

擬音語入力によるクラシック音楽検索手法の提案

石原 健司[†] 木村 文則[†] 前田 亮[†]

[†] 立命館大学情報理工学部 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: [†] is003085@ed.ritsumei.ac.jp, fkimura@is.ritsumei.ac.jp, amaeda@media.ritsumei.ac.jp

あらまし 現在の楽曲検索システムには、アーティスト名や歌詞といった楽曲の情報を利用するものが多い。しかしユーザがこれらの情報を持っていない場合、曲を聴いただけではクラシック音楽やBGMなどの歌詞を用いない楽曲に対する検索は困難である。本論文では擬音語による問合せを入力とした楽曲検索手法を提案する。本手法ではクラシック音楽のメロディを擬音化し、入力文字列とのDPマッチングによりランキングを行う。メロディの擬音化にはMIDI形式で表現されたクラシック音楽のデータを対象としている。また本論文では提案手法の評価実験を行い、考察を行う。

キーワード 楽曲検索, 擬音語, DP マッチング

A Method for Classical Music Retrieval by Onomatopoeia

Kenji ISHIHARA[†] Fuminori KIMURA[†] and Akira MAEDA[†]

[†] College of Information Science and Engineering, Ritsumeikan University

1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577 Japan

E-mail: [†] is003085@ed.ritsumei.ac.jp, fkimura@is.ritsumei.ac.jp, amaeda@media.ritsumei.ac.jp

Abstract When we use a music retrieval system, we need information of music such as singers or music lyrics. However, if we do not know these music information, it is difficult to search music, especially for instrumental music. In this paper we propose a method for retrieving music by onomatopoeia. With the proposed method, this system shows ranking based on the result of DP Matching between user's query by onomatopoeia and classical music represented by onomatopoeia. This method uses onomatopoeia converted from MIDI of classical music. This paper shows the result of evaluation by the proposed method.

Keyword Music Retrieval, Onomatopoeia, DP matching

1. はじめに

現在の楽曲検索システムには歌詞やアーティスト名などのキーワードを利用して曲を絞り込むキーワード検索や、マイクを通して自らが歌うことで目的の曲を検索する音声検索がある。しかしキーワードを入力するシステムではユーザが歌手や歌詞などの楽曲に関する情報がわからない場合や、歌詞が用いられないインストゥルメンタル曲を検索することは困難である。また、歌って検索を行うシステムでは音程情報が利用できるが公共の場で歌うことが憚られる場合やユーザが音程を正しく歌えない場合には目的の曲を検索することは困難である。

このような問題の改善としてMP3の録音データを入力に利用し、類似した曲を検索する手法[1]や、感性語や楽曲の構成を利用して楽曲を絞り込む検索手法

[2]など様々な既存研究が存在する。しかし、このような曖昧な入力データを利用した検索手法では、ユーザが楽曲の特徴を明確に選択できず、検索目的である楽曲が決まっているような場合には不向きであるといえる。

そこで本研究では、楽曲の特徴を感覚で表現することの出来る擬音語に着目し、擬音語文字列でメロディを入力することで曲を検索する手法を提案する。擬音語とは一般的にパチパチやシトシトなどの物が発する音やその状態を字句で表現したものである。他人に曲を伝える際に、曲のフレーズを擬音語により表現することが多いため、本論文ではユーザが曲をハミングした際、その音を字句で表現したものを擬音語とする。擬音語で表現可能な楽曲情報はリズムのため、正しく音程を表現できないユーザの利用が考えられる。

検索対象にはインストゥルメンタル曲で、入手が容易であるクラシック音楽の MIDI データ[3][4]を利用する。

2. 擬音語での楽曲検索手法

2.1 概要

本手法では、Web から収集した MIDI データを擬音語文字列へと変換し、そのデータ群に対してクエリである擬音語文字列とのマッチングを行う。文字列の類似度を基にランキングを行い、曲の試聴や作曲者などの楽曲情報を付与することでユーザが目的の曲を判断しやすい形にして提示する。

