情報処理学会論文誌 Vol.61 No.11 1718–1728 (Nov. 2020)

ダンスゲーム譜面の特性分析とクラスタリングに基づく特徴的な譜面の自動生成

辻野 雄大[1,a)] 山西 良典[2,b)] 山下 洋一[3,c)] 井本 桂右[4,d)]

受付日 2020年1月28日, 採録日 2020年9月10日

概要：ダンスゲーム譜面は，操作を要求される頻度，リズム構成の複雑さなど，複数の要素によって特徴づけられる．これらの要素によって，ダンスゲーム譜面の難しさおよび面白さに影響する特性が形成される．この特性を客観的かつ多次元で表現する特徴量を提案する．提案した特徴量に基づいて k-means 法によるクラスタリングを実施し，特性が類似した譜面のクラスタを得る．クラスタごとに楽曲と譜面の関係を深層学習させることによって，クラスタの特性を備えた譜面の自動生成が可能となり，多様な面白さに対応した特徴的な譜面の提供を実現した．

キーワード：ダンスゲームの自動生成，ダンスゲームの分析，Procedural Content Generation

Generation of Characteristic Charts Based on Analyzing and Clustering Charts of Dance Games

Yudai Tsujino[1,a)] Ryosuke Yamanishi[2,b)] Yoichi Yamashita[3,c)] Keisuke Imoto[4,d)]

Received: January 28, 2020, Accepted: September 10, 2020

Abstract: Multiple factors, e.g., the frequency of action and the complexity of rhythm, may characterize the charts on dance games. Those factors affect the difficulty level and fun of the charts. This paper designs objective and multi-dimensional features concerning such factors. As conducting the k-means clustering with the designed features, the characteristically similar charts structure the clusters. The relationships between songs and charts are leaned for each cluster by the deep neural network. It was confirmed that our approach realized to automatically generate the charts with the characteristics corresponding to the learning cluster. The characteristic charts for each of diverse “fun” could be provided.

Keywords: automatic generation of dance games, analysis of dance games, procedural content generation

1 立命館大学大学院情報理工学研究科
Graduated School of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525–8577, Japan
2 関西大学総合情報学部
Faculty of Informatics, Kansai University, Takatsuki, Osaka 569–1095, Japan
3 立命館大学情報理工学部
College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525–8577, Japan
4 同志社大学理工学部
Faculty of Science and Engineering, Doshisha University, Kyotanabe, Kyoto 610–0394, Japan
a) is0221rs@ed.ritsumei.ac.jp
b) ryama@kansai-u.ac.jp
c) yyama@is.ritsumei.ac.jp
d) keisuke.imoto@ieee.org

1. はじめに

今日のビデオゲーム市場で，大きなシェアを占めているジャンルに「音楽ゲーム」が存在する．音楽ゲームの代表例として，*Dance Dance Revolution*® (DDR) などのダンスゲームがあげられる．**図 1** に，ダンスゲーム画面の例を示す．多くの音楽ゲームでは，画面に表示される指示符に沿って，プレイヤが何らかの動作を行い，その動作のリズムに対する成否を競う．ダンスゲームでは，プレイヤは「ステップ」と呼ばれる指示符に従って，足元のパネルを踏む．ある楽曲に対して用意されたステップの配置パターンは「譜面」と呼ばれている．ダンスゲームのプレイヤのな

© 2020 Information Processing Society of Japan

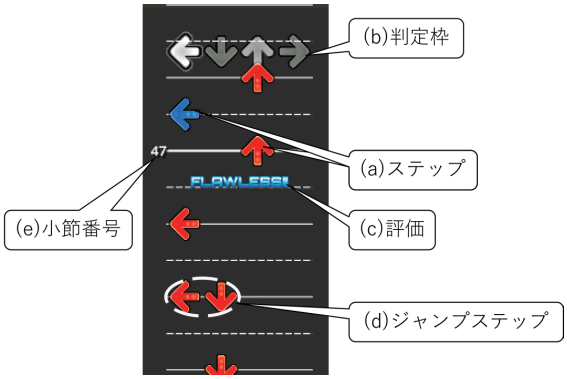


図 1　ダンスゲーム *Stepmania* におけるゲーム画面の例．楽曲に合わせて (a) ステップが画面下部から上部へ移動し，(b) 判定枠とステップが重なったタイミングでプレイヤが対応するパネルを踏むという動作を行う．(c) 評価は動作を正しいリズムで行えたかどうかを表す．ステップの中には，同一タイミングで 2 カ所以上のパネルを踏むことを指示する (d) ジャンプステップも存在する．(e) 小節番号は，その時刻が楽曲先頭から数えて何小節目にあたるかを示す．ステップの色は，そのステップが属する拍階層（2 章に述べる）によって変化する

Fig. 1　A screenshot of the dance game *Stepmania*. The game flow is as follows: (a) the steps scroll upwards from the bottom to the top of the screen, (b) the player hits the corresponding panel with his/her feet as the steps reach the receptor, and (c) how accurate the timing of the player's action is judged. Jump steps remarked as (d) require the player to hit two or more panels simultaneously. The number remarked as (e) shows the serial number of measures in the track. The color of the step indicates the beat layer the step belongs to (see Section 2).

かには，何年も継続して遊ぶ上級者もいれば，1 度だけ遊ぶような初心者もいる．

幅広い実力のプレイヤが面白いと感じるゲームとするためには，簡単な内容から始まり，少しずつ難しくなっていくように難易度バランスを構成することが重要である [1]．多くのダンスゲームでは，難易度の異なるたくさんの譜面を用意することで，多くのプレイヤが面白いと感じるよう配慮されている．譜面の難易度に影響する要素は複数あるとされている．仮に，リズムの複雑さが譜面の難易度に影響する唯一の要素であるならば，同程度の難易度とされている譜面は楽曲によらず同程度に複雑なリズムを含んでいることになる．しかし，実際には同程度の難易度の譜面であってもリズムの複雑さは様々であり，動作の要求頻度やジャンプステップの頻度なども難易度に影響すると考えられる．このように，同程度の難易度の譜面であっても，プレイヤに要求する動作の組合せによって，それぞれ異なる「難しさの個性」を持つといえる．この難しさの個性が，ダンスゲーム譜面の「面白さ」にも大きく影響する．同程度の難易度の譜面であっても，どの譜面をより面白いと感じるかはプレイヤによって異なる．その事例として，ゲーム制作者によって同程度の難易度であるとラベル付けされた譜面間でプレイヤのゲームログを比較したとき，プレイヤによって好んでプレイしている譜面は異なることが多い．以上から，プレイヤは譜面の特性を「難易度」という単一の基準のみではなく，「面白さ」という異なる観点でも評価しているといえる．これら 2 つの観点から譜面を評価した例として，「難しいがやりがいがあり，面白い譜面である」，「簡単で遊びやすく，面白い譜面である」，「簡単だがつまらない譜面である」などがあげられる．

