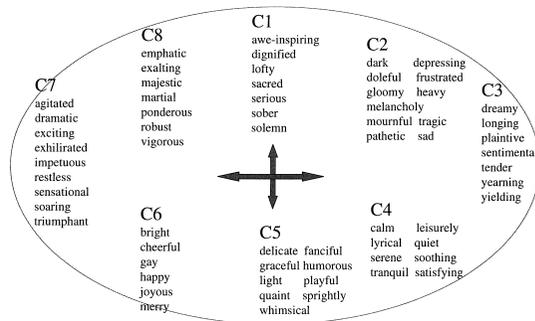


図 9 Hevner 音楽形容詞群
Fig. 9 The cyclic adjectives list.

3.1 (プロセス1) 本提案方式による楽曲データからの印象メタデータ生成

本提案方式による楽曲データからの印象メタデータ生成について、2章に示した本提案方式のステップごとに次に示す。

3.1.1 Step-1 楽曲のセクション設定

楽曲のセクション設定として、以下のセクション設定実現方法(セクションの境界を決定する要素)を示す。

- セクション設定実現方法 1 楽章の境界をセクションの境界とする方法
- セクション設定実現方法 2 楽曲形式の要素をセクションとする方法
- セクション設定実現方法 3 調が転調する時点をセクションの境界とする方法
- セクション設定実現方法 4 拍数が変わった点をセクションの境界とする方法

本実現では、事例として、セクション設定実現方法 1: 楽章を対象としたセクション設定を行う。

このセクション設定実現方法の選択については、対象楽曲がクラシック楽曲であり、楽章が作曲者の意図として最も強く客観的なセクション設定であることを前提として行った。

3.1.2 Step-2 セクションごとの印象生成

各セクションごとの印象生成方式として、以下の方式を示す。

印象生成実現方式 1 楽曲メタデータ自動生成方式^{4),10),11)}

印象生成実現方式 2 発想標語を用いたクラシック音楽メタデータ生成方式³⁾

本実現では、事例として印象生成実現方式 1: 楽曲メタデータ自動生成方式による印象生成を行う。対象楽曲が、クラシック楽曲中において、楽譜の付加記述(発想標語など)の少ない分野であること、および、生成される印象が印象とその相関量の数値で表され、重みとしての計算が可能であることからこの方式を選択

した。

3.1.3 Step-3 セクションごとの印象より、隣り合う前後のセクションから生じる印象の差に応じて重みを算出

本実現では、時間的に隣り合う印象間の印象変化を扱う方法として、Hevner の定義した音楽形容詞群間の距離を印象の変化量(差)とする印象の差の計量方式を示す。

Hevner 音楽形容詞群は、図 9 に示すように Hevner が心理学実験をもとに、楽曲を表現するために必要な形容詞セットとして 66 語を設定し、さらにそれらを意味的に 8 形容詞群 C_1, C_2, \dots, C_8 へ分類した形容詞群である。各形容詞群は、6~11 語から成り立つ。8 つの形容詞群は、環状に位置し、それぞれ円上での距離が近いほど意味的に近いと定義されている。さらに、円の対称にあるグループどうし(たとえば、 c_1 と c_5)は、逆の印象であると定義されている。たとえば、 c_1 を中心に考えると、 $c_2 \rightarrow c_3 \rightarrow c_4$ の順に意味的に遠くなり、同様に $c_8 \rightarrow c_7 \rightarrow c_6$ の順に意味的に遠くなり、 c_5 が意味的に c_1 の逆となる。

心理学者である Hevner は、すべての形容詞から、楽曲を顕著に表現し効果的に特徴付ける形容詞を示している。これらの形容詞の選択は、被験者にクラシック楽曲を聞かせ、音楽を表現するのにふさわしい語を選択する方法によって行われ、最終的に 66 語が選択され、それらの 66 語により 8 つの群が形成された。各群に含まれる形容詞どうしは、互いに近い印象を表す形容詞として使われていることを示したものである。この実験および形容詞の選択については、文献 1), 2) に詳しく述べられている。

本実現方式における印象の差の計量において、印象の重みを次のように設定する。

$$(\text{印象の差} = x : \min < x < \max)$$

印象の重みとは、セクション間の印象の差を考慮し

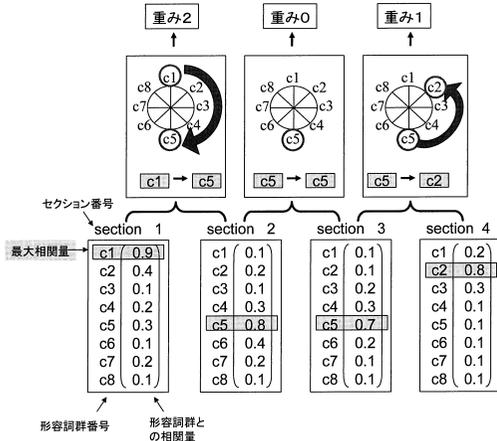


図 10 Step3: セクションごとの印象より、隣り合う前後のセクションから生じる印象の差に応じて印象の重みの算出方法
Fig. 10 The implementation method of Step-3.

ない場合を $x = 1$ とし、これに対する比として表現する指標である。当該セクションの印象の重みは、楽曲全体の印象への影響度が、セクション間の印象変化を考慮しない場合の x 倍の影響を与える印象として扱うことを意味している。

以下では、設定値を印象の重みとし、前セクションと当該セクションの間での印象変化による当該セクションの楽曲全体の印象への影響度を表す指標とする。

本実現方式では、この印象の重みの設定として、印象の変化を楽曲全体の印象へ反映できる。たとえば、 x の値として、 $x = 0, 1, 2$ の値を用いた場合には、 $max(最大値) = 2$ 、 $mean(中間値) = 1$ 、 $min(最小値) = 0$ となる。すなわち当該セクションの最大の影響度を 2 倍、最小の影響度を 0 倍、中間の影響度を 1 倍と設定することになる。この値 x の設定により、印象変化による楽曲全体の印象メタデータへの影響を制御することができる。

本実現方式では、印象の重み $x = 0, 1, 2$ を設定した場合について実現を行った。重みの算出の例(印象の重み $x = 0, 1, 2$ を設定した場合)を図 10 に示す。この図のセクション 1 の中心形容詞群は c1 であり、セクション 2 の中心形容詞群は C5 である。c1 と c5 は Hevner 音楽形容詞群上で最大変化の関係にあるため、この場合は重み 2 となる。また、セクション 3 の中心形容詞群は c5 であり、c5 と c5 は最小変化の関係にあるため、この場合は重み 0 となることを示している。