2.2 全体の流れ

図 1 に提案手法の流れを示す。Web から収集した MIDI データに対してあらかじめ擬音語文字列へと変換する処理を施しておき、ユーザの入力した擬音語文字列に対して一曲毎にマッチングを行う。MIDI データの擬音化、マッチングについてはそれぞれ 3 節、4 節で述べる。マッチングにより出した類似度をランキング作成の基準とする。

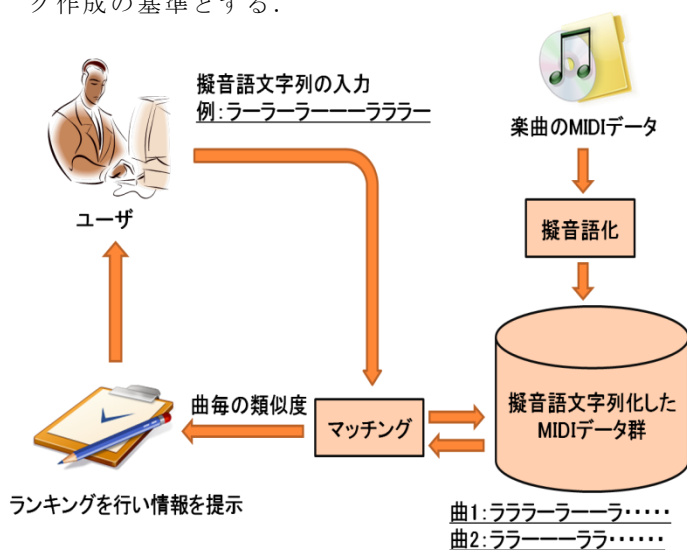


図 1 システムの流れ

3. MIDI データの擬音化

3.1 MIDI データのテキスト形式への変換

MIDI データを擬音語文字列へと変換する過程で、処理を行いやすいように必要な情報を抽出する必要がある[5]。MIDI データには「どの位置でどの音がどれだけの強さで鳴るか」といった情報が順に記述されている。ある MIDI データの例を図 2 に、その例をテキスト形式に変換したものを図 3 に示す。図 2 は Domino¹という MIDI 再生ソフトウェアで MIDI を見やすい形式で

表示したものである。縦軸が音程を表しており、横軸が時間を表している。4 分音符なら「480」、8 分音符なら「240」のように音の長さで値が設定されており図 3 ではその値を元に曲の情報を時系列順に記述している。利用する MIDI データは楽器毎にトラックが分けられているため主旋律を含むトラックを抽出し利用した。

