

音楽ゲームにおける類似譜面検索システムの開発

Development of Similar Musical Sheet Search System for Music Game

光武 聖悟^{*1}, 山口 真之介^{*1}, 近藤 秀樹^{*1}, 大西 淑雅^{*1}, 西野 和典^{*1}
 Seigo MISTUTAKE^{*1}, Shin-nosuke YAMAGUCHI^{*1}, Hideki KONDO^{*1},
 Yoshimasa ONISHI^{*1}, Kazunori, NISHINO^{*1}

^{*1}九州工業大学

^{*1}Kyusyu Institute of Technology

Email: q231076s@mail.kyutech.jp

あらまし: 音楽ゲームにおいて, プレイヤーは, 習熟のために, 曲に合わせたさまざまな体の動きを習得しなければならない. 目的の動きを練習するためには, その動きを要求する類似した譜面を見つける必要がある. そこで, 本研究では, 既存の曲の譜面から, 練習したい類似の譜面を検索するシステムを開発した.

キーワード: 音楽ゲーム, 譜面, 検索システム

1. はじめに

プレイヤーは習熟のために曲に合わせて, さまざまな体の動きを習得しなければならないが, 単純にゲームをプレイするだけでは上達は見込めない. pop'n music は 1500 曲を超え, 各曲に最大 4 つの難易度が存在するため, 5000 を超える譜面数が存在する. ある譜面に対応した動きを練習するためには, その動きを要求する類似した譜面を見つける必要があるが, 数多くの譜面の中からネット上で一つ一つの譜面画像を見て, 目的の譜面を探し出すことは容易ではない. そこで, 本研究では, 既存の曲の譜面から, 練習したい類似の譜面を検索するシステムを開発する.

2. システム概要

本システムは, 図 1 で示す通り, 譜面画像をシステムのウィンドウに入れドラッグ&ドロップすることで, その画像からテンプレートマッチングでポップ君を認識し, ポップ君(図 2 の楕円一つ)の落ちてくる順番をテキストデータとして取得する. その後, データベースに登録されているテキストデータを Bi-gram で分割し, Bi-gram での出現回数からベクトル化(Bag of Words)を行い, Cos 類似度を用いて, ポップ君の落ちてくる順番が類似した譜面を検索し, その譜面をウィンドウに表示を行うことで, 類似譜面の検索を実現する.

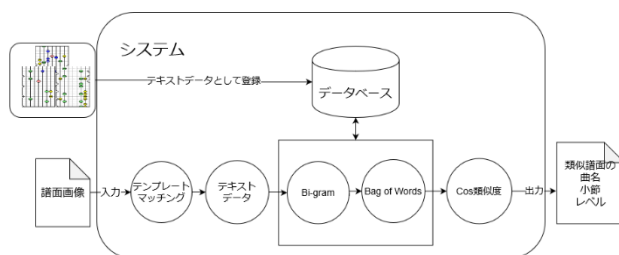


図 1 システムの全体図

2.1 テンプレートマッチング

テンプレートマッチングとは, 入力画像に対して, テンプレート画像と同じサイズの検索窓をスライドしていきながら, 照らし合わせて類似度を算出し, 類似度が高い検出対象があると判断する画像処理の一種である.

本研究では, 譜面画像をレーンごとに分割し, レーン番号を 1~9 に振り分ける. それぞれのレーンに対応したテンプレート画像を用いて, テンプレートマッチングを行い, マッチングした座標とレーン番号を取得する.

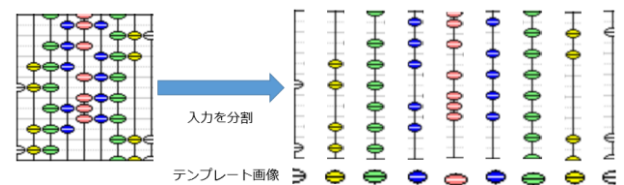


図 2 テンプレートマッチング概要

2.2 テキスト化

9 つのレーンを左から 1~9 までの番号を振り分け, ポップ君が落ちてくる順番をレーンの番号と合わせて, テキスト化を行う. 違うタイミングで押すかどうかを判別するために, 区切り文字“, ”を用いる. 同じタイミングで押す場合は区切り文字は入れない. 図 3 はテキスト化を行うと, ‘1, 2, 3, 4, 5, 6, 7, 8, 9’ となる.

