

複数の楽器を用いた楽曲を対象とする メタデータ自動生成方式とその実現

大前 寛子^{†1} 石橋 直樹^{†2}
清木 康^{†3} 安西 祐一郎^{†4}

広域ネットワーク上に散在する多種多様なメディアデータからの適切なデータを検索する方式の実現が課題となっている。メディアデータの中の楽曲データに着目すると、複数の楽器を用いた楽曲データから、メタデータとなる印象を抽出する手法はほとんど存在しない。そこで、本論文では、複数の楽器を用いた楽曲を対象とするメタデータ自動生成システムを示す。複数の楽器から構成される楽曲データから楽器ごとの印象を抽出し、それぞれの楽器ごとの印象を楽器の重要度に応じて重み付けを行ない、それらと合成して楽曲としての印象をメタデータとして抽出する。また、実際のアンケート結果を用いて、複数の楽器を用いた楽曲データによる印象抽出の有効性と重み付けの有効性を示す。

An Automatic Metadata Creation Method for Music with Multiple Musical Instruments

HIROKO OMAE,^{†1} NAOKI ISHIBASHI,^{†2} YASUSHI KIYOKI^{†3}
and YUICHIRO ANZAI^{†4}

With a rapid popularization of the global-area network, realizing an information retrieval method for media data is essentially important today. As focusing to the information retrieval methods for music data, absolute methods to automatically extract impressions, which becomes metadata, of music data do not exist. In this paper, we propose an automatic metadata creation method for music with multiple musical instruments. The impression for each musical instrument is extracted from the music data which consist of multiple instruments and weighted according to each importance of the musical instrument. Our method extract the impression of a music datum by compounding the impression of each musical instrument with others. We clarify the effectiveness of the method in a view point of compounding the impression of each musical instrument by comparing to present methods, and the effectiveness of weighting by showing the result of the actual questionnaire.

1. はじめに

近年、広域ネットワーク上に散在する多種多様なデータ(テキスト、画像、音声など)を検索対象として、適切なデータの検索を可能とする方式の実現が重要な課題となっている。

メディアデータ群を対象とした既存の検索方式は、部分的なパターンマッチングによる直接的な方法²⁾が主流であった。パターンマッチングによる検索方式では、同一メディア間かつ同一形式間での検索は行えるが、異なる形式の異種メディア間では検索を行うことができない。

一方、メディアデータを対象とした検索に抽象情報(例: 作品名、作者名、感情表現など)である「メタデータ」を用いることで、異種メディア間の検索を行うことを可能に

する研究⁴⁾が行なわれている。印象によるメディアデータの検索方式としては、拡張された意味の数学モデルを用いたメディアデータの意味的連想検索方式⁶⁾が提案されている³⁾⁴⁾。

本研究では、メディアデータの中でも、特に楽曲データを対象としたメタデータ生成方式を示す。楽曲データのメタデータ生成において、楽曲の印象表現を用いた「楽曲のメタデータ自動生成システム」¹⁴⁾¹⁵⁾に関する研究がある。この研究は、楽曲の中でもピアノ曲という、単一の楽器による楽曲データのみを対象としていたが、現在我々がよく耳にする楽曲は、複数の楽器から構成される形式の方がはるかに多い。また、楽曲データから印象表現によるメタデータを抽出する他の研究においても、複数の楽器によって構成された楽曲を扱ったものや、その際に楽器ごとの特徴を考慮したものはほとんど存在しない。

そこで本稿では、複数の異なる楽器を用いた楽曲データを対象とするメタデータ自動生成方式を提案する。本方式は、上記の「楽曲のメタデータ自動生成システム」に関する研究を拡張したもので、メタデータをその楽曲中で使用されている楽器ごとに取り出し、楽器の重要度順に適切な重み付けを行い、1つのメタデータに統合し、曲全体としての印象を抽出するというものである。重み付けを行う際、楽器の重要度は音量と音高により決め、重

†1 慶應義塾大学大学院 理工学研究科
Graduate School of Science and Technology, Keio University

†2 慶應義塾大学大学院 政策・メディア研究科
Graduate School of Media and Governance, Keio University

†3 慶應義塾大学 環境情報学部
Faculty of Environmental Information, Keio University

†4 慶應義塾大学 理工学部
Faculty of Science and Technology, Keio University

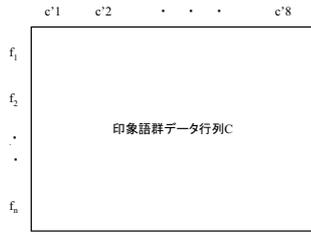


図 3 印象語群データ行列 C

$$c'_k = (f_1, f_2, \dots, f_{n-1}, f_n)^t$$

そして印象語群データ c'_k を列ベクトルとして構成される印象語群データ行列 C (図 3) が作成される。

$$C = (c'_1, c'_2, \dots, c'_8)$$

2.2.3 楽曲構造基礎データの抽出

Standard MIDI File (以下、SMF) を入力として与え、楽曲の構造と形態について解析し、楽曲の特徴を決定するために必要な楽曲の基礎データを抽出する。