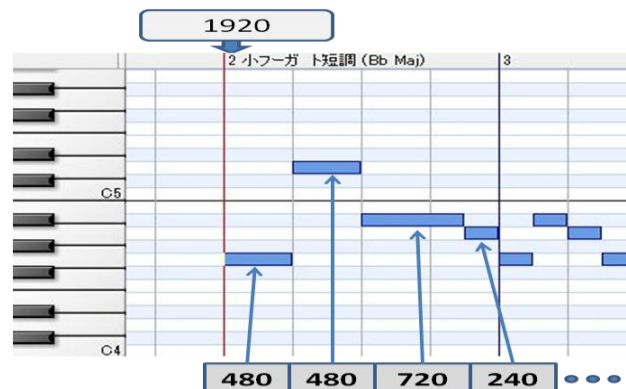


図 2 MIDI データの例

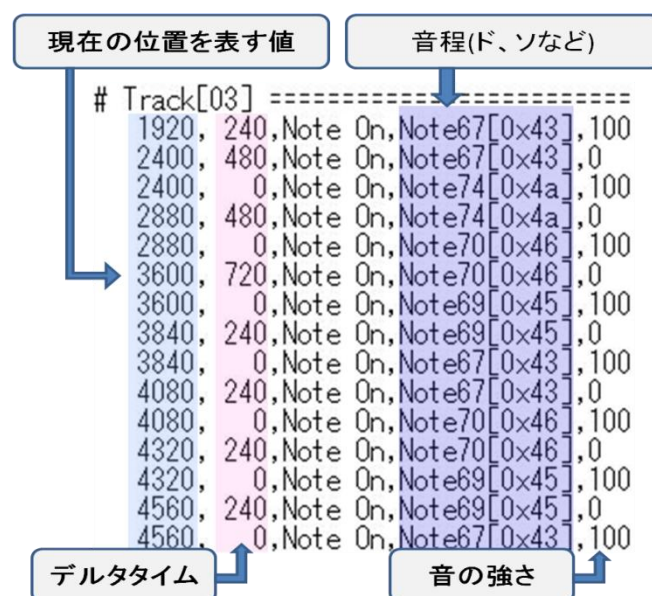


図 3 MIDI データをテキスト形式に変換した結果の例

3.2 MIDI イベントのテキスト形式による表示例

図 3 では、音が鳴り始める、鳴り終わるといったイベントが起こる毎に 1 行ずつその情報を記述している。1 行に記述されているコンマで区切られたそれぞれの値は「累計時間」、「デルタタイム」、「イベントの状態」、「音程」、「音の強さ」を表している。デルタタイムとは、直前に発生したイベントの累計時間から現在のイベントが発生するまでの差である。この値を利用することで音符の長さを判断し、擬音化を行う。

¹ <http://takabosoft.com/domino>

3.3 擬音語文字列化

3.3.1 アンケート

曲を聴いた際に人がどのような擬音語を利用するかを調査するため、日本人 10 人に対してアンケートを行った。アンケートではメロディがわかりやすい 5 曲を選択し、実際に擬音で曲を検索する場合を想定して入力をしてもらった。

3.3.2 擬音語の割り当て

アンケートにより収集した入力データを基に音符の長さに対する擬音語の対応を決定した。対応の一覧を表 1 に示す。本論文では、濁点のつく文字なら低い音、というように入力された文字の種類による音程の判別は行わず、「ラ」と「ー」（長音）に置き換えて擬音語文字列を作成している。「ー」の数に関しては音符が長いほどユーザによる入力に違いが表れるため、3 つ以上入力された場合は 3 つに置き換えている。曲のテンポにより各音符と擬音語の対応を変更している。この対応を基にフーガト短調の MIDI データを擬音語化した例を図 4 に示す。

表 1 擬音対応表

音の長さ	擬音文字
2分音符以上 (例)	ラ——
付点4分音符～2分音符 (例)	ラ——
付点8分音符～付点4分音符 (例)	ラー
32分音符～付点8分音符 (例)	ラ
32分音符以下	擬音として入力可能な長さでないため省略

*この例はテンポ90程度の曲の場合

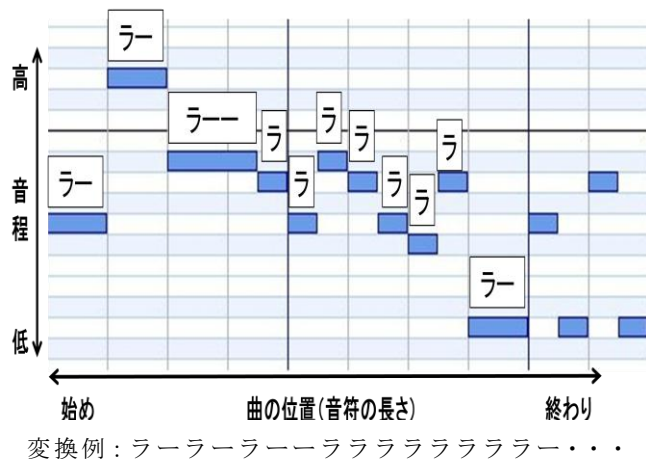


図 4 MIDI データの擬音語化の例

4. 擬音語文字列のマッチング

4.1 マッチングの流れ

第 3 節の要領で自動で擬音語文字列へと変換した曲に対してマッチングを行う。本論文では DP マッチングと呼ばれるマッチング手法を用いる。曲毎にユーザが検索に利用するであろうメロディを含むトラックに絞りマッチングを行い類似度を算出する。

