

複数の音符列から構成される音楽データを対象とした 印象メタデータの自動抽出方式

石橋 直樹[†] 清木 康^{††} 中神 康裕^{†††} 佐藤 聡^{††††}

[†] 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

^{††} 慶應義塾大学環境情報学部 〒252-8520 神奈川県藤沢市遠藤 5322

^{†††} 慶應義塾大学 SFC 研究所 〒252-8520 神奈川県藤沢市遠藤 5322

^{††††} 筑波大学学術情報処理センター 〒305-8577 茨城県つくば市天王台 1-1-1

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

†††akira@cc.tsukuba.ac.jp

あらまし 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抽出方式を提案する。提案方式は、個々の楽器に対応する音符列から、印象を表す形容詞群、および、それらと個々の音符列の相関量を動的に抽出し、また、音符列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、実現システムを用いて、提案方式の実現可能性、および、有効性を示す。

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 Data with Multiple Note Streams

Naoki ISHIBASHI[†], Yasushi KIYOKI^{††}, Yasuhiro NAKAGAMI^{†††}, and Akira SATO^{††††}

[†] Graduate School of Media and Governance, Keio University 5322, Endo, Fujisawa,
Kanagawa, 252-8520 Japan

^{††} Faculty of Environmental Information, Keio University 5322, Endo, Fujisawa, Kanagawa,
252-8520 Japan

^{†††} Keio Research Institute at SFC, Keio University 5322, Endo, Fujisawa, Kanagawa,
252-8520 Japan

^{††††} Science Information Processing Center, University of Tsukuba 1-1-1, Tennoudai,
Tsukuba City, Ibaraki, 305-8577 Japan

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

†††akira@cc.tsukuba.ac.jp

Abstract For music data on the global area network, we propose an automatic extraction method of metadata that represent impression. 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 with an implemented system.

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。

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] [16]、楽曲名、作曲者名などの楽曲情報を静的に定義する [6] [7] ことで楽曲検索を実現する方式など

が提案されている。前者は、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。後者は、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、作曲家、演奏者、曲名など、楽曲をあらゆる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、具体的な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、未知楽曲の検索を実現する方法論として、SD法、因子分析を用いた検索空間の生成と、GA、および、ニューラルネットワークを用いた自動インデキシングシステム [5]、SD法を用いた印象尺度の設定 [10] と、楽曲印象値の自動付与システム [11]、音楽の感情価評価尺度 [15] を用いた感情価の自動生成方式 [14]、および、その感情価を用いた類似度計算方式 [12] などが提案されている。

また、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目的として、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] [17]、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、その印象語との相関量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] [17] において示した印象メタデータの自動抽出方式は、1) 単一の音符列から構成され、かつ、2) 多くの主観において印象が一点に定まる短い音符列を対象として、楽曲データから、印象メタデータを自動的に抽出する。これは、これらの方式で用いている音楽心理学研究 [1] [2] [2] が、1、および、2の条件を満たす楽曲群を対象として、印象語との相関を求めていることによる。このため、文献 [8] [17] 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、かつ、時間的に構造、または、印象が変化す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、かつ、時間的に構造、または、印象が変化す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、各時区間、各楽器毎に、形容詞、および、その形容詞との相関量で表現された印象メタデータを自動抽出し、それら部位毎の印象メタデータを動的に合成することで、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、第1に、楽曲データ群を楽器毎に分割する。ここで、分割後、得られる単一の音符列を、『単位音符列』と定義する。第2に、単位音符列を

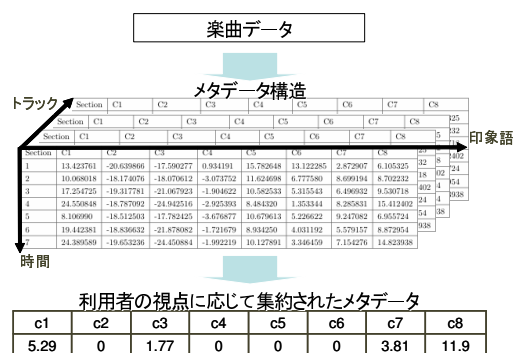


図 1 メタデータ生成の流れ

Fig. 1 A Data Flow for the Metadata Extraction

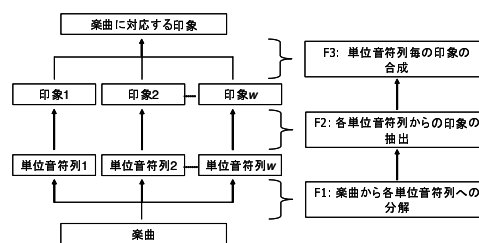


図 2 提案方式概要

Fig. 2 An Overview of the Proposed Method

対象として、印象メタデータを動的に抽出する。第3に、単位音符列毎の印象メタデータを動的に合成することで、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、複数楽器から構成される楽曲データを対象として提案方式、および、主旋律のみから構成される楽曲データ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群を評価することで、本方式の実現可能性、および、有効性を示す。

