

アニメーション制作のための背景音楽選定支援システム*

小野田裕 (法政大学情報科学部), 伊藤克亘 (法政大学情報科学部)

1 まえがき

映画やテレビのような、映像メディアは、映像だけでは成り立たない。ほとんど音を伴っている。そして、映像に加えられるのは、映像に表現された対象から発せられる音だけではない。特殊な効果音や音楽は、映像が視聴者に与える印象を高めるために用いられる。それらは、映像の一部と言ってもいいほど、映像表現の重要な側面を担っている [1]。

先行研究 [2] では、制作者により互いに調和するように組み合わせられた音楽と映像の印象と、調和を崩した音楽と映像の印象を SD 法によって被験者に評価させている。その結果、調和した音楽と映像は、調和を崩した音楽と映像と比較して、音楽と映像と組み合わせることにより音楽と映像の評価が上昇する効果大きいことを見いだしている。また、先行研究 [3] では、「短調の音楽がスノーマンを人さらいにする」という章において、アニメーション作品「スノーマン」にて主人公の少年とスノーマン（雪だるま）が空を飛び回るシーンに着目し、長調の BGM を短調の BGM に置き換えた。そうする事により、楽しそうに一緒に飛び回っているシーンから一転、スノーマンが少年をさらっていったかのように見えるという調査結果を示している。このように、あるシーンにあてる BGM 選定を誤ると、与えたいものとは違った印象を視聴者に与えてしまう。

そこで、本研究では BGM 選定に不安を覚えるアニメーション制作の初心者がシーンの状況にあった BGM を検索し、適切な BGM をあてる補助を可能にするシステム完成を目指す。そのために、アニメーション作品の BGM と、BGM の使われているシーンの印象評価をし類似度を求め、BGM とシーンがどれくらいマッチしているかを調査した。

2 楽曲検索システムに関する従来研究

現在、音楽を検索するシステムについて様々なアプローチで研究がされている。

楽曲の音響特徴量に着目している研究としては、先行研究 [4] において、楽曲の旋律の音高推移に着目し、音高推理特徴量と感情価との関係を分析している。その結果、音高推移特徴量とは高揚・抑鬱の感情価との間に相関関係があるという結果が得られている。また、先行研究 [5] では、リズムパターンに着目し、人間が感じる感情反応の分析を行っている。曲調変化の感知者数と主旋律のリズムパターンのマッチングに強い相関があり、リズムパターンの変化が曲調の変化に重要な要因であることを突き止めている。また、感性に基づ

いた着目して楽曲検索法に関する研究も行われている。先行研究 [6] においては、画像に見合う曲、曲に見合う画像の異種メディア相互検索の研究が行われている。

以上の従来研究では、主に音楽をクラシックに絞り込んで研究を行い、特徴を明らかにしたり、検索システムの構築をしたりしているが、アニメーションの BGM に対する調査はなされていない。アニメーション作品では、BGM の印象によって真逆の印象を受けてしまうことがあるという事例が先行研究 [3] で示されている。なので次章から、アニメーション作品に使用されている BGM に関する印象について調査を進め、BGM 選定支援システムについて記述する。

3 BGM 選定支援システム

本研究では、印象語対を評価し、それを印象得点として距離を求め、その距離の小さいものを探索することで各シーンにあった BGM を検索する。

3.1 実験材料

実験に使用するものは、BGM と、実際にある印象を与えようとして作られたアニメーションのワンシーンである。BGM 収集に関しては、アニメーション作品のオリジナルサウンドトラックに収録されているものから、オープニングやエンディング、アイキャッチ以外で使用されている曲の中から、歌が入っていない楽曲をサンプルとした。6 作品のオリジナルサウンドトラックから 194 曲の評価用音源を収集した。なお、楽曲の平均時間は 1 分 55 秒であった。アニメーションのシーンは、主観評価により、違う印象を受けたシーンを 10 シーン収集した。なお、収集したアニメーションのシーンに使われている BGM はサンプルとして収集していない。

3.2 実験方法

映像と音の相互検索に関する論文 [5] を参考に印象語対を決定、表 2 に示す。これらの印象語対について、SD 法を用い、-2~2 の 5 段階評価を行い、そのデータの評価得点とする。なお、複数人から印象調査を行った得点に関しては、平均をとってそれを評価得点とした。

4 検索システム

4.1 検索手法

前節で評価した 194 曲の楽曲と 10 シーンを対象として、類似の印象を持つデータの検索システムを作成する。検索システムでは、検索したデータの評価得点をもとにする。あるデータが与えられたとき、その i 番目の評価得点を tm_i とし、検索対象のデータの i 番目の評価得点を m_i とし、次式により距離を求める。

* Background music selection support system for animation making by Yutaka Onoda. (Hosei Univ.) et. al.