4.2 DP マッチング

DP マッチング[6]の「DP」とは、動的計画法 (Dynamic Programming) の略であり、文字の一致やずれに対してコストを累積し、文字列同士の類似度をはかる手法である。今回は文字一致の場合はコストの減少はなし、文字不一致の場合では「+3」、字のずれでは「+1」といったようにコストを加算していく。よって累計コストが低いほど類似度が高い部分とみなす。一例を図 5 に示す。図 5 では「ラーラーララ」という文字列に対して「ララーラ」を入力擬音語文字列とした場合である。①であらかじめ文字の不一致部分を確認しておき、②で移動のコストも加えた累計コストを導き出す。②の表の左上端の「0」の位置を原点とし、隣接する 3 マスに進んだ場合最もコストが低くなる方向へと進み右下端まで計算を行う。図 5 の場合は累計コストが 4 となる。

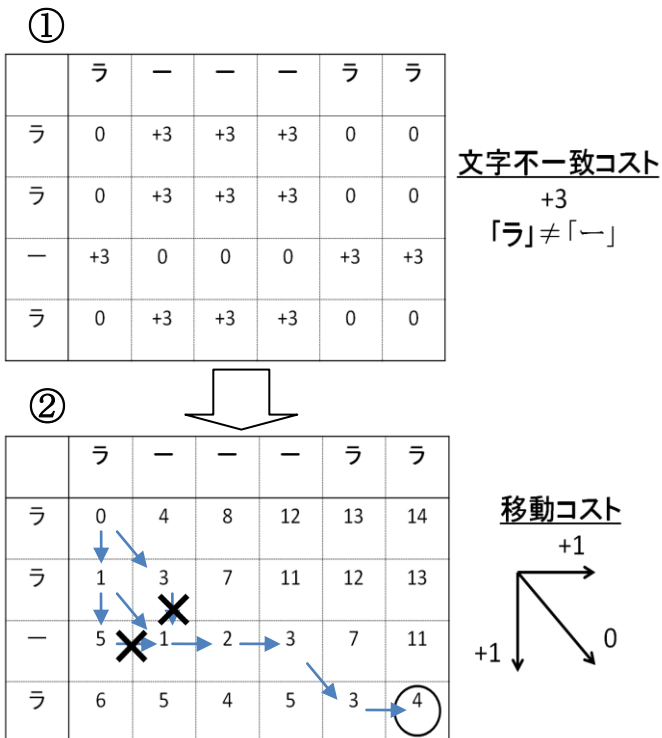


図 5 DP マッチング例

4.3 マッチングの手順

擬音語文字列へと変換した曲のトラックに対して DP マッチングを行う。ユーザが検索の対象としたメロディ部と、そのメロディ部を擬音語文字列で表現した検索クエリに含まれる音符の数は同じはずである。そのため図 6 に示すように入力された擬音語文字列の「ラ」の数を音符の数とし、曲の擬音語文字列に対して 1 音符ずつずらしながら DP マッチングを繰り返し、類似度の高くなる部分を探す。また特徴的なフレーズは楽曲中に複数出現することがある。ランキングで他曲との比較の際に曲中の最小累計コストが同じ場合には同順とはせず、図 7 のように 2 番目以降の低い値を比較することで低い値を持つ楽曲のランキング順位を高くする。

またマッチングを行う前準備として、付点 4 分休符以上の長さの休符をメロディの区切りとし擬音語文字列を分割している。これはユーザがメロディの途切れる部分の前後を同時に入力に利用することは考えられないためであり、また入力音符数より少なくノイズとなる部分を除去することが目的である。これらの理由から区切れを跨いでマッチングを行わないように擬音語文字列を分割している。その例を図 8 に示す。図 8 では入力クエリに含まれる音符数が 7 である場合、②のマッチングで休符前の音符がマッチングに利用される。休符を考慮しない場合、次のマッチングでは休符前の 6 音符と休符後の 1 音符でマッチングを行ってしまうため、図 8 の③に示すように休符後からマッチングを再開する。

