

## 音楽データベース群を対象とした印象メタデータの自動抽出方式

石橋 直樹<sup>†</sup> 清木 康<sup>††</sup>

<sup>†</sup> 慶應義塾大学大学院政策・メディア研究科 〒252-8520 神奈川県藤沢市遠藤 5322

<sup>††</sup> 慶應義塾大学環境情報学部 〒252-8520 神奈川県藤沢市遠藤 5322

E-mail: †{naoki,kiyoki}@mdbl.sfc.keio.ac.jp

あらまし 日々増加する高域ネットワーク上の音楽データを対象とした検索方式として、メロディを入力として用いた直接的な検索方式、また、曲名、作曲者名等といった楽曲情報を用いた検索方式などが提案されている。これらの方式は、それぞれ、既聴楽曲の検索、または、既知の楽曲情報を用いた検索を実現する方法論として有効だが、高域ネットワークにおいて急増する未知の楽曲を対象とした検索の実現は困難である。これら未知の楽曲を対象とした検索において、楽曲の与える印象に応じた検索等、楽曲と利用者の主観を結びつける検索方式の実現が望まれる。本論文では、高域ネットワークに散在する未知楽曲を対象とした検索を実現するために、楽曲の与える印象を表すメタデータの自動抽出方式を提案する。提案方式は、複数の楽器を用いた楽曲データから、楽器毎の印象を表す形容詞群、および、それらと楽曲の相関量を動的に抽出し、また、楽器毎の印象を合成することで、楽曲のメタデータとして動的に定義する。本論文では、提案方式、および、主旋律のみを用いた方法論を比較することで、提案方式の実現可能性を示す。

キーワード 音楽データベース、マルチメディアデータベース、メタデータ

## An Automatic Extraction Method for Impressionistic Metadata of Music Databases

Naoki ISHIBASHI<sup>†</sup> and Yasushi KIYOKI<sup>††</sup>

<sup>†</sup> Graduate School of Media and Governance, Keio University 5322, Endo, Fujisawa, Kanagawa, 252-8520 Japan

<sup>††</sup> Faculty of Environmental Information, Keio University 5322, Endo, Fujisawa, Kanagawa, 252-8520 Japan

E-mail: †{naoki,kiyoki}@mdbl.sfc.keio.ac.jp

**Abstract** For querying music data that drastically increase on the global area network, number of methods has been proposed. For examples, the methods which directly match melodies, and the methods that apply pattern-matching to the abstract information like names of songs or authors have been proposed. These methods are commonly used for executing queries for the songs that the user has already heard or knows the abstract information, though it is difficult to search when the user does not know any information. For realizing such queries to the songs which user has never heard, it is effective to realize connectivity between users and music according to the subjective of users such as impression. In this paper, we propose an automatic extraction method of metadata that represent impression for music data on the global area network. The proposed method dynamically extracts adjectives and correlations according to each instrument of a datum, and it dynamically integrates the results of each instrument to generate the metadata of the datum. We clarify feasibility of the proposed method by comparing to a method that evaluates the main instrument.

**Key words** Music Database, Multimedia Database, Metadata

## 1. はじめに

近年、ネットワークの広域化、高速化、記憶媒体の大容量化などを背景に、多数の音楽データベース群が広域ネットワークに散在している。

日々増加する音楽データベース群を対象に、音符の配列自体を直接検索する方式 [9] [13] (以下 A 方式と記す)、楽曲名、作曲者名などの楽曲情報を静的に定義する [5] [6] ことで楽曲検索を実現する方式 (以下 B 方式と記す) などが提案されている。A 方式は、既聴楽曲の検索を目的としているため、利用者の興味に応じた未聴楽曲の検索の実現が困難である。B 方式は、A 方式より抽象的なデータを扱い楽曲を検索する手法として位置づけられ、作曲者、演奏者、曲名など、楽曲をあらわす具体的な抽象情報を検索対象のデータ構造、および、検索語に要求する。このため、抽象情報が不明である楽曲の検索の実現が困難である。