2. 印象メタデータの自動抽出方式

提案方式は、本節で示す基本機能群を用いて、単位音符列と形容詞群の相関量を格納したマトリクス構造のメタデータを自動的に生成する。さらに、そのメタデータを対象として、利用者固有の集約演算を適用することで、楽曲に対応する印象メタデータを抽出する (図 1)。

複数の音符列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、単位音符列へ分割する。F2は、単位音符列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、単位音符列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 2)。

2.1 楽曲分割機能

$$(F1) \text{division}(M_h, S_j, U_k) \rightarrow \{T_{[h,1,1]}, \dots, T_{[h,f,g]}\}$$

F1は、楽曲 M_h を、 $f \times g$ 個の単位音符列 $T_{[h,m,n]}$ に分

解する。 h は楽曲の識別子, m は時区間, n はトラックの識別子を表す。 S_j は, 字区間に分割するためのセパレータの識別子を表す。 具体的には, 小節単位, 楽章単位などといったセパレータを, S_j として定義する。 U_k は, 楽曲をトラックへ分割するセパレータで, Standard MIDI File のトラック単位などのセパレータを, T_k として定義する。

2.2 メタデータ抽出機能機能

$$(F2) f_{extract}(T_{[h,m,n]}, P_i) \rightarrow \{I_{[h,1,m,n]}, \dots, I_{[h,e,m,n]}\}$$

F2 は, 音符列 $T_{[h,n,m]}$ を, 印象メタデータの表現に変換する機能で, e 個の形容詞群との相関量 $I_{[h,l,m,n]}$ を出力する。 l 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。 印象空間 P_i は, 印象語群, および, 音符列との相関量を有する。 すなわち, 提案方式は, 既存の音楽心理学研究をこの P_i として実現することにより, 任意の音楽心理学研究に応じた印象メタデータ抽出を行う枠組みを提供する。 例えば, 文献 [17] で示される方式は文献 [1] [2] [3] [12] の成果を, 文献 [12] で示される方式は文献 [15] の成果を用いて, 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

$$(F3) f_{aggregate}(Q_s, R_t, \{I_{[h,1,1,1]}, \dots, I_{[h,e,f,g]}\}) \rightarrow \{I_{[h,1]}, \dots, I_{[h,e]}\}$$

F3 は, マトリックス構造で表された印象メタデータを合成する関数で, 形容詞との相関量 $I_{[h,l]}$ を出力する。 この機能を用いることで, 単一の音符列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, 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, 印象メタデータクリーニング機構の4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用いる。

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, 以下の楽曲分割機構を実現した。 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。 時区間の分割は, 手動により行うものとした。

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] [17] を拡張した印象メタデータの抽出機構

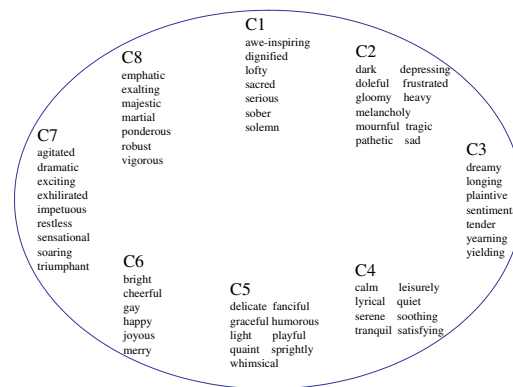


図3 Hevnerにより定義された形容詞群

Fig.3 Adjectives Defined by Hevner

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3)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3)。

提案方式は, Hevner の研究を用いたシングルトラック方式 [8] [17] を拡張し, メタデータクリーニング機能を加えることで, 複数の楽器, 時間的な構造変化を有する複雑な音楽データを対象とした印象メタデータの抽出へ, 適用可能とした。

3.3 メタデータクリーニング機能

相関量 $I_{[h,l,m,n]}$ が低い印象語群を, 印象メタデータから排除し, また, 全印象語群との相関量が低い音符列群を, 本実現への非適合音符列として排除する機構を, 印象メタデータクリーニング機構として実現する。