ゲームの難易度や面白さに影響する要素を調査した研究はいくつか存在している．Wehbe らは，シンプルなアクションゲームに対して，画面スクロールの速さやジャンプ動作の複雑さなどのパラメータを変化させることで，難易度およびプレイヤ体験が変化するかを調査している [2]．Spiel らは，パズルゲーム *TETRIS* を題材にして，自明に難易度に影響するゲームの速度ではなく，ピースの選択アルゴリズムを変化させることで，難易度および面白さが変化するかを調査している [3]．一方で，音楽ゲームに関しては，人類学的な研究 [4] であるか，譜面の自動生成にかかわる研究が主流であった．譜面の自動生成は，ゲーム内のコンテンツを自動生成する Procedural Content Generation（PCG）[5] の一課題として位置付けられる．PCG では，アクションゲームのステージ [6] やシューティングゲームの武器 [7] などが自動生成されている．音楽ゲームにおける PCG として，Donahue らの *Dance Dance Convolution*（DDC）[8] が代表的な研究としてあげられる．DDC は，Long Short Term Memory（LSTM）モデルを用いて，プレイヤが望む任意の楽曲に対して譜面を自動で生成することを実現した．

DDC では，データセット内の譜面すべてを用いてモデルを学習させている．ところが，前述のとおり譜面データセットの中には様々な特性を持った譜面が混在している．多様性のあるデータを，各データが持つ特性を明示することなく同一のモデルに学習させると，特性が混在した譜面が生成されてしまう．このような生成結果の一例として，30 秒近くにわたっていっさいの指示符が存在しない区間と，指示符が極端に多い区間が混在する譜面が出力されることを確認した．DDC は指示符が少ない譜面と多い譜面を両方学習しているため，指示符の頻度に関するこれらの特性を共に反映させようとした結果，生成譜面のある区間では指示符が少なく，またある区間では指示符が多いという，複数の特性が混在した中途半端な生成結果になったと考えられる．このような生成結果は，音響信号からダンスゲーム譜面を生成するという Donahue らの目標自体は達成していても，多様な面白さの感性に対して十分に対応できているとはいえない．また，大多数のデータが共通して持つ特性を学習する過程で，中程度の事例数がある特性が学習されず，本来あるべき特性が獲得されないことがある．

© 2020 Information Processing Society of Japan

たとえば，ジャンプステップを特性として持つ譜面は一定数の存在が認められるが，全体の傾向としてはジャンプステップの事例数は多くはない．そのため，すべての譜面を学習したモデルでは，ジャンプステップを特性とする譜面は生成できない．極端にデータ数が少ない特性は，統計的機械学習のアプローチでは学習が困難である．しかしながら，一定数の事例が存在する特性は，大多数の特性と分離することで学習可能になるため，難しさの個性として扱い，譜面生成モデルを構築できると考えた．これらの問題を解決するためには，音響信号のみではなくゲーム譜面そのものの特性に着目し，学習データとなる譜面を特性ごとに分類したうえで学習することが有効であろう．

我々の先行研究において，ダンスゲーム譜面を表現する複数の特徴量を設定し，クラスタリングによって譜面を特性ごとに分類できることを確認した [9]．特定の特性を持つ譜面のみをモデルに学習させることによって，意図した特性を持つ譜面が生成されることが期待される．これにより，多様な面白さの感性に対応した特徴的な譜面をプレイヤに提供可能になると考えた．データを分類したうえでモデルを学習させることの効用は，ゲーム以外の生成モデルにおいても指摘されている．Sarroff らは，自動作曲を行う Autoencoder モデルを 10 ジャンルの楽曲を用いて学習させた [10]．これに対して Colonel らは，文献 [11] p.5 において

> Including several genres of music in a corpus trains an autoencoder to be a jack of all trades, but master of none. By focusing the corpus of our exercise on tonal sounds, we encourage the net to master representations of those sounds.

と，学習データに複数のジャンルの楽曲を含めず特定のジャンルに集中させるべきであると指摘している．

本稿では，先行研究 [9] の発展として，ゲーム譜面データセットを分析することで，異なる面白さの特性を持つダンスゲーム譜面の生成手法を提案する．譜面の難易度と面白さに影響する要素として考えられる特徴をあげ，各特徴に基づいてゲーム譜面をクラスタリングする．これによって，データセット内に面白さの特性が異なるデータが含まれていることを示し，特性に基づくデータセットの整理を実施する．そして，クラスタごとに譜面生成モデルを学習させることによって，一貫した面白さの特性を備えた特徴的な譜面の自動生成を実現する．

2.　環境設定

本稿では，先行研究 [8], [9], [12], [13] にならって，ITG データセットを使用する[*1]．**表 1** にデータセットの情報を示す．このデータセットには，オープンソースのダンスゲーム *Stepmania* [*2]で遊べるように作られた 133 曲分の譜面データが含まれている．1 曲につき 4〜5 種類の譜面が収録されており，難易度が低い順に Beginner，Easy，Medium，Hard，Challenge というラベルが付与されている．本稿では，“曲名”（難易度ラベル名）という書式で，データセット内に含まれる特定の譜面を指すものとする．

表 1　データセットの情報（文献 [8] を参考に作成）
Table 1　The statistics of the dataset (with reference to the existing paper by Donahue et al. [8]).

データセット名	ITG
制作者人数	8
パックの数	2
楽曲数	133
譜面数	652
公開年	2004〜2005

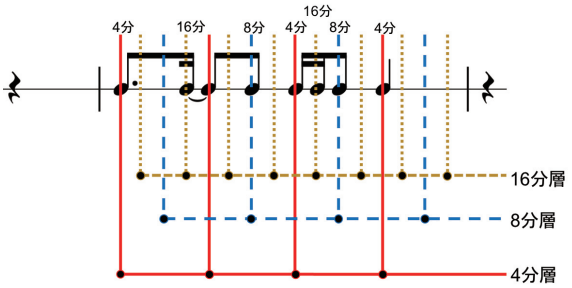


図 2　拍階層の概念．音符の上に書かれた「4 分」「8 分」が各音符の拍階層表現である．4 分層に属する音符は，8 分層や 16 分層にも属するが，そのなかで最も数字が小さい 4 分層に存在するものとする
Fig. 2　The concept of the beat layer. The note should belong to the lowest-resolution layer. For example, the last note in this figure is assumed as the step belongs to the 4th layer, though it belongs to the 8th and 16th layers too.