楽音データ：Note

楽曲データに含まれる楽音数を notenum とする。楽音データ $Note_a$

($a = 1, 2, \dots, notenum$) は次のように定義される。

$$Note_a = (npitch_a, nvel_a, nsdel_a, nedel_a, ndur_a)$$

Note の各要素はそれぞれ、音程番号、音量、発音開始相対時間、発音終了相対時間、楽音の長さである。

和音データ：Chord

同時に発音している楽音の集合を和音とし、楽曲データに含まれる和音数を chnum とする。和音データ $Chord_b$ ($b = 1, 2, \dots, chnum$) は次のように定義される。

$$Chord_b = (chsdel_b, chedel_b, chdur_b)$$

Chord の各要素は、それぞれ和音の開始相対時間、和音の終了相対時間、和音の長さである。

調名データ：kn

SMF における調データから、楽曲の調名を抽出する。調名データ kn は、調の名前を表す文字列 keyname で定義される。

$$kn = (keyname)$$

演奏時間データ：pt

楽曲全体の演奏時間 pt は次のように定義される。

$$pt = nedel_{notenum} - nsdel_1$$

拍子データ：mt

SMF における拍子データから、楽曲の拍子を抽出する。拍子の分母および分子を表す値 bb 、 bc により、拍子データ mt は次のように定義される。

$$mt = (bb, bc)$$

テンポデータ：tem

SMF におけるテンポデータから、1 分間に演奏される 4 分音符の数 Mt を抽出する。 Mt の値により、テンポデータ tem は次のように定義される。

$$tem = Mt$$

平均音高値：mnap

楽曲の音高値の平均は $mnap$ とし、次のように定義し

ている。

$$mnap = \frac{\sum_{i=1}^{notenum} (ndur_i)(npitch_i)}{nedel_{notenum} - nsdel_1}$$

固定/流動リズム値：unac/ac

システムでは、「固定的リズム」とは、和音が 4 分音符以上の長さの楽音で構成されるとしている。また、「流動的リズム」とは、和音を構成する楽音中に 4 分音符未満の長さの楽音が含まれているとしている。固定的なリズムにおける和音の長さの和は固定リズム値とし、流動的なリズムにおける和音の長さの和は流動リズム値としている。ここで、「固定的リズム」の Chord を示す指標を $Cchord$ 、「流動的リズム」の Chord を示す指標を $Achord$ とすると、固定/流動リズム値 $unac/ac$ は次のように定義される。

$$unac = \sum_{i \in Cchord} (chdur_i)$$

$$ac = \sum_{i \in Achord} (chdur_i)$$

単純/複雑/ノーハーモニー値：tc/oc/nc

主要和音により構成される Chord を示す指標を $Tchord$ 、それ以外の和音により構成される Chord を示す指標を $Ochord$ としている。単純/複雑/ノーハーモニー値 $tc/oc/nc$ は次のように定義される。

$$tc = \sum_{i \in Tchord} (chdur_i)$$

$$oc = \sum_{i \in Ochord} (chdur_i)$$

$$nc = \sum_{i=1}^{chnum} (chdur_i) - (tc + oc)$$

上昇/下降/水平メロディ値：um/dm/nm

上昇/下降/水平メロディはそれぞれ、次に発音される楽音との音高差が正、負、0 であることを指すものとしている。ここで、上昇、下降、水平メロディを示す指標をそれぞれ、 $Umnote$ 、 $Dmnote$ 、 $Nmnote$ とすると、上昇/下降/水平メロディ値は次のように定義される。

$$um = \sum_{i \in Umnote} (ndur_i)$$

$$dm = \sum_{i \in Dmnote} (ndur_i)$$

$$nm = \sum_{i \in Nmnote} (ndur_i)$$

2.2.4 楽曲構造要素量解析

楽曲構造データを解析して得られた楽曲構造基礎データを用いて、楽曲構造要素量を求めている。また、それらの特徴量により楽曲を特徴付けした楽曲構造要素ベクトル (以下、MFV) を生成している。楽曲データ i の MFV $_i$ は次のように定義される。

$$MFV_i = (key_i, tempo_i, pitch_i, rhythm_i, harmony_i)^t$$

MFV の各要素はそれぞれ基準が異なる。そこで、各特

表 2 長調の調性度

ハ (C)	ト (G)	ニ (D)	イ (A)	ホ (E)	ロ (B)
0.9375	1.0000	0.7500	0.7500	0.4375	0.6250
嬰ハ (F#)	ヘ (F)	変ロ (Bb)	変ホ (Eb)	変イ (Ab)	変ニ (Db)
0.5000	0.6875	0.1250	0.4375	0.3125	0.0000