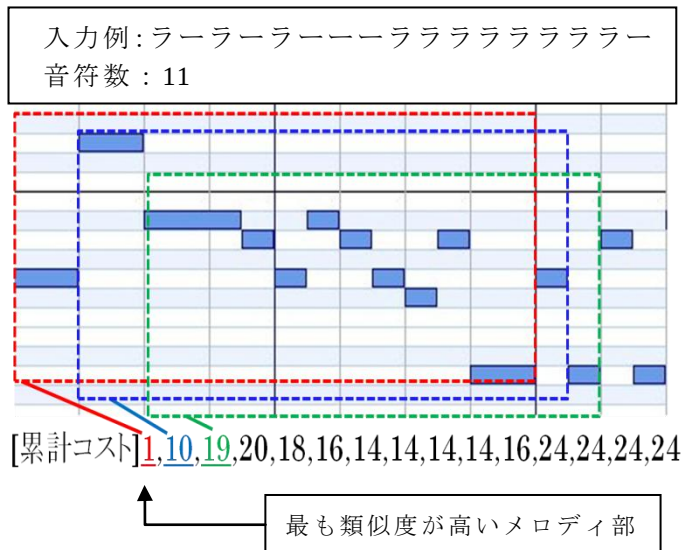


図 6 マッチングの手順の例

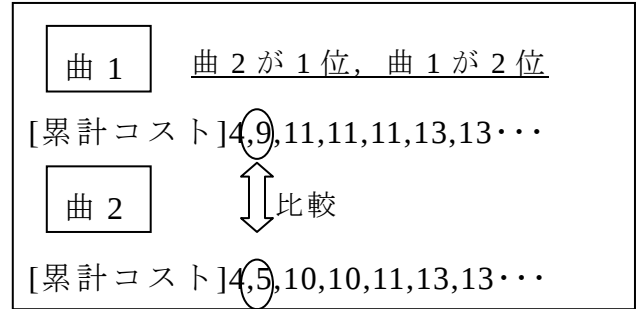


図 7 ランキングの決定

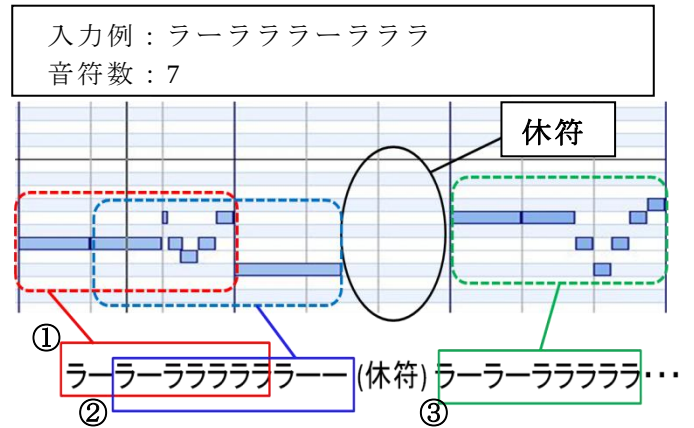


図 8 メロディが途切れる場合のマッチング

5. 評価実験

表 2 の 10 曲のデータに対して 3 人のユーザに入力データを作成してもらい、合計 30 件の入力データを用意し、マッチングにより類似度を計算した。1 音符ずつにマッチングを行うことにより、1 曲から複数の値が結果として出力されるため、各曲の最小コストに着目しランキングを作成した。その結果を表 3 に示す。「曲 1」、「曲 2」といった番号は曲を表しており、行の左に記している①、②などの数字はユーザの番号である。「曲 1」を例に説明すると、「曲 1」の入力文字列として①、②、③の 3 人のユーザが作成した入力データでマッチングを行っている。最小コストが同じ場合は、それ以降のコストを見ることで順位を決定している。またランキングとは別に検索精度の評価指標として MRR (Mean Reciprocal Rank, 平均逆順位) を用いて評価を行う。

$$RR_i = \frac{1}{n} \quad (n \text{ は正解が出現した順位})$$

$$MRR = \frac{1}{N} \sum_{i=1}^N RR_i \quad (N \text{ は検索件数})$$

検索クエリ毎に初めて正解が出現した順位の逆数を求め、全検索クエリの RR を平均することで、システムの検索精度を評価する。最大値が 1、最小値が 0 であるため、1 に近いほど検索精度が良いと言える。