先行研究 [9], [12], [13] に基づいて，拍階層表現を用いる．小節を n 等分（ただし $n \in \mathbb{N}$ かつ $n \geq 4$）することで表現できる時刻集合を「n 分層」と定義する．あるステップが存在する時刻において，ある n（≥ 4）が以下の 2 つの条件をともに満たすとき，そのステップは n 分層に存在するものとし，「n 分（のステップ）」と表現する．

条件 A　ステップが存在する時刻が n 分層の時刻集合に含まれている．

条件 B　n が条件 A を満たす最小の整数である．

Stepmania においては，ステップが存在する拍階層によって，図 1 のように異なる色でステップが表示される．赤色のステップは 4 分，青色のステップは 8 分，黄色のステップは 16 分であることを表す．**図 2** に，実際の楽器の楽譜

[*1] In The Groove, http://east.stepmania-online.com/In%20The%20Groove%201.zip および In The Groove 2, http://east.stepmania-online.com/In%20The%20Groove%202.zip（最終閲覧日：2020 年 1 月 17 日）

[*2] News - Stepmania, https://www.stepmania.com（最終閲覧日：2020 年 1 月 17 日）

© 2020 Information Processing Society of Japan

と拍階層の概念の対応付けを示す．ダンスゲームにおける「ステップに合わせてパネルを踏む」という動作は，楽器の演奏における「音符に合わせて発音を開始する」という動作に対応する．拍階層表現は，動作を行う時刻が楽曲聴取上で知覚しやすいタイミングであるかどうかを表現している．ステップが存在している階層の n が大きいほど，そのステップは複雑なリズムで構成されており，知覚しにくいといえる．反対に，n が小さい層のステップが連続しているとき，それらのステップは知覚しやすいリズムであるといえる．

3.　ダンスゲーム譜面の多様性

同程度の難易度とされているダンスゲーム譜面であっても，それぞれの譜面が持つ特性は大きく異なる．**図 3**，**図 4** は，それぞれ “*Disconnected*”（Hard），“*Sweet World*”（Hard）の一部である．図中の時間軸は 0.1 秒ごとに目盛を付している．図 3 の “*Disconnected*”（Hard）は，テンポが 200 BPM（Beats Per Minute）と速く，図の範囲で 1 秒あたりに約 6.1 ステップが存在しているため，動作を素早く行う必要がある．一方で，ステップどうしの時間間隔が一定であり，4 分と 8 分のステップのみで構成されているため，リズムが単純である．図 4 の “*Sweet World*”（Hard）は，テンポが 132 BPM と遅めであり，図の範囲で 1 秒あたり約 2.8 ステップとステップ頻度が低いため，素早い動作はあまり求められない．しかし，4 分のステップが少なく 8 分や 16 分のステップが多いという，いわゆる裏拍を主としたリズムであるため，動作を行うタイミングが難解である．この 2 つの譜面は，ともに Hard 難易度として用意された譜面であるが，ステップ頻度とリズムの複雑さという特徴に注目すると，譜面の特性が異なることが分かる．そのため，プレイヤによって，この 2 つの譜面のうちどちらをより難しいと感じるか，どちらをより面白いととらえるかは，異なると考えられる．たとえば，体を動かすことは好きだが，リズム感に自信がないプレイヤにとっては，前者の方が遊びやすく楽しい譜面であると感じるであろう．

同程度の難易度とされている譜面どうしで特性が異なる事例がある一方で，難易度が異なる譜面どうしで特性が類似している事例も存在する．**図 5** は “*Remember December*”（Challenge）の一部であり，図 3 および図 4 よりも難しい譜面として制作されている．テンポは 192 BPM であり，図の範囲のステップ頻度は 1 秒あたり約 8.3 ステップである．リズムは 4 分と 8 分のステップを中心に，時折 16 分のステップが混ざる構成となっており，図 3 の “*Disconnected*”（Hard）よりは複雑だが図 4 の “*Sweet World*”（Hard）ほど難解なリズムではないと考えられる．これらの特徴から形成される図 5 の “*Remember December*”（Challenge）の特性は，図 4 の “*Sweet World*”（Hard）よりも図 3 の “*Disconnected*”（Hard）に近いと考えられる．このように

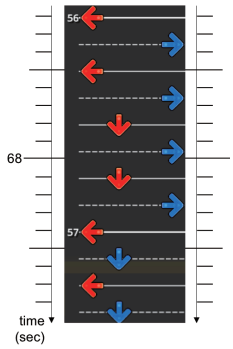


図 3　“*Disconnected*”（Hard）の一部．テンポは 200 BPM
Fig. 3　A part of Hard chart of “*Disconnected*.” The tempo of the track is 200 BPM.

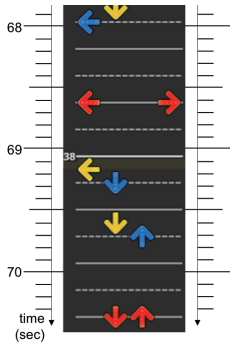


図 4　“*Sweet World*”（Hard）の一部．テンポは 132 BPM
Fig. 4　A part of Hard chart of “*Sweet World*.” The tempo of the track is 132 BPM.

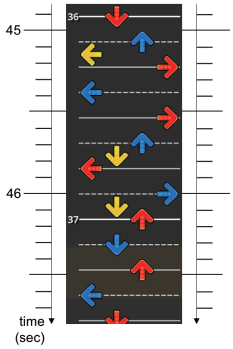


図 5　“*Remember December*”（Challenge）の一部．テンポは 192 BPM
Fig. 5　A part of Challenge chart of “*Remember December*.” The tempo of the track is 192 BPM.

譜面の特徴を考慮することで，データセットを類似した特性を持つ譜面の集合に分割し，相反する特性を持つ譜面どうしを分離することができる．

ITG データセットの中には，他の譜面とは大きく異なる特性を持つ譜面が存在する．極端な特性を持つ譜面の一例として，**図 6** に “*Bend Your Mind*”（Hard）の一部を示す．図中の白点線で囲んだステップは，プレイヤに足元のパネルを 3 カ所同時に踏むことを指示している．人間の足は 2 本しかないため，手と足を併用するなどの方法で無理

© 2020 Information Processing Society of Japan

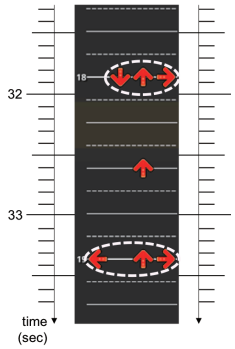


図 6　“*Bend Your Mind*”（Hard）の一部．テンポは 140 BPM
Fig. 6　A part of Hard chart of “*Bend Your Mind*.” The tempo of the track is 200 BPM.

やりパネルを反応させる必要がある．このようなステップを含む譜面は，ダンスゲーム本来の「音楽に合わせて，単純な動作を連続して行う」という面白さとは異なる，「困難な動作の指示に対して，攻略法を考え出す」という新たな面白さをプレイヤに提供するものと位置づけることができる．しかしながら，このような特性を持つ譜面はデータセット内に少数しか存在しない．その他の極端な特性を持つ譜面の例として，64 分のステップが譜面中の 14.8%を占め，非常に繊細なリズム把握能力が求められる “*Go 60 Go*”（Challenge）や，譜面中の約 23.7 秒間にわたってステップの頻度が 1 秒あたり約 10.7 ステップと一定であり，素早い動作を長時間持続することを求められる “*Delirium*”（Challenge）などがあげられる．これらのような譜面はクラスタリングや譜面生成モデルの学習においてノイズとなることが懸念される．実際に，先行研究において，極端な特性を持つ譜面が単独で 1 つのクラスタを占有し，クラスタ間のデータ数の不均衡を引き起こすことが確認された [9]．これらのような極端な特性を持つ譜面をあらかじめ検出し，データセットから取り除くことによって，多くの譜面に共通する特性をモデルに学習させる．