表 3 短調の調性度

イ (a)	ホ (e)	ロ (b)	嬰ヘ (f#)	嬰ハ (c#)	ニ (d)
0.0000	-0.6250	-0.6250	-0.6250	-0.8750	-0.7500
ト (g)	ハ (c)	ヘ (f)	変ロ (Bb)		
-0.5000	-1.0000	-0.7500	-0.8750		

微量が-1~1の値をとるように正規化している。

各特徴量は、次のように定義される。

key

key は、楽曲データの調名データ kn および表 2、表 3 に示す調性度を用いて決定している。これらの表は、梅本・中島の研究¹¹⁾によるもので、各調の長調らしさ、短調らしさを表す値を-1~1に正規化して作成されている。

tempo

tempo は、楽曲データのテンポデータ tem を用いて次のように定義される。

$$tempo = \begin{cases} \frac{tem - tstnd}{tmax - tstnd} & tem - tstnd \geq 0 \\ \frac{tem - tstnd}{tstnd - tmin} & tem - tstnd < 0 \end{cases}$$

ここで、 $tmax$ 、 $tmin$ 、 $tstnd$ はそれぞれ、M.M(メルツェル・メトロノーム)の最大値(=184)、最小値(=40)、基準値(=88)である。

pitch

pitch は、楽曲データの平均音高値 $mnap$ を用いて次のように定義される。

$$pitch = \begin{cases} \frac{mnap - pstnd}{pmax - pstnd} & mnap - pstnd \geq 0 \\ \frac{mnap - pstnd}{pstnd - pmin} & mnap - pstnd < 0 \end{cases}$$

ここで、 $pmax$ 、 $pmin$ 、 $pstnd$ はそれぞれ、88 鍵のピアノの最高音、最低音および中央ラ音に対応する、SMF における音程番号である。

rhythm

rhythm は、楽曲データの固定/流動リズム値: $unac/ac$ を用いて次のように定義される。

$$rhythm = \frac{unac - ac}{unac + ac}$$

harmony

harmony は、楽曲データの単純/複雑/ノーハーモニー値: $tc/oc/nc$ を用いて次のように定義される。

$$harmony = \frac{tc - oc}{tc + oc + nc}$$

melody

melody は、楽曲データの上昇/下降/水平メロディ値: $um/dm/nm$ を用いて次のように定義される。

$$melody = \frac{um - dm}{um + dm + nm}$$

2.2.5 楽曲印象語群ベクトル生成

変換行列 T と MFV を用いて、楽曲を 8 つの印象語

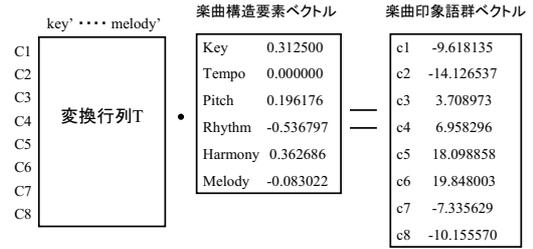


図 4 楽曲印象語群ベクトル生成例

群によって特徴付けした楽曲印象語群ベクトル (以下、MCV) を生成している (図 4)。楽曲データ i の MCV_i は、楽曲データと各印象語群との相関量 mcv により、以下のように定義される。

$$MCV_i = (mcv_{i1}, \dots, mcv_{i8})^t$$

2.2.6 楽曲メタデータ生成

印象語群データ行列 C と楽曲印象語群ベクトル MCV を入力として与え、楽曲データを意味の数学モデル⁴⁾³⁾におけるデータ行列と同じ特徴で特徴づけした楽曲メタデータ (以下、MMD) を生成している。楽曲データ i の楽曲メタデータ MMD_i は、楽曲データと n 個の特徴との相関量 mmd により次のように定義される。

$$MMD_i = (mmd_{i1}, \dots, mmd_{i8})^t$$

MMD は意味の数学モデルにおけるデータ行列と同じ特徴で特徴づけされているため、意味の数学モデルによるメディアデータ検索方式の検索対象として適用可能である。

3. 複数の楽器を用いた楽曲を対象とするメタデータ自動生成方式

本章では、本稿で提案する複数の楽器を用いた楽曲を対象としたメタデータ自動生成システムの設計と実装について述べる。

3.1 本システムの概要

システムの概要を図 5 に示す。本システムでは、複数の楽器を用いた楽曲データを入力とし、各トラックごとに楽器の特徴を考慮した印象語群ベクトルを抽出する。次に重み付け関数を適用し、各トラックごとの印象語群ベクトルを合成することで、入力された楽曲全体としての印象語群ベクトルを抽出、これを出力とする。

3.2 入力/出力データ形式

デジタル化された楽譜データとして Standard MIDI File ("SMF") が一般に普及していることから、入力として与える楽曲データには SMF を用いる。SMF は、ヘッダ情報やトラックデータ (時間情報 + MIDI イベント) が含まれるいくつかのブロックから構成される。