本実現方式では、Hevner 音楽形容詞群がクラシック楽曲に適した形容詞群であるためこの方法を選択した。2 章に示した提案方式の Step-3 は、Hevner 音楽形容詞群に依存するものではない。

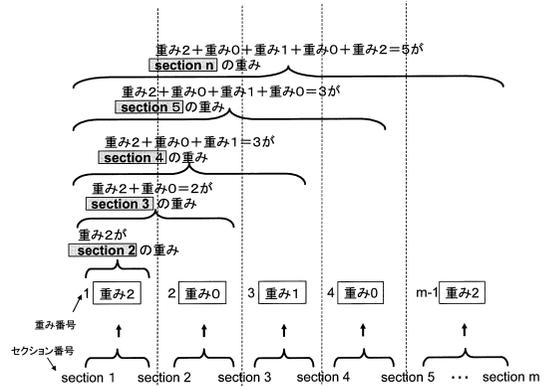


図 11 Step-4 印象重み付け実現方法 1 の重みの反映
Fig. 11 The implementation method of Step-4 Method1.

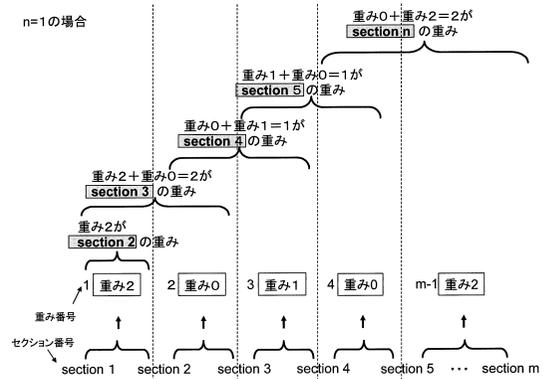


図 12 Step4 印象重み付け実現方法 2 の重みの反映
Fig. 12 The implementation method of Step-4 Method2.

3.1.4 Step-4 セクションごとの印象に重み付け

セクションごとの印象の重み付けを実現するために、時間的に前後に隣り合うセクションから算出される印象の差を、セクションごとの印象に重み付けすることにより、前後関係から生じる各セクションの印象の強度を判定し、印象へ反映させる方法を示す。

印象重み付け実現方法 1 当該セクションより前のすべての前後差を重みとして反映させる方法

印象重み付け実現方法 2 当該セクション前の指定セクション数の前後差を重みとして反映させる方法

図 11, 図 12 に示すように、印象重み付け実現方法 1 は、人が当該セクションの印象を決定させる場合、最初のセクションから当該セクションまでのすべての印象の推移を印象の強さとして反映させている。印象重み付け実現方法 2 は、人が当該セクションの印象を決定させる際、指定セクション数前までの印象の推移を印象の強さとして反映させている。

印象重み付け実現方法 1 では、セクション m の印象には、セクション 1 からセクション m までの差の合

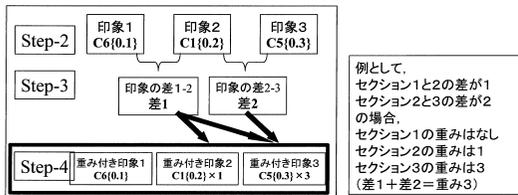


図 13 Step-4 印象重み付け実現方法 1: 当該セクションより前のすべての前後差を重みとして反映する例

Fig. 13 An example of the implementation of Step-4 Method 1.

計が重みとして与えられる。たとえば、セクション 5 の重みは、セクション 1 と 2 の差、セクション 2 と 3 の差、セクション 3 と 4 の差、セクション 4 と 5 の差のすべての合計である。印象重み付け実現方法 2 では、セクション m の印象には、指定セクション数 n として、セクション $m-n$ からセクション m までの差の合計が重みとして与えられる。たとえば、指定セクション数が 3 の場合、セクション 5 の重みは、セクション 5 から 3 を引いたセクション 2 までが対象であり、セクション 2 と 3 の差、セクション 3 と 4 の差、セクション 4 と 5 の差のすべての合計が重みとなる。

本実現方式では、図 13 において Step-4 に示すように、事例として印象重み付け実現方法 1: 当該セクションより前のすべての前後差を重みとして反映により、重みを反映させる方法を用いる。

たとえば、楽曲としてクラシックの単一楽器の楽曲を対象とした場合、Step-1 において、楽曲をセクションとして設定することができる。クラシックの単一楽器の楽曲のうち、楽章を持つ楽曲の楽章数は 3~4 であることが多いので、セクションを楽章に対応させた場合、セクション数は少なく、当該セクションより前のすべての前後差が当該セクションの印象に影響を与えたと考えることができる。本実現方式においては、比較的楽章数の少ない楽曲へ適用することを前提として、印象重み付け実現方法 1 を選択する。

印象重み付け実現方法 1, 2 以外に、印象の重みの反映方法として、以下の 2 方法が考えられる。

- 印象重み付け実現方法 3: 現在着目しているセクションが前のセクション群の印象に与える影響を重みとして反映させる方法
- 印象重み付け実現方法 4: 現在着目しているセクションが他のどのセクションからも影響を受けず、影響を重みとして反映させない方法

本実現方式は、前セクション群が、現在着目しているセクション (印象を決定する対象となっているセクション) の印象に与える影響を重みとして反映させる

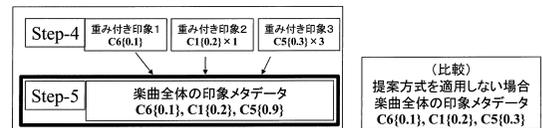


図 14 Step-5: 重み付けされたセクションごとの印象より、選択および合成により、楽曲全体の総合的な印象メタデータを生成

Fig. 14 An example of the implementation of Step-5.