表 2 使用したクラシック音楽

	曲名	作曲者
曲 1	小フーガト短調 BWV578	ヨハン・ゼバスティアン・バッハ
曲 2	前奏曲 作品 28 第 7 番 イ長調	フレデリック・フランソワ・ショパン
曲 3	前奏曲 作品 28 第 15 番 変ニ長調	フレデリック・フランソワ・ショパン
曲 4	パヴァーヌ	ガブリエル・ユルバン・フォーレ
曲 5	アヴェ・マリア	シャルル・フランソワ・グノー
曲 6	ペール・ギュント第 1 組曲 朝の気分	エドヴァルド・ハーゲルupp・グリーグ
曲 7	アヴェ・ヴェルム・コルプス ニ長調 K.618	ヴォルフガング・アマデウス・モーツァルト
曲 8	白鳥の歌 セレナーデ	フランツ・ペーター・シュューベルト
曲 9	アルルの女 第 2 組曲 メヌエット	ジョルジュ・ビゼー
曲 10	ベルガマスク組曲 月の光	クロード・アシル・ドビュッシー

表 3 ランキングの結果

	曲1	曲2	曲3	曲4	曲5	曲6	曲7	曲8	曲9	曲10
曲1	① 1位(0)	9位(21)	4位(5)	2位(4)	7位(10)	8位(11)	10位(22)	5位(5)	3位(5)	6位(6)
	② 4位(8)	9位(22)	7位(12)	3位(6)	8位(15)	2位(6)	10位(30)	5位(9)	1位(6)	6位(9)
	③ 2位(4)	10位(20)	4位(4)	7位(5)	8位(7)	1位(2)	9位(18)	5位(4)	3位(4)	6位(4)
曲2	① 4位(8)	1位(2)	6位(9)	3位(4)	9位(13)	2位(2)	5位(8)	10位(17)	7位(10)	8位(11)
	② 6位(3)	1位(1)	5位(3)	3位(2)	9位(6)	2位(1)	8位(5)	10位(6)	4位(3)	7位(4)
	③ 4位(7)	2位(3)	5位(9)	3位(5)	9位(12)	1位(1)	7位(9)	10位(16)	6位(9)	8位(10)
曲3	① 6位(11)	2位(7)	1位(2)	3位(8)	7位(11)	4位(9)	8位(11)	10位(18)	9位(11)	5位(10)
	② 5位(5)	10位(23)	8位(5)	6位(5)	3位(4)	2位(3)	9位(21)	1位(1)	4位(5)	7位(5)
	③ 4位(3)	3位(3)	6位(4)	2位(2)	9位(7)	1位(1)	7位(5)	10位(10)	5位(3)	8位(6)
曲4	① 7位(7)	9位(15)	8位(8)	1位(2)	2位(4)	6位(7)	10位(15)	3位(4)	5位(6)	4位(5)
	② 6位(7)	9位(13)	5位(6)	1位(1)	4位(6)	7位(7)	10位(13)	2位(2)	8位(7)	3位(5)
	③ 5位(4)	10位(18)	1位(1)	3位(3)	8位(9)	2位(2)	9位(16)	6位(4)	4位(4)	7位(4)
曲5	① 4位(2)	8位(6)	7位(4)	3位(2)	1位(0)	9位(7)	2位(2)	10位(7)	6位(2)	5位(2)
	② 5位(2)	9位(6)	6位(4)	4位(2)	1位(1)	10位(6)	2位(1)	3位(1)	8位(4)	7位(4)
	③ 2位(5)	9位(21)	6位(8)	3位(6)	1位(4)	5位(7)	10位(23)	4位(6)	7位(9)	8位(9)
曲6	① 3位(8)	9位(71)	7位(21)	4位(12)	8位(24)	1位(5)	10位(84)	5位(14)	2位(8)	6位(17)
	② 2位(8)	9位(68)	7位(18)	6位(16)	8位(26)	1位(8)	10位(81)	4位(12)	3位(8)	5位(14)
	③ 4位(8)	5位(10)	6位(10)	3位(8)	8位(14)	1位(0)	10位(16)	9位(14)	2位(0)	7位(13)
曲7	① 9位(21)	2位(6)	5位(13)	3位(12)	7位(16)	4位(13)	1位(1)	10位(24)	8位(21)	6位(14)
	② 2位(2)	6位(6)	3位(2)	5位(5)	9位(10)	4位(3)	1位(0)	10位(13)	8位(8)	7位(7)
	③ 4位(5)	10位(7)	7位(5)	3位(5)	1位(4)	5位(5)	9位(7)	2位(4)	6位(5)	8位(6)
曲8	① 6位(12)	9位(37)	8位(13)	2位(4)	7位(13)	5位(11)	10位(47)	1位(0)	3位(8)	4位(10)
	② 5位(4)	9位(16)	7位(7)	4位(4)	2位(1)	8位(10)	10位(16)	3位(2)	1位(1)	6位(7)
	③ 1位(2)	9位(32)	6位(6)	3位(5)	8位(10)	2位(4)	10位(38)	4位(5)	5位(6)	7位(7)
曲9	① 5位(8)	9位(15)	4位(6)	3位(5)	8位(11)	1位(0)	10位(16)	7位(10)	2位(0)	6位(9)
	② 2位(4)	9位(50)	8位(12)	4位(7)	6位(10)	3位(4)	10位(51)	5位(8)	1位(0)	7位(10)
	③ 1位(0)	9位(52)	8位(12)	4位(7)	7位(12)	2位(0)	10位(55)	5位(12)	3位(2)	6位(12)
曲10	① 4位(8)	9位(22)	6位(11)	3位(8)	1位(5)	7位(12)	10位(22)	2位(6)	8位(16)	5位(8)
	② 1位(4)	10位(27)	6位(6)	3位(5)	8位(9)	5位(5)	9位(26)	7位(6)	2位(5)	4位(5)
	③ 2位(2)	9位(32)	6位(6)	5位(5)	7位(8)	4位(4)	10位(32)	1位(1)	3位(3)	8位(10)