4.　ダンスゲーム譜面から得る特徴量の提案と分析

先行研究 [9] と同様に，第 1 著者の知見に基づき，譜面データから得る特徴量を設定した．該当著者は，ダンスゲームの制作メーカが主催する公式大会で複数回の入賞経験があり，ダンスゲームのプレイに対する実力と譜面理解に関する知見を十分に有している．

特徴量は，大きく分けて「譜面特徴量」と「楽曲特徴量」に分類される．譜面特徴量は，同じ曲に対して用意された譜面であったとしても，得られる値は譜面ごとで異なる．一方で，楽曲特徴量は，同じ曲に対して用意された譜面であれば，異なる譜面であっても同一の値が得られる．いくつかの特徴量は Donahue らの論文 [8] 中でも触れられている．特徴量の計算には，譜面中の最初のステップから最後のステップまでの部分を用いた．

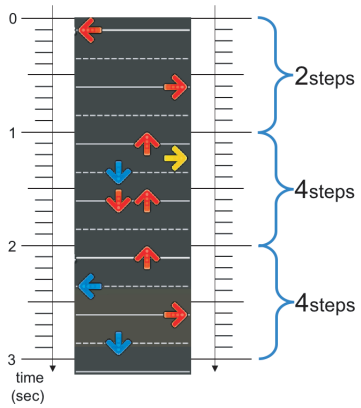


図 7　1 秒あたりステップ数の集計例
Fig. 7　An example of counting steps per second.

4.1　譜面特徴量

4.1.1　1 秒あたりステップ数

譜面全体を 1 秒ごとに切り分ける．切り分けられた 1 秒ごとの各区間に含まれるステップ数を数えることで，以下の特徴量を得る．

n_l　最もステップ数が少ない区間のステップ数．
n_m　最もステップ数が多い区間のステップ数．
n_μ　譜面中のすべての区間のステップ数の平均．
n_σ^2　譜面中のすべての区間のステップ数の分散．

例として，**図 7** のような 3 秒間のみの譜面を考える．0 秒から 1 秒の区間に 2 ステップ，1 秒から 2 秒の区間に 4 ステップ，2 秒から 3 秒の区間に 4 ステップが存在する．この譜面に対する特徴量の値は，

$$n_l = 2,$$
$$n_m = 4,$$
$$n_\mu = \frac{2+4+4}{3} \approx 3.33,$$
$$n_\sigma^2 = \frac{(2-n_\mu)^2 + (4-n_\mu)^2 + (4-n_\mu)^2}{3} \approx 0.89,$$

となる．

4.1.2　ジャンプステップ

同一のタイミングに 2 つ以上の矢印が存在するステップがある．これは，プレイヤに指定されたすべてのパネルを同時に踏むことを要求している．一般的に，ジャンプして両足で 2 カ所のパネルを踏むことで対処するため，このようなステップを「ジャンプステップ」と定義する．ある譜面中のジャンプステップの回数を集計し，以下の特徴量を得る．

j_r = ジャンプステップの回数/総ステップ数．
j_d = ジャンプステップの回数/譜面の長さ [秒]．

4.1.3　拍階層割合

2 章の定義に基づいて，譜面中のそれぞれのステップが

© 2020 Information Processing Society of Japan

どの拍階層に存在するかを集計し，以下の特徴量を得る．

$l_4 = 4$ 分のステップ数/総ステップ数．

$l_8 = 8$ 分のステップ数/総ステップ数．

$l_{12} = 12$ 分のステップ数/総ステップ数．

$l_{16} = 16$ 分のステップ数/総ステップ数．

$l_{24} = 24$ 分のステップ数/総ステップ数．

$l_{32} = 32$ 分のステップ数/総ステップ数．

$l_oth = 1 - \sum_{q \in \{4,8,12,16,24,32\}} l_q$．

4.1.4 ステップ向き 3-gram の $tf - idf$

連続する 3 ステップの向きの組合せについて出現頻度を集計し，3-gram とする．**表 2** に，出現頻度上位 3 種類の 3-gram を示す．ITG データセット内には，6,910 種類の 3-gram が存在し，出現頻度の総和は 176,279 であった．このうち，総出現頻度の 90%を占める出現頻度上位 979 種類の 3-gram を，分析の対象とした．以下，979 種類の 3-gram の集合を G とし，$g \in G$ とする．

それぞれの $g \in G$ に対して，ある 1 譜面中の出現頻度 tf_g と，データセット内で出現している譜面の数 df_g から，式 (1) を用いて $tf - idf$ 値を計算し，s_g とする．

$$s_g = \frac{tf_g}{\sum_{k \in G} tf_k} \times \log_2 \frac{m}{df_g}. \quad (1)$$

式 (1) で，m はデータセット内の譜面の数 652 である．s_g は，g がその譜面にどの程度固有であるか，すなわちどの程度特徴付けているかを示す．

4.2 楽曲特徴量

4.2.1 テンポ

ITG データセットに含まれる譜面データには，楽曲のテンポ値の情報が記録されている．テンポ値の単位には BPM（Beats Per Minute）が用いられている．楽曲の途中でテンポが変化する場合は，すべての変化点のタイミングと変化後のテンポ値が記録されている．

Stepmania においては，テンポが速いほど画面上のステップの移動速度が早くなり，プレイヤに素早い判断を要求することになるため，一般的にテンポが速い曲の方が難しいとされている．また，楽曲途中でのテンポの変化は，通常は等速運動をするステップの移動速度の変化を引き起こし，プレイヤにとって予測しづらいギミックとして作用する．

以下に，テンポに関して取得する特徴量をあげる．単位は BPM である．t_μ を除き，小数第一位を四捨五入し，整数値として扱う．

t_l 譜面中で最も長い時間を占めるテンポ値．

t_s 譜面中で最も遅い瞬間のテンポ値．

t_f 譜面中で最も速い瞬間のテンポ値．

t_μ ＝総拍数/楽曲の長さ [秒] × 60．

さらに，テンポの変化に関する以下の特徴量を得る．

t_r $= t_f / t_s$．

表 2 出現頻度上位 3 種類の 3-gram

Table 2 The top three step sequences of 3-gram frequency.