方式であり、着目しているセクションが、前セクション群の印象に変化を与えるという場合は扱っていない。直観的には、初めて聞く楽曲について受ける印象をメタデータとして生成する場合に対応する方式である。過去に聞いた楽曲の印象や、過去に同じ楽曲を聞いた印象などが、現在聞いている印象に影響を与えない場合に対応する方式である。したがって、本実現方式は、印象重み付け実現方法 3 の場合には対応しないので、印象重み付け実現方法 4 との比較によって、その有効性を検証する。

3.1.5 Step-5 重み付けされたセクションごとの印象より、選択および合成により、楽曲全体の総合的な印象メタデータを生成

Step-4 において重み付けされたセクションごとの印象より、選択および合成する方法として、次の 3 方法を示す。

印象合成実現方法 1 最後のセクションの印象を楽曲全体の印象メタデータとする方法

印象合成実現方法 2 印象の重みに応じて、上位 m 個を抽出し、それらを楽曲全体の印象メタデータとする方法

印象合成実現方法 3 すべてのセクションの印象を、楽曲全体の印象メタデータとする方法

本実現方式では、図 14 に示すように、事例として印象合成実現方法 3: すべてのセクションの印象を楽曲全体の印象メタデータとし、楽曲全体の印象メタデータ生成を行う。対象楽曲は、Step-1 において楽章によるセクション設定をする。楽章数として 3~4 楽章を持つ楽曲が多く、それらではセクション数に応じた印象も少ないので、本実現方式では、楽章やクラシック楽曲の特性に対応すべく、印象合成実現方法 3: すべてのセクションの印象を、楽曲全体の印象メタデータを選択した。

3.2 (プロセス 2) 抽出された印象 (メタデータ) と問合せとの間の意味的連想検索

図 8 において示したように、音楽メディアデータを検索するプロセスのうち、第 2 のプロセスは、意味的連想検索方式^{5),6)}を用いて実現する。2 章において示した印象 (メタデータ) 生成方法により生成された

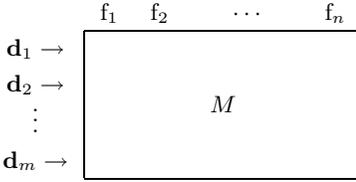


図 15 データ行列 M によるメタデータの表現

Fig. 15 Metadata represented in data matrix M .

印象 (メタデータ) を対象として, ここでは意味の数学モデルによる意味的連想検索方式^{5),6)}を適用する. この意味的連想検索方式を適用することにより, 検索者は, 独自の印象表現を表す単語列 (文脈) を与えると, それと意味的に相関の強い印象 (メタデータ) を持つ楽曲を検索することが可能となる. ここでは, その意味的連想検索方式について概説する. 本実験では, 文献 5), 6) に示された Longman 英英辞典⁸⁾による 2000 次元の意味空間を実現し, 本提案方式によって生成されたメタデータを用いて楽曲データ群をその空間上にマッピングし, 意味的連想検索を適用する. 詳細は, 文献 5), 6) に述べられている.

(1) メタデータ空間 MDS の設定

初めに, m 個の基本データについて各々 n 個の特徴 (f_1, f_2, \dots, f_n) を列挙した特徴付ベクトル d_i ($i = 1, \dots, m$) が与えられているものとし, そのベクトルを並べて構成する $m \times n$ 行列を M とおく (図 15) このとき, M は, 列ごとに 2 ノルムで正規化されている.

(a) データ行列 M の相関行列 $M^T M$ を計算する.

(b) $M^T M$ を固有値分解する.

$$M^T M = Q \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_\nu & \\ & & & 0 \dots 0 \end{pmatrix} Q^T,$$

$$0 \leq \nu \leq n.$$

ここで行列 Q は,

$$Q = (q_1, q_2, \dots, q_n)$$

である. この q_i ($i = 1, \dots, n$) は, 相関行列の正規化された固有ベクトル (以下, “意味素”) である. 相関行列の対称性から, この固有値はすべて実数であり, その固有ベクトルは互いに直交している.

(c) メタデータ空間 MDS を以下で定義する. 非ゼロ固有値に対応する固有ベクトル (以下, “意味素” と呼ぶ) によって形成される正規直交空間をメタデータ空間 MDS と定義する. この空間の次元 ν は, データ行列のランクに一致する. この空間は, ν 次元ユークリッド空間となる.

$$MDS := span(q_1, q_2, \dots, q_\nu).$$

$\{q_1, \dots, q_\nu\}$ は MDS の正規直交基底である.

(2) メディアデータのメディアデータベクトルの作成方式

ここでは, メディアデータを表現するメディアデータベクトルを形成する方法を示す.

(a) Step-1: メディアデータの特徴づけ

t 個の印象語 (あるいは, t 個のオブジェクト) o_1, o_2, \dots, o_t から成るメディアデータ P を次のように特徴付ける.

$$P = \{o_1, o_2, \dots, o_t\}.$$

ここで, 各印象語 o_i は, データ行列の特徴と同一の特徴を用いて表現される特徴付きベクトルである.

$$o_i = (o_{i1}, o_{i2}, \dots, o_{in})$$

(b) Step-2: メディアデータ P のベクトル表現
メディアデータ P を構成する t 個の印象語 o_1, o_2, \dots, o_t が, それぞれ n 次元のベクトルで定義されている. 印象語 o_1, o_2, \dots, o_t は, 合成することで n 次元ベクトル表現され, メディアデータベクトル p を形成し, メタデータ空間 MDS に写像される. これにより, 同じ空間上に言葉とメディアデータが配置されることになり, 言葉とメディアデータの関係を空間上の距離として動的に計算することが可能となる.

(3) メタデータ空間 MDS の部分空間 (意味空間) の選択

検索者は与える文脈を複数の単語を用いて表現する. 検索者が与える単語の集合をコンテキストと呼ぶ. このコンテキストを用いてメタデータ空間 MDS に各コンテキストに対応するベクトルを写像する. これらのベクトルは, メタデータ空間 MDS において合成され, 意味重心を表すベクトルが生成される. 意味重心から各軸への射影値を相関とし, 閾値を超えた相関値 (以下, 重み) を持つ軸からなる部分空間 (以下, 意味空間) が選択される.

(4) メタデータ空間 MDS の部分空間 (意味空間) における相関の定量化

選択されたメタデータ空間 MDS の部分空間 (意味空間) において, メディアデータベクトルのノルムを検索語列との相関として計量する. これにより, 与えられたコンテキストと各メディアデータとの相関の強さを定量化している. この意味空間における検索結果は, 各メディアデータを相関の強さについてソートしたリストとして与えられる.

表 1 Step-1: セクション設定
Table 1 Results of applying Step-1: Division of music into sections.

musicID	section1	section2	section3	section4	section5
1	1	2	3	4	null
2	1	2	null	null	null
3	1	2	3	null	null
4	1	2	3	null	null
5	1	2	3	null	null
6	1	2	3	4	5

表 2 Step-2: セクションごとの印象
Table 2 Results of applying Step-2: Impression extraction for each section.

