第 28 回年次大会予稿

物語のシーンにおける登場人物の感情状態とBGMの関係性抽出

Extraction of Relationship between Character's Emotional State and BGM in Story Scene

根本さくら^{1*}, 石川一稀¹, 宇田朗子¹, 白石智誠¹,
Sakura NEMOTO^{1*}, Kazuki ISHIKAWA¹, Akiko UDA¹, Tomonari SHIRAISHI¹,
中村祥吾¹, 長野恭介¹, 山内拓真¹,
Shogo NAKAMURA¹, Kyosuke NAGANO¹, Takuma YAMAUCHI¹,
村井源¹, 平田圭二¹, 迎山和司¹, 田柳恵美子¹
Hajime MURAI¹, Keiji HIRATA¹, Kazushi MUKAIYAMA¹, Emiko TAYANAGI¹

1 公立はこだて未来大学

Future University HAKODATE

〒041-8655 北海道函館市亀田中野町116-2

E-mail: b1017015@fun.ac.jp

*連絡先著者 Corresponding Author

物語のシーンにおける登場人物の感情状態と, BGM の関係性は未だ明らかではない. しかし, この関係性が明らかになれば, 物語のシーンに適した BGM の自動選択システムや BGM の自動生成システムの構築に活用できると考えられる. そこで本研究では, AI による自動生成が特に進んでいるゲームを題材とし, その中でも物語性が強くかつ人気を博しているという理由から RPG を対象とし, BGM 自動選曲システムの構築を目標としたデータ作成を行った. 具体的には, あるシーンの登場人物の感情状態への評価付けと, BGM に対する音響特徴量の抽出を行うことで 477 シーンをデータ化し対応関係を分析した.

The relationship between the emotional state of a character and the background music in a scene of a story is not yet clear. However, if this relationship is clarified, it can be used to build an automatic BGM selection system and automatic BGM generation system suitable for a story scene. This research focused on computer role playing games because of popularity and nature of story, and created data set for an automatic BGM selection system. In particular, evaluated emotional state of characters and acoustic features in scenes were digitized as 477 data set, and correspondences of those data set were analyzed.

キーワード: 感情, BGM, RPG, BPM, スペクトルセントロイド, クロマベクトル

Emotional, BGM, RPG, BPM, Spectral centroid, Chroma vector

1 はじめに

映像を伴う作品のほとんどには, 各シーンにBGMが付与されている. コンピューターゲームもそのようなメディアのひとつである. 例えば, コンピューターゲームの楽しいシーンには, 明るく楽しい雰囲気のあるBGMが流れる. シーンの雰囲気に合ったBGMは, そのシーンをより魅力的に見せたり, その作品への没入感を高めたりする役割を担っている.

現在筆者らは, ロールプレイングゲーム(以下, RPGという)の自動生成システムを開発中である[1]. このシステムの実現にあたり, シーンに適したBGMを自動選択するシステムが必要である. このシステムを実現しシーンに適したBGMを付与するために, 本研究ではシーンの登場人物の感情状態によってBGMを決定できると仮定し, 既存のRPGのシーンにおける登場人物の感情状態とBGMの関係性を抽出した.

2 関連研究

物語の動画に対しシーンに合ったBGM付与を目的とした研究事例は多数存在する. 清水ら(2017)は, 動きや色, 被写体のキーワードといった動画特徴量から印象を推定し, その結果に基づいて楽曲を生成することで動画の印象に合った楽曲を付与する手法を提案している[2]. また小野ら(2011)は, 楽曲を印象によって分類することで, ホームビデオに対するBGM付与を支援するシステムを提案している[3]. 上述の事例のような動画のシーンと, RPGのシーンが異なる点の1つは, シーンに影響を与えるプレイヤーの有無である. 例えば動画では, 動画特徴量を利用し, 映像の盛り上がりに合わせてBGM