Step 1	Step 2	Step 3	出現頻度	出現譜面数
↑	↓	→	4,457	600
↓	↑	↓	4,238	553
←	↑	↓	4,216	589

t_n テンポ変化が記録されている回数．

t_t 変化前と変化後のテンポの比が 2^m ($m \in \mathbb{Z}$ *and* $m \neq 0$) であるテンポ変化が記録されている回数．ステップの移動速度変化のみを目的としたテンポ変化の回数と見なすことができる．

4.2.2 譜面停止ギミック

ITG データセット内のいくつかの楽曲には，指定された時間の間ステップの移動を一時停止する「譜面停止ギミック」が設定されている．これは，楽曲途中のテンポ変化と同様，ステップの移動速度に変化を与えることでプレイヤを騙すギミックとして設定されている．譜面停止ギミックに関する特徴量は，以下の 1 つである．

p_n 譜面停止ギミックが設定されている回数．

4.3 特徴量に基づく譜面の分析と考察

ITG データセット内のすべての譜面について特徴量を取得し，特徴量が譜面の難易度に影響していること，および他の譜面とかけ離れた値を持つ譜面が存在することを確認した．**図 8** に，n_μ の集計結果を表す箱ひげ図を示す．Beginner，Easy，Medium，Hard，Challenge はそれぞれの難易度ラベルがつけられた 133 譜面（Challenge のみ 120 譜面）の集計結果，All はデータセット内の全 652 譜面についての集計結果を示す．また，**図 9** に，データセット内の 133 楽曲に対する t_μ，t_l，t_s，t_m の集計結果を表す箱ひげ図を示す．両方の図において，Q_1，Q_3 をそれぞれ第 1 四分位数，第 3 四分位数として，$Q_1 - 1.5(Q_3 - Q_1)$ を下回る値，および $Q_3 + 1.5(Q_3 - Q_1)$ を上回る値を外れ値として示している．

4.3.1 1 秒あたりのステップ数平均

図 8 より，難易度が上がるにしたがって，平均ステップ数が多い範囲に分布しており，ステップ数は難易度が高いほど多くなる傾向が分かる．また，難易度が上がるにしたがって，四分位範囲 $Q_3 - Q_1$ が広がっている．このことから，簡単な譜面はどの楽曲でも似たような頻度でステップが現れる一方，難しい譜面では楽曲によってステップ頻度に差があるといえる．さらに，Medium 難易度と Hard 難易度には，それぞれ 1 つ上の難易度の第 3 四分位数よりも高い値を示す譜面が存在が外れ値として存在する．これらの譜面は，少なくとも n_μ の観点においては，該当の難易度としては不適切に難しい譜面であると見なすことができる．

© 2020 Information Processing Society of Japan

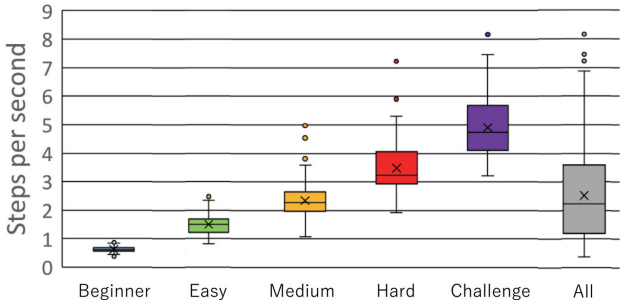


図 8　n_μ の分布を示す箱ひげ図

Fig. 8　A box plot concerning the distribution of n_μ for each difficulty level.

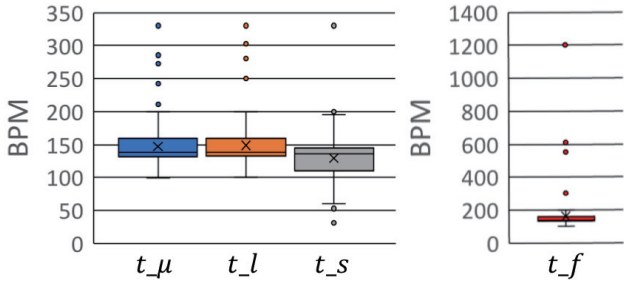


図 9　t_μ，t_l，t_s，t_m の分布を示す箱ひげ図

Fig. 9　Box plots concerning the distribution of t_μ, t_l, t_s and t_f.

4.3.2　テンポ

図 9 中の t_μ，t_l より，ITG データセットの楽曲のうち半分が 130〜160 BPM の範囲に該当することが分かった．一方で，例外的に速いテンポの楽曲も存在する．“*Pandemonium*” は楽曲を通して 330 BPM であるため，$t_s = t_f = 330$ という値を示した．この値は，特に t_s において極端な外れ値となっている．“*Robotix*” は $t_l = 150$ だが，楽曲中盤に徐々に加速し $t_f = 1{,}200$ を記録したのちに減速するという構成になっており，t_f において他の楽曲と大きくかけ離れた値が観測されている．以上から，ITG データセットには 140 BPM 前後のテンポの楽曲が多く収録されている一方で，200 BPM を超えるような高速テンポの楽曲もわずかに存在していることが確認された．

5.　クラスタリングに基づく譜面の自動生成

4 章で設定した特徴量を用いることで，譜面の特性を数値で表現することを実現した．特徴量の傾向が近い譜面どうしをクラスタリングすることによって，データセット内の譜面を特性に基づいて分類し，共通の特性を持つ譜面の集合を得ることができる．クラスタごとに譜面生成モデルを学習させることによって，そのクラスタの特性を反映した特徴的な譜面の自動生成を実現する．

5.1　前処理：外れ値を持つ譜面の検出

4.3 節より，ある特徴量において他の譜面とかけ離れた値を持つ譜面がデータセット内に存在することが確認された．このような譜面は面白いと評価される一方で，類似した性質を持つ他の譜面がほぼ存在しない特殊な事例と見なされる．統計的な機械学習のアプローチにおいて，データ数が少ない特殊事例は学習の対象として適切ではないうえ，大多数の譜面が共有している特性の学習を阻害するノイズデータとして作用してしまう．ノイズデータが引き起こす問題の一例として，クラスタリングにおいて少数のデータのみで 1 つのクラスタを占めてしまうという問題を確認している [9]．また，譜面生成モデルがノイズデータを学習すると，生成される譜面にノイズデータの性質が中途半端に反映されてしまう．

これらの問題に対処するため，外れ値を持つ譜面をあらかじめデータセットから除外する．データセット内の全 652 譜面を母集団として，各特徴量に対して増山の棄却検定（$p = 0.01$）[14] を実施し，外れ値を検出した．ただし，特徴量の次元数が 979 と非常に大きい s_g については，主成分分析（PCA）を実施し，寄与率が高い上位 7 つの主成分にのみ棄却検定を実施した．652 譜面のうち，少なくとも 1 つの特徴量が外れ値と見なされた 205 譜面を，データセットから除外した．

5.2　クラスタリング

クラスタリングを行うにあたって，4 章にあげた特徴量を用いて各譜面を表す特徴ベクトルを生成する．譜面ごとに，以下の 4 種類の特徴ベクトルを生成する．

$\boldsymbol{n} = (n_l, n_m, n_\mu, n_\sigma^2, j_r, j_d)$,

$\boldsymbol{l} = (l_4, l_8, l_{12}, l_{16}, l_{24}, l_{32}, l_oth)$,

$\boldsymbol{s} = (s_g | g \in G)$,

$\boldsymbol{t} = (t_\mu, t_l, t_s, t_f, t_r, t_n, t_t, p_n)$,

ただし，$\boldsymbol{s}$ は 979 次元のベクトルとなる．

それぞれの特徴ベクトルに対して PCA を実施した．$\boldsymbol{n}$，$\boldsymbol{l}$，$\boldsymbol{t}$ については，すべての主成分を利用し，それぞれ同一の次元数をもつベクトル $\boldsymbol{pc_n}$，$\boldsymbol{pc_l}$，$\boldsymbol{pc_t}$ に変換した．$\boldsymbol{s}$ は次元数が非常に大きいため，累積寄与率が 90%を超えるように寄与率が高い主成分から順に選択し，208 次元のベクトル $\boldsymbol{pc_s}$ に変換した．