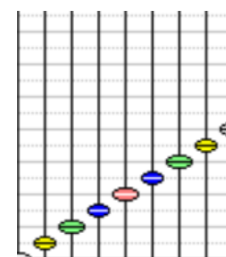


図 3 譜面画像

2.3 Bi-gram

Bi-gram とは一般的に任意の文字列を 2 文字単位に分割する手法であり、一般的に文書を、形態素解析を行わず見出し語を作成するために用いられる。

本研究では、同タイミングで押すポップ君は一つの文字として扱い、2 文字単位で分割を行った。

テキスト : "5,37,37,37,37,34567,19,5,5,5"
 Bi-gram : "5,37","37,37","37,37", ..., "5,5"

図 4 Bi-gram の手法

2.4 Bag of Words

Bag of Words は文章に出現する文字の出現回数を特徴とする手法である。本研究では、Bi-gram によって分割されたデータを出現回数によってベクトル化を行う。

テキスト : " 5,37,37,37,34567,19,5,5,5"

Bi-gram : "5,37","37,37","37,37","37,34567","34567,19","19,5","5,5","5,5"

ベクトル : [1, 2, 1, 1, 1, 2]

図 5 Bag of Words の手法

2.5 Cos 類似度

ベクトル化したテキストデータから類似度を出力するアルゴリズムとして Cos 類似度を使う。Cos 類似度は次の式で定義される。

$$\text{Cos}(A, B) = \frac{A \times B}{|A| \times |B|}$$

2.6 データベース

データベースには、pop'n music に収録されている 20 種類の譜面をプログラムと同様に手動でテキスト化したものが登録されている。手動で作成したテキストデータがプログラムで作成したテキストデータとの差異がないかを調べるために、データベースに登録されている 5 つの譜面を譜面画像からプログラムでテキスト化を行い、比較したものを表 1 に示す。

表 1 : プログラムによるテキスト化の認識率

曲	総ポップ君数	不認識数	認識率
A	543	0	1
B	496	1	0. 998
C	446	1	0. 998
D	592	0	1
E	735	5	0. 993

3. 実行結果

入力画像から類似度が高い順に 5 つの検索結果が出力される。1 つの検索結果に、類似度、検索されたテキストデータ、曲名とその譜面の難易度、検索結果が現れる。小節、曲のレベル、曲の bpm(beats per minute) の 6 項目が順に表示される。

図 6 は、図 3 の譜面の類似譜面を検索した結果であり、類似度、曲名とその譜面の難易度、小節、曲のレベルのみを抽出した出力結果である。図 7 は図 6 の 1 行目のテキストの譜面画像、図 8 は図 6 の 3 行目のテキストの譜面画像である。

図 7 が図 8 よりも類似度が高い理由はテンプレートマッチングしてテキストデータに変換したときに、拍子の違い (ポップ君の間隔) を考慮していないからである。拍子の違いを考慮しなかった要因の一つに bpm の違いがある。bpm が 2 倍違うと譜面画像ではまったく同じであっても実際のプレイでは速度が 2 倍異なる。そして、Cos 類似度以外の新たなアルゴリズムによる評価が必要とされるため、今回は実装しなかった。

類似度:0.935, 猫侍の逆襲_H, 小節:4, level:38
 類似度:0.935, めうめうべったんたん!!_H, 小節:63, level:38
 類似度:0.926, 打打打打打打打打打打_H, 小節:38, level:38
 類似度:0.866, TOXIC VIBRATION_N, 小節:8, level:30
 類似度:0.802, お天気とチョコレート_H, 小節:39, level:30

図 6 図 3 の譜面画像を入力した検索結果

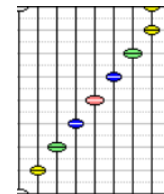


図 7 猫侍の逆襲_H、4 小節目

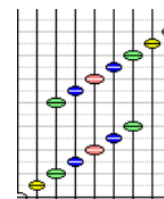


図 8 打打打打打打打打打打_H、38 小節目

4. おわりに

本研究では、プレイヤーの習熟を支援することを目的とした類似譜面検索システムの開発を行った。このシステムは bpm、ボタンの押しリズム、非類似性の重み付けを考慮していないため、入力と検索結果の難易度の差が大きく異なる可能性がある。類似した譜面画像が実際のプレイで似ているかどうかの評価が必要である。