の盛り上がりのタイミングを合わせる手法が用いられている[3]. しかしRPGのシーンでは, ある物語展開に対し即座に反応することあれば長時間熟考することもあり, シーン長や聴取するBGM長はプレイヤーによって異なる. 従って盛り上がりの同期といった時間軸方向でシーンに合わせる手法はRPGのシーンには適していない. また, RPGはプレイヤーが自身を主人公に投影して操作する没入型の物語体験を提供するジャンルである. したがって, BGMは主人公の感情状態と適切に調和する必要がある. そのためRPGのシーンに対するBGM付与では, 動画特徴量などの直接的な特徴量だけを指標にするのではなく感情状態などの抽象的な特徴量を用いる必要があると考えられる.

3 評価手法の検討

この章では, 登場人物の感情状態への評価付けの方法や, BGMから抽出した特徴量について述べる. はじめに既存のRPGから分析対象を選定した. 次に, 選定作品の各シーンにおける登場人物の感情状態を評価付けし, データを収集した. そして, そのシーンで鳴っているBGMの音響特徴量を取り出したのち, 感情状態と音響特徴量の関係性を抽出した.

3.1 既存作品の分析

3.1.1 分析対象の選定

シリーズの売上やシリーズの継続性, 一般的な認知度を基準とし, 社会的評価の高いRPGを分析対象とするため作品の選定を行った. 選定の結果「ドラゴンクエスト」[4]シリーズ3作品, 「FINAL FANTASY」[5]シリーズ3作品の計6作品を分析対象とした. 分析作品を表1に示す.

表1 分析対象リスト

作品名
ドラゴンクエスト IV 導かれし者たち
ドラゴンクエスト VI 幻の大地
ドラゴンクエスト VII エデンの戦士たち
FINAL FANTASY
FINAL FANTASY III
FINAL FANTASY V

3.1.2 分析対象シーンの限定

RPGは主に「戦闘シーン」、「マップ移動シーン」、「イベントシーン」の大まかに3つのシーンが繰り返されることで構成されている。「戦闘シーン」は、敵との戦闘が行われるシーンである。「マップ移動シーン」は、プレイヤーが「戦闘シーン」や「イベントシーン」に移行するためにマップ状のキャラクターを移動させるシーンである。「イベントシーン」は、シナリオに基づき物語が展開するシーンである。例として、主人公キャラクターと仲間キャラクターとの別れのシーンなどが挙げられる。「戦闘シーン」では敵の強さによってBGMが変更される場合もあるが、ほとんどは固定されたBGMが流れることが多い。「マップ移動シーン」ではマップごとに固定されたBGMが流れることが多い。また、「戦闘シーン」と「マップ移動シーン」がゲーム中に現れる回数はプレイヤーの操作によって異なるため、同一シーンが複数回データ化されてしまう可能性がある。以上の理由により、シーンごとにBGMが固定されていないかつ、繰り返し行われることのない「イベントシーン」のみを分析対象とした。

3.1.3 シーンの区切り方

分析を行うにあたり、何をもって1シーンとして扱うかを定義した。本研究では登場人物の感情状態とそのシーンで鳴っているBGMに着目するため、2通りの分け方を考案した。1つ目は、BGMの区切りを1シーンとする方法である。この分け方では、登場人物の感情状態に変化があ

った場合でもBGMが切り替わっていなければ1シーンとして扱う。2つ目は、登場人物の感情状態の変化を1シーンとする方法である。この分け方では、登場人物の感情に変化があったところまでを1シーンとするので、同一BGM内に複数のシーンが含まれる可能性がある。今回作成したBGMの自動選択システムは、ゲームシナリオに付属している登場人物の感情特徴量を読み取りBGMを自動選択するシステムであるため、登場人物の感情状態の変化が基準である後者の区切り方で分析を行った。