$\boldsymbol{pc_n}$，$\boldsymbol{pc_l}$，$\boldsymbol{pc_s}$，$\boldsymbol{pc_t}$ のうち任意数を結合することで，譜面ごとに 1 つのベクトルを生成する．本稿では，4 つすべてを結合した 229 次元のベクトルを用いてクラスタリングした結果について報告する．

5.1 節で外れ値が検出されなかった 447 譜面を対象に，k-means 法によるクラスタリングを行った．クラスタ数 k を，4 から 12 の範囲で変更しながらクラスタリングを試行した．本稿では，試行した中で各クラスタの特性が最も明確に現れていると第 1 著者が判断した，$k = 8$ の結果

© 2020 Information Processing Society of Japan

表 3　クラスタリング結果
Table 3　Result of the dance charts clustering.

ID	譜面数	特徴量のクラスタ内平均値							クラスタの特性
		n_μ	j_r	j_d	l_4	l_8	l_{16}	t_μ	
1	133	3.512	7.8%	0.268	53.9%	34.0%	11.9%	140.265	ステップ数が多い．
2	122	1.229	5.8%	0.070	95.4%	4.5%	0.1%	134.592	ステップ数が少ない．
3	82	1.868	13.8%	0.256	86.6%	12.9%	0.5%	143.252	ジャンプステップが多い．
4	58	2.493	8.1%	0.200	74.0%	21.1%	4.7%	140.788	ステップ数が中程度．
5	26	2.397	6.2%	0.150	84.5%	14.9%	0.6%	185.993	テンポが速い．
6	24	0.758	5.8%	0.070	99.7%	0.3%	0.1%	135.177	4 分のステップの割合が高い．
7	1	4.374	11.2%	0.492	36.4%	30.7%	26.2%	107.004	“*Zodiac*” (Challenge) のみ．
8	1	2.293	1.3%	0.030	47.4%	40.8%	8.3%	127.994	“*Hustle Beach*” (Hard) のみ．

を代表として報告する．**表 3** に，クラスタリング結果を示す．表内の「クラスタ内の特性」は，特徴量の平均値に基づいて第 1 著者が考察し与えたものである．クラスタ 1，2，4，6 は，ステップの頻度 n_μ とリズムの複雑さ l_k ($k \in \{4, 8, 16\}$) に差が見られた．クラスタ 3 は j_r と j_d が他に比べて高く，ジャンプステップが多い譜面が属した．クラスタ 5 は t_μ が高く，テンポが速い楽曲の譜面が属した．以上のように，クラスタによってそれぞれ異なる特性を示すことが確認された．

外れ値特徴量を持つ譜面を除外したにもかかわらず，クラスタ 7 とクラスタ 8 は 1 譜面のみが属するクラスタとなった．これらは，単一の特徴量で外れ値となることはないが，高い値を示す特徴量の組合せが他の譜面とは異なっているために起きた現象と考えられる．今後の課題として，Minimum Covariance Determinant [15] のような多変量データに対する異常検出手法を適用し，この問題を解決することがあげられる．また，本稿ではクラスタリング数の妥当性を第 1 著者の主観に基づいて評価した．客観的な指標を用いた妥当性の評価についても今後の課題としたい．

5.3　譜面の学習と自動生成

クラスタごとに，音楽データを入力，譜面データを出力とする LSTM モデルを学習させた．モデルの構造は，Donahue らの先行研究 DDC [8] と同一のものを用いた．

譜面数が比較的多い，クラスタ 1，2，3 を学習の対象とした．クラスタごとにランダムに，64 譜面を学習データ，9 譜面を検証データ，9 譜面をテストデータに振り分けた．クラスタ n を学習したモデルを，Model n と名付けた．また，比較対象として，上記 3 クラスタすべての学習データ 192 譜面と検証データ 27 譜面を用いて学習したモデルを用意し，Model all とした．

図 10，**図 11**，**図 12**，**図 13** に，それぞれ楽曲 “*Walking on Fire*”（クラスタ 1 テストデータ），“*Hillbilly Hardcore*”（クラスタ 2 テストデータ），“*Bubble Dancer*”（クラスタ 2 テストデータ），“*Twilight*”（クラスタ 3 テストデータ）を入力音楽データとした際に，各モデルが生成した譜面の一部を示す．同一図中の 4 つの譜面は，それぞれ入力楽曲の同一部分に対して生成された譜面である．破線丸で囲ったステップはジャンプステップである．Model 1，2，3 に注目すると，配置されたステップ数について，Model 2 が最も少なく，Model 3，Model 1 の順に多くなっていくことが確認される．また，Model 3 が生成した譜面にはジャ

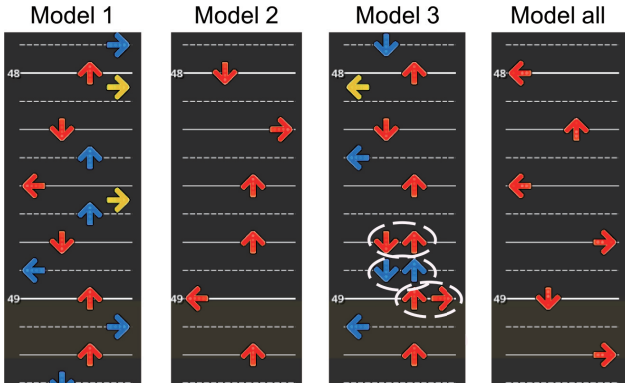


図 10　楽曲 “*Walking on Fire*” を各モデルに入力し，出力された譜面の一部
Fig. 10　Comparison of each model's prediction for the track “*Walking on Fire*.”

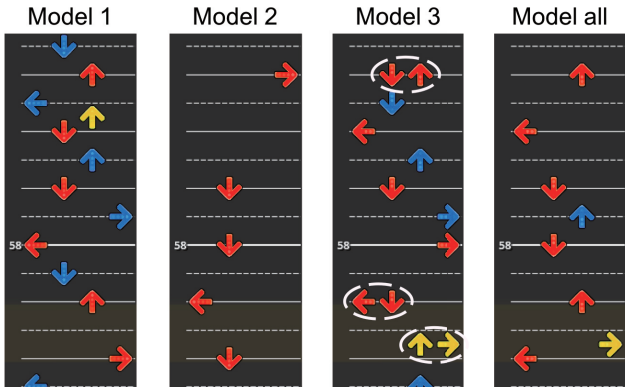


図 11　楽曲 “*Hillbilly Hardcore*” を各モデルに入力し，出力された譜面の一部
Fig. 11　Comparison of each model's prediction for the track “*Hillbilly Hardcore*.”