本実現の出力する印象語群との相関量 $I_{[h,l,m,n]}$ は, Hevner の研究 [1] [2] [3] において, 被験者が行った投票数を用いて計量する。 したがって, 特定の印象語群との相関量の高い楽曲は, 多くの利用者の主観において, その印象語群との相関が認められる楽曲である。 逆に, 相関量が相対的に低い印象語群は, 多くの利用者の主観において, その楽曲との相関が認められない楽曲である。

提案方式では, Hevner の研究を用いて複数の音符列から得られた印象メタデータを合成し, 楽曲の印象メタデータを生成するため, 相関量の低い印象語群が, メタデータ合成において累積し, 結果の精度を低下させてしまう。 したがって, 本実現では, 有意と認められる相関量の下限を, 閾値 r として定め, 相関量が r より低い印象語群は, 印象

メタデータから排除する。結果として得られる印象メタデータ $MCV'_{[h,m,n]}$ の各相関量 $I'_{[h,l,m,n]}$ は、式 (1) から求める。

$$I'_{[h,l,m,n]} = \begin{cases} I_{[h,l,m,n]}, & I_{[h,l,m,n]} \geq r \\ 0, & I_{[h,l,m,n]} < r \end{cases} \quad (1)$$

3.4 印象メタデータ合成機構

(IM1) トラック間合成機構

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、各時区間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、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、また、演奏の時間が長いトラックほど重要度は高く、短いトラックほど重要度は低いと考え、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、 $0 \sim a$ の範囲で正規化する。本論文の実験では、 $a = 1$ として検証した。

- (1) 音量の平均値, $v_{[h,m,n]}$
- (2) 音高の平均値, $mn_{[h,m,n]}$
- (3) 演奏時間, $pt_{[h,m,n]}$

重み付け決定要素は、トラック毎に抽出し、音量、音高に関しては、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 $v'_{[h,m,n]}$ 、音高の平均値を $mn'_{[h,m,n]}$ 、演奏時間を $pt'_{[h,m,n]}$ と表す。

また、これらの重み付け決定要素を用いて、次の重み付け関数を設定した。

- (1) 線形関数, $I''_{[h,l,m,n]} = I'_{[h,l,m,n]} \times W_{[h,m,n]}$
- (2) n 次関数, $I''_{[h,l,m,n]} = I'_{[h,l,m,n]} \times W^g_{[h,m,n]}$
- (3) 指数関数, $I''_{[h,l,m,n]} = I'_{[h,l,m,n]} \times c^{W_{[h,m,n]}}$
- (4) 対数関数, $I''_{[h,l,m,n]} = I'_{[h,l,m,n]} \times \log_c(W_{[h,m,n]} + 1)$

(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 $I''_{[h,l,m,n]}$ から、時区間 m に対応する印象メタデータ $I_{[h,l,m]}$ を生成する。合成関数に関しても、線形関数、指数関数など、多様な関数の検討が望まれるが、本論文の実験では、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 (2))。

$$I_{[h,l,m]} = \sum_{n=1}^g I''_{[h,l,m,n]} \quad (2)$$

(IM2) 時区間合成機能

時区間合成機能は、トラック間合成機能を用いて生成された時区間毎の印象メタデータ ($I_{[h,1,1]}, \dots, I_{[h,8,f]}$) を合成し、楽曲に対応するメタデータ ($I_{[h,1]}, \dots, I_{[h,8]}$) を生成する。本実験では、各時区間の印象の総和が、楽曲の印象に相当すると仮定し、時区間毎に生成された印象メタデータを合算する方式をとった (式 (3))。

$$I_{[h,l]} = \sum_{m=1}^f I_{[h,l,m]} \quad (3)$$

4. 実験

本実験では、複数のトラックから構成される実験データを対象として、提案方式、シングルトラック方式を適用し、それぞれの結果を比較する事で、提案メタデータ自動抽出方式の実現可能性、および、有効性を検証する。具体的には、次の実験を示す。

[実験1] 提案方式とシングルトラック方式を比較することで、トラック間合成機構の有効性、および、実現可能性を評価する。

[実験2] 時区間の分割を行わず、1曲単位でメタデータ抽出を行う方式、および、提案方式を比較することで、時区間合成機構の有効性を評価する。

また、各実験を行うために、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15曲を選択した。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3.2個) 選択し、正解メタデータとして定義した。さらに、これらの15曲を、楽曲の構造に応じて分割し、計154の区間楽曲データを作成し、主観により3個選択し、正解メタデータとして定義した。

4.1 実験1

4.1.1 実験手法

本実験では、本方式が有するトラック間合成機構の有効性を検証するため、区間毎に分割された楽曲集合を対象として、メタデータクリーニング機構の用いる閾値 r を、 $0 \sim 9$ の範囲で連続的に推移させ、提案方式、および、シングルトラック方式を用いてメタデータ抽出を行う。