3.2 感情状態の評価付け

物語の登場人物の感情状態については、4枚の絵で構成された物語の登場人物の感情を数値化した例などが存在する[6]。また、RPGの登場人物の行動から想起される感情状態に6つのカテゴリで評価を付けた例も存在する[7]。しかし、RPGの登場人物の感情状態への評価付けとして適切であるカテゴリは未だ明らかでない。今回は上述の課題を解決するための出発点として、Hevner(1936)が定義した円環状モデルを採用した[8]。Hevnerは66種類の印象形容詞を意味ごとに8つのグループに分類し円環状に配置した。その分類結果をもとに、シーンの登場人物の感情状態を8次元のベクトルで表現した。各グループに属する単語は6～10個と多いので、大串(2006)と同様に代表的な表現語3つをそのグループの表現語とした[9]。表現語を表2に示す。

分析者は各シーンに対し8つの感情グループ全てに0.0(弱)から1.0(強)の値で評価を付けた。

3.3 音響特徴量の抽出

シーンの登場人物の感情状態とBGMの情報を結びつけるため、分析対象としたシーンに現れたBGMの音響特徴量を抽出した。

表2 表現語リスト

グループ名	表現語
グループ 1	spiritual, solemn, serious (宗教的, 荘厳, 真面目)
グループ 2	sad, dark, frustrated (悲哀, 暗さ, 失意)
グループ 3	sentimental, tender, dreamy (感傷, 優しさ, 幻想)
グループ 4	quiet, satisfying, soothing (平穏, 満足, 鎮静)
グループ 5	humorous, fanciful, graceful (ユーモア, 奇抜, 優美)
グループ 6	happy, bright, cheerful (喜び, 陽気, 明るさ)
グループ 7	exciting, dramatic, passionate (興奮, 劇的, 激情)
グループ 8	vigorous, robust, majestic (力強さ, たくましさ, 堂々)

今回用いた音響特徴量は、BPM、スペクトルセントロイド、クロマベクトルの3種類である。先行研究[10]に準じ、音響特徴量抽出の標準的ライブラリの一つであるLibrosaを用いて抽出を行った[11]。BPMは1分間の拍数を示す値である。得られた値をそのまま特徴量とした。スペクトルセントロイドは、周波数成分をパワーによって重み付けした平均(スペクトル平均)であり、音の明るさに関する特徴量である[12]。今回は、30秒間でフレームごとに得られるスペクトルセントロイドの平均を特徴量とした。クロマベクトルは音楽音響信号に含まれる各ピッチクラスのパワーを表す[13]。楽曲からクロマベクトルの抽出を行った際、ピッチクラスそれぞれのパワーが0.0から1.0までの値としてフレームごとに得られる。その値が0.5以上であるようなピッチクラスをそのフレームにおいて生じているとみなしてカウントし配列に格納した。そしてこの配列を、音名を度数としたヒストベクトルにし、そのヒストベクトルの分散を特徴量とした。しかしクロマベクトルは円環上になっておりC, C#, D, D#, E, F, F#, G,

G#, A, A#, Bの順番, C#, D, D#, E, F, F#, G, G#, A, A#, B, Cの順番というように横に1音シフトして得られる12通り全ての分散を考慮する必要がある。よってその12通りのうち分散が最も小さくなるような順番のものを選択し、小数第3位を四捨五入して特徴量とした。

4 結果と考察

RPGの各シーンにおける表現語のグループごとの出現回数は、グループ1が149回、グループ2が71回、グループ3が59回、グループ4が236回、グループ5が11回、グループ6が84回、グループ7が96回、グループ8が97回であった。全体を通しグループ5の出現回数が11回と最小であった。また、反対にグループ4の出現回数は236回と最多であった。1シーンに2つ以上グループが含まれている際は1グループずつを1回としてカウントした。

音響特徴量の各範囲における各感情状態の平均値を抽出しレーダーチャートにしたものを図1から図3に示す。各図に記されている1から8までの数字は、表2の表現語リストのグループ番号の感情状態と対応している。また、今回セントロイド6の範囲である2357以上に値する楽曲は見つからなかったため除外した。