© 2020 Information Processing Society of Japan

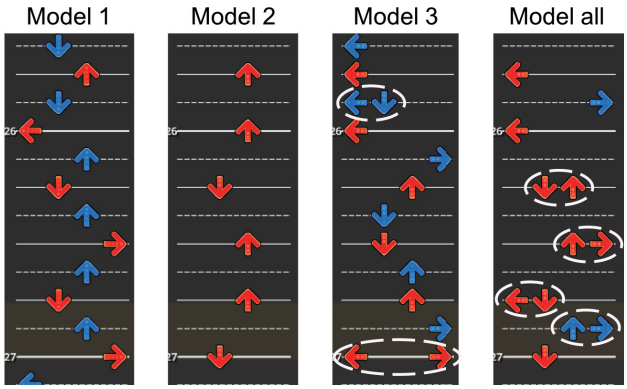


図 12　楽曲 “*Bubble Dancer*” を各モデルに入力し，出力された譜面の一部

Fig. 12　Comparison of each model's prediction for the track “*Bubble Dancer*.”

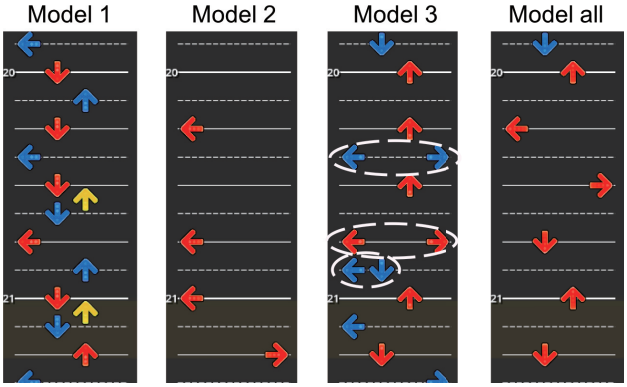


図 13　楽曲 “*Twilight*” を各モデルに入力し，出力された譜面の一部

Fig. 13　Comparison of each model's prediction for the track “*Twilight*.”

ンプステップが多く存在するが，Model 1 と Model 2 が生成した譜面にはジャンプステップが見られなかった．以上から，Model 1，2，3 が生成した譜面は，5.2 節に記述したクラスタ 1，2，3 の特性をそれぞれ強く反映しており，それぞれが異なる面白さの特性に指向した譜面であると考えられる．その一方で，3 つのクラスタすべてを学習した Model all の生成譜面におけるステップ数は，図 11 と図 12 では Model 2 と同程度である一方で，図 10 と図 13 では Model 2 と 3 の中間程度となっている．図 10，図 11，図 13 の Model all の生成譜面にはジャンプステップが見られなかったが，図 12 中にはいくつかのジャンプステップが確認された．Model all の生成譜面にはそれぞれのクラスタの特性が混在しており，一貫した特性を持つ譜面を生成していないと考えられる．

5.4　生成譜面の評価

5.4.1　*F*-score

それぞれのクラスタでテストデータに分類された楽曲を，各モデルに入力することによって，譜面を生成した．

表 4　各モデルの F-score

Table 4　F-scores for each model.

	Test 1	Test 2	Test 3
Model 1	**0.678**	0.364	0.474
Model 2	0.546	**0.455**	0.508
Model 3	0.676	0.401	0.506
Model all	0.662	0.444	**0.517**

表 5　各モデルが生成した譜面の n_μ, j_r, j_d の平均値

Table 5　Average of n_μ, j_r and j_d for all charts generated by each model.

	n_μ	j_r	j_d
Model 1	4.481	1.7%	0.067
Model 2	1.736	1.2%	0.020
Model 3	3.816	11.2%	0.423
Model all	2.998	6.7%	0.181

生成された譜面でステップが配置された時刻と，正解データにおいてステップが存在する時刻を比較し，F-score を計算した．**表 4** に結果を示す．ここで，Test n はクラスタ n から選ばれたテストデータを表す．

Test 1 および Test 2 については，それぞれ対応する Model 1 と Model 2 が，他のモデルよりも高い F-score を示した．一方で，Test 3 に対して最も高い数値を示したのは，Model 3 ではなく Model all であった．このことから，Model 1 および Model 2 は，ステップを配置する時刻についての特徴をうまくとらえられているといえる．しかし，Model 3 については，ステップ配置時刻の特徴をとらえていない．

5.4.2　データセットと自動生成譜面の統計比較

各モデルに Test 1，Test 2，Test 3 の楽曲データを入力し，計 24 譜面を生成した．**表 5** に，生成された各譜面に対して n_μ，j_r，j_d を計算し，24 譜面の平均を求めた結果を示す．表 5 と表 3 を比較すると，n_μ について，Model 1 および Model 2 はそれぞれある程度クラスタ 1 およびクラスタ 2 に近い値を示したが，Model 3 はクラスタ 3 よりも明らかに高い値を示した．j_r と j_d について，Model 3 は他のモデルよりも高い値を示した．以上より，Model 1 および Model 2 については，対応するクラスタのステップ頻度についての特性を，ある程度学習できたといえる．Model 3 は，ジャンプステップが多いというクラスタ 3 の特性を反映できている一方で，ステップ頻度についてはうまく学習できなかったのではないかと考えらえる．

6.　おわりに

本稿では，ダンスゲーム譜面の難易度の個性，すなわち面白さの特性に影響する特徴量を設定した．設定した特徴量は，難易度によって確かに差があることを確認した．また，いくつかの特徴量で例外的に高い値を示す譜面が存在

© 2020 Information Processing Society of Japan

することを確認した．このような極端な譜面をデータセットから除外したうえで，クラスタリングを実施した．クラスタリングを通して，ステップの頻度，リズムの複雑さ，ジャンプステップの頻度などの特徴に基づいて，特性ごとに譜面を分類することができた．クラスタごとに譜面生成モデルを学習させることで，各クラスタの特性を反映した，異なる面白さの特性を持つ譜面が生成できた．客観評価実験を通して，各譜面生成モデルが生成した譜面の特性が，学習したクラスタに含まれる譜面の特性と一致していることを確認した．

本稿の提案手法において，それぞれの譜面生成モデルが学習に用いたデータ数は決して多くはないが，生成されたダンス譜面は面白さの各特性に指向していることが確認された．従来研究においては，譜面の特性を考慮せずに学習データセットを構築し，譜面生成モデルを学習させていた．その結果，学習データ量は十分にあるものの，「生成譜面に特性が混在している」，「本来あるべき特性が反映されていない」といった問題が生じており，プレイヤにとっては面白いと感じられない譜面が生成されていたと考察される．これは，エンタテインメント分野におけるコンテンツ生成タスクにおいて，データセットからノイズを取り除き，コンテンツの特性に指向した学習データセットを用意することで，エンタテインメント性の高いコンテンツの生成が可能になることを示唆している．提案手法のアプローチは，本来存在するデータ量を減らすことになるが，面白さの観点でより高品質なコンテンツ生成を目指すうえで有効であると考える．一方で，本稿で外れ値として除外した 205 譜面は，それぞれが他の譜面とは異なる面白さの特性を備えている譜面である．このような独自の特性を備えた譜面を生成することは我々の最終的な目標の 1 つである．譜面のデータ数が増加し，ITG データセットのみでは特殊事例であった特性が譜面の傾向として形成されれば，クラスタリングなどを通してこれらの特性を獲得し生成譜面に反映させることが可能になると考えられるが，これは今後の課題としたい．