musicID	sectionID	impression (相関量の大きい順, 中心的形容詞群 (最大値) は先頭に示す)
1	1	(0.943833)c7, (0.556886)c8, (0.344942)c6, (0.004992)c5
	2	(0.911841)c4, (0.607810)c3, (0.339127)c2
	3	(0.845277)c7, (0.640228)c8, (0.279776)c1, (0.047236)c6
	4	(0.893347)c7, (0.525898)c8, (0.324326)c6, (0.019427)c1, (0.003479)c5
2	1	(1.000000)c1, (0.389815)c8, (0.155606)c4, (0.036401)c2
	2	(0.964553)c6, (0.632214)c7, (0.541761)c5
3	1	(0.750157)c7, (0.517903)c2, (0.493212)c4, (0.352798)c3, (0.010716)c8
	2	(0.835024)c1, (0.374259)c6, (0.096786)c8
	3	(0.880076)c7, (0.023542)c2, (0.322126)c4, (0.197224)c8, (0.374259)c6
4	1	(0.565800)c6, (0.560122)c7, (0.411083)c8, (0.314871)c1, (0.293489)c5
	2	(0.823173)c1, (0.616680)c4, (0.030150)c8, (0.289698)c2, (0.283440)c3
	3	(0.349715)c5, (0.288381)c1, (0.274808)c6, (0.091652)c4, (0.024332)c8
5	1	(0.482665)c4, (0.359013)c1, (0.353201)c3, (0.189203)c2
	2	(0.855797)c1, (0.563867)c4, (0.409650)c2, (0.280966)c3, (0.075656)c8
	3	(1.000000)c1, (0.524677)c8, (0.165502)c7, (0.016474)c5
6	1	(0.487484)c3, (0.421902)c5, (0.183421)c4, (0.181398)c6
	2	(0.374532)c5, (0.319723)c3, (0.303257)c6, (0.032955)c7, (0.031589)c4
	3	(0.940428)c3, (0.719452)c4, (0.445776)c2, (0.009393)c1
	4	(0.864319)c4, (0.812407)c5, (0.659748)c3, (0.207848)c6
	5	(0.768683)c7, (0.497271)c6, (0.299690)c5, (0.010949)c8

4. 実 験

3章において示した実現方式に従って実験システムを構築し, 提案する印象メタデータ生成方式の実現可能性, 楽曲データ検索への適用可能性を検証するために, 次の2実験を行った.

実験 1 提案方式による印象メタデータ生成実験

実験 2 提案方式による印象メタデータ生成, および生成された印象メタデータを対象とした意味的連想検索適用実験

実験 1 は, 提案方式による印象メタデータ生成を行い, 提案方式の実現可能性の検証を行う.

実験 2 は, 提案方式により生成された印象メタデータを意味的連想検索へ適用し, 提案方式の実際の楽曲データ検索への応用における適用可能性の検証を行う.

4.1 実 験 1

実験 1 では, 2章で示した提案方式により, ベートーベンソナタ集 6 曲を対象に印象メタデータを生成した. ここでは, 3章に示した各 Step において, クラッシ

クピアノ曲に適した方法を選択し, 実現した.

Step-1 方法 1 楽章によるセクション設定を用い, 対象楽曲を, 楽章をセクションとして設定した. そのセクション設定を表 1 に示す.

Step-2 方法 1 楽曲メタデータ自動生成方式によるセクションごとの印象生成楽曲メタデータ自動生成方式により, セクションごとの印象を生成した. 楽曲メタデータ自動生成方式により生成される印象メタデータは, 楽曲データ (およびその各 section) と Hevner 形容詞群 (c_i) との相関量 (-1 から $+1$ を値の範囲とする) によって表される. 本実験では, 各楽曲データ (その各セクション) について, 楽曲データ (その各セクション) との相関量が 0 以上の値を持つ Hevner 形容詞群 (c_i) を楽曲データ (その各セクション) の印象メタデータとして抽出する.

そのように生成されたセクションごとの印象を表 2 に示す. 表の, musicID には楽曲番号を, sectionID には楽章の番号を, impression は Hevner

表3 Step-3: セクションごとの印象の前後差

Table 3 Results of applying Step-3: Computation of impression-strength of impression.

musicID	sectionID→sectionID	1→2	2→3	3→4	4→5
1	中心的形容詞群名→中心的形容詞群名	c7→c4	c4→c7	c7→c7	null
	印象の前後差	1	1	0	null
2	中心的形容詞群名→中心的形容詞群名	c1→c6	null	null	null
	印象の前後差	1	null	null	null
3	中心的形容詞群名→中心的形容詞群名	c7→c1	c1→c7	null	null
	印象の前後差	1	1	null	null
4	中心的形容詞群名→中心的形容詞群名	c6→c1	c1→c5	null	null
	印象の前後差	1	2	null	null
5	中心的形容詞群名→中心的形容詞群名	c4→c1	c1→c1	null	null
	印象の前後差	1	0	null	null
6	中心的形容詞群名→中心的形容詞群名	c3→c5	c5→c3	c3→c4	c4→c7
	印象の前後差	1	1	1	1

表4 Step-4: セクションごとの印象の前後差より、重みの算出

Table 4 Results of applying Step-4: Calculation of the degree of a gap.

musicID	sectionID	1	2	3	4	5
1	重みの算出内容	影響を受けない	1	1+1	1+1+0	null
	印象の重み	1	1	2	2	null
2	重みの算出内容	影響を受けない	1	null	null	null
	印象の重み	1	1	null	null	null
3	重みの算出内容	影響を受けない	1	1+1	null	null
	印象の重み	1	1	2	null	null
4	重みの算出内容	影響を受けない	1	1+2	null	null
	印象の重み	1	1	3	null	null
5	重みの算出内容	影響を受けない	1	1+0	null	null
	印象の重み	1	1	1	null	null
6	重みの算出内容	影響を受けない	1	1+1	1+1+1	1+1+1+1
	印象の重み	1	1	2	3	4

形容詞群名と括弧内にその相関量を表す。

Step-3 方法 Hevner 音楽形容詞群上での意味的な距離による、セクションごとの印象の時間的前後関係からの影響（強調・軽減）の計量

各セクションごとに、最大の相関量を持つ印象の形容詞群を求め、それを各楽章の中心的形容詞群とした。

初めのセクション（楽章）の中心的形容詞群と、次のセクション（楽章）の中心的形容詞群との差を計量した。この計量においては、時間的に隣り合うセクション（前段セクションと自セクション）間において、それらの中心的形容詞群が、Hevner 音楽形容詞群上において、逆の印象の形容詞群のときの重みを 2、同じ印象の形容詞群のときの重みを 0、それ以外の印象の形容詞群のときの重みを 1 として設定した。これらの重みは、Hevner 音楽形容詞群上において、印象変化が最大となったときに、自セクションにおける元々の印象の形容詞群の相関量の値を倍とすること、すなわち、前段セクションとの間の印象変化の波及により、

自セクションが元々計算された印象の 2 倍の強さの印象を与えるがことを示している。また、前段セクションと自セクション間で印象が変化しない場合には、自セクションの印象について、前段セクションとの間の印象変化が波及しないことを示している。