このような背景に応じて、我々の研究プロジェクトでは、楽曲の与える印象を用いた楽曲検索の実現を研究課題に定め、ピアノ曲の印象に対応するメタデータの自動抽出方式、および、その意味的連想検索方式 [8] [14]、発想標語を用いたクラシック音楽の意味的連想検索方式 [4] を提案した。これらの方式は、楽曲が聴者に与える心理的影響を『印象』、印象を表す単語を『印象語』、また、印象語、および、その重みにより定義されたメタデータを『印象メタデータ』と定義し、その印象メタデータを用いた楽曲検索を実現することで、利用者の視点・興味に応じた未聴楽曲の検索を実現する方式として位置づけられる。

文献 [8] [14] において示した印象メタデータの自動抽出方式は、クラシック音楽の類型におけるピアノ曲を対象として、楽曲データから、印象メタデータを自動的に抽出する。ここで対象とされるピアノ曲は、原則的に単一のピアノ奏者による演奏を前提としていたため、同印象メタデータの自動抽出方式は、単一楽器から構成される楽曲のみを対象としていた。このため、文献 [8] [14] において示した印象メタデータの自動抽出方式 (以下単にシングルトラック方式と記す) は、多重奏曲、ポップスなど、複数の楽器から構成される楽曲群への適用が困難である。以下、楽曲を構成する各楽器に対応する音符配列、または、可聴波を、『トラック』と定義し、以下参照する。

本論文では、複数の楽器から構成される楽曲を対象とした、印象メタデータの自動抽出方式を提案する。

本論文で提案する印象メタデータの自動抽出方式の本質は、各楽器毎に、重み付きの形容詞で表現された印象メタデータを自動抽出し、重み付き形容詞の表現で自動生成された印象メタデータの合成により、楽曲に対応する印象メタデータの自動抽出を実現している点にある。具体的に、本方式は、1) 楽曲データ群を楽器毎に分割し、2) 楽器毎に分割された楽曲データ (以下単にトラックと記す) を対象として印象メタデータを動的に抽出し、3) トラック毎に自動抽出された印象メタデータを動的に合成する

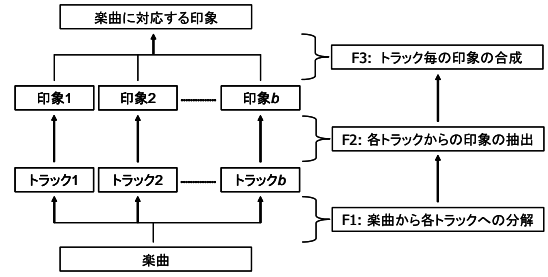


図 1 提案方式概要

Fig. 1 An Overview of the Proposed Method

ことで、楽曲に対応する印象メタデータを自動的に抽出する。

本論文では、本方式、および、主旋律のみを対象としてシングルトラック方式を適用し、自動抽出されたメタデータそれぞれの相違を評価することで、本方式の実現可能性を示す。

## 2. 印象メタデータの自動生成方式

複数楽器から構成される楽曲を対象とした印象メタデータの自動生成方式は、楽曲分割機能 (F1)、印象メタデータ抽出機能 (F2)、印象メタデータ合成機能 (F3) の 3 機能から構成される。F1 は、楽曲データを、トラック毎の楽曲データへ分割する。F2 は、トラック毎の楽曲データから、形容詞、および、その形容詞との相関量から構成される印象メタデータを抽出する。F3 は、トラック毎に抽出された印象メタデータを合成し、楽曲に対応する印象メタデータを生成する。これらの 3 機能は、まず楽曲を対象に F1 を適用し、以降 F2, F3 の順で適用される (図 1)。