今後の展望として，ダンスゲームのプレイヤが生成された譜面の特性を判別できるかどうかを，被験者実験によって評価することを検討している．この被験者実験を通して，クラスタリングによって獲得した譜面の特性が，難しさや面白さの個性としてプレイヤが認識している譜面の特性であることを確認できると考える．また，クラスタリングに用いる特徴量の精査，各クラスタに含まれた譜面に対する考察などを通して，より譜面の特性を詳細に分析することも今後の課題としたい．本稿では，学習データセットの変化による効用を検証するため，譜面生成には Donahue ら [8] と同じ構造のモデルを用いた．モデル構造や入力特徴量を変更することによって，入力楽曲やプレイヤの嗜好に応じて，難易度および面白さの特性を調整したダンス譜面を生成する方法についても検討する．

謝辞 本研究は，一部，立命館大学ゲーム研究センターの支援のもと行われた．また，実験の実施にあたり，立命館大学情報理工学部の西原陽子准教授の協力を得た．記して謝意を示す．

参考文献

[1] Hunicke, R.: The case for dynamic difficulty adjustment in games, *Advances in Computer Entertainment Technology*, pp.429–433 (2005).

[2] Wehbe, R.R., Mekler, E.D., Schaekermann, M., Lank, E. and Nacke, L.E.: Testing Incremental Difficulty Design in Platformer Games, *Proc. 2017 CHI*, pp.5109–5113 (2017).

[3] Spiel, K., Bertel, S. and Kayali, F.: “Not another Z piece!” Adaptive Difficulty in TETRIS, *Proc. 2017 CHI*, pp.5126–5131 (2017).

[4] Behrenshausen, B.G.: Toward a (Kin)Aesthetic of Video Gaming: The Case of Dance Dance Revolution, *Games and Culture*, Vol.2, No.4, pp.335–354 (2007).

[5] Shaker, N., Togelius, J. and Nelson, M.J.: *Procedural Content Generation in Games: A Textbook and an Overview of Current Research*, Springer (2016).

[6] Pedersen, C., Togelius, J. and Yannakakis, G.N.: Modeling player experience for content creation, *IEEE Trans. Computational Intelligence and AI in Games*, Vol.2, No.1, pp.54–67 (2010).

[7] Hastings, E.J., Guha, R.K. and Stanley, K.O.: Evolving content in the galactic arms race video game, *2009 IEEE Symposium on Computational Intelligence and Games*, pp.241–248 (2009).

[8] Donahue, C., Lipton, Z.C. and McAuley, J.: Dance Dance Convolution, *Proc. ICML 2017*, pp.1039–1048 (2017).

[9] Tsujino, Y., Yamanishi, R. and Yamashita, Y.: Characteristics Study of Dance-charts on Rhythm-based Video Games, *Proc. 2019 IEEE Conference on Games* (*CoG*), pp.1–4 (2019).

[10] Sarroff, A.M. and Casey, M.: Musical audio synthesis using autoencoding neural nets, *Joint 40th International Computer Music Conference* (*ICMC*) *and 11th Sound & Music Computing Conference* (*SMC*) (2014).

[11] Colonel, J., Curro, C. and Keene, S.: Improving Neural Net Auto Encoders for Music Synthesis, *Audio Engineering Society Convention 143* (2017).

[12] 辻野雄大，山西良典，西原陽子，福本淳一：時系列深層学習に基づく難易度間関係モデルを用いたダンスゲーム譜面難易度の自動調整，情報処理学会論文誌，Vol.59, No.11, pp.1953–1964 (2018).

[13] Tsujino, Y. and Yamanishi, R.: Dance Dance Gradation: A Generation of Fine-Tuned Dance Charts, *Entertainment Computing – ICEC 2018*, pp.175–187 (2018).

[14] Kobayashi, K. and Pillai, K.: *A Handbook of Applied Statistics Pharmacology*, CRC Press (2012).

[15] Rousseeuw, P.J. and Driessen, K.V.: A fast algorithm for the minimum covariance determinant estimator, *Technometrics*, Vol.41, No.3, pp.212–223 (1999).

© 2020 Information Processing Society of Japan

辻野 雄大　（学生会員）

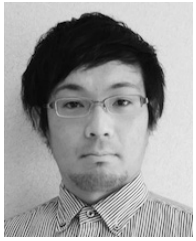
2017年立命館大学情報理工学部メディア情報学科卒業．2019年同大学大学院情報理工学研究科情報理工学専攻博士課程前期課程修了．2019年同博士課程後期課程入学，現在に至る．エンタテインメントコンピューティング，音楽情報処理に興味を持つ．ダンスゲームの公式世界大会に出場した実力と経験を基に，ダンスゲーム譜面の自動生成手法に関する研究に従事．World Pump Festival 2016 Speed Male 部門日本代表，The 8th KAC Dance Dance Revolution ベスト6，Dance Dance Revolution ROUND1 頂上決戦 Foot God's Division 第3位および解説者．IEEE Graduate Student 会員および Young Professionals 会員．

山西 良典　（正会員）

2007年名古屋工業大学工学部知能情報システム学科卒業．2009年同大学院工学研究科情報工学専攻博士前期課程修了．2012年同博士後期課程修了．博士（工学）．同年立命館大学情報理工学部助手，2013年同特任助教．2014年同助教，この間，UBC（カナダ） 客員助教．2018年同講師．2020年関西大学総合情報学部准教授．現在に至る．感性情報処理，Web インテリジェンス，マルチメディア情報処理，音楽情報処理等のコンテンツ処理研究に従事．電子情報通信学会，人工知能学会，日本感性工学会，日本音響学会，芸術科学会，ACM，ACL，各会員．

山下 洋一　（正会員）

1984年大阪大学工学研究科電子工学専攻前期課程修了．同年4月大阪大学産業科学研究所文部技官，1993年1月同助手，1994年8月同講師，1997年4月立命館大学理工学部助教授，2001年4月同教授，2004年4月同大学情報理工学部教授，現在に至る．博士（工学）．音声情報処理に関する研究に従事．電子情報通信学会，日本音響学会，人工知能学会，ISCA，IEEE 各会員．



井本 桂右

2010年京都大学大学院修士課程修了．同年4月日本電信電話（株）入社．2017年3月総合研究大学院大学博士課程修了．博士（情報学）．2017年4月立命館大学情報理工学部助教．2020年4月同志社大学理工学部准教授，現在に至る．音響イベント検出，アレイ信号処理に関する研究に従事．電子情報通信学会，日本音響学会，IEEE 各会員．

© 2020 Information Processing Society of Japan