この方法で、最初のセクションから最後のセクション（最終楽章）までのセクション間の印象変化の計量を行った。生成された印象の前後差を表 3 に示す。この表において、musicID は楽曲番号、sectionID→sectionID は前後差の対象となる楽章の組、印象の前後差は計量した印象の前後差を表す。

Step-4 方法 1 自セクションより前のすべての前後差を重みとして反映による印象の設定

ここでは、自セクションより前のすべてのセクション間の印象の前後差を重みとして、自セクションの印象に反映させることにより、自セクション印象を設定する。初めのセクションの印象はそのまま保持し、2 番目以降のセクションの印象の重み

表5 Step-4: 印象の重みを反映させたセクションごとの印象

Table 5 Results of applying Step-4: Weighting impression according to impression-strength.

music ID	section ID	1	2	3	4	5
1	計算内容	$(0.943833 \times 1)c7$	$(0.911841 \times 1)c4$	$(0.845277 \times 2)c7$	$(0.893347 \times 2)c7$	null
	計算結果	$(0.943833)c7$	$(0.911841)c4$	$(1.690554)c7$	$(1.786694)c7$	null
2	計算内容	$(1.000000 \times 1)c1$	$(0.964553 \times 1)c6$	null	null	null
	計算結果	$(1.000000)c1$	$(0.964553)c6$	null	null	null
3	計算内容	$(0.750157 \times 1)c7$	$(0.835024 \times 1)c1$	$(0.880076 \times 2)c7$	null	null
	計算結果	$(0.750157)c7$	$(0.835024)c1$	$(1.775200)c7$	null	null
4	計算内容	$(0.565800 \times 1)c6$	$(0.823173 \times 1)c1$	$(0.349715 \times 3)c5$	null	null
	計算結果	$(0.565800)c6$	$(0.823173)c1$	$(1.049145)c5$	null	null
5	計算内容	$(0.482665 \times 1)c4$	$(0.855797 \times 1)c1$	$(1.000000 \times 1)c1$	null	null
	計算結果	$(0.482665)c4$	$(0.855797)c1$	$(1.000000)c1$	null	null
6	計算内容	$(0.487484 \times 1)c3$	$(0.374532 \times 1)c5$	$(0.940428 \times 2)c3$	$(0.864319 \times 3)c4$	$(0.768683 \times 4)c7$
	計算結果	$(0.487484)c3$	$(0.374532)c5$	$(1.880856)c3$	$(2.592957)c4$	$(3.074732)c7$

については、自セクションより前のセクション間の印象の前後差の合計を重みとして設定した。各セクションの重みを算出したデータを表4に示す。また、前後差を反映させた各セクションの印象を表5に示す。表のmusicIDは楽曲番号を示し、各sectionIDについて、元々計算された中心的形容詞群との相関量の計算内容、および、印象に関する前後差の重みを反映させた印象の強さを表す計算結果を示す。

Step-5 方法3 すべてのセクションの印象を対象に合成したものを楽曲全体の印象メタデータとするによる、楽曲全体の印象メタデータ生成。

Step-4により生成された各セクションの印象の強さについて、セクション間において形容詞群ごとに印象の強さを合計し(sum)、その合計値の最大値によって、各形容詞群の合計値を割ることにより、各形容詞群の印象の強さを0~1に正規化(normalization)し、その値を楽曲全体(全セクション)の印象メタデータとして抽出した。

生成された印象メタデータを表6に示す。この表において、musicIDは楽曲番号、functionは合計値sumおよび正規化normalizationを示し、impression metadataは、楽曲と相関の強い形容詞群に関する印象の強さ、すなわち、本方式における楽曲の印象メタデータを示す。

本方式との比較対象として、従来方式を次のように設定した。

従来方式

楽曲全体を対象とした印象メタデータ生成方式。セクション設定やセクションごとの印象生成を行うのではなく、楽曲全体から楽曲全体の印象メタデータを生成する方式。この方式では、セクショ

表6 Step-5: 楽曲全体の印象メタデータ

Table 6 Results of applying Step-5: Combining impression of adjusting themselves to a query structure.

musicID	function	impression metadata
1	sum	$c4(0.911841),c7(4.421081)$
	normalization	$c4(0.2062484),c7(1.0)$
2	sum	$c1(1.0),c6(0.964553)$
	normalization	$c1(1.0),c6(0.964553)$
3	sum	$c7(2.525357),c1(0.835024)$
	normalization	$c7(1.0),c1(0.3306558)$
4	sum	$c1(0.823173),c5(1.049145),c6(0.565800)$
	normalization	$c1(0.7846131),c5(1.0),c6(0.5392962)$
5	sum	$c1(1.855797),c4(0.482665)$
	normalization	$c1(1.0),c4(0.260085)$
6	sum	$c3(2.36834),c4(2.592957),c5(0.770259),c7(3.074732)$
	normalization	$c3(0.770259),c4(0.8433115),c5(0.1218096),c7(1.0)$

ン間の印象の差による印象の重みの反映はされない。提案方式により生成された印象メタデータと、従来方式により生成された印象メタデータの比較を、表7に示す。

4.2 実験2

実験2は、提案方式による印象メタデータの意味的連想検索への適用結果を示し、提案方式の実現可能性を検証した。実験2の対象データとして、次の3種類のメタデータ生成を行った。

対象メタデータ1 セクションを持つ楽曲6曲に対して、従来方式(印象の前後差による重みの反映なし)を適用した印象メタデータ。楽曲データとしては、実験1で用いた6曲と同曲であり、ベートーベンソナタ集から6曲選択した。

対象メタデータ2 セクションを持つ楽曲6曲(対象

表7 従来方式による印象メタデータ(印象の前後差による重みの反映なしの印象メタデータ)と、提案方式による印象メタデータ(印象の前後差による重みの反映ありの印象メタデータ)の比較

Table 7 Impression-metadata of the usual method and the proposal method.

musicID	従来方式	提案方式
1	(0.656532)c1 (0.059458)c4 (0.484544)c5 (0.303777)c6 (0.056466)c7 (0.271880)c8	(0.2062484)c4 (1.0000000)c7
2	(0.865085)c1 (0.280935)c3 (0.901062)c4 (0.198674)c5	(1.000000)c1 (0.964553)c6
3	(0.302013)c1 (0.345525)c4 (0.463276)c5 (0.229731)c6	(1.0000000)c7 (0.3306558)c1
4	(0.670740)c1 (0.017362)c2 (0.278432)c3 (0.726591)c4 (0.093318)c5	(0.7846131)c1 (1.0000000)c5 (0.5392962)c6
5	(0.571166)c1 (0.480767)c4 (0.398878)c5	(1.000000)c1 (0.260085)c4
6	(0.900959)c1 (0.030146)c2 (0.185254)c3 (0.737469)c4 (0.023023)c5	(0.7702590)c3 (0.8433115)c4 (0.1218096)c5 (1.0000000)c7

メタデータ 1 と同曲) に対して, 提案方式(印象の前後差による重みの反映あり)を適用した印象メタデータ. 実験 1 で用いた 6 曲と同曲であり, ベートーベンソナタ集から 6 曲選択した.