表 1. 評価に使用する印象語対

美しい	醜い
自然な	不自然な
潤いのある	渴いた
うれしい	悲しい
暖かい	冷たい
明るい	暗い
緊張した	ゆったりした
単純な	複雑な
強い	弱い
大きい	小さい
大胆な	繊細な
重い	軽い
動的な	静的な
澄んだ	濁った
かたい	柔らかい
新鮮な	古くさい

$$\sqrt{\sum_{i=1}^n (tm_i - m_i)^2} \quad (n = \text{印象語対数})$$

この式により求めた距離の小さいもの（類似度が高いもの）を検索結果とし、上位5曲を検索結果とした。

4.2 類似度調査

上記の手法により導きだされた5曲のBGMを、各サンプルシーンとともに流したものを視聴してもらい、シーンとBGMがマッチしているかどうかを-2~0~2の5段階（-2: マッチしていない, -1: どちらかというマッチしていない, 0: どちらでもない, 1: どちらかというマッチしている, 2: マッチしている）で評価してもらおう。なお、見てもらう前に、そのシーンにいたるまでの経緯を説明する。そうする事によって、そのシーンの情報をまったく知らないより、アニメーション制作者の立場に近づくと考えたからである。評価は、映像の音声は流さず、無音の映像と検索結果のBGMを同時に流す、という形式で行った。これは、映像に元から入っているBGMから感じる印象をカットし、純粋に映像から受けた印象評価を求めたからである。現段階で、2人に評価を求めたところ、54%の組み合わせでシーンとBGMが適合しているという結果を得られた。また、各シーンで一番類似度の高かった曲のみの評価は50%となっていた。

4.3 考察

結果を受けて、検索結果がマッチしやすいシーンとそうでないシーンがあることがわかった。シリアスな、緊張感のある会話をしているシーンでは、シーンとBGMがマッチしているという回答が100%だった。そこで流れ、マッチしたと評価されたBGMにはやはり「冷たい」「暗い」「緊張した」「重い」「静的な」と、ネガティブな意味を持つ印象語に評価が集中した。しか

し、反対に、先輩の昔の思い出を語るシーンでは、マッチしているという回答が10%と、結果が良くない。この原因は、本当は良い思い出を語る暖かいシーンなので、映像への評価は「暖かい」「美しい」「うれしい」と、ポジティブな印象語に評価が集中するだろうと予測したのだが、暗い曲ばかりが検索結果として出てしまった。これは、映像の色がセピアになるという、回想シーン特有の手法によって全体の色が暗くなっていたこと、また、映像の中で雨が降っていたことにより、そのシーンだけ見たときでは暗い印象のほうが強く出てしまったのが原因と推測される。前後の物語の流れをしっかりと伝えなくてはならない。

類似度が一番高いはずの楽曲であるにもかかわらず、全体の評価と比較してマッチしているという回答が少ない。そのうえ、まったくマッチしていないと評価される組み合わせがあった。今後サンプル数を増やしていくことで、主観評価による誤差ならば、解決すると考える。

5 あとがき

SD法を用いて映像と楽曲の印象評価得点をそれぞれ求めた。その得点を用いて、シーンに適合するBGMの検索を行った。満足度調査によって導出された、適合率は、全体で54%となっている。印象調査中、シーンに対する評価を付ける印象語対の中で【潤いのある 渴いた】の印象語対において得点が付けづらいという意見があった。実際にデータを見ても、サンプルの10シーン中、上記の印象語対に値がついたのは3シーンのみとなっている。各データへの評価用印象語対の見直しが必要がある。また、今回は評価値の平均値を利用して検討を行ったのだが、感性には個人差がある。その個人差に対する対応として、個人適応に関する検討が報告されている。感性の個人差に対する対応が、今後の課題である。

参考文献

- [1] 岩宮, "映像作品における視聴覚コミュニケーション", 電子情報通信学会技術研究報告, vol.102, No.533, pp.39-46, 2002
- [2] 岩宮, "オーディオ・ヴィジュアル・メディアによる音楽聴取行動における視覚と聴覚の相互作用", 日本音響学会, vol.48, No.3, pp146-153, 1992
- [3] 堀野 他, "旋律の音高変化に基づく楽曲特徴量についての検討", 情報処理学会研究報告. [音楽情報科学] 2001(45), 29-34, 2001-05-23
- [4] 川野邊 他, "楽曲により喚起される感情反応とRhythmに基づいた曲調変化との関係", 情報処理学会研究報告. [音楽情報科学] 2001(103), 27-34, 2001-10-26
- [5] 宝珍 他, "印象に基づく映像と音楽の相互検索に関する一考察", 情報処理学会研究報告. データベース・システム研究会報告 2002(41), 97-104, 2002-05-21
- [6] 川崎 他, "動画映像から受ける印象の因子分析と映像再生速度の各因子に与える影響" 電子情報通信学会論文誌, A, 基礎・境界 J85-A(9), 1022-1025, 2002-09-01
- [7] 岩宮, "よくわかる最新音響の基本と仕組み" 秀和システム, 2007年, 93-96頁