全クエリでのMRR = 0.67

6. 考察

ランキングの結果では、30 件中 16 件で正解の楽曲が 1 位、1 位から 3 位以内では 22 件が正解の楽曲という結果となった。残りの順位の低かった 8 件には大きく分けて二つの傾向が見られた。

一つ目の傾向は、正解の楽曲自体の類似度も高いが他曲の類似度がより高くなるような場合である。これは正解の楽曲の類似度が問題ではなく、他曲の類似度

が高いことが問題となっている。この原因としては、本手法では「ラ」と「ー」の 2 種類で擬音語文字列を作成しているため、類似する部分が表れやすく、DP マッチングの文字不一致によるコストの増減の影響が少ないことが挙げられる。またユーザによる入力擬音語文字列の長さや特徴の少なさが、他曲でも類似度の高いと判定される部分が表れる要因となっている。

順位の低かった入力データのもう一つの傾向は、正

解の楽曲自体の類似度が低い場合である。正解の楽曲自体の類似度が低くなるのは、ユーザが「ラ」と記述しているがシステム側の擬音語の割り当て（表 1）では「ラー」と記述しているといった場合が多くみられた。この問題が顕著に表れた場合、図 9 のように類似度が極端に低下する例も見られた。この例ではユーザによって 4 分音符部の入力に「ラ」の場合と「ラー」の場合の 2 通りが見られたことで累計コストが大きく変化している。このようにユーザの音の長さの感じ方の違いがマッチングに大きく影響を与えている。

ランキングの結果がよかった入力データに関して、1 位の正解楽曲と 2 位の他曲との類似度の差があまりない場合がある。そのため正解の楽曲以外の曲の類似度が高くなる問題は最も精度に影響を与えていることが伺える。