対象メタデータ 3 セクションを持たない楽曲 5 曲に対する(セクションがないため重みの反映はできない)印象メタデータ. 楽曲データとしては, ブルグミュラー 25 の練習曲より 5 曲選択した.

対象とした各楽曲と, 印象メタデータ生成の適用方式(従来方式,あるいは,提案方式)を表 8 に示す. 対象メタデータ 1 と対象メタデータ 2 の印象メタデータは, 実験 1 の表 7 に示しており, 対象メタデータ 3 は, 表 9 に示す.

実験 2.1 では, 従来方式(印象の前後差による重みの反映なし)により生成された“対象メタデータ 1”を対象として意味的連想検索に適用した場合, および, 提案方式(印象の前後差による重みの反映あり)により生成された“対象メタデータ 2”を対象とし, 同様に意味的連想検索に適用した場合における検索結果を比較した.

さらに, 音楽専門家である被験者 10 名を対象に行った同様の楽曲の印象に関するアンケート結果との比較を行い, 提案方式によって得られた検索結果の検証を行った. 被験者選択について, 本実験では, 被験者間での楽曲に対する知識の差を少なくするため, また, 過去に該当楽曲を聞いた際の印象を考慮せずに印象抽出作業を行える技術を必要とするために, 音楽演奏活

表 8 楽曲名と印象メタデータ生成の適用方式

Table 8 Titles of music and its applied methods to extract impression-metadata.

ID	楽曲名	musicID	楽章	適用方式
メ タ 対 象 タ 1	ソナタ 1 番	usual1	複数	従来
	ソナタ 13 番	usual2	複数	従来
	ソナタ 14 番	usual3	複数	従来
	ソナタ 17 番	usual4	複数	従来
	ソナタ 23 番	usual5	複数	従来
	ソナタ 29 番	usual6	複数	従来
メ タ 対 象 タ 2	ソナタ 1 番	proposal1	複数	提案
	ソナタ 13 番	proposal2	複数	提案
	ソナタ 14 番	proposal3	複数	提案
	ソナタ 17 番	proposal4	複数	提案
	ソナタ 23 番	proposal5	複数	提案
	ソナタ 29 番	proposal6	複数	提案
メ 対 象 タ 3	op.100-5	single1	単一	提案
	op.100-24	single2	単一	提案
	op.100-20	single3	単一	提案
	op.100-25	single4	単一	提案
	op.100-15	single5	単一	提案

表 9 対象メタデータ 3 の印象メタデータ

Table 9 Impression-metadata of target data 3.

musicID	impression metadata	
Single1	(0.533519)c5 (0.962478)c6 (0.669144)c7 (0.119148)c8	
	Single2	(0.019427)c1 (0.003479)c5 (0.324326)c6 (0.893347)c7 (0.525898)c8
		Single3
Single4	(0.541761)c5 (0.964553)c6 (0.632214)c7	
Single5	(1.000000)c7 (0.499072)c8	

動を行う専門家を対象とした実験を行った. これにより, 無作為に被験者を選択する場合と比べ, 被験者間および 1 被験者内での楽曲内および楽曲間における印象指定のぶれを軽減することが可能となることを想定した.

実験 2.2 では, “対象メタデータ 1”を持つ楽曲群と“対象メタデータ 3”を持つ楽曲群(“対象メタデータ 1”+“対象メタデータ 3”)を対象とした意味的連想検索を行った. また, “対象メタデータ 2”を持つ楽曲群と“対象メタデータ 3”を持つ楽曲群(“対象メタデータ 2”+“対象メタデータ 3”)とを対象とした意味的連想検索を行った. さらに, 専門家を対象に行ったアンケート(被験者 10 名)との比較を行った.

“対象メタデータ 3”の抽出対象として用いた単一のセクションから成る楽曲群については, 楽曲全体を通じて単一に近い印象から成っているため, 複数の部分の印象から成る楽曲に比べて, 本方式により生成された印象メタデータは, よりシャープにそれらの楽曲の印象を表しているものと考えられる.

本実験では、これら単一のセクションから成る楽曲の印象メタデータと、“対象メタデータ2”の抽出対象として用いた複数のセクションから成る楽曲の印象メタデータを対象とした検索結果を比較し、複数のセクションから成る楽曲の印象メタデータ生成方式としての提案方式の有効性を検証する。

各実験対象楽曲について、10名の専門家の被験者を対象としたアンケートは、楽曲を各被験者に聞かせた後、その楽曲の印象を総合的に表す形容詞群を Hevner 音楽形容詞群の中から選択させる方法、すなわち、被験者が、その楽曲の印象として最も合致した形容詞群を、形容詞群のリスト図9から選ぶ方法によって行った。

被験者にとっては、各楽曲について、印象として合致する形容詞群の順に、形容詞群をランキングすることはきわめて困難であるため、各楽曲について、印象として最も合致する1形容詞群の選択を求めた。各楽曲についての10名分の結果から、各楽曲に対し、最も多くの被験者が選択した形容詞群を、アンケート結果として、その楽曲の印象として合致する形容詞群とした。

このアンケートは、本提案方式によって生成した印象メタデータを対象として行った検索の結果として得られた楽曲ランキングとアンケート結果との対応により、本提案方式の有効性を判定するために行ったものである。

アンケート結果は、提案方式により生成された印象メタデータが検索において有効に機能するかを検証するためであり、検索結果の上位ランキングの楽曲とアンケート結果が合致する場合に、本方式の有効性が認められることになる。

4.2.1 実験2.1の結果

対象メタデータ1、対象メタデータ2を対象とし、検索語としてc7を用いた場合における意味的連想検索結果を表10、表11に示す。これらの表において、rankは検索語と楽曲間の相関量によってソートした順位、musicIDは楽曲の番号、normは検索語であるc7形容詞群との相関量を示す。

本提案方式のStep-2によって抽出される各楽曲の各セクションの印象は表2のようになり、c7は各楽曲(musicID 1・2・3・4・5・6)のセクションに出現している。これにより、各楽曲について、c7との間で相関量が計量されることになる。また、他の検索語に適用した場合の結果を例として表12、表13に示す。