4.1 感情状態とBPMの関係性

感情状態とBPMの関係を示したレーダーチャートを図1に示す。また、BPMは以下の範囲に分割した。

- BPM1, 99.4以下
- BPM2, 99.4より大きく112.3以下
- BPM3, 112.3より大きく117.5以下
- BPM4, 117.5より大きく129.2以下
- BPM5, 129.2より大きく136.0以下
- BPM6, 136.0より大きい

スペクトルセントロイド、クロマベクトルともに、各範囲におけるグループ1からグループ8の感情状態の平均値における最大値と最小値の差は、4グループ以上で0.4より大きい値になった。それに対し、BPMの各範囲におけるグループ1からグループ8の感情状態における平均値の最大値と最小値の差のうち、6グループが0.4以下の値であった。以上より、BPMの値では感情状態に大きく差が見られないことが判明した。

図1 登場人物の感情状態とBPMのレーダーチャート

4.2 感情状態とスペクトルセントロイドの関係性

感情状態とスペクトルセントロイドの関係を示したレーダーチャートを図2に示す。また、スペクトルセントロイドは以下の範囲に分割された。

- ・スペクトルセントロイド1, 823未満
- ・スペクトルセントロイド2, 823以上1239未満
- ・スペクトルセントロイド3, 1239以上1593未満
- ・スペクトルセントロイド4, 1593以上1881未満
- ・スペクトルセントロイド5, 1881以上2357未満

- ・スペクトルセントロイド6, 2357以上

セントロイドは、グループ1の平均値が1, グループ2の平均値が0.575である。スペクトルセントロイドの値が低いとき、つまり暗い音が多いBGMは真面目なシーンや悲しいシーンで使用されていることが明らかとなった。また、グループ3, 4, 5, 6ではセントロイド1が出現することがなかった。よって、感傷のシーン、平穩のシーン、ユーモアのシーン、喜びのシーンにおいて極端に暗い音が鳴ることはないということが明らかになった。

図2 登場人物の感情状態とスペクトルセントロイドのレーダーチャート

4.3 感情状態とクロマベクトルの関係性

感情状態とクロマベクトルの関係を示したレーダーチャートを図3に示す。また、クロマベクトルは以下の範囲に分割した。

- ・クロマベクトル1, 7.582未満

- ・クロマベクトル2, 7.582以上8.481未満
- ・クロマベクトル3, 8.481以上9.157未満
- ・クロマベクトル4, 9.157以上9.784未満
- ・クロマベクトル5, 9.784以上10.38未満
- ・クロマベクトル6, 10.38以上

クロマベクトルは、グループ4と6において、クロマベクトル3から大幅に値が大きくなる傾向が見られた。したがって、平穏なシーンや喜びのシーンにおいて分散が小さいBGMが採用される可能性は低いということが明らかになった。

図3 登場人物の感情状態とクロマベクトルのレーダーチャート

5 まとめと今後の展望

本研究では、既存のRPGを分析し、各シーンの登場人物の感情状態と音響特徴量の関係

を抽出した。対象とした作品や登場人物の感情への評価付け、抽出した音響特徴量に関して今後の課題を3つ挙げる。

5.1 分析作品の選定

本研究では「ドラゴンクエスト」シリーズ、「FINAL FANTASY」シリーズの2つのシリーズを対象とし感情状態と音響特徴量の抽出を行った。「ドラゴンクエスト」シリーズはゲームの特性上、より感情移入しやすくするために主人公が発話をしないなど、特に主人公の感情状態がわかりにくくなっている。したがって、評価が「平穏」を示すグループ4に偏ってしまうという問題が発生した。ゆえに、選定する作品の登場人物、特に主人公の発話量なども考慮した上で分析作品を選定する必要がある。

5.2 感情状態への評価付け

今回はより多くのデータを収集するため、1つのシーンの感情状態に対し1人の評価者が評価付けをする形でデータを収集した。しかし、評価者ごとに評価基準が一定でなかった。したがって、複数人の評価の平均をそのシーンの感情状態として用いる必要がある。

