

## SVM を用いた音楽ゲーム譜面自動生成の検討 Study on auto-generation of a music game score using SVM

柴崎 大地<sup>†</sup>      酒井 充<sup>†</sup>      丸山 博<sup>†</sup>  
Daichi Shibasaki   Mitsuru Sakai   Maruyama Hiroshi

山下 和也<sup>†</sup>      広瀬 貞樹<sup>†</sup>  
Kazuya Yamashita   Sadaki Hirose

### 1. はじめに

音楽ゲームとは、リズムや音楽に合わせて画面に表示される指示に従って、プレイヤーがボタンを押したり、画面をタッチすることで、得点を獲得するゲームである。音楽ゲームにはプレイヤーに指示するアクション等の情報(ノート)を記述したゲーム譜面が使用される。ゲーム譜面を自動生成するプログラムはすでに開発されているが、生成された譜面に対するプレイヤーの評価は良いとは言えず、現状ではほとんど人間が作った譜面がプレイに使用されている。

ゲーム譜面の自動生成には音楽音響信号を用いる自動採譜の技術が応用できると考えられる。自動採譜に関する問題は大きく分けると、音響信号から各楽音の発音開始時刻と基本周波数を推定する多重音解析の問題、そして多重音解析で得られた情報から音符列に変換するリズム解析の 2 つの問題がある。ゲーム譜面自動生成の場合、音程の出力をしないため、楽音の発音開始時刻を推定する問題と、その情報からノート列を構成する問題になると考え、本稿では、後者の問題について取り組む。

自動採譜では、音高と発音時刻を同時に推定する必要があり、非負値行列因子分解(以下、NMF)を使った推定が提案されている[1]。これはスペクトログラムを行列とみなし、スペクトルが非負値スペクトルパターンとの重ね合わせで表現できるという仮定で分解を行うことで、基底スペクトルとその音量の変化の行列に分解できる。

本研究では、人間がノートありと判断しやすい音量の変化パターンがあると考え、人間が作った譜面と音響信号をもとに Support Vector Machine(以下、SVM) の学習を行い、その結果を用いて譜面を生成する方法を検討し、さらにその有効性の検討を行う。

### 2. 提案手法

ゲーム譜面の先頭には、楽曲の速さを表す Beat Per Minute (以下、BPM)が設定されている。これは、BPM から楽曲の拍の長さがわかり、ノートの入力位置を絞り込めるためである。本研究では各楽曲の BPM 情報とオフセットをあらかじめ入力しておき、ノートの入力位置候補を決め

たうえで実験を行う。また、ノート間の最短時間は 32 分音符の長さとする。

一般にプレイヤーへの操作のタイミングとその時のアクション内容という二つの情報の組をノート、一楽曲分のノート列を譜面と呼ばれるが、本研究では自動譜面生成の第一段階としてノートのタイミングのみの生成を行う。

入力する音響信号  $\mathbf{w}$  は、サンプリング時刻を  $t_n (n = 1, \dots, N)$  とすると、次式で表わされる。

$$\mathbf{w} = (w_n)_{n=0}^{N-1}$$

なお、 $\mathbf{w}$  は各楽曲をサンプリング周波数  $f_s = 44.1$  [KHz] でサンプリングしたものである。また、ノート入力可能位置  $i$  は、ノートの入力可能位置の数を  $M$  とすると、次式で表わされる。

$$\mathbf{i} = (i_m)_{m=0}^{M-1}$$

$$i_m = (m \times L + a) / f_s$$

ここで  $L$  は 32 分音符の長さ、 $a$  はオフセットである。図 1 は横軸を時間としたときの音響信号  $\mathbf{w}$  とノート入力可能位置  $\mathbf{i}$  の例である。また、譜面  $\mathbf{y}$  は次式のように表す。

$$\mathbf{y} = (y_m)_{m=0}^{M-1}$$

ここで、

$$y_m = \begin{cases} 1 & (i_m \text{ にノートあり}) \\ 0 & (i_m \text{ にノートなし}) \end{cases}$$

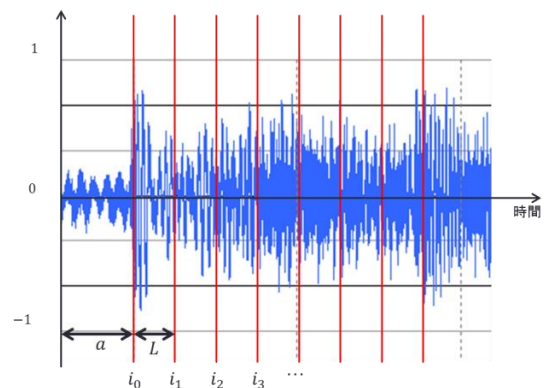


図 1 入力音響信号とノート入力可能位置

<sup>†</sup> 富山大学, University of Toyama

学習には人間が作った譜面  $y_h$  をもとに正例 ( $y_{h,m} = 1$ ), 負例 ( $y_{h,m} = 0$ ) として SVM の学習を行う. また,  $i_m$  における特徴ベクトル  $x_m$  は次式で与える.

