

1R-7

場面の情感を考慮した BGMの自動生成システム

中村 順一, 加来 徹也, 乃万 司, 吉田 将

九州工業大学 情報工学部

1 はじめに

コンピュータ・アニメーションなどにより生成された動画に対して、BGMと効果音を付加することは、各場面の情感や登場人物の動作を強調するために有効である。そこで、BGMと効果音を自動的に生成するシステムを開発した。生成システムへの入力は、(1)各場面の長さ、(2)その場面の情感(楽しい、悲しいなど)、(3)基本となるメロディ、(4)登場人物の動作を「場面変化の時刻表」として指定する。演奏はMIDIインタフェースのシンセサイザにより行う。場面変化にあわせて時間調整を行ったBGMを人手で作成することは、短い動画であっても容易ではない。そこで本システムによる自動生成はコンピュータ・アニメーションの総合的な質を向上するのに有効である[1]。

2 システムの概要

システムは、図1に示すように、Unix WS, Macintosh, シンセサイザで構成した。WS上のPrologプログラムによりBGMと効果音を生成し、これをMIDI形式に変換してMacintoshに送る。そのデータを用いてMacintosh上の市販プログラムであるPerformerにより、シンセサイザを演奏する。

映像に対するBGMおよび効果音の付加は以下のように行う。

1. 映像を見て、場面の列に分割し、時間をストップウォッチなどで測定する。主として映像が大きく変化する部分で分割すればよい。対象の映

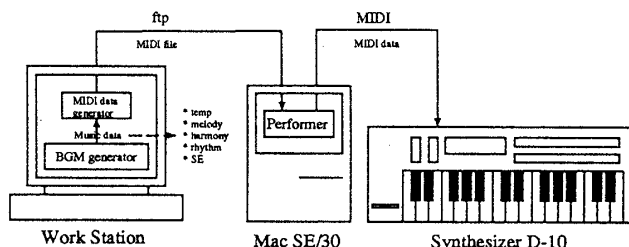


図1: システムの構成

```
[1, % 場面番号
[[a,1],[a,2],[b,1],[a,2]], % モチーフ列
[glad,1], 16.9, % 情感タイプと継続時間
[m(acou_piano_1), % 音色
h(harpsi_1), r(elec_bass_1)]]
```

図2: 各場面に対するBGM用の指定の例

像がコンピュータ・アニメーションであれば、自動化可能である。

2. 各場面に対して以下のものを決定する。

情感タイプ: 嬉しい, 楽しい, 眠たい, 悲しい, 腹立たしい, 疲れた, の6種類から選択し, 強弱を1-5の数値に決定する。

モチーフ列: その場面にあわせたモチーフ(1小節の旋律の断片で指定)と2-8個のモチーフの列を決定する。なお音色も同時に指定する。情感タイプとモチーフ列の指定の例を図2に示す。これは4つのモチーフの列であることを指定している。[a,1]等はモチーフのIDで, 実際のモチーフは別途定義する。また情感は「少し嬉しい」で, 継続時間は16.9秒である。

動作データ: 登場人物の動作の内, 効果音を付

⁰Automatic Background Music Generation

Based on Emotional Situation in a Scene

Jun-ichi NAKAMURA, Tetsuya KAKU, Tsukasa NOMA,
Sho YOSHIDA

Department of Artificial Intelligence
Kyushu Institute of Technology

```
[1,                               % 場面番号
 [usagi, fall, 3,                 % 登場人物, 動作と程度
 5.6, 6.2]]                       % 開始時間と終了時間
```

図 3: 動作データの例

加する部分を決定する。これは、登場人物の ID, 動作名 (例: 殴る, 落ちる, 驚く, まばたきする), その程度 (1-5), 開始時間および終了時間, により指定する。指定の例を図 3 に示す。これは兎が中程度の速度で落ちる動作が, 5.6-6.2 秒の間継続することを示している。

3. 各場面に対して以上のパラメータを指定し, WS 上で MIDI 形式の BGM を生成する。次にこのファイルを Performer に送り, 演奏する。

3 BGM と効果音の生成

各場面に対して与えられたパラメータを用いて, 以下のようにして BGM を生成する [2].

1. 情感タイプに従って基本となるテンポを決定する。「非常に悲しい」には「非常に遅い」とする。値は実験により決定した。
2. 場面の長さに合わせてテンポを調整する。この場合, 場面の変わり目でテンポが急激に変化しないように, 場面の最後で徐々にテンポを変化させる。
3. テンポと場面の長さから必要な小節数を求め, コード進行規則を用いてコード進行を決定する。コード進行規則は, コードの機能 (tonic, dominant, subdominant) 間の進行に対して, 文献および実験により決定したコストを与え, コストができるだけ小さくなるようにした。
4. 求めたコード進行に従ってモチーフ列を変形し, 必要な長さだけ繰り返して, 旋律を生成する。変形は, コードに合うようにモチーフの音程を変化させるが, その際, 音符に重みを与え, できるだけコードから外れる音を少なくするようにした。

5. 伴奏を生成する。伴奏は, 和声とリズムから構成し, これらは情感タイプをキーとしてデータベースを検索することにより行う。データベースには, 情感タイプで分類した 151 の和声と 163 のリズム・パターンを登録している。伴奏により, 同じ旋律であっても印象を変化させることができる。

6. 情感タイプに従って調と音量を決定する。たとえば, 「悲しい」に対しては, 短調と小音量を選択する。

7. 登場人物の動作に従って効果音を生成する。動作名をキートンとして効果音のパターンをデータベースから検索し, タイミングを合わせる。同時に動作の種類に従って, BGM を中断するようにする。これは効果音をより印象的にするためであり, 実験により決定した。効果音は, 「殴る」のように単独音のもの, 「落ちる」のように連続音のもの, 「まばたく」のように短いメロディのもの, の 3 つに分類した。

4 おわりに

本稿では, 場面の長さや情感に従って BGM と効果音を自動生成するシステムについて述べた。このシステムは, コンピュータ・アニメーションの総合的な質を向上するのに有効である。今後は, 実際にアニメーション生成システムに接続し, データベースの内容の改良を行う予定である。

参考文献

- [1] T.Noma, K.Kai, J.Nakamura, N.Okada: *Translating from Natural Language Story to Computer Animation*, First Singapore International Conference on Intelligent Systems, Raffles Convention City, Singapore (1992).
- [2] Jun-ichi NAKAMURA, Tetsuya KAKU, Tsukasa NOMA, Sho YOSHIDA: *Automatic Background Music Generation based on Actors' Emotion and Motions*, Pacific Graphics '93, Korea (1993).