### 2.1 楽曲分割機能

$$(F1) \quad f_{divi}(M_n) \rightarrow \{T_{[n,1]}, \dots, T_{[n,b]}\}$$

F1 は、楽曲  $M_n$  を、 $b$  個のトラック  $T_{[n,m]}$  に分解する。 $n$  は楽曲の識別子、 $m$  はトラックの識別子を表す。具体的には、波形データの混声音の分離 [7], [11], Standard MIDI File からのトラック別データ抽出などが相当する。

### 2.2 印象メタデータ抽出機能

$$(F2) \quad f_{extr}(T_{[n,m]}) \rightarrow \{\{A_{[n,m,1]}, W_{[n,m,1]}\}, \dots, \{A_{[n,m,c]}, W_{[n,m,c]}\}\}$$

F2 は、トラック  $T_{[n,m]}$  を印象メタデータの表現に変換する機能で、 $c$  個の形容詞  $A_{[n,m,o]}$ 、および、その重み  $W_{[n,m,o]}$  を出力する。 $o$  は形容詞の識別子を表す。これはすなわち、音楽心理学の研究成果などを用いて、トラックのデータを形容詞、および、形容詞との相関量の形式へ変換する過程に対応する。例えば、[14] [10] などの方式は、それぞれ、[1] [2] [3]、および、[12] を用いて、単一楽器で表現される音楽データを、形容詞群、および、その形容詞への相関量の構造へ変換する。

### 2.3 印象メタデータ合成機能

$$(F3) \quad f_{comp}(\{\{A_{[n_1,1]}, W_{[n_1,1]}\}, \dots, \{A_{[n_b,c]}, W_{[n_b,c]}\}\}) \rightarrow$$

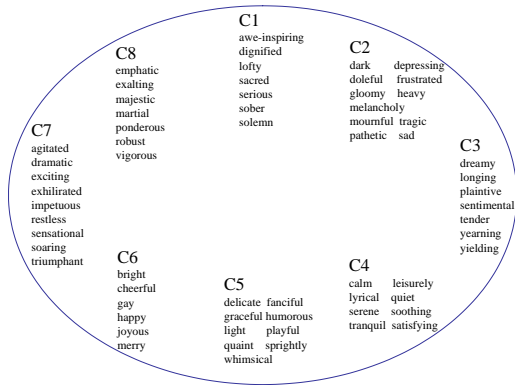


図 2 Hevner により定義された形容詞群

Fig. 2 Adjectives Defined by Hevner

$$\{\{A_{[n,1]}, W_{[n,1]}\}, \dots, \{A_{[n,c]}, W_{[n,c]}\}\}$$

F3 は、全トラックの印象メタデータ群を合成する関数で、 $c$  個の形容詞  $A_{[n,o]}$ 、および、その形容詞との相関量  $W_{[n,o]}$  を出力する。この機能を用いることで、単一のトラックのみから構成されていた楽曲を対象としていた心理学研究の成果を応用して、交響楽、ポップス、ロックなどの、複数トラックから構成される楽曲データを対象とした印象メタデータの抽出が実現可能となる。

### 3. 印象メタデータ抽出機構の実現方式

複数のトラックから構成される楽曲データを対象とした印象メタデータの抽出システムを実現するため、楽曲分割機構、印象メタデータ抽出機構、印象メタデータ合成機構の 3 機構を実現した。これらの機構は、Standard MIDI File(以下単に SMF と記す) を対象楽曲データの構造として使い、実装を行った。

#### 3.1 楽曲分割機構

任意個数のトラックから構成された SMF を対象として、各トラックに対応する音符配列を取り出す楽曲分割機構を実現した。波形データで構成される音楽データを対象とした分割方式としては、文献 [7] [11] 等有効な方法論が提案されているが、本研究の実装では対象としていない。

#### 3.2 印象メタデータ抽出機構

楽曲分割機構により生成されたトラック毎 SMF 群を対象として、Hevner の研究 [1] [2] [3] を用いたシングルトラック方式 [8] [14] を拡張した印象メタデータの抽出機構を実現した。

Hevner の研究では、楽曲の構造を規定する要素 (以下楽曲構造要素と記す) として、調性 (key)、テンポ (tempo)、音高 (pitch)、リズム (rhythm)、和声 (harmony)、旋律 (melody) の 6 つに着目し、この 6 つの楽曲構造要素と、8 つの印象語群 (図 2) によって表現される印象との相関関係を検証した。8 つの印象語群は、図のように、印象語間で類似性があるものを隣接するよう、円形に配置された。Hevner は、楽曲構造要素と、各印象語群によって表現さ

表 1 Hevner による各印象語群に対する楽曲構造要素の相対重要性

Table 1 The Relative Importances of the Musical Structure Elements Corresponding to the Impressionist Words Defined by Hevner

	c1	c2	c3	c4	c5	c6	c7	c8
key	長調 14	短調 20	短調 12	長調 3	長調 21	長調 24	—	—
tempo	速い 14	速い 12	速い 16	速い 20	速い 6	速い 20	速い 21	速い 6
pitch	低い 10	低い 19	高い 6	高い 8	高い 16	高い 6	高い 9	低い 13
rhythm	固定 18	固定 3	流動 9	流動 2	流動 8	流動 10	固定 2	固定 10
harmony	単純 3	複雑 7	単純 4	単純 10	単純 12	単純 16	複雑 14	複雑 8
melody	上昇 4	—	—	上昇 3	下降 3	—	下降 7	下降 8

表 2 変換行列  $C$

Table 2 The Conversion Table  $C$

	Key'	Tempo'	Pitch'	Rhythm'	Harmony'	Melody'
C1	4	-14	-10	18	3	4
C2	-20	-12	-19	3	-7	0
C3	-12	-16	6	-9	4	0
C4	3	-20	8	-2	10	3
C5	21	6	16	-8	12	-3
C6	24	20	6	-10	16	0
C7	0	21	-9	2	-14	-7
C8	0	6	-13	10	-8	-8

れる印象との相関関係として、印象語群によって表現される印象との相関関係を調べ、印象語群によって表現される 8 つの印象に対する各楽曲構造要素の相対重要性を示した (表 2)。

以上の Hevner の研究を用いたシングルトラック方式 [8] [14] では、次の 3 つの処理を行うことで、印象メタデータの抽出を実現した。

(Step-1) 楽曲構造基礎データ抽出

Step-1 では、SMF から、楽音データ (Note)、和音データ (Chord)、調名データ (kn)、演奏時間データ (pt)、拍子データ (mt)、テンポデータ (tem)、平均音高値 (mnap)、固定/流動リズム値 (unac/ac)、単純/複雑/ノーハーモニー値 (tc/oc/nc)、上昇/下降/水平メロディ値 (um/dm/nm) の各データを抽出する。

(Step-2) 楽曲構造要素量解析

更に、Step-1 で得られた楽曲構造基礎データを、Hevner の定義した楽曲構造要素に写像するため、楽曲構造要素量の解析を適用する。この解析により、楽曲  $M_n$  に対応し、調性値 (key)、テンポ値 (tempo)、音高値 (pitch)、リズム値 (rhythm)、和声値 (harmony)、旋律値 (melody) の 6 値から構成される楽曲構造要素ベクトルを生成する。楽曲データ  $M_n$  のトラック  $T_{[n,m]}$  に対応する楽曲構造要素ベクトル  $MFV_{T_{[n,m]}}$  は、式 (1) のように表される。

$$MFV_{T_{[n,m]}} = (key_{T_{[n,m]}}, tempo_{T_{[n,m]}}, pitch_{T_{[n,m]}}, rhythm_{T_{[n,m]}}, harmony_{T_{[n,m]}}, melody_{T_{[n,m]}}) \quad (1)$$

(Step-3) 印象メタデータ抽出

Step-3 では、 $MFV_{T_{[n,m]}}$ 、および、Hevner の示した各印象に対する楽曲構造要素の相対重要性 (表 1) を応用した変換行列  $C$  (表 2) を用いて、式 (2) で表される印象語

群ベクトル  $MCV_{T_{[n,m]}}$  を生成する。  $mcv$  は、トラック  $T_{[n,m]}$  と各印象語群との相関量を表す。

$$MCV_{T_{[n,m]}} = (mcv_{T_{[n,m]1}}, \dots, mcv_{T_{[n,m]8}}) \quad (2)$$

### 3.3 印象メタデータ合成機構

楽曲分割機構、印象メタデータ抽出機構を適用して得られたトラック毎の印象メタデータを対象に、印象メタデータ合成機構は、トラック毎の重要度を反映させ、重要度に応じた印象メタデータの合成を行うことで、楽曲に対応する印象メタデータを生成する。

(Step-A) トラック重要度の反映

トラックの重要度は、音量の平均が大きいトラックほど重要度は高く、音量の小さいトラックほど重要度は低いと考え、また、音高の平均が高いトラックほど重要度は高く、音高が低いトラックほど重要度は低いと考えられることから、以下の3種類の要素を重み付け決定要素とした。なお、ここで述べるトラックの重要度とは、トラックによる楽曲の印象に与える影響の大小を示すもので、重み付け決定要素はそれぞれ、0~1の範囲で正規化しておく。

- (1) 音量の平均値
- (2) 音高の平均値
- (3) 上記2つの両方

重み付け決定要素は、トラック毎に抽出し、そのトラック内での最大値で割ることにより求めた。正規化された音量の平均値を  $v'_{T_{[n,m]}}$ 、音高の平均値を  $p'_{T_{[n,m]}}$  と表す ( $T_{[n,m]}$  は楽曲  $M_n$  のトラックを表す)。また、音量の平均値と音高の平均値の両方を用いたものは、式(3)で示すものとする。

$$d_v v_{T_{[n,m]}} + d_p p_{T_{[n,m]}} \quad (3)$$

ここで、 $v_{T_{[n,m]}}$ 、 $p_{T_{[n,m]}}$  は、それぞれ、正規化する前の音量と音高の平均値、 $d_v$  と  $d_p$  は、それぞれ、音量と音高の和を求める際にかかる係数である。今回の実験では、これらの係数はともに1とする。

また、トラックの重要度を表す上記の重みを用いて、次の重み付け関数を設定した。

- (1) 線形関数

式(4)は、特定のトラック  $T_{[n,m]}$  と、特定の形容詞群  $o$  (Hevner の成果を用いた場合、これは C1~C8 に相当する) の相関量  $mcv_{T_{[n,m]o}}$  を対象とした線形関数で、相関量に重み  $W_{T_{[n,m]}}$  を掛ける。重み  $W_{T_{[n,m]}}$  は、上記3種のいずれかの重みが相当する。

$$mcv'_{T_{[n,m]o}} = mcv_{T_{[n,m]o}} \times W_{T_{[n,m]}} \quad (4)$$

- (2) n 次関数

式(5)は、特定のトラック  $T_{[n,m]}$  と、特定の形容詞群  $o$  の相関量  $mcv_{T_{[n,m]o}}$  を対象とした n 次関数で、相関量に

表3 実験データ (抜粋)

Table 3 Data Samples for the Experiments

Song Name	Artist Name
Piano Man	Billy Joel
Wonderful Tonight	Eric Clapton
My Heart Will Go On	Celine Dion
Daniel	Elton John
In Your Eyes	Peter Gabriel
Virtual Insanity	Jamiroquai
Stairway to Heaven	Led Zeppelin
A Winter Shade of Pale	Procol Horum
Under the Bridge	Red Hot Chili Peppers
Satisfaction	Rolling Stones
:	:

重みとして  $W_{T_{[n,m]}}$  の  $g$  乗を掛ける。重み  $W_{T_{[n,m]}}$  は、上記3種のいずれかの重みが相当する。今回の実験では、 $g = 2$  とした。

$$mcv'_{T_{[n,m]o}} = mcv_{T_{[n,m]o}} \times W_{T_{[n,m]}}^g \quad (5)$$

- (3) 指数関数

式(6)は、特定のトラック  $T_{[n,m]}$  と、特定の形容詞群  $o$  の相関量  $mcv_{T_{[n,m]o}}$  を対象とした指数関数で、相関量に重みとして、 $g$  の  $W_{T_{[n,m]}}$  乗を掛ける。重み  $W_{T_{[n,m]}}$  は、上記3種のいずれかの重みが相当し、今回の実験では、 $g = 2$  とした。

$$mcv'_{T_{[n,m]o}} = mcv_{T_{[n,m]o}} \times g^{W_{T_{[n,m]}}} \quad (6)$$

- (4) 対数関数

式(7)は、特定のトラック  $T_{[n,m]}$  と、特定の形容詞群  $o$  の相関量  $mcv_{T_{[n,m]o}}$  を対象とした対数関数で、相関量に重みとして、 $g$  を底とした  $W_{T_{[n,m]}}$  の対数を掛ける。重み  $W_{T_{[n,m]}}$  は、上記3種のいずれかの重みが相当し、今回の実験では、自然対数を用いた。

$$mcv'_{T_{[n,m]o}} = mcv_{T_{[n,m]o}} \times \log_g W_{T_{[n,m]}} \quad (7)$$

(Step-B) トラック毎の印象メタデータの合成

トラック毎の重要度を反映し、重み付けられたトラック毎の印象語群ベクトル  $MCV'_{T_{[n,m]}}$  から、楽曲に対応する印象語群ベクトル  $MD_n$  を生成する。合成関数に関して、線形関数、指数関数など、多様な関数の検討が望まれるが、今回の実験では、各語群毎の和を求めることで、印象語群ベクトルを生成するものとした(式(8))。

$$MD_n = \left( \sum_{m=1}^b mcv'_{T_{[n,m]1}}, \dots, \sum_{m=1}^b mcv'_{T_{[n,m]8}} \right) \quad (8)$$

## 4. 実験

ポップス、ロックの音楽的類型から選択されたサンプルデータを対象として、シングルトラック方式、マルチトラック方式を適用し、そのメタデータ群の妥当性を評価する事で、提案メタデータ自動抽出方式の実現可能性を検証する。

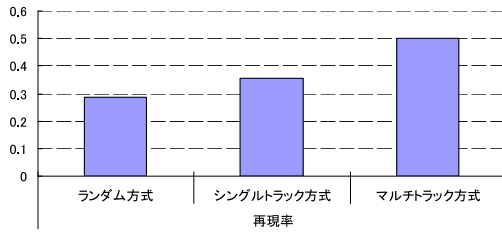


図 3 全 62 曲を対象とした実験結果  
Fig. 3 Results for Computing all 62 Songs

#### 4.1 実験手法

本実験では、1960 年代から現在に至るポップス、ロックの楽曲群から、ジャンル、作成年代を均等に分散させるという観点において、代表曲 62 曲を選択した (表 3)。これらの楽曲は、すべて、複数のトラックから構成される SMF で、主旋律、および、伴奏を含む。

各楽曲に対して、対応する形容詞群 (C1~C8) を主観により 3 個程度 (平均 3.016 個) 選択し、正解メタデータ  $C_M$  として定義した。本実験では、提案方式、および、シングルトラック方式を用いて抽出した印象メタデータのうち、各楽曲  $M$  において、Hevner の形容詞群 C1~C8 の中で最も相関の高かった形容詞群で、 $C_M$  とほぼ同数であることから、上位 3 件の形容詞群のみをメタデータ  $R_M$  として採用するものとした。ここで、各楽曲における  $R_M$  と  $C_M$  が、2 件以上一致する楽曲を本実験における適合楽曲と定義すると、適合楽曲の集合  $X$  は式 (9) で表される。本実験における提案方式、および、シングルトラック方式は、この適合楽曲集合  $X$ 、および、楽曲の総数  $a$  を用いて、式 (10) で表す再現率で評価する。なお、シングルトラック方式においては、実験データが有する主旋律のみを選択し、適用するものとした。

$$X = \{M | \text{count}(R_M \cap C_M) \geq 2\} \quad (9)$$

$$P = \frac{\text{count}(X)}{a} \quad (10)$$

提案方式は、複数の楽器から構成される楽曲と、Hevner の形容詞群の間の相関量を求め、印象メタデータとして採用する。これはすなわち、各楽曲が、最も適合する形容詞群を、C1~C8 の形容詞群から選択する作業であるといえる。ここで、各楽曲に対して 8 個の形容詞群から無作為に 3 個を選択し、各楽曲のメタデータとして採用する方式をランダム方式として定義する。ランダム方式を用いて無作為に選択された 3 つの形容詞群のうち、2 つ以上が正解と重なる確率  $R$  は、式 (11) で表される。

$$R = \frac{{}_3C_3 \times {}_5C_5 + {}_3C_2 \times {}_5C_1}{{}_8C_3} = \frac{2}{7} \quad (11)$$

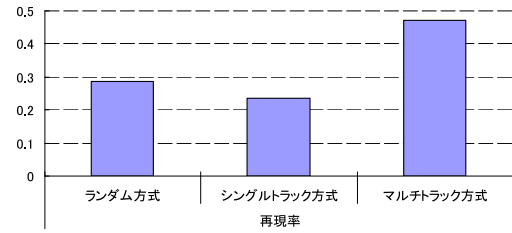


図 4 シングルトラック方式およびマルチトラック方式の結果が異なった 34 曲を対象とした実験結果  
Fig. 4 Results for Computing 34 Songs which the Single-Track Method and the Multi-Track Method Returned Different Metadata

本実験では、提案方式、シングルトラック方式、および、ランダム方式を比較することで、提案方式の実現可能性を示す。

#### 4.2 実験結果

図 3 は、全 62 曲を対象としたメタデータ抽出の再現率を表し、図 4 は、提案方式、および、シングルトラック方式が異なる結果を出力した 34 曲のみを対象とした結果を示す。なお、本実験では、トラック毎の重要度を表す重みとして音高と音量を用い、合成関数としては  $n$  次関数を用いた。これは以上の構成を用いることで、予備実験の結果が良好であったことによる。

第 1 に、図 4 より、ポップス、ロックの音楽的類型において、主旋律以外の楽器が楽曲全体の印象に大きく寄与する楽曲を対象とした場合、提案方式が特に高い再現率を示すことが確認された。図 4 において、シングルトラック方式は、無作為にメタデータを付与するランダム方式をも下回る結果を示した。これは、主旋律の印象と、他の旋律を含む楽曲全体の印象が極端に乖離しているため、シングルトラック方式を用いたメタデータの抽出が適切に行われず、極端に低い再現率を示したことを意味する。シングルトラック方式は、単一のトラックから構成される楽曲データを対象とした場合、高い再現率を示すことが確認されている [8] [14] が、図 4 の結果より、マルチトラックから構成される楽曲に単純に適用することが困難であることが確認された。一方、提案方式は、主旋律の印象、および、楽曲全体の印象が乖離している場合においても、従来方式、ランダム方式より高い再現率を示した。

第 2 に、提案方式は、シングルトラック方式、ランダム方式に比べ、全 62 曲を対象とした場合、より正解メタデータに近い結果を示した (図 3)。この結果において、提案方式は、シングルトラック方式、ランダム方式の双方より高い再現率を示した。また、図 4 と異なり、全曲を対象データとした場合、シングルトラック方式が、ランダム方式より高い再現率を示した。これはすなわち、実験データの半数近くが、曲全体の印象と同様の印象を、主旋律から与えるということに起因すると考えられ、シングルトラック方式がランダム方式に比べ有効である点も改めて確認さ

れた。

これより、複数のトラックから構成される楽曲データを対象とした印象メタデータの抽出に関し、シングルトラック方式、ランダム方式との比較において、提案方式を適用することが有効であることが確認された。

## 5. 結 論

本論文では、複数の楽器から構成された楽曲データを対象に、印象を表す形容詞群、および、その重みを動的に抽出するメタデータの自動抽出方式を提案した。本方式の特徴は、複数の楽器表現によって構成される楽曲を対象として、楽器毎の音符配列に音楽心理学の成果を適用することで、印象語との相関量を動的に計算し、また、楽器毎に計算された印象語との相関量を合成することで、楽曲に対応するメタデータを動的に生成する。さらに、実験により、提案方式と従来方式を比較することで、提案方式の実現可能性を示した。

今後の課題としては、トラック毎の印象語群ベクトル合成関数の検証、ならびに、同合成関数に適用する重みの係数の検証、提案方式の心理学的側面からの有効性の検証、リズム、音色に応じたメタデータの重み付け機能を実現する予定である。

## 6. 謝 辞

本研究に当たって貴重な御助言を頂いた慶應義塾大学 SFC 研究所の中神康裕氏、筑波大学の北川高嗣先生、佐藤聡先生、慶應義塾大学の吉田尚史先生、伊地智麻子氏、また、実験システムの構築において御支援頂いた慶應義塾大学の大前寛子氏に感謝の意を表します。

## 文 献

- [1] Hevner, K.: "Expression in Music: a Discussion of Experimental Studies and Theories," *Psychological Review*, Vol.42, pp.186-204 (1935).
- [2] Hevner, K.: "Experimental Studies of the Elements of Expression in Music," *American Journal of Psychology*, Vol.48, pp.246-268 (1936).
- [3] Hevner, K.: "The Affective Value of Pitch and Tempo in Music," *American Journal of Psychology*, Vol.49, pp.621-639 (1937).
- [4] 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," 第 12 回データ工学ワークショップ (DEWS2001) 論文集, 電子情報通信学会 (2001).
- [5] ISO: "Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 3: Audio," ISO 11172-3 (1993).
- [6] ISO: "Generic coding of moving pictures and associated audio information - Part 3: Audio," ISO 13818-3 (1998).
- [7] 木下 智義, 坂井 修一, 田中 英彦: "周波数成分の重なり適応処理を用いた複数楽器の音源同定処理," 電子情報通信学会論文誌 D-II Vol.J83-D-II, No.4, pp.1083-1081 (2000).
- [8] Kitagawa, T. and Kiyoki, Y.: Fundamental Framework for Media Data Retrieval Systems Using Media-lexico Transformation Operator - In the Case of Musical MIDI Data, *Proceedings of THE TENTH*

*EUROPEAN-JAPANESE CONFERENCE ON INFORMATION MODELLING AND KNOWLEDGE BASES* (2000).

- [9] 小杉 尚子, 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," 情報処理学会論文誌, Vol.43 No.02, pp.287-298 (2002).
- [10] 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," DBWeb2000, pp.229-234 (2000).
- [11] 小野 徹太郎, 斎藤 英雄, 小沢 慎治: "自動採譜のための GA を用いた混合音推定法," 計測自動制御学会論文集, Vol.33, No.5, pp.417-423 (1997).
- [12] 谷口 高士: "音楽作品の感情測定尺度の作成および多面的感情状態尺度との関連検討," 心理学研究, Vol.65, No.6, pp.463-470 (1995).
- [13] 柳瀬 隆史, 高須 淳宏, 安達 淳: "メロディからの特徴抽出による曲検索システム," 電子情報通信学会 第 10 回データ工学ワークショップ (DEWS'99), 2B-5 (1999).
- [14] 吉野 太智, 高木 秀幸, 清木 康, 北川 高嗣: "楽曲データを対象としたメタデータ自動生成方式とその意味的連想検索への適用," データベースシステム, No.7, pp.109-116 (1998).