また、出力されるデータは、楽曲における印象語群の相関量を表す、印象語群ベクトルである。

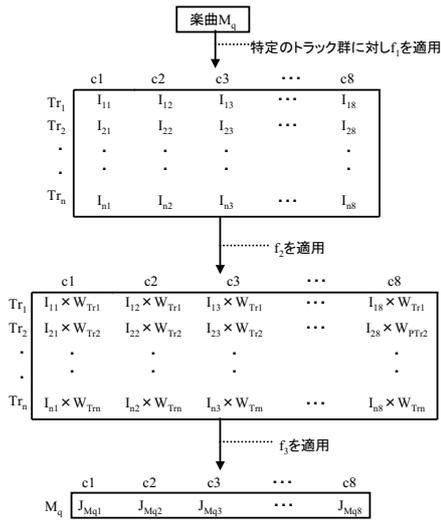


図 5 複数の楽器を用いた楽曲のメタデータ抽出システム

3.3 特定の単一トラックを対象とした印象の抽出関数

f_1

本節では、トラックごとの印象抽出を行うための関数 f_1 について述べる。この関数は、背景で述べた楽曲メタデータ自動生成システムを拡張することにより実装する。従来のシステムを複数の楽器に対応させるために、楽器種別の相違より生じる、各基準値等のパラメータの変更も行う。

3.3.1 関数の定義

関数 f_1 は、特定の楽曲 M_q の、特定のトラック Tr_r についての印象を抽出する。以下のように定義する。

$$f_1(M_q, Tr_r) \rightarrow [I_{r1}, I_{r2}, \dots, I_{r8}]$$

ここで、 $M_q (q = 1, 2, \dots, k)$ は楽曲、 $Tr_r (r = 1, 2, \dots, n)$ はトラック、 $I_{r1} \sim I_{r8}$ は Tr_r と印象語群との相関量を表す。

3.3.2 複数の楽器を用いた楽曲データに適用する際の変更点

複数の楽器を用いた楽曲データに適用するために、従来のシステムにおける以下のパラメータを変更・追加する。

フォーマット

このシステムで扱う SMF のフォーマットは 1 であるものとする。

調性度

従来のシステムでは、調性度は -1 ~ 1 の値をとるよう正規化されていた。しかし、調性度を -1 ~ 1 の値に設定して予備実験を行ったところ、複数の楽器を用いた楽曲データに適用する場合には、調性以外の楽曲構造要素の値に比べて、調性度の値が大きくなってしまい、印象語群の相関量が、ある特定の印象語群に偏って大きな値を示すという問題が生じた。そこで、本研究では調性度を -0.5 ~ 0.5 の値をとるよう修正する。

テンポの基準値

従来のシステムでは、テンポの基準値は 88 と設定され

ていた。これは、通常ピアノ曲で用いられる中庸の速さとして設定されている値であるため、ピアノ以外の楽器による演奏に用いるには不適切であると考えられる。

従来のシステムがクラシック曲を対象にしているのに対し、本研究ではロック・ポップスなど、比較的テンポの速い曲を対象とするため、従来の基準値では小さすぎると考えられ、今回は基準値を 120 に設定する。

音高の最小値・最大値・基準値

従来のシステムは、ピアノの最低音、最高音を音高の最小値、最大値とし、基準値を中央のラ音としていたが、ピアノ以外の楽器による演奏に用いるには不適切であると考えられる。そこで、新たな設定として、SMF で設定可能な最低音、最高音とその中央の音を音高の最小値、最大値、基準値とする。

楽器名: *inst*

SMF のコントロールチェンジデータから、トラック内で使用されている楽器名 *inst* を抽出する。

平均音量: *vel*

楽曲の音量の平均を *vel* とし、次のように定義する。

$$vel = \frac{\sum_{i=1}^{notenum} (ndur_i)(nvel_i)}{nedel_{notenum} - nsdel_1}$$

3.4 楽器の重要度に対応する重み付け関数 f_2

楽曲の印象を決定する際の影響力は、楽器によって異なると考えられる。本稿では、この影響力を楽器の重要度と呼ぶ。重要度によって重み付けを行なうことで、楽器毎の印象の適切な合成を行ない、楽曲としての印象を抽出する。本節では、まず重み付け関数の決定方法を説明し、次に、関数の定義について述べる。

3.4.1 重み付け関数の決定

本方式では、楽曲の違いやユーザの感性の違いで最適な重み付け関数は変化すると考え、複数の決定要素と関数を用意する。

まず、重み付けを行う前の段階で、各トラックごとに取り出した印象語群との相関量を -1 ~ 1 に正規化しておく。相関量を正規化した後、楽器の重要度として重みを掛け合わせ、それらを合成して楽曲としての印象 (印象語群との相関量) を抽出する。