4.2.2 実験2.2の結果

“対象メタデータ1”を持つ楽曲群と“対象メタデー

表10 実験2.1結果 対象メタデータ1(従来方式)による印象メタデータを対象とした意味的連想検索結果: 検索語c7

Table 10 Experiment 2.1: Results of applying semantic associative search to target data 1 (query: c7).

rank	musicID	norm
1	usual1	0.281198
2	usual2	0.189557
3	usual3	0.187106 (アンケート結果)
4	usual5	0.187106
5	usual4	0.179305
6	usual6	0.178416

表11 実験2.1結果 対象メタデータ2(提案方式)による印象メタデータを対象とした意味的連想検索結果: 検索語c7

Table 11 Experiment 2.1: Results of applying semantic associative search to target data 2 (query: c7).

rank	musicID	norm
1	proposal3	0.468157 (アンケート結果)
2	proposal1	0.392431
3	proposal6	0.303124
4	proposal5	0.215250
5	proposal2	0.189799
6	proposal4	0.187244

表12 実験2.1結果 対象メタデータ1(従来方式)による印象メタデータを対象とした意味的連想検索結果: 検索語c1

Table 12 Experiment 2.1: Results of applying semantic associative search to target data 1 (query: c1).

rank	musicID	norm
1	usual2	0.696650
2	usual5	0.696650 (アンケート結果)
3	usual3	0.694899
4	usual4	0.693666
5	usual6	0.693666
6	usual1	0.693131

表13 実験2.1結果 対象メタデータ2(提案方式)による印象メタデータを対象とした意味的連想検索結果: 検索語C1

Table 13 Experiment 2.1: Results of applying semantic associative search to target data 2 (query: c1).

rank	musicID	norm
1	proposal5	0.705471 (アンケート結果)
2	proposal1	0.703330
3	proposal3	0.702067
4	proposal6	0.695215
5	proposal4	0.694117
6	proposal2	0.690877

タ3”を持つ楽曲群(“対象メタデータ1”+“対象メタデータ3”)について、検索語としてc7を用いた場合の意味的連想検索結果を表14に示す。

“対象メタデータ2”を持つ楽曲群と“対象メタデータ3”を持つ楽曲群(“対象メタデータ2”+“対象メタデータ3”)について、検索語としてc7を用いた場合

表 14 対象メタデータ 1 (従来方式) + 対象メタデータ 3 (単一セクション楽曲) の検索結果: 検索語 c7

Table 14 Experiment 2.2: Results of applying semantic associative search to target data 1 and 3 (query: c7).

rank	musicID	norm
1	single5	0.468157
2	single2	0.294786
3	single3	0.294786
4	single4	0.293734
5	single1	0.293734
6	usual1	0.281198
7	usual2	0.189557
8	usual3	0.187106 (アンケート結果)
9	usual5	0.187106
10	usual4	0.179305
11	usual6	0.178416

表 15 対象メタデータ 2 (提案方式) + 対象メタデータ 3 (単一セクション楽曲) の検索結果: 検索語 c7

Table 15 Experiment 2.2: Results of applying semantic associative search to target data 2 and 3 (query: c7).

rank	musicID	norm
1	proposal3	0.468157 (アンケート結果)
1	single5	0.468157
3	proposal1	0.392431
4	proposal6	0.303124
5	single2	0.294786
6	single3	0.294786
7	single4	0.293734
8	single1	0.293734
9	proposal5	0.215250
10	proposal2	0.189799
11	proposal4	0.187244

の意味的連想検索結果を表 15 に示す。

これらの表において, rank は検索語と楽曲間の相関量によってソートした順位, musicID は楽曲の番号, single は対象メタデータ 3 の楽曲 (単一セクション楽曲), usual は対象メタデータ 1 の楽曲 (従来方式), proposal は対象メタデータ 2 の楽曲 (提案方式) を示す. norm は検索語である c7 形容詞群との相関量を示す。

4.3 考 察

実験 1 の結果より, 6 曲の各対象楽曲データを対象として, 提案方式は, 印象の変化を重みとして反映させた印象メタデータを生成を実現できたことが分かる. 実験 1 では, Step-1 ~ Step-5 までのプロセスを順に表 1 ~ 表 6 で示している. 表 7 では, 提案方式 (印象の前後差による重みの反映あり) による印象メタデータと, 従来方式 (印象の前後差による重みの反映なし) による印象メタデータの比較において, 印象変化の大きい楽曲, つまり印象の差の大きい楽曲において, 提

案方式では, 印象メタデータの重みが大きくなっていることが分かる.

実験 2.1 では, 提案方式により生成された印象メタデータを対象に意味的連想検索を適用した場合において, アンケート結果により正解とされた楽曲が, より上位に検索されている. 表 10, 表 11 において, 提案方式, および, 従来方式により得られた印象メタデータを対象とした意味的連想検索の結果を示した. 従来方式による印象メタデータを対象とした検索の結果では, アンケート結果により正解とされた楽曲が rank3 にあり, また, 提案方式による印象メタデータを対象とした検索の結果では, アンケート結果により正解とされた楽曲が rank1 にあり, 本提案方式により, より正しく楽曲が検索されることが分かる (アンケートでは, 被験者には, 各楽曲において最も相関が高いと思う Hevner 音楽形容詞群名 (c1 ~ c8) の選択を求めた. たとえば, 検索語 c7 における検索結果の楽曲ランキングにおいて, アンケート結果として, Hevner 音楽形容詞群名 (c1 ~ c8) から c7 が選ばれた楽曲が検索結果ランキングの上位にあれば, その楽曲に割り当てられた印象メタデータは正当性が高いと判断できる). 実験 2.1 の結果より, 従来方式との比較において, 印象の前後差を重みとして反映させた本提案方式の印象メタデータの生成が有効であることが分かる.

実験 2.2 では, 複数セクション (複数楽章) 楽曲と単一セクション (単一楽章) 楽曲の両者を対象として印象メタデータを生成し, それらを対象に意味的連想検索を適用した. 実験 2.2 における従来方式との比較において, 提案方式 (印象の前後差による重みの反映あり) により生成された印象メタデータを対象とした場合, アンケート結果として示された正解楽曲がより上位に検索された. 表 14, 表 15 において, 従来方式, および, 提案方式により得られた印象メタデータ, および, 単一楽章により構成される楽曲の印象メタデータを対象とした意味的連想検索の結果を示した.

従来方式 (印象の前後差による重みの反映なし) による印象メタデータを用いた場合の検索結果では, アンケートによる正解楽曲は rank8 にある. これに対して, 提案方式 (印象の前後差による重みの反映あり) を用いた場合の検索結果では, アンケートによる正解楽曲は rank1 にある.