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楽曲 M_h の部分区間の印象メタデータ $I_{[h,l,m]}$ 、および、正解印象メタデータ $A_{[h,l,m]}$ が、2件以上一致する楽曲を本実験における適合楽曲とすると、適合楽曲の集合 X は式 (4) で表される。また、C1~C8全ての語群において相関量が0となる楽曲を、本実験の対象外楽曲と定義すると、本実験の適用可能楽曲の集合 Y は、式 (5) で表される。

$$X = \{M_h | \text{count}((I_{[h,l,m]} > 0) \wedge (A_{[h,l,m]} = \text{true}) \geq 2)\} \quad (4)$$

$$Y = \{M_h | \text{count}(I_{[h,l,m]} > 0)\} \quad (5)$$

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、第1に、この適合楽曲集合 X 、および、適用可能楽曲集合 Y を用いて、式(6)で表す適合楽曲率で評価する。

$$P = \frac{\text{count}(X)}{\text{count}(Y)} \quad (6)$$

さらに、再現率、および、適合率を用いて、提案方式、および、シングルトラック方式が自動的に生成する印象メタデータの評価を行う。

なお、シングルトラック方式については、実験データが有するヴォーカル・メロディのみを選択的に使用し、適用するものとした。また、文献[13]において示した実験において結果が良好だったことから、提案方式は、重み付け決定要素として音量、重み付け関数として指数関数を用いた。

4.1.2 実験結果と考察

複数楽曲から構成される楽曲データを対象として、提案方式が、シングルトラック方式に比べ、高い再現率、適合率を示すことが確認された(図5)。図4、図5では、提案方式の結果をMT、シングルトラック方式の結果をSTとしてプロットしている。図5では、印象メタデータの有する印象語群との相関量の閾値 r を増加させることで(すなわちメタデータクリーニング機構の作用を強めることで)、誤って抽出されたメタデータを排除し、より正解に近いメタデータが抽出されることが確認された。

以上のことから、提案方式の有するトラック間合成機構が、既存の方式に比べ、複数のトラックから構成される楽曲を対象として、印象メタデータを、有効に抽出することを確認した。

4.2 実験 2

4.2.1 実験手法

本実験では、1) 区間毎に分割された楽曲集合を対象として、提案方式を適用し、また、2) 楽曲全体を対象として印象メタデータ抽出機構、および、トラック間合成機構を適用する。すなわち、提案方式の3次元マトリクス構造を有するメタデータ構造、および、時間の軸を排除した2次元マトリクス構造のメタデータ構造を設定し、時区間合成機構の有効性を検証する。評価には、実験1で用いた再現率、および、適合率を用いる。また、クリーニング機構の用いる閾値 r は、実験1の結果より、トラック合成機構により再現率と適合率が交差する12とした。

4.2.2 実験結果と考察

区間毎のメタデータを生成し、それらを合成することで、提案方式は、時間構造を有さないトラック間合成機能より多くの正解を、メタデータとして抽出した(図6)。これは、時間的に構造が変化する楽曲データ全体を対象として、相対的に短い単位音符列を前提とした Hevner の研究成果を適用することが、困難であることを示している。すなわち、楽曲全体を対象としてトラック間合成機能を適用することで、時間的に変化する調、リズム、テンポ、ピッチ

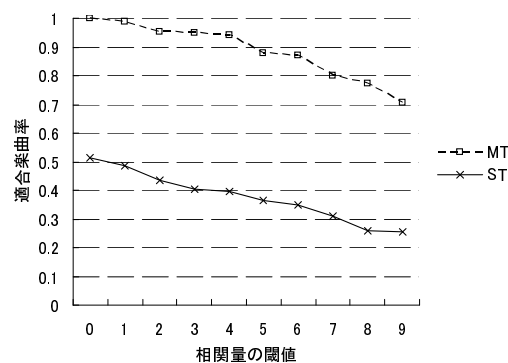


図4 音量を重み付け決定要素 ($0 \leq W \leq 1$) に用いた結果
Fig. 4 Results using the Velocity for the Weights ($0 \leq W \leq 1$)

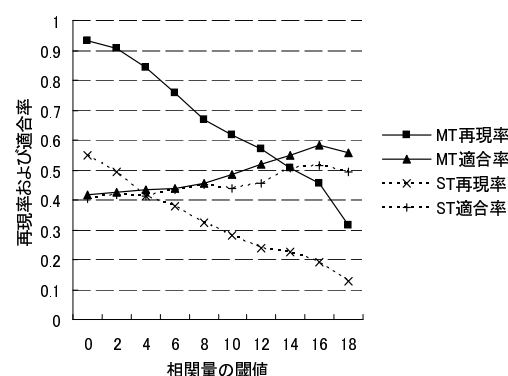


図5 音量を重み付け決定要素 ($0 \leq W \leq 1$) に用いた再現率と適合率
Fig. 5 Precision and Recall using the Velocity for the Weights ($0 \leq W \leq 1$)