楽器の重要度は、音量の平均が大きい楽器ほど重要度は高く、音量が小さい楽器ほど重要度は低いと考え、また、音高の平均が高い楽器ほど重要度は高く、音高が低い楽器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、ここで述べる楽器の重要度とは、楽器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、トラックごとに抽出し、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'_i 、音高の平均値を p'_i と表す (t : トラッ

ク数)。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、

$$a_v v_t + a_p p_t$$

とする。ここで、 v_t 、 p_t はそれぞれ、正規化する前の音量と音高の平均値、 a_v と a_p はそれぞれ、音の強さと音高の和を求める際にかける係数である。今回の実験では、これらの係数はどちらも 1 とする。

また、上記の重み付けの要素に着目した重み付け関数をそれぞれ、

- (1) $p(v'_t)$
- (2) $q(p'_t)$
- (3) $r(a_v v_t + a_p p_t)$

とする。そして、これらを用いて、重み付けのための関数は以下の 4 通りとした。

- 線形関数

v'_t 、 p'_t 、 $a_v v_t + a_p p_t$ をそのまま重みとして掛ける。

- 2 次関数

v'_t 、 p'_t 、 $a_v v_t + a_p p_t$ をそれぞれ 2 乗したものを重みとして掛ける。

- 指数関数

v'_t 、 p'_t 、 $a_v v_t + a_p p_t$ を指数として用いた結果を重みとして掛ける。今回の実験では底を 2 とする。

- 対数関数

v'_t 、 p'_t 、 $a_v v_t + a_p p_t$ を対数として用いた結果を重みとして掛ける。今回の実験では自然対数を用いる。

これらの関数は、いずれも増加関数である。

3.4.2 関数の定義

重み付け関数 f_2 は以下のように定義する。

$$f_2(M_q, Tr_r, [I_{r1}, I_{r2}, \dots, I_{r8}]) \\ \rightarrow [I_{r1} \times W_{Tr_r}, I_{r2} \times W_{Tr_r}, \dots, I_{r8} \times W_{Tr_r}]$$

ここで、 $W_{Tr_r} \sim W_{Tr_r}$ は Tr_r において用いられる楽器の重要度を表す。

この関数は、 f_1 によって得られる相関量に、前節で述べた方法により求められた楽器の重要度によって与えられる相関量を、重みとして掛け合わせる関数である。

3.5 楽曲に対応する印象の合成関数 f_3

本節では、楽曲全体としての印象を抽出するための合成関数について述べる。

3.5.1 相関量の合成

重み付け関数により重み付けをした後、それぞれの相関量を合成する。重み付け決定要素に対応して、合成関数を以下のように定義する。

$$\sum_{i=1}^t p(v_t) \cdot c_{it} \\ \sum_{i=1}^t q(p_t) \cdot c_{it} \\ \sum_{i=1}^t r(a_v v_t + a_p p_t) \cdot c_{it}$$

ただし、 c_{it} ($i = 1, 2, \dots, 8$) は、 i 番目の印象語群と t 番目のトラックとの相関量を表す。

表 4 実験用サンプル楽曲

楽曲名	アーティスト名
Piano Man	Billy Joel
Riders of the Storm	Doors
Killing me Softly	Roberta Flack
Bombaleo	Gypsy Kings
Babylove	Supremes
I Still Believe	Mariah Carey
Wonderful Tonight	Eric Clapton
My Heart Will Go On	Celine Dion
Stairway To Heaven	Led Zepplin
Bohemian Rhapsody	Queen
Africa	Toto
Detroit Rock City	Kiss
Bombtrack	Rage Against the Machine
Dance To The Music	Sly and the Family Stone

3.5.2 関数の定義

合成関数 f_3 は以下のように定義する。

$$f_3(M_q, [\{Tr_r, [I_{r1} \times W_{Tr_r}, \dots, I_{r8} \times W_{Tr_r}]\}, \\ \vdots \\ \{Tr_r, [I_{n1} \times W_{Tr_n}, \dots, I_{n8} \times W_{Tr_n}]\}]) \\ \rightarrow [J_{M_q1}, \dots, J_{M_q8}]$$

ここで、 $J_{M_q1} \sim J_{M_q8}$ は M_q と印象語群との相関量を表す。この関数は f_2 より得た各トラックの相関量を合成する。

4. 実験

本方式の有効性を検証するため、前章で示した方式を実現する実験システムを構築し、以下の検証実験を行った。
実験 1: 楽曲の旋律のみを取り出した楽曲データをピアノ曲を対象としたメタデータ自動生成システムに入力として与え、複数の楽器を用いることの実効性を検証する実験