また、シーンの感情状態を表すカテゴリとしてHevner(1936)が定義した音楽心理を表現する8つの感情円環モデルを用いている(3.2節)。一般にゲームの1シーンではさまざまな状況でさまざまなキャラクタが登場する。BGMがどのような効果を狙ってどの状況やキャラクタに対応して選曲・付与されるかも考慮する必要がある。ゆえに、音楽心理を表現する8つの感情で十分かどうか、あるいは何らかの拡張が必要かを検討する必要がある。

5. 3 ゲーム BGM の音響特徴量

今回は BPM, スペクトルセントロイド, クロマベクトルの 3 つの音響特徴量を用いている(3.3 節). これらの特徴量は, 音源分離や推薦システムなどで一般的に用いられているものである. ゲームプレイ中に聴取する音楽であることを考えると, ゲームシーンとの連想を作りやすい楽曲, ユーザの集中力を高めるような楽曲という観点からも識別する必要がある, そのための特徴量を検討する必要がある. また, これらの音響特徴量以外に, メロディラインや使用されている楽器に着目することでより感情状態との関係性が見えてくるのではないかと推察する.

参考文献

- [1] 中村祥吾, 斉藤勇璃, 白石智誠, 太田和宏, 根本さくら, 石川一稀, 宇田朗子, 小川卓也, 山内拓真, 西川和真, 宍戸建元, 長野恭介, 蓬畑旺周, 稲垣武, 村井源, 迎山和司, 田柳恵美子, 平田圭二, 角薫, 松原仁 : 「公立はこだて未来大学 プロジェクト報告書(最終)」, クリエイティブ A.I., 2019.
- [2] 清水柚里奈, 菅野沙也, 伊藤貴之, 嵯峨山茂樹, 高塚正浩 : 「動画特徴量からの印象推定に基づくBGMの自動素材選出」, 情報処理学会研究報告, Vol. 2017-MUS-114, No. 17, 2017.
- [3] 小野佑大, 石先広海, 帆足啓一郎, 小野智弘, 甲藤二郎 : 「音楽のムード分類結果を利用したホームビデオへのBGM付与システム」, 情報処理学会研究報告, Vol. 2011-MUS-89, No. 16, 2011.
- [4] ドラゴンクエスト, 株式会社スクウェア・エニックス (1986).
- [5] FINAL FANTASY , 株式会社スクウェア・エニックス (1987).
- [6] 石塚 賢吉, 鬼沢 武久 : 「物語のシーンの印象に基づいたテーマ音楽の変奏曲の生成」, 第22回ファジィシステムシンポジウム講演論文集, pp. 587-592, 2006.
- [7] 小林右京, 伊藤淳子, 宗森純 : 「人気ロールプレイングゲームのシーンを基とした プロット作成システムの提案」, 情報処理学会シンポジウムシリーズ, Vol. 2014, pp. 1447-1455, 2014.
- [8] Hevner, K : Experimental studies of the elements of expression in music, American Journal of Psychology, 48, pp. 246-268 , 1936.
- [9] 大串健吾 : 「音楽と感情」, バイオメカニズム学会誌, Vol. 30, No. 1, pp. 3-7, 2006.
- [10] Shun Sawada, Satoru Hukayama, Masataka Goto, Keiji Hirata: Transdrums: A Drum Pattern Transfer System Preserving Global Pattern Structure , ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom, 2019, pp. 391-395, 2019.
- [11] Brian McFee, Colin Raffel, Dawen Liang, Daniel P. W. Ellis, Matt McVicar, Eric Battenberg, Oriol Nieto: librosa: Audio and music signal analysis in python: PROC. OF THE 14th PYTHON IN SCIENCE CONF. (SCIPY2015), 2015.
- [12] 吉田幸平 : 「自然な音声合成実現に向けた音響的特徴の分析」, 日本工業大学研究報, Vol. 46, No. 1, pp. 129-132, 20016.
- [13] Realtime chord recognition of musical sound: A system using common lisp music. ICMC 1999, pp. 464-467, 1999.