チなどの楽曲構造要素が、統計的に集約されてしまうことで、メタデータの誤認を引き起こしていると考えられる。また、適合率に大きな変化は観測されなかった(図7)ことから、提案方式が、不正解を増加させず、正解を増加させたことが確認された。

以上のことから、提案方式の有する時区間合成機構が、印象メタデータの動的な抽出において、有効であることを確認した。

5. 結論

本論文では、複数の楽器から構成され、かつ、時間的に構造が変化する楽曲データを対象として、印象を表す形容詞群、および、その形容詞群との相関量を動的に抽出する印象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、時区間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、各音符列毎に計算された印象語との相関量を合成することで、楽曲に対応する印象メタデータを動的に生成する。また、実験により、提案方式と従来方式を比較することで、提案方式の有効性、および、実現可能性を示した。

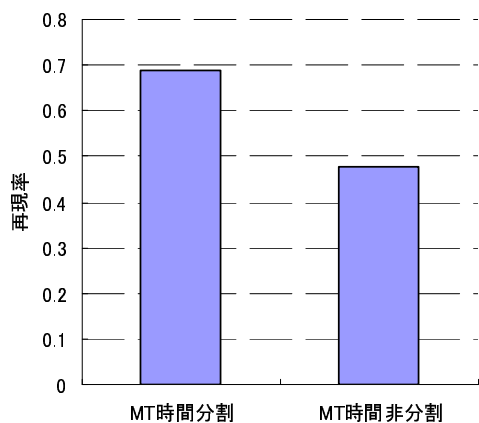


図 6 楽曲に対応する印象メタデータの再現率
Fig. 6 Recall of Impressionistic Metadata among Sections

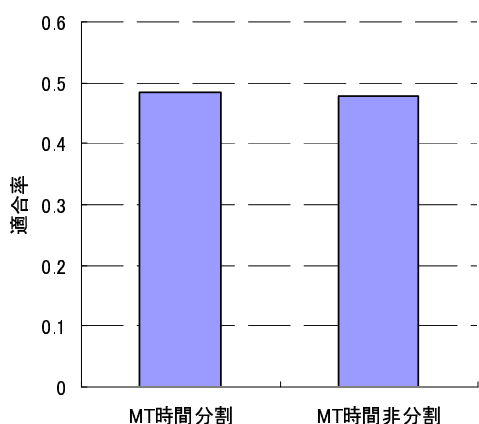


図 7 楽曲に対応する印象メタデータの適合率
Fig. 7 Precision of Impressionistic Metadata among Sections

今後の課題としては、リズム、音色に応じたメタデータの重み付け機能、トラック間の相互作用に応じた印象メタデータの重み付け方式を実現する予定である。

謝 辞

本研究に当たって貴重な御助言を頂いた筑波大学の北川高嗣先生、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筑波大学の高木秀幸氏、慶應義塾大学の大前寛子氏に感謝の意を表します。本研究の一部は、日本学術振興会学術創生研究プロジェクト「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。

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, 4B-4 (2001).
- [5] 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," 情報処理学会論文誌, Vol.42, No.12, pp.3201-3212 (2001).
- [6] 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).
- [7] ISO: "Generic coding of moving pictures and associated audio information - Part 3: Audio," ISO 13818-3 (1998).
- [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] 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," 情報処理学会研究報告, NL-147-6, pp35-40 (2002).
- [11] 熊本 忠彦, 太田 公子: "印象に基づく楽曲検索のための音楽作品データベース構築システム," 人工知能学会全国大会 (第 16 回) 論文集, 1E4-04, pp.1-4 (2002).
- [12] 小川 潤, 佐藤 聡, 北上 始: "感性に基づく音楽作品のための類似度計算方式," 情報処理学会 データベースと Web 情報システムに関するシンポジウム (DBWeb2000) 論文集, pp.229-234 (2000).
- [13] 大前 寛子, 石橋 直樹, 清木 康, 安西 裕一郎: "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," 情報処理学会研究報告, DBS-125-84, pp.145-152 (2001).
- [14] 佐藤 聡, 菊池 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," 情報処理学会研究報告, DBS-118-8, FI-54-8, pp.57-64 (1999).
- [15] 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).
- [16] 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 (DEWS1999) 論文集, 2B-5 (1999).
- [17] 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).