$$x_m = (x_{m,d})_{d=0}^D$$

ここで,  $x_{m,d}$  を求めるときのサンプル数を  $J$  とし,

$$x_{m,d} = \frac{1}{J} \sum_{j=0}^{J-1} w_{k_{m,d}+j}^2$$

$$k_{m,d} = i_m - \left\lfloor \frac{D}{2} \right\rfloor \times J - \left\lfloor \frac{J+1}{2} \right\rfloor + J \times d$$

よって  $x_m$  を求めるためのデータ区間長は  $J \times D \times fs$  である. 次の図 2 は  $J \times D \times fs = L$ ,  $D = 5$  の時の  $x_m$  を求めるイメージである. 図の各  $J$  の範囲における  $w$  の二乗平均が  $x_m$  の各要素となる.

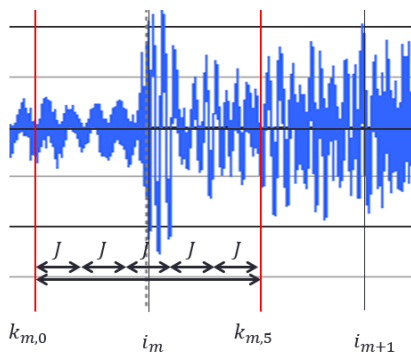


図2  $x_m$  を求める際の状況

### 3. 実験

実験に使用する楽曲は4分の4拍子であり, BPMや曲の長さはばらばらである. 使用する譜面数は25個である.

#### 3.1 同一譜面内の cross validation

1つの譜面を10分割し, 譜面ごとに cross validation を行う. データ区間長は楽曲によって変動するように  $60/BPM/8$  (32分音符の長さ) と  $60/BPM \times 4$  (1小節の長さ) で与え, 次元数  $D = 32$  で実験を行う. なお, 学習するときに入力位置の候補をすべて譜面に記述しているため, 負例が正例よりかなり多くなるので, オーバーサンプリングで正例を増やして学習を行う. また, SVM は Weka[2] を使用する.

評価項目として, 人が作った譜面を正解として, 適合率, 再現率, F値を求める. 表1は25個の譜面それぞれに対し, cross validation を行い, 適合率, 再現率, F値の平均をとったものである.

表1 各譜面の評価の平均

データ区間長	適合率(%)	再現率(%)	F値(%)
60/BPM/8	64.60	64.69	64.64
60/BPM $\times$ 4	74.06	73.97	73.90

#### 3.2 全譜面の cross validation

25個の譜面を4グループに分け, cross validation を行う. パラメータ設定は3.1と同一である. 次の表2はその結果である.

表2 全譜面での結果

データ区間長	適合率(%)	再現率(%)	F値(%)
60/BPM/8	55.45	55.40	55.20
60/BPM $\times$ 4	44.45	44.98	44.10

### 4. 考察

表1を見るとデータ区間長  $60/BPM \times 4$  の方が良い結果となっている. このことからノートを決定する時, 32分音符の短い時間のみで判断せず, 広い時間で前後の関係をみたうえでノートを決定する方が良いと考えられる. 一方, 表2では,  $60/BPM \times 4$  の方が悪くなっている. これは, 楽曲ごとの変動が大きいためと考えている. 今後はこの原因を検討していき, 他の特徴も取り入れていこうと考えている.

### 5. おわりに

本稿では, 音楽ゲーム譜面自動生成は, 楽音の発音開始時刻を推定する問題と, その情報からノート列を構成する問題になると考え, 後者について取り組んだ. 人間の作った譜面と音響信号を元に SVM の学習を行い, 譜面を生成したが, F値 55.2% が得られた.

今後は, 他の特徴も検討していき, よりよいゲーム譜面生成を目指したい.

#### 参考文献

- [1] 落合 和樹, 中野 允裕, 小野 順貴, 嵯峨山 茂樹 "時間周波数分解能の異なるスペクトログラムの並列 NMF による多重音解析" 情報処理学会研究報告, 2011
- [2] Weka3.8.1 (閲覧日: 2017/06/28)  
<http://www.cs.waikato.ac.nz/ml/weka/>