単一楽章楽曲は, 複数楽章楽曲と比べて, ストレートに印象を表しているため, 印象の数が少なく, 各印象の重みが大きい.

これに対し, 従来方式による複数楽章楽曲の印象メタデータは, 印象の数が多く, 各印象の重みが小さく

なってしまうため、両者を対象とした検索の結果、単一楽章楽曲 (single) が複数楽章楽曲 (usual) に比べて、上位にランキングされる傾向にあり、アンケート結果 (usual3) が、rank8 になっているものと考えられる。

提案方式による複数楽章楽曲の印象メタデータは、印象の前後関係から生じる重みを波及させているため、印象が強調されるセクションに対応する印象の重みが大きくなり、単一楽章楽曲 (single) が複数楽章楽曲 (usual) が混在してランキングされる傾向が生まれる。

検索結果において、本提案方式により、アンケートによる正解 (proposal3) が上位に検索されていることを示している。

これらの結果より、印象の時間的推移をとまなう音楽メディアデータの印象メタデータ生成方式により生成される印象メタデータが、意味的連想検索機構において、有効に機能しており、提案方式の意味的連想検索機構への適用可能性、適用の有効性を示している。

本実験は、提案方式の本質的な特徴である section 間単位での印象の前後差の反映を実験システムが実現できていることを検証すること、およびその特徴が具体的にどのような検索結果として表れるかを示すために行ったものである。

本実験により、本論文において提案した印象の前後関係から生じた重みを波及させた印象メタデータの生成方式、および、生成された印象メタデータを対象とした意味的連想検索が、印象の時間的推移をとまなう音楽メディアデータの検索において有効に機能する可能性を検証することができた。より多くの検索対象による有効性評価については、今後の課題とする。

5. ま と め

本論文では、楽曲を時間軸によってセクションに区切り、各セクションごとの印象を抽出し、さらに印象の前後の差を抽出し、それを後のセクションの印象に重みとして波及させる印象メタデータ生成方式を提案した。

また、本方式により印象の前後差により生じた重みを波及させた印象メタデータ生成を行う実現方式を示した。

さらに、本方式により生成される印象メタデータを対象として、意味的連想検索を適用することにより、検索者が、印象などの感性的な表現を言葉によって与えることによって、その印象に相関の強い楽曲を検索する方法を示し、印象メタデータの従来方式と比較により、本方式の有効性を示した。

今後の課題として、楽曲のセクションへの分割の自動化があげられる。本論文における実験では、楽曲形式・楽章を単位として分割を行ったが、両者とも、楽曲中での特徴により分割されたセクションである。この特徴を自動的に抽出することにより、セクションへの分割の自動化が可能となるもの考えられる。

また、本提案方式の、クラシック音楽における最適化、および、他ジャンルへの適用が考えられる。

本実験において、中間生成される各セクションの印象群は、時間的に連続して推移する。このように時間的に推移する印象に対して、その推移をとまなう問合せを対象とした検索の実現も可能と考えられる。さらに、提案方式は、さまざまな時間的変化をするメディアデータに適用可能な方式であり、他連続メディアデータ (映像、動画、文章) に応用することも考えられ、これを今後の研究の課題とする。

謝辞 本論文の方式設計、実験にあたり、貴重なご助言をいただいた石橋直樹氏 (慶應義塾大学政策・メディア研究科)、中神康裕氏 (慶應義塾大学 SFC 研究所) に感謝いたします。また、本研究の一部は、日本学術振興会学術創成研究プロジェクト「人文社会科学と自然科学を連携するメタレベル知識ベースシステムの開発」によるものである。ここに記して謝意を表します。

参 考 文 献

- 1) Hevner, K.: Experimental Studies of the Elements of Expression in Music, *American Journal of Psychology*, Vol.48, pp.246-268 (1936).
- 2) Hevner, K.: Expression in Music: A Discussion of Experimental Studies and Theories, *Psychological Review*, Vol.42, pp.186-204 (1935).
- 3) 伊地智麻子, 清木 康: 発想標語を用いたクラシック音楽メタデータ生成による意味的連想検索方式, データ工学ワークショップ (DEWS2001) (2001).
- 4) Kitagawa, T. and Kiyoki, Y.: Fundamental Framework for Media Data Retrieval System Using Media Lexco Transformation Operator, *Information Modeling and Knowledge Bases*, IOS Press (2000).
- 5) 清木 康, 金子昌史, 北川高嗣: 意味の数学モデルによる画像データベース検索方式とその学習機構, 電子情報通信学会論文誌, Vol.J79-D, No.4, pp.509-519 (1996).
- 6) Kiyoki, Y., Kitagawa, T. and Hayama, T.: A Metadatabase System for Semantic Image Search by a Mathematical Model of Meaning, *Multimedia Data Management Using Metadata*

to Integrate and Apply Digital Media, Chapter 7, McGrawHill (1998).

- 7) 小杉尚子, 西原祐一, 細谷精一, 山室雅司, 串間和彦: ハミングを用いた音楽検索システム, 情報処理学会研究報告: データベースシステム, 119-9, pp.49-54 (1999).
- 8) Longman Dictionary of Contemporary English, Longman (1987).
- 9) 小川 潤, 佐藤 聡, 北上 始: 感情に基づく音楽作品のための類似度計算方式, アドバンスデータベースシンポジウム (DBWeb2000), pp.229-234 (2000).
- 10) 高木秀幸: 楽曲メディア・データベースを対象とした意味的連想検索に関する研究, 筑波大学大学院修士課程理工学研究科修士論文 (1998).
- 11) 吉野太智, 高木秀行, 清木 康, 北川高嗣: 楽曲データを対象としたメタデータの自動生成とその意味的連想検索への適用, 情報処理学会研究報告: データベースシステム, 98-DBS-116(2), pp.109-116 (1998).

(平成 15 年 3 月 25 日受付)

(平成 15 年 10 月 7 日採録)

(担当編集委員 河野 浩之)



伊地智麻子 (学生会員)

1979 年生. 2002 年慶應義塾大学環境情報学部卒業. 同大学院政策・メディア研究科修士課程に在学. データベースシステム, 感性情報システムの研究に従事.



清木 康 (正会員)

1978 年慶應義塾大学工学部電気工学科卒業. 1983 年同大学大学院工学研究科博士課程修了. 工学博士. 同年日本電信電話公社武蔵野電気通信研究所入所. 1984 年~1995 年筑波大学電子・情報工学系講師, 助教授を経て, 1996 年慶應義塾大学環境情報学部助教授, 1998 年同大学教授. データベースシステム, 知識ベースシステム, マルチメディアシステムの研究に従事. ACM, IEEE-CS, 電子情報通信学会, 日本データベース学会各会員.