実験 2: 複数の楽器を用いた楽曲データをシステムに入力として与え、重み付けの実効性を検証する実験

4.1 実験条件

4.1.1 実験データ

楽曲メディアデータとして、比較的有名な楽曲 14 曲を用意した。それらの楽曲名とそのアーティスト名を表 4 に示す。

4.1.2 アンケート

被験者 16 人を対象に、アンケートを行った。アンケートは、楽曲を聞いて、その印象に最も適していると思われる印象語群を 2 つ選ぶという形式で行なった。

4.2 評価方法

本システムの評価方法は以下の通りである。

- (1) アンケートの結果、投票数の多かった上位 2 つの印象語群を正解とする。
- (2) システムの出力の結果、楽曲と相関の高かった上位 2 つの印象語群と正解を比較する。

アンケートを用いることによって、システムを客観的に評価することができる。

4.3 実験 1

4.3.1 実験方法

表 4 の最初の 11 曲について、旋律と思われる楽器演奏の部分のみを取り出し、それを単一の楽器による楽曲データとして、ピアノ曲を対象としたメタデータ自動生

表 5 単一の楽器による楽曲データによる結果

楽曲名	アンケート	単一の楽器	複数の楽器
Piano Man	c3,c4	c6, c5	c4,c3
Riders of the Storm	c1,c5	c1,c8	c1,c2
Killing me Softly	c4,c2	c2,c1	c1,c2
Bombaleo	c7,c6	c6,c5	c6,c5
Babylove	c5,c6	c1,c5	c4,c1
I Still Believe	c3,c4	c5,c4	c3,c4
Wonderful Tonight	c4,c3,c6	c6,c5	c1,c4
My Heart Will Go	c1,c3	c1,c6	c1,c2
Stairway To Heaven	c2,c1	c5,c4	c4,c1
Bohemian Rhapsody	c4,c3	c1,c2	c1,c4
Africa	c5,c6,c3	c4,c6	c5,c4

成システムに入力した。また、同じ表 4 中の最初の 11 曲について、複数の楽器を用いた楽曲データをシステムに入力した。そして、同じサンプル曲を対象にアンケートを行い、その結果とシステムの出力結果を同様に比較した。複数の楽器を用いた楽曲データをシステムに入力する際には、重み付け決定要素として音量を用い、関数として線形関数を用いた。

4.3.2 実験結果

旋律として取り出した楽曲データと複数の楽器を用いた楽曲データをシステムに与えた結果とアンケート結果を表 5 に示す。なお、アンケート結果において、2 位となる印象語群が 2 つある場合には、上位 3 つまでを正解とした。

4.3.3 考察

表 5 より、旋律と思われる楽器のみの楽曲において、アンケート結果の上位 2 つとシステムの出力結果の上位 2 つが完全に一致したものはなく、1 つだけ一致したのは 8 曲、全く一致しなかったのは 3 曲であった。この結果から、比較的有名な曲をサンプル曲としたため、知っている曲を聞いた場合、旋律だけを聞いているにもかかわらず、人間はその伴奏と一緒にイメージしてしまっている可能性があると考えられる。また、知らない曲であっても旋律だけでは印象を決めるのに十分ではなく、曲を聞く際にある程度のイメージを自分の中で作り上げてしまっているものと思われる。

この実験を行う際の問題点としては、旋律とみなして取り出した楽器が、その楽曲全体を通して旋律を演奏するとは限らないということである。楽曲中で旋律を演奏する楽器が変化するという事は十分に考えられる。

また、複数の楽器を用いた楽曲において、アンケート結果の上位 2 つとシステムの出力結果の上位 2 つが完全に一致したのは 2 曲、1 つ一致したのは 8 曲、全く一致しなかったのは 1 曲であった。

以上の結果と問題点から、複数の楽器で演奏されている楽曲の印象は旋律だけで決められるのではなく、故に複数の楽器による印象抽出を考えることが必要であるということが分かる。

4.4 実験 2

4.4.1 実験方法

表 4 の全曲について、複数の楽器を用いた楽曲データをシステムに入力として与えた。この際、楽器ごとに抽出した印象を様々な重み付け関数にかけて合成した場合と、重み付けを行わずに合成した場合の 2 通りの結果を出力した。また、同じサンプル曲で同様のアンケートを

表 6 実験 2:重み付けなしで合成した結果

楽曲名	アンケート	重みなし
Piano Man	c3,c4	c3,c4
Riders of the Storm	c1,c5	c1,c2
Killing me Softly	c4,c2	c1,c2
Bombaleo	c7,c6	c5,c6
Babylove	c5,c6	c4,c1
I Still Believe	c3,c4	c3,c4
Wonderful Tonight	c4,c3,c6	c1,c4
My Heart Will Go	c1,c3	c1,c2
Stairway To Heaven	c2,c1	c4,c1
Bohemian Rhapsody	c4,c3	c1,c4
Africa	c5,c6,c3	c5,c4
Detroit Rock City	c7,c8	c7,c6
Bombtrack	c8,c7	c2,c1
Dance To The Music	c6,c5	c6,c5