MRR での評価結果は全クエリで 0.67 の精度であるがユーザ毎に見た場合、①が 0.87, ②が 0.70, ③が 0.43 と精度に大きな差が見られる。精度の低かったユーザの入力データを参照すると、前述した問題の他にメロディを擬音語で入力する際に正しい音符の数が入力できていない入力データが存在した。メロディの中間の音符が抜けている場合、それ以降の音符のマッチングがうまくいかないと考えられる。

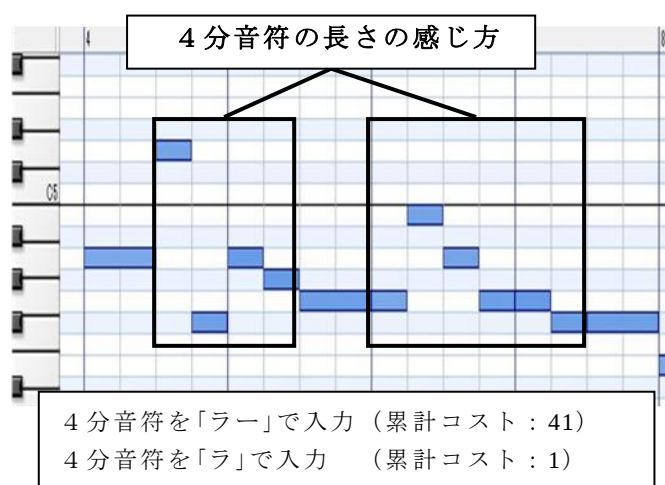


図 9 ユーザの感じ方の違い

7. おわりに

本論文ではメロディは思い出せるが楽曲の情報がわからないという場合でも目的の曲を検索するために、擬音語文字列を入力とする曲の検索手法を提案し評価実験を行った。

本研究の今後の課題としては、ユーザの音の長さに対する感じ方の問題に柔軟に対応する手法の実装が挙げられる。現状の手法では、入力データを作成するユ

ーザの音の感じ方次第で類似度に影響が表れるため、様々な入力パターンを想定しマッチングを行うことで結果の向上が期待できる。また今回利用していない音程などの特徴を用いることで、複数表れる曲のメロディ部や入力に利用されないであろう部分を特定し重みをつけることで、類似度にあまり違いが表れない問題の改善が期待できる。今回行った評価実験では入力パターンや対象とする楽曲の数が少ないため、さらにアンケートを行い、ユーザの入力パターンを収集する必要がある。またキーワード検索の要領での擬音語入力では入力が難しいと感じるため、感覚的に問合せを入力できるインタフェースの導入が必要であると思われる。

また音声の入力を行わずとも、前の音より高いか低いかというような音程の入力方法も考えられる。音程の情報を用いる場合、他の曲との類似度の差を明確にすることが可能であるが、ユーザの問合せの負担や難しさが増化すると考えられる。そのため、ユーザ毎に問合せに利用する情報を選択可能なシステムの実現が必要である。

参考文献

- [1] 大西 正志, 辻 紗千, 獅々堀 正幹, 北 研二. 時系列アクティブ探索による類似音楽検索システムの改良. 第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM 2011) 論文集, 2011
- [2] 梶 克彦, 長尾 確. 楽曲に対する多様な解釈を扱う音楽アノテーションシステム. 情報処理学会論文誌 48(1), -pp. 258-273, 2007.
- [3] Classic Music MIDI Nocturne: <http://nocturne.vis.ne.jp/>
- [4] Classical Music on Classical Archives: <http://www.classicalarchives.com/index.html?navID=0>
- [5] 斧山 青矢, 吉田 真一. フレーズ毎の音楽特徴を用いた楽曲検索システムの提案. 高知工科大学情報システム工学科 学士学位論文 2011 <http://www.kochi-tech.ac.jp/library/ron/2010/2010info/1110381.pdf>
- [6] 内田誠一: DP マッチング概説: 基本と様々な拡張, 電子情報通信学会技術研究報告(特別講演), PRMU2006-166, pp. 31-36, 2006.