

BGM 自動生成システムのためのシーンの分割手法

5S-4

濱地 真由美 山下 真史 甲斐 郷子 中村 順一

九州工業大学 情報工学部

1 はじめに

これまで筆者らは、BGMの生成に要する時間やコストの削減を目的として、コンピュータアニメーション等の映像中の各場面の情感を基に、音楽を自動生成するBGM自動生成システムを作成してきた[1]。このシステムにおいて、BGMの付加は、手作業により映像をいくつかのシーンに分割して行なっている。そこで、物語からシーンの時間が得られると、映像に対するBGMの付加がより容易になる。本稿では、映像のストーリーとBGMとの関係を参考にしたシーンの分割について述べる。

2 BGM自動生成システムの概要

システムの入力は、モチーフファイル、場面情報ファイル、動作情報ファイルである。モチーフファイルには、BGM生成に必要なモチーフを記述する。ここで、モチーフとは1,2小節程度のメロディのことを指す。また、場面ファイルには、各シーン毎に、使用するモチーフ名、情感、時間、音色を記述し、動作ファイルには、効果音を必要とする動作に関する情報を記述する。システムはこれらのファイルを入力とし、BGM、および効果音の曲データ形式のファイルを出力する。

入力ファイルにおける情報は、現在すべてユーザが与えている。特に各シーンの時間は、実際にユーザが映像を見て、どこでシーンを区切るかを決定して計っているために、手間がかかる。この手間を省

くためにシーンの分割手法を考察する。

3 映画の構造

映像とBGMとの関係の調査における対象として映画を考えた。しかし、実際に映画の映像とBGMとの関係を調査する前に、映画がどのような構造をしているか知る必要がある。そこで、文献[2]を調べた結果、次のことが得られた。

映画は木構造になっている。映画は複数のシークウェンスからなっている。また、シークウェンスは複数、または単一のシーンからなっており、シーンは複数、または単一のショットからなっている。ここでショットとは、1台のカメラによって連続して撮影される映像のことを指す。

4 調査

文献[2]より得られた映画の構造を基にして、実際の映画を調査した。調査対象には、「スターウォーズ」(1977)を選択した。

また、調査は以下の方法で行なった。まず、クライマックスシーン20分を抜き出し、それを筆者らの判断でショットに分割した。分割した各ショットとBGMにどのような関係があるかを調査した。

調査を行なった結果、BGMの変化とショットの変化はほぼ同時であり、BGMが変化する箇所多くは以下のようにショットが変化する時であった。

- 場所が変化する時：場所の変化に伴い、音量が変化する。
- 会話が始まる時、または終る時：会話の始まりには音量は小さくなり、終わりには音量が大きくなる。
- 新たな視点(登場人物)が導入される時：新たな視点の導入に伴い、音量が変化する。

A Method to Divide Story into Scenes for Automatic BGM Generation

Mayumi HAMACHI, Masafumi YAMASHITA, Kyoko KAI, Jun-ichi NAKAMURA

Kyushu Institute of Technology

680-4 Kawazu, Iizuka, Fukuoka 820, JAPAN

- 中心人物に心理的変化が起こる時: 悲しくなる等の心理的変化が起った時, 曲想が変化する.
- 大きな出来事が発生する時: 爆発等の大きな出来事が発生した時に曲想が変化する.

なお, このようなショットの変化が起こった場合に, 必ずしも BGM が変化するわけではない. 例えば, 場所が変わったからといって, 常に音量が変化するわけではなく, BGM の変化が起こらない場合もあった.

また, BGM の変化には, ショットの変化と同時に起こらない場合があった. BGM の変化が, ショットの変化より僅かに遅れる場合は, 余韻を残しながら次へ進ませる時である. また, BGM の変化がショットの変化より先に, 徐々に変化していく場合は, BGM によって次の状況を導かせようとする時であった.

5 アルゴリズム

以上の結果を考慮して, 映像をシーンに分割するアルゴリズムを考えた. 映像をショットに分割したものを入力とし, シーンの分割箇所を出力する.

具体的には, 各ショット毎に以下のものを記述したファイルを入力とする.

- 場所のラベル
- 会話の有無
- 発生する出来事について
 - 度合 (5 段階)
 - ショット開始からの時間
- 心理変化について
 - 人物
 - 情感名とその度合 (5 段階)
 - ショット開始からの時間
- ショットの継続時間

シーンの分割は, まずシーンの切れ目となる可能性のある箇所を取り出す. シーンの切れ目となる可能性のある箇所は, (1) ショットの終り, (2) 出来事が発生する時, (3) 心理状態が変化する時である. 取り出した箇所について, 入力情報を基に, 以下の規則により点数を付加していく.

- 場所の変化: 距離により付加
- 会話の有無の変化: 会話の始まり, 終り, 継続時に付加
- 出来事の発生: その度合により付加
- 心理状態の変化: 前後の情感と度合の変化により付加
- シーン開始からの時間: シーンの時間が極端に短い場合は BGM の付加は困難, 長い場合は BGM の変化が乏しくなる

映像の始めから順に, 以上のような場合に点数を付加した結果, 一番点数の高い箇所をシーンの区切れとする. この後も同様に, 区切った箇所をシーンの始まりとして点数を付加してシーンに分割していく. このようにして分割していった場合, 最後のシーンの時間が, 極端に短くなることがある. その場合は, 1つ前のシーンの始まりから映像の最後までを最後のシーンとする.

6 おわりに

映画「スターウォーズ」における BGM とショットの関係を調べ, その結果より, アルゴリズムを考察した. 今後はインプリメントを行ない, 適切な点数付けの方法を決定する予定である.

参考文献

[1] Nakamura, J., et al: Automatic Background Music Generation based on Actors' Mood and Motions, The Journal of Visualization and Computer Animation, vol.5 pp.247-264 (1994).

[2] Daniel Arijon: 映画の文法,(1980).