表 7 実験 2:重み付け決定要素に音量を用いた結果

楽曲名	アンケート	重み付け関数			
		線形	2 次	指数	対数
Piano Man	c3,c4	c4,c3	c4,c3	c4,c3	c4,c1
Riders of the Storm	c1,c5	c1,c2	c1,c2	c1,c2	c1,c2
Killing me Softly	c4,c2	c1,c2	c1,c2	c1,c2	c1,c2
Bombaleo	c7,c6	c6,c5	c6,c5	c5,c6	c5,c6
Babylove	c5,c6	c4,c1	c4,c1	c4,c1	c4,c1
I Still Believe	c3,c4	c3,c4	c3,c4	c3,c4	c3,c4
Wonderful Tonight	c4,c3,c6	c1,c4	c1,c4	c1,c4	c1,c4
My Heart Will Go	c1,c3	c1,c2	c1,c2	c1,c2	c1,c2
Stairway To Heaven	c2,c1	c4,c1	c4,c1	c4,c1	c4,c1
Bohemian Rhapsody	c4,c3	c1,c4	c1,c4	c1,c4	c1,c4
Africa	c5,c6,c3	c5,c4	c5,c4	c5,c4	c5,c4
Detroit Rock City	c7,c8	c7,c6	c7,c6	c7,c6	c7,c6
Bombtrack	c8,c7	c2,c1	c2,c1	c2,c1	c2,c1
Dance To The Music	c6,c5	c6,c5	c6,c5	c6,c5	c6,c5

表 8 実験 2:重み付け決定要素に音高を用いた結果

楽曲名	アンケート	重み付け関数			
		線形	2 次	指数	対数
Piano Man	c3,c4	c4,c3	c4,c3	c4,c1	c4,c1
Riders of the Storm	c1,c5	c1,c2	c1,c2	c1,c2	c1,c2
Killing me Softly	c4,c2	c1,c2	c1,c2	c1,c2	c1,c2
Bombaleo	c7,c6	c5,c6	c5,c6	c5,c6	c5,c6
Babylove	c5,c6	c4,c1	c1,c4	c4,c1	c4,c1
I Still Believe	c3,c4	c3,c4	c3,c4	c3,c4	c3,c4
Wonderful Tonight	c4,c3,c6	c1,c4	c1,c8	c1,c4	c1,c4
My Heart Will Go	c1,c3	c1,c2	c1,c2	c1,c2	c1,c2
Stairway To Heaven	c2,c1	c4,c1	c4,c1	c4,c1	c4,c1
Bohemian Rhapsody	c4,c3	c1,c4	c1,c4	c1,c4	c1,c4
Africa	c5,c6,c3	c5,c4	c5,c4	c5,c4	c5,c4
Detroit Rock City	c7,c8	c7,c6	c7,c8	c7,c6	c7,c6
Bombtrack	c8,c7	c2,c1	c2,c1	c2,c1	c2,c1
Dance To The Music	c6,c5	c6,c5	c6,c5	c6,c5	c6,c5

表 9 実験 2:重み付け決定要素に音量と音高を用いた結果

楽曲名	アンケート	重み付け関数			
		線形	2 次	指数	対数
Piano Man	c3,c4	c4,c3	c4,c3	c4,c3	c4,c1
Riders of the Storm	c1,c5	c1,c2	c1,c2	c1,c2	c1,c2
Killing me Softly	c4,c2	c1,c2	c1,c2	c1,c2	c1,c2
Bombaleo	c7,c6	c5,c6	c5,c6	c5,c6	c5,c6
Babylove	c5,c6	c4,c1	c4,c1	c4,c1	c4,c1
I Still Believe	c3,c4	c3,c4	c3,c4	c3,c4	c3,c4
Wonderful Tonight	c4,c3,c6	c1,c4	c1,c4	c1,c4	c1,c4
My Heart Will Go	c1,c3	c1,c2	c1,c2	c1,c2	c1,c2
Stairway To Heaven	c2,c1	c4,c1	c4,c1	c4,c1	c4,c1
Bohemian Rhapsody	c4,c3	c1,c4	c1,c4	c1,c4	c1,c4
Africa	c5,c6,c3	c5,c4	c5,c4	c5,c4	c5,c4
Detroit Rock City	c7,c8	c7,c6	c7,c8	c7,c6	c7,c6
Bombtrack	c8,c7	c2,c1	c2,c1	c2,c1	c2,c1
Dance To The Music	c6,c5	c6,c5	c6,c5	c6,c5	c6,c5

行い、システムの出力結果と比較した。重み付けに使用する要素と関数は 3 章で述べた全ての場合についてそれぞれ行った。

4.4.2 実験結果

サンプル曲をシステムに入力として与え、重み付けを行わずに合成した結果を表 6 に、重み付け決定要素として、音量 v_t を用いた結果を表 7 に、音高 p_t を用いた結果を表 8 に、音量と音高の両方を用いた結果を表 9 に示す。

4.4.3 考察

表 10 に各関数についての評価結果を示す。それぞれ、アンケート結果と出力結果が 2 つとも一致、1 つだけ一

表 10 実験 2 の評価

一致した数	曲数												
	重みなし	音量				音高				音量と音高			
		線形	2次	指数	対数	線形	2次	指数	対数	線形	2次	指数	対数
2	2	3	3	3	2	3	4	2	2	3	4	3	2
1	10	9	9	9	10	9	7	10	10	9	8	9	10
0	2	2	2	2	2	2	3	2	2	2	2	2	2

致、全く一致しなかった曲数を表す。

このことより、重み付けを行った方が、重み付けを行わない場合に比べ、ほとんどの場合でアンケート結果に近い結果が出るのが分かる。また、結果に大差は出なかったものの、重み付け関数には線形関数、2次関数、指数関数のいずれかを用いるとよりアンケート結果に近い結果が出ること、対数関数は不適切であることが分かった。中でも音量と音高で2次関数を用いた場合が最もよく楽曲の印象を表していることが分かる。

また、表7、表8、表9のうちで、システムによって出力された結果の上位2つが Hevner の円形チェックリスト(図1)において隣り合う印象語群であるものは11曲あった。Hevner の円形チェックリストは印象語群間で類似性があるものを隣接するように配置しているので、楽曲としての印象を抽出するという目的は十分に達成できたと言える。残りの4曲において、出力結果の上位2つが互いに類似性の持たない印象語群となった原因には、データ自身に問題があった、もしくはパラメータの設定が適切でなかったことが考えられる。

5. おわりに

本稿では、メディアデータの中の楽曲データに着目し、楽曲の印象をメタデータとして生成する方式を示した。

ここでは、¹⁴⁾¹⁵⁾のシステムを拡張し、複数の楽器を用いた楽曲データを対象とするメタデータ自動生成方式の提案を行ない、そのシステムの設計と実現を行った。

このシステムでは、楽器ごとの特徴を反映させるためにパラメータチューニングを行ない、楽器の重要度により重み付けし、重み付け関数を用いて楽器ごとの印象を楽曲の印象として統合した。

また、アンケート結果とシステムの出力結果を比較することによって、実験の評価を行ない、複数の楽器を用いた楽曲データへの提案手法の有効性、楽器ごとの印象を合成する際の重み付けの有効性、重み付け関数の有効性を示した。

今後の課題として、以下のことが挙げられる。

- 提案システムの意味的連想検索方式への適用
- 楽器の音色による重み付けの検討
- リズム楽器や歌詞による印象を考慮したシステムの開発
- 楽器の特徴を考慮した、適切なパラメータ設定の検討
- 重み付け係数 a_v と a_p の設定の検討
- 重み付け関数のパラメータ設定の検討
- 楽曲名、アーティスト名などによる直接的検索方式のシステムへの組み込み

参考文献

- 1) 木本晴夫: “感性語による画像検索とその精度評価,” 情報処理学会論文誌, Vol.40, No.3, March, pp.886-898, 1999.
- 2) 蔭山哲也、高島洋典: “ハミング歌唱を手掛りとするメロディ検索,” 電子情報通信学会論文誌 D-II, J77-D-II, No.8, August, pp.1543-1551, 1994.
- 3) T.Kitagawa and Y.Kiyoki: “The mathematical model of meaning and its application to multi-database systems,” Proceedings of 3rd IEEE International Workshop on Research Issues on Data Engineering, pp.130-135, April, 1993.
- 4) Y.Kiyoki and T.Kitagawa: “A semantic associative search method for knowledge acquisition,” Information Modelling and Knowledge Bases, Vol.6, No.7, pp. 121-130, 1995.
- 5) 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価の自動生成,” データベースシステム, Vol.118, No.8, pp.57-64, May, 1999.
- 6) 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探索方式とその学習機構,” 電子情報通信学会論文誌, J79-D-II, No.4, pp.509-519, April, 1990.
- 7) 井口征士ほか共著: “感性情報処理” 電子情報通信学会, 1994.
- 8) K.Hevner: “expression in music: a discussion of experimental studies and theories,” Psychological Review, Vol. 42, pp. 186-204, 1935.
- 9) K.Hevner: “experimental studies of the elements of expression in music,” American Journal of Psychology, Vol.48, pp.246-268, 1936.
- 10) K.Hevner: “the affective value of pitch and tempo in music,” American Journal of Psychology, Vol.49, pp. 621-630, 1937.
- 11) 梅本堯夫(編): “音楽心理学,” 誠信書房, 1996.
- 12) Radocy Rudolf E., and Boyle David J. 共著, and 徳丸吉彦ほか共訳: “音楽行動の心理学,” 音楽之友社, 1985.
- 13) 谷口高士: “音楽と感情,” 北大路書房, 1998.
- 14) 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, 1998.
- 15) 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,” 情報処理学会研究報告, MUS, No.36, pp